Title
Electrophilic Nucleosides: Tools to Explore Protein Kinases

Author
Gushwa, Nathan Ness

Publication Date
2009

Peer reviewed|Thesis/dissertation

Electrophilic Nucleosides: Tools to Explore Protein Kinases

by

Nathan Ness Gushwa

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

[redacted]

‘

‘

In dedication:

For D.A.G. – hard work, tenacity, and intelligence

I would like to acknowledge my advisor, Jack Taunton, as well as the following people who have contributed directly or indirectly to the work presented here:

Juan Oses-Prietro, Alma Burlingame, Sumin Kang, Jing Chen, Haralambos Hadjivassiliou, Ville Paavilainen, Jennifer Liu, Jimmy Blair, Raynard Bateman, Chao Zhang, Kevan Shokat, Mari Nishino, Arthur Baca, James McKerrow, Shalini Singh, Todd Miller, Michael Eck, Jonathan Choy, Stephane Gourguechon, and Ching C Wang.

Special thanks to Michael Cohen, Jennifer Garrison, and Stephane Gourguechon for their unending help and tutelage.

Abstract

The protein kinase family represents a significant challenge in medicinal chemistry. This important class of enzymes shares a structurally conserved catalytic core, making selective inhibition of individual members difficult. Covalent inhibition offers the potential of additional selectivity based on chemical reactivity and here I have designed and synthesized a panel of electrophilic nucleosides to investigate select protein kinases. These compounds share a nucleoside scaffold that combines elements of the broad spectrum tyrosine kinase inhibitor pyrazolopyrimidine-1 (PP1) and an electrophilic nucleoside, *p*-fluorosulfonylbenzoyl 5' adenosine. This combination increases the scaffold's reversible binding affinity to target a subset of protein kinases. I tested the effect of several different 5' electrophiles on the activity of Src-kinase family members. A 5'-vinyl sulfonate ester reacted with a cysteine in the "glycine-rich" loop of Src. This cysteine is also found in FGFR tyrosine kinases, but is present in only three of the eleven Src-family kinases (Src, Yes, and Fgr). A pyrrolopyrimidine nucleoside bearing a 5'-vinyl sulfonate ester inhibited autophosphorylation and FGFR3 dependent cell proliferation in a cysteine dependent manner.

Kinases of the eukaryotic parasite *Trypanosoma brucei* were profiled using the electrophilic nucleoside **47**, which targets an essential catalytic lysine found in nearly all protein kinases. This compound binds to a subset of the *T. brucei* kinases, five of which are essential: TbPK50 and homologues of casein kinase 1, Gsk3, NIMA, and Clk protein kinases. I used compound **47** in conjunction with quantitative mass spectrometry as a tool for chemical proteomics. Using this technique, covalently modified *T. brucei* kinases were profiled for binding with the reversible inhibitor PD0166326. Exchanging the 5' electrophile of **47** for an acrylate ester changes the compound's effect on *T. brucei*. This new electrophile still preferentially reacts with protein kinases, but with higher selectivity.

One unique target of the acrylate is the *T. brucei* homologue of Bud32 kinase. This atypical kinase may have important functions in *T. brucei* biology and virulence.

Table of Contents

List of Figures and Schemes

List of Tables

Introduction

Covalent inhibition is an important strategy for drug development and the design of small molecule probes of biological processes. A familiar example of covalent inhibition is the β-lactam antibiotics that block cell-wall synthesis in gram-positive bacteria. Many other drugs similarly form covalent bonds to their targets[1,2]. Covalent inhibitors inactive their target enzyme(s) in a time-dependent manner. Initial association is non-covalent, and in a second rate determining step, the inhibitors react with nearby nucleophilic residues[3,4]. The kinetic and thermodynamic components of this two-step process can be resolved[5,6], but the end result is potential for sub-picomolar potencies and selectivity based on chemical principles[7,8].

Covalent inhibitors have been used to probe the function of individual proteins as well as whole classes of enzymes. Formation of a stable covalent bond allows direct identification of the protein components responsible for a covalent inhibitor's phenotypic effects. In this manner, the natural product, trapoxin, helped to define mammalian histone deacetylases[9]. Covalent interaction can also significantly boost the potency of otherwise specific, yet weakly binding inhibitors. This advantage is exemplified in many protease inhibitors. Appending electrophiles to peptides and peptidomimetics yields covalent mechanism-based inhibitors for several protease families[10]. Broadly targeting reactive classes of reactive proteins can yield tools profile families of enzymes and speed inhibitor development[11]. Covalent probes for hydrolases allowed concurrent optimization of the potency and selectivity of inhibitors of fatty acid amide hydrolases[12].

Protein kinases constitute a large family of enzymes with a central regulatory role in all eukaryotes and many prokaryotic organisms. Protein kinases represent 2% of human genes and share significant structural homology in their catalytic domains[13]. Specific inhibition of members of this family holds the promise of providing therapeutics for cancer and other human diseases, but represents a challenge due to this high structural homology[14,15]. The kinase active site consists of a highly conserved ATP

binding pocket located between the N and C-terminal lobes of the catalytic domain and a poorly defined peptide binding surface[16]. The majority of kinase inhibitors are ATP competitive and must exploit subtle variations in the nucleotide binding site for selectivity among members of this enzyme family[15].

There is a vast literature of natural and synthetic kinase inhibitors, but very few examples of covalent inhibitors. Unlike proteases, kinases form no covalent intermediates with their substrates[16]. Covalent inhibition of protein kinases must rely on residues within the active site that do not function as nucleophiles during catalysis. One such residue is a essential catalytic lysine. In a subset of lipid and protein kinases, this lysine is targeted by wortmannin, an ATP competitive natural product with an electrophilic furan ring[17]. This lysine is a ubiquitous electrophile in protein kinases and Patricelli and coworkers targeted its nucleophilic character to develop a kinase selective tool for proteomics. Quantitative mass spectrometry of covalently bound kinases identified enzymes expressed within different cell lines, and competitive labeling with other reversible inhibitors yielded binding profiles in cellular lysates[18]. Non-catalytically conserved nucleophiles offer an oportunity to boost compounds' selectivity among members of the kinase family. A series of electrophilic 4- anilinoquinazolines inhibit members of the EGFR kinase family by targeting a non-conserved cysteine[19]. The homologous cysteine is the target of an inhibitor of Btk[20]. The most dramatic example of selective covalent inhibition is the Rsk kinase inhibitor fmk. This compound combines two selectivity elements to target a single kinase with nanomolar affinity[21]. These compounds serve as a precedent for covalent inhibition of protein kinases and illustrate the advantages that it offers.

In this document we detail our design and testing of a series of electrophilic nucleosides to target select protein kinases. In Chapter one, I describe the design of a nucleoside scaffold, which binds the kinase active site and test different 5' electrophiles

for their effect on enzyme inhibition. In Chapter two, I show how one of these electrophiles targets a previously unexploited non-conserved cysteine in the active site of Src kinase. This nucleophile-electrophile pairing is recapitulated in the FGFR kinase family which contains an identical cysteine. In Chapter three we use electrophilic nucleosides to investigate kinases in the parasite *Trypanosoma brucei* for potential drug targets. Using quantitative mass spectrometry, we identify kinases bound by the nucleoside scaffold with two different 5' electrophiles and apply this technology to profile binding of a reversible inhibitor. This research highlights the use of covalent inhibition as a tool for selectivity in protein kinase binding as well as for proteomic investigation of this enzyme family.

References:

1) Yocum RR, Rasmussen JR, Strominger JL. The mechanism of action of penicillin. Penicillin acylates the active site of Bacillus stearothermophilus D-alanine carboxypeptidase. J Biol Chem. 1980, 10: 3977-3986.

2) Potashman MH, Duggan ME. Covalent modifiers: an orthogonal approach to drug design. J Med Chem. 2009, 52: 1231-1246.

3) Marangoni AG. Enzyme kinetics: a modern approach. John Wiley & Sons Inc, New Jersey 2003.

4) Nishiyama Y, Mitsuda Y, Taguchi H, Planque S, Hara M, Karle S, Hanson CV, Uda T, Paul S. Broadly distributed nucleophilic reactivity of proteins coordinated with specific ligand binding activity. J Mol Recognit. 2005, 18: 295-306.

5) Maurer TS, Fung H. Comparison of methods for analyzing kinetic data from mechanism-based enzyme inactivation: application to nitric oxide synthase. AAPS Pharmsci. 2000, 2: E8.

6) Maurer TS, Tabrizi-Fard MA, Fung H. Impact of mechanism-based enzyme inactivation on inhibitor potency: implications for rational drug discovery. J Pharm Sci. 2000, 89: 1404-1411.

7) Smith AJT, Zhang X, Leach AG, Houk KN. Beyond picomolar affinities: quantitative aspects of noncovalent and covalent binding of drugs to proteins. J Med Chem. 2009, 52: 225-33.

8) Cohen MS, Zhang C, Shokat KM, Taunton J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. Science 2005, 308: 1318-1321.

9) Taunton J, Hassig CA, Schreiber SL. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. Science 1996, 272: 408-411.

10) Powers JC, Asgian JL, Ekici OD, James KE. Irreversible inhibitors of serine, cysteine, and threonine proteases. Chem Rev. 2002, 102: 4639-4750.

11) Jessani N, Young JA, Diaz SL, Patricelli MP, Varki A, Cravatt BF. Class assignment of sequence-unrelated members of enzyme superfamilies by activity-based protein profiling. Angew Chem Int Ed. 2005, 44: 2400-2403.

12) Leung D, Du W, Hardouin C, Cheng H, Hwang I, Cravatt BF, Boger DL. Discovery of an exceptionally potent and selective class of fatty acid amide hydrolase inhibitors enlisting proteome-wide selectivity screening: concurrent optimization of enzyme inhibitor potency and selectivity. Bioorg Med Chem Let. 2005, 15: 1423-1428.

13) Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science 2002, 298: 1912-1934.

14) Zhang J, Yang PL, Gray NS. Targeting cancer with small molecule kinase inhibitors. Nature Reviews Cancer 2009, 9: 28-39.

15) Knight ZA, Shokat KM. Features of selective kinase inhibitors. Chem Biol 2005, 12: 621-637.

16) Adams JA. Kinetic and catalytic mechanisms of protein kinases. Chem Rev. 2001, 101: 2271-2290.

17) Wymann, MP, Bulgarelli-Leva G, Zvelebil MJ, Pirola L, Vanhaesebroeck B, Waterfield MD, Panayotou G. Wortmannin inactivitates phosphoinositide 3-kinase by covalent modification of lys-802, a residue involved in the phosphate transfer reaction. Molecular and Cellular Biology 1996, 16: 1722-1733.

18) Patricelli MP, Szrdenings AK, Liyangage M, Nomanbhoy TK, Wu M, Weissig H, Aban A, Chun D, Tanner S, Kozarich JW. Functional interrogation of the kinome using nucleotide acyl phosphates. Biochemistry 2007, 46: 350-358.

19) Fry DW, Bridges AJ, Denny WA, Doherty A, Greis KD, Hicks JL, Hook KE, Keller PR, Leopold WR, Loo, JA, McNamara DJ, Nelson JM, Sherwood V, Smaill JB, Trumpp-Kallmeyer S, Dorbrusin EM. Specific, irreversible inactivation of the epidermal growth factor receptor and erB2, by a new class of tyrosine kinase inhibitor. Proc Natl Acad Sci USA 1998, 95: 12022-12027.

20) Pan Z, Scheerens H, Li S, Schultz BE, Sprengerler PA, Burrill LC, Mendonca RV, Sweeney MD, Scott KCK, Grothaus PG, Jeffery DA, Spoerke JM, Honigberg LA, Young PR, Dalrymple SA, Palmer JT. Discovery of selective irreversibly inhibitors for burton's tyrosine kinase. ChemMedChem 2007, 2: 58-61.

21) Cohen MS, Zhang C, Shokat KM, Taunton J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. Science 2005, 308: 1318-1321.

Chapter 1: Design and Testing of Electrophilic Nucleosides

Introduction:

The electrophilic compound *p*-fluorosulfonylbenzoyl 5' adenosine (FSBA) is an ATP mimetic that reacts with nucleophilic residues such as lysine[1], tyrosine[2], and cysteine[3] in purine nucleotide binding proteins. When bound to protein kinases, FSBA reacts with an invariant catalytic lysine to form a stable covalent adduct[4]. Reaction with FSBA at lys295 helped to define the catalytic mechanism of the first defined oncogenic gene product, Src tyrosine kinase[5]. More recently, FSBA has found limited use as an activity-based probe for protein kinases and other nucleoside binding proteins[2,5].

In addition to FSBA, the catalytic lysine in the active site of protein kinases is targeted by at least two other electrophilic compounds. A class of pyrrol-5-carboxaldehyde inhibitors of the insulin-like growth factor receptor kinase form a reversible imine with the catalytic lysine, and the natural product wortmannin modifies an analogous lysine in the phosphoinositide lipid kinases as well as the atypical PIKK protein kinase family using an α,β unsaturated ketone[7,8]. The potency of all three electrophilic compounds depends significantly on their reaction with the catalytic lysine.

We sought to develop a suitable scaffold to mount different electrophiles and investigate how changing chemical reactivity can affect a compound's potency against a select kinase *in vitro* and *in vivo*. FSBA's nucleoside core offers an excellent starting point to develop such a scaffold. Nucleoside chemistry is well developed, and different reactive groups can be easily attached at the 5' position similar to FSBA without disturbing many of the elements that govern nucleoside binding in the active site[9]. In spite of these advantages, FSBA also suffers from two distinct disadvantages than must be addressed for use with kinases in complex protein mixtures or *in vivo*. First, most kinases have a high K_m for ATP and therefore weakly bind the adenosine portion of FSBA[10]. The near millimolar concentration of ATP *in vivo* effectively eliminates covalent modification of kinases unless it is also used at millimolar concentrations[6]. Second, at

the high concentrations required to compete effectively with cellular ATP, the fluorosulfone group is likely to react nonspecifically with abundant and irrelevant proteins. Third, FSBA binds and reacts with diverse and more abundant nucleotide binding proteins in the proteome[2,11].

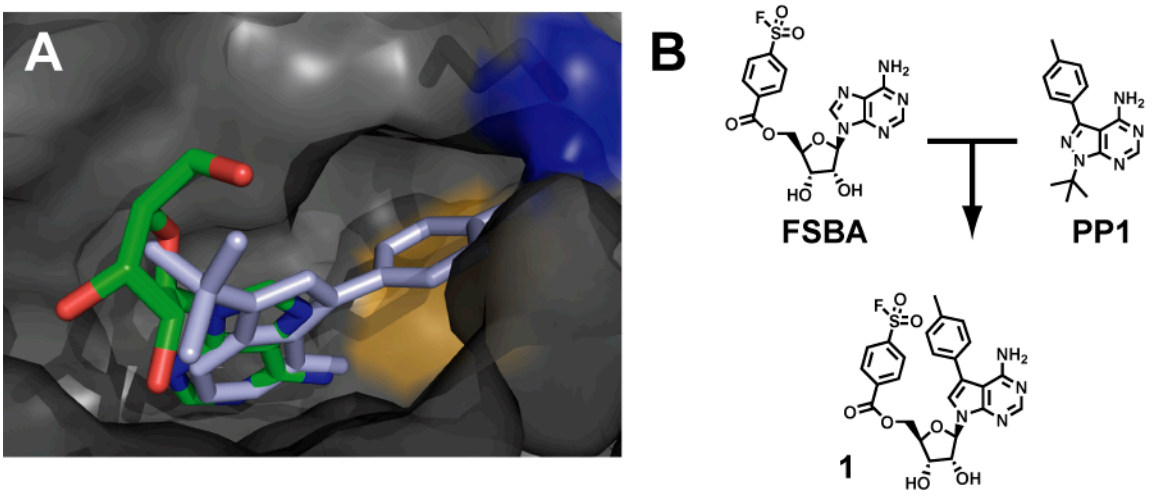


Figure 1-1: Targeting the conserved lysine of Src-family protein kinases with a fluorosulfone electrophile. (A) Superposition of the nucleoside portion of AMP-PNP (colored by atom) and *p*-chlorophenyl pyrazolopyrimidine PP2 (blue) in the active site of Src-family member Lck. The *p*-chlorophenyl substituent of PP2 extends into a hydrophobic pocket, defined in part by the threonine gatekeeper (orange surface). The figure was constructed with PyMOL from PDB files 1QPC and 1QPE. (B) Combination of fluorosulfonyl benzoyl adenosine (FSBA) and PP1 yields compound **1**.

We hypothesize that the disadvantages of FSBA can be addressed by adding an aryl substituent at N7 postion of adenine. Such a compound is expected to have increased reversible binding affinity with at least a subset of the kinome in preference to other nucleotide binding proteins which lack a hydrophobic pocket adjacent to N7. This hypothesis is based on crystal structures of Src-family kinases in complex with ATP analogues and pyrazolopyrimidine PP1 and PP2[13] (Figure 1-1). The *p*-tolyl group of PP1 occupies a hydrophobic pocket at the back of the ATP binding cleft. Access to this pocket is governed by the "gatekeeper" residue. Small amino acid sidechains, like threonine, permit PP1 binding, and larger residues cause a steric clash[12]. Figure 1-1b suggests replacing adenine with a *p*-tolyl pyrrolopyrimidine, analogous to the

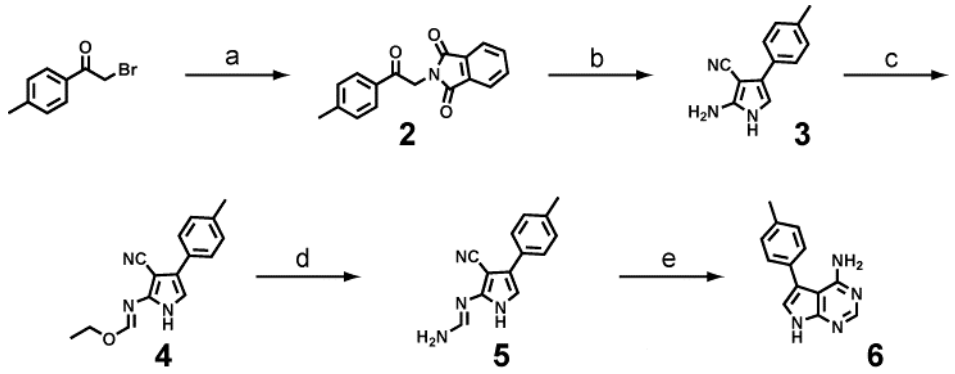
pyrazolopyrimidine of PP1 would preserve critical hydrogen bonds to the hinge region of the kinase as well as the dihedral angle of the nucleoside's glycosidic bond[14,15,16]. The *p*-tolyl ring simultaneously increases a nucleoside scaffold's binding affinity for kinases with a small gatekeeper while eliminating binding to many off-target ATP utilizing enzymes. Additionally the nucleoside structure preserves the location of the 5' electrophile in FSBA. The nucleophilic lysine lies less than 6 Å from the 5' hydroxyl. Figure 1-1b illustrates our concept for combining FSBA and PP1 to produce compound **1**.

In the following chapter, we present data related to the synthesis and testing of compound **1** and related electrophilic nucleosides. I show evidence that the structural hybrid of FSBA and PP1 reversibly binds to Src kinase with much higher affinity than FSBA and retains its ability to covalently modify the kinase *in vitro* and *in vivo*. I present several combinations of a nucleoside scaffolds and electrophiles and highlight their potential for potent and selective inhibition of protein kinases.

Results and Discussion:

Synthesis of a *p*-fluorosulfonylbenzoyl 5' nucleoside, compound 1

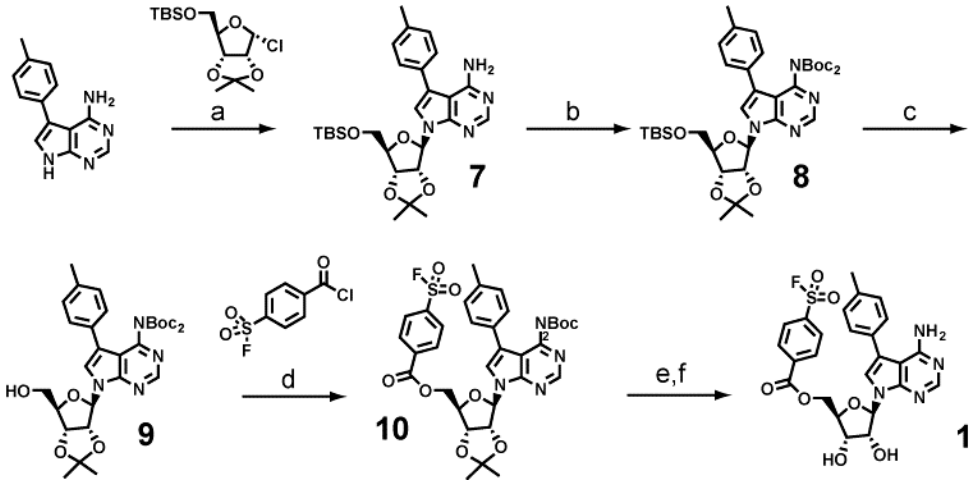
The pyrrolopyrimidine **6** was synthesized in good yield from commercially available 2-bromoacetophenone based on published methods[17]. The route is illustrated in Scheme 1-1. We achieved higher yields of **6** in greater purity if annulation proceeded by treating the isolated amino-imine intermediate **5** with sodium methoxide in methanol rather than via direct conversion from **4** at higher temperatures and concentrations of ammonia. Compound **6** was thereby produced in 61% overall yield with no need for purification before use in subsequent reactions.



Scheme 1-1: Synthesis of pyrolo[2,3-d]pyrimidine **6**. Reagents and conditions: (a) potassium phthalimide, DMF (90%); (b) NaOH, malononitrile, MeOH (93%); (c) triethyl orthoformate, acetic anhydride; (d) NH_3, MeOH (93% - 2 steps); (e) NaOMe, MeOH (78%).

The fully protected nucleoside **8** was prepared in two steps from a chlorinated ribose derivative and pyrrolopyrimidine **6** (scheme 1-2). The glycosidic bond was formed in moderate yield in toluene with potassium hydroxide and phase-transfer catalyst Tris[2-(2methoxyethoxy)ethyl]amine (TDA-1)[18]. Reactions with NaH in DMF led to a mixture of α and β anomers, while use of acetonitrile as solvent yielded a side product linked through N1 of the pyrimidine ring. Following purification, the exocyclic amine was Boc-protected in near quantitative yield to produce compound **8**. From this intermediate we could access the 5' and/or the 2' and 3' hydroxyls by using different deprotection conditions.

The electrophilic fluorosulfone **1**, was obtained by a selective deprotection of the 5' hydroxyl and reaction with the commercially available acid chloride, followed by a final deprotection (Scheme 1-2). The final product was a white amorphous solid that was sparingly soluble in water and most other solvents.



Scheme 1-2: Synthesis of 5'-fluorosulfonyl benzoyl nucleoside **1**. Reagents and conditions: a) KOH, TDA-1, toluene (42%); b) Boc_2O, THF (92%); c)TBAF, THF (95%); d) DIPEA, DCM (46%); e) TFA, THF, H_2O (82%).

Compound 1 is a potent time-dependent inhibitor of Src kinase

Measurements of full length active Src kinase activity after treatment with **1** showed the dramatic effect of incorporating the *p*-tolyl ring on the nucleoside scaffold (Figure 1-2a). Inhibitors were pre-incubated with enzyme and 250 μM ATP before initiating the kinase reactions. With this concentration of ATP competing for the active site, approximately three-fold higher than the K_M[10], FSBA did not inhibit the enzyme, even at concentrations as high as 10 μM. Fitting a curve to data from compound **1** gave a Hill coefficient of 1.3, suggestion that the compound may react with the catalytic lysine. A Hill coefficient greater than one is indicative of an irreversible binder.

Under the conditions of this *in vitro* assay, PP1 was more potent than compound **1**. Its calculated IC_{50} is ~500 nM versus ~1 μM for **1**. This difference is more significant considering the time-dependent nature of covalent inhibition. If compound **1** is forming a

covalent bond, the enzyme activity in this assay represents the losses from formation of an irreversible complex as well as the population that is reversibly bound by the scaffold prior to covalent bond formation. It is always difficult to directly compare reversible and irreversible inhibitors without careful kinetic studies, but we can say with confidence that there is a cost to replacing the *tert*-butyl group of PP1 with a 5' substituted ribose.

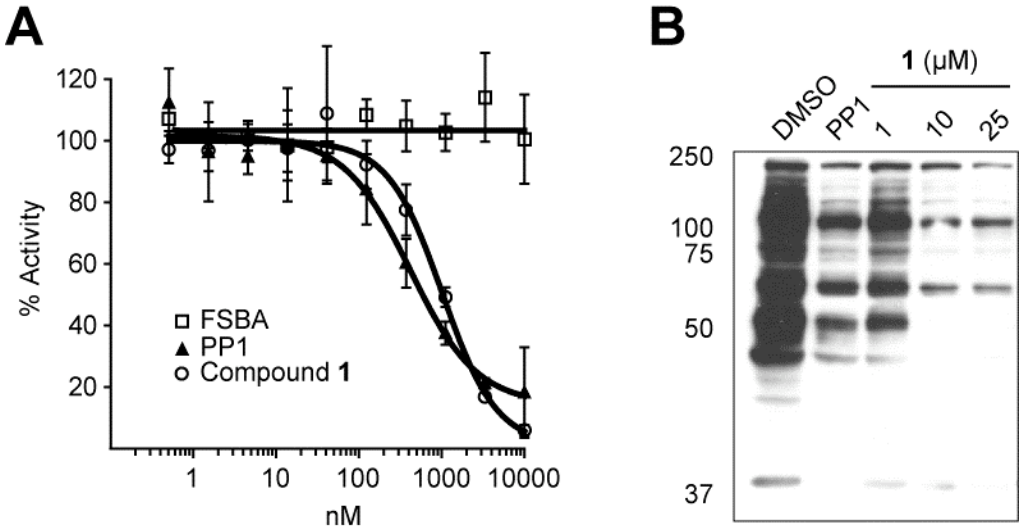


Figure 1-2: Inhibition of Src kinase by the electrophilic nucleoside **1**. (A) Recombinant Src was incubated indicated concentrations of compounds with 250 μM ATP. After 30 minutes the reactions were diluted 1:1 into a solution containing ^{32}P-ATP and substrate peptide and incubated an additional 30 min. ^{32}P incorporation was measured and kinase activity was plotted as a percentage of DMSO control activity. Points are the mean of three independent experiments with error bars representing one standard deviation from the mean. (B) NIH-3T3 cells expressing I338T v-Src were treated with DMSO, 10 μM PP1 or compound **1** for 1.5 hr. Whole cell lysates were resolved by SDS-PAGE and blotted for phospho-tyrosine.

One of our goals in synthesizing compound **1** was to have a scaffold capable of delivering electrophiles to the kinase active site *in vivo*. To test this, we assayed compound **1**'s ability to inhibit tyrosine kinase activity in CZ9 cells. These cells express v-Src with an I338T gatekeeper mutation that permits binding of PP1-like scaffolds[19]. Western blots for total cellular phospho-tyrosine showed a dramatic and dose dependent decrease in cells that had been treated with **1** (Figure 1-2b). Interestingly, compound **1** was more effective than PP1 in this assay. This shift in the relative potency may be due

to the cells' longer exposure to the compounds, 1.5 hr vs 30 min in the *in vitro* assay. If compound **1** is acting as a covalent inhibitor, a longer incubation time means a larger fraction of the total enzyme will be sequestered in irreversible complexes with **1**.

We obtained evidence for time-dependent inhibition by measuring Src kinase activity at different times after addition of compound **1** (Figure 1-3a). Activity drops sharply from a value of ~25% of control at the initial 5 min time point to near background levels after 40 min. This inhibition is not reversible. Src treated with compound **1** was extensively dialyzed but recovered none of its enzymatic activity (Figure 1-3b). Src treated with the scaffold alone (**5OH**, with a free 5' hydroxyl) recovered its full activity after an identical treatment. Time-dependent inhibition and failure to recover activity after dialysis provide strong evidence that compound **1** acts by covalent inhibition. The lack of irreversible inhibition by **5OH** strongly suggests that the fluorosulfone is key to irreversible inhibition, likely by forming a covalent bond to the catalytic lysine as observed with FSBA.

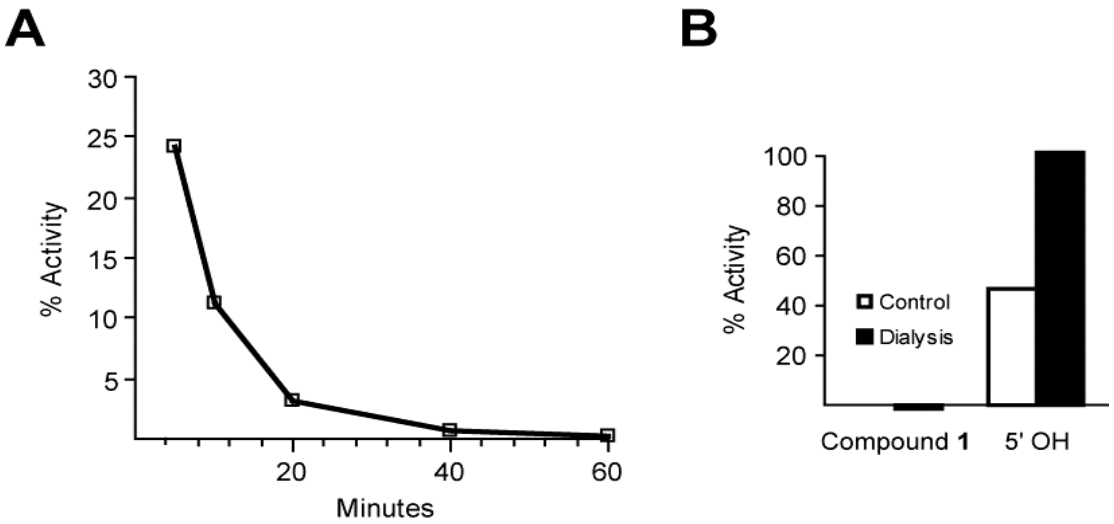
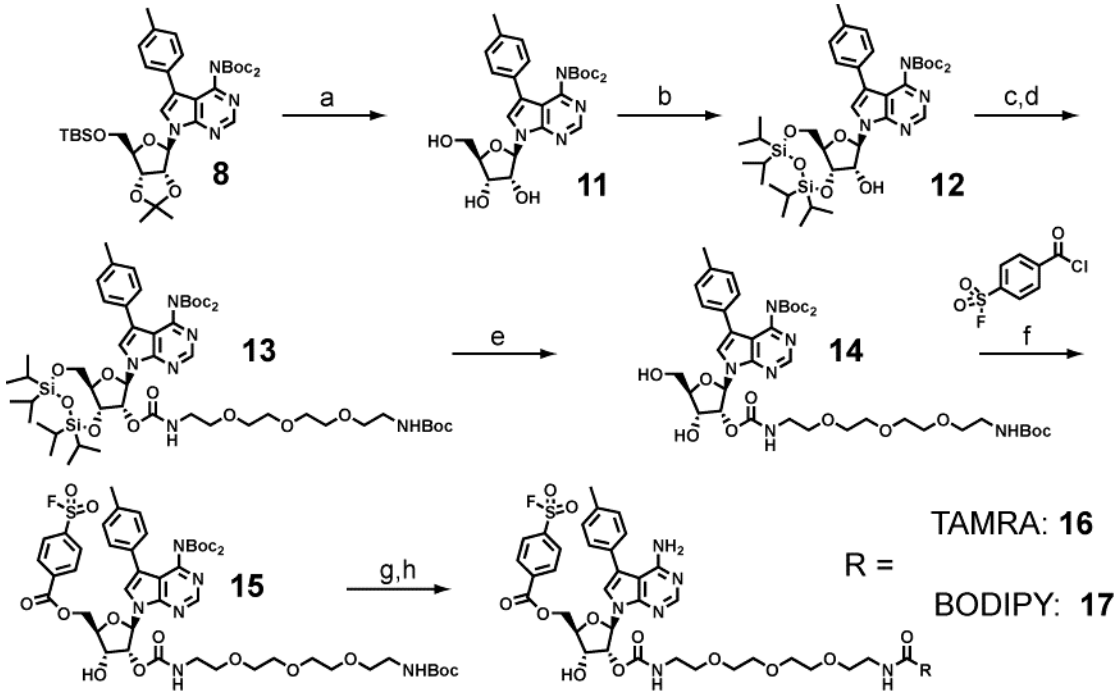


Figure 1-3: The 5' fluorosulfonyl benzoyl nucleoside **1** inhibits Src kinase in a time-dependent irreversible manner. (A) Recombinant Src was incubated ±1 μM **1** with 100 μM ATP for the indicated time then diluted 1:12.5 into a mix of 100 μM ATP, ^{32}P-ATP and substrate peptide. After 32 min ^{32}P incorporation was measured. Values are expressed as a percent of no compound controls. (B) Recombinant Src was treated with **1**, a related nucleoside lacking the 5' electrophile or DMSO. Samples were either stored or dialyzed then diluted 1:5 for reaction as in (A). Values are normalized to DMSO controls.

Fluorescent variants of compound 1 form a specific stable bond to Src

I synthesized two different fluorescent versions of **1** by attaching either a TAMRA (**16)** or BODIPY (**17**) tag to a 2' carbamate-anchored linker (Scheme 1-3). The ribose protecting groups were removed, and a short diamino-PEG was added using carbonyldiimidazole after reprotection of the 3' and 5' hydroxyls as silyl ethers. Attempts to directly attach the linker after removal of the isopropylidene from **10** resulted in rapid formation of a stable 2'-3' carbonate. This necessitated the synthesis of intermediates **12** - **14**. After installing the electrophile and Boc-deprotection, the fluorophores were attached to the primary amine using commercially available NHS esters to make fluorescent fluorsulfones **16** and **17**.



Scheme 1-3: Synthesis of Fluorescent 5'-fluorosulfonyl benzoyl nucleosides **16** & **17**. Reagents and conditions: (a) *aq*HCl, THF (85%); (b) 1,3dichloro-1,1,3,3-tetraisopropyldisiloxane, pyridine (57%), (c) CDI, DIPEA, 1,11-diamino-3,6,9-trioxaundecane, DCM; (d) Boc_2O, DCM (90% - 2 steps); (e) TBAF, THF (90%); (f) DMAP, DIPEA, DCM (21%); (g) TFA, DCM; h) TAMRA-NHS or BODIPY-NHS, DIPEA, DMF (**16**: 41% - 2 steps, **17**: 62% - 2 steps).

TAMRA derivative **16** forms a stable bond to recombinant Src. After incubation of recombinant Src (1μM) was incubated with **16**, a fluorescent band at ~50 kDa was observed (Figure 1-4a). Labeling was completely blocked when Src was pretreated with **1** demonstrating that binding and covalent bond formation is specific. BSA (1 mg/mL) was included in the labeling reaction to stabilize the kinase and can be seen as a minor fluorescent smear just above Src. The labeling of this lysine-rich protein was no competed by excess **1** and was far weaker than Src, despite being present at a 20-fold higher concentration.

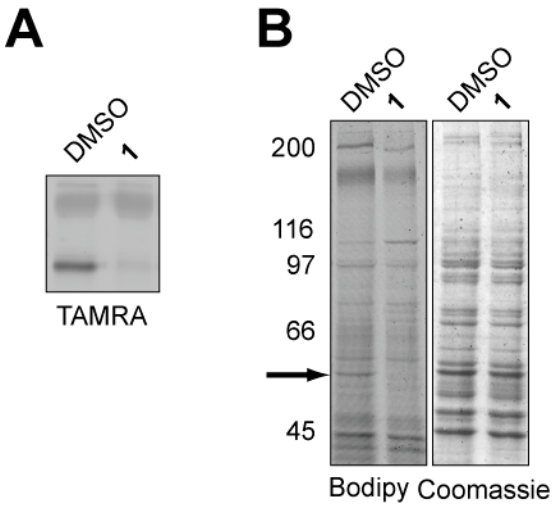
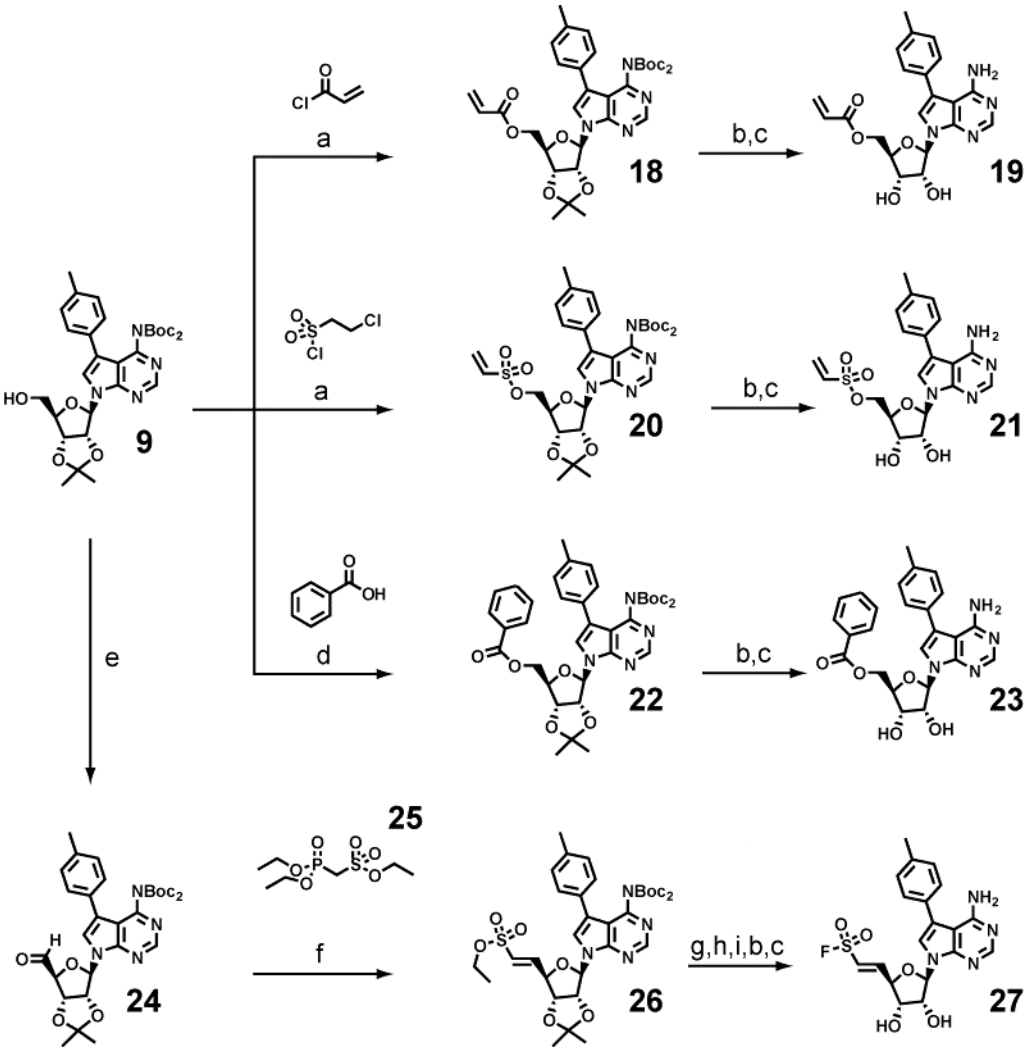


Figure 1-4: The 5' fluorosulfonyl benzoyl nucleoside **1** forms a specific covalent bond to Src kinase. (A) Fluorescent image of recombinant Src labeled with TAMRA tagged compound **16**. The enzyme was pretreated with either compound **1** (5 μM) or DMSO, labeled with **16** then resolved by SDS-PAGE. (B) Fluorescent and Coomassie stained image of lysates from cells treated with compound **17**. NIH-3T3 cells expressing I338T v-Src were incubated ± compound **1** then treated in fresh media with **17**. The membrane fraction of lysates were resolved by SDS-PAGE and scanned. Arrow indicated the expected molecular weight of v-Src.

We next tested whether this specific interaction could be detected *in vivo*. The BODIPY tag of compound **17** is uncharged and membrane permeable making it ideal for cellular studies. CZ9 cells were treated with **1** then **17**, and the membrane fraction was analyzed. These lysates have multiple fluorescent bands, but only one was competed by pretreatment with compound **1** (Figure 1-4b). This band matches the expected molecular weight of the v-Src construct. The labeling experiment and v-Src inhibition by **1** (Figure 1-2b), show that our new nucleoside scaffold is cell permeable and of high enough affinity to deliver the fluorosulfonyl benzoyl electrophile to a cellular kinase target.

Compound **17** is the first FSBA-like activity-based probe capable of acting in cells and competing with the high levels of intercellular ATP[5,6,11,20].



Scheme 1-4: Synthesis of *p*-tolyl pyrolo[2,3-d]pyrimidine nucleosides with different 5' substituents. Reagents and conditions: (a) DIPEA, DCM (52%); (b) TFA, DCM; (c) TFA, H_2O, THF (**19**: 99% - 2 steps, **21**: 96% - 2 steps, **23**: 95% - 2 steps, **27**: 15% - 4 steps); (d) EDCI, DMAP, DCM (71%); (e) EDCI, dichloroacetic acid, DMSO; (f) *n*-BuLi, THF (61% - 2 steps); (g) TBAI, acetone; (h) triphosgene, DMF; (i) CaF_2, KF, MeCN.

5'-Michael acceptors are time-dependent Src inhibitors

A small panel of 5'-electrophiles was synthesized from intermediate **9** to test for other potential lysine targeting groups (Scheme 1-4). The 5'-acrylate and vinylsulfonate esters **19** and **21** were made with the commercially available chlorides, and the non-

electrophilic benzoyl nucleoside **23** was similarly prepared. Vinylsulfonate **26** was produced by a Horner-Wadsworth-Emmons reaction after oxidation of the 5' alcohol. As expected[21], reactions with the TBA sulfonate salt version of **25** gave a mixture of E and Z isomers, but the ethyl sulfonic ester yielded only the trans product. Vinyl sulfonyl fluoride **27** was produced by hydrolysis of the ester and subsequent chlorination and fluorination[22].

The new compounds were tested for their ability to block labeling of recombinant Src byTAMRA fluorosulfone **16**. Compound **16** did not label Src that had been pretreated with the 5'-Michael acceptors **19** and **21** (Figure 1-5a). At the concentration tested (5μM) these compounds were as effective as compound **1**. Compound **27** faied to block labeling and was not investigated further.

Competitive binding experiments provided insight into the binding of compound **1**. Recombinant Src was incubated with 5 μM of various untagged inhibitors then labeled with TAMRA fluorosulfone **16** (5μM). FSBA has no effect on labeling by **16** whereas PP1 inhibited labeling nearly as well as **1**. The fully reversible nucleoside inhibitor **23**, in which a 5'-benzoyl replaces the 5'-fluorosulfone of **1**, had or no effect on labeling by **16**. These results support our interpretation of data in Figure 1-2. The hybrid nucleoside structure of **1** significantly increases binding affinity relative to FSBA, but the scaffold does not have as high of affinity for the active site as PP1. The difference between compounds **1** and **23** in their ability to block labeling suggest a significant portion of compound **1**'s potency is derived from its ability to form a covalent bond.

In vitro assays revealed that vinylsulfonate **21** is a potent Src inhibitor. The compound appeared to titrate the enzyme and was far more potent than compounds **1** or **19** (Figure 1-5b). A time course of inhibition revealed more information about the different electrophiles. All three compounds displayed a time-dependent inhibition of recombinant Src (Figure 1-5c), but the enzyme is fully inactivated by vinylsulfonate **21**

before the first time point. These results could be explained by an extremely tight non-covalent interaction, but given the structural similarity of compounds **19** and **21**, a difference in the rate of covalent bond formation seems more plausible.

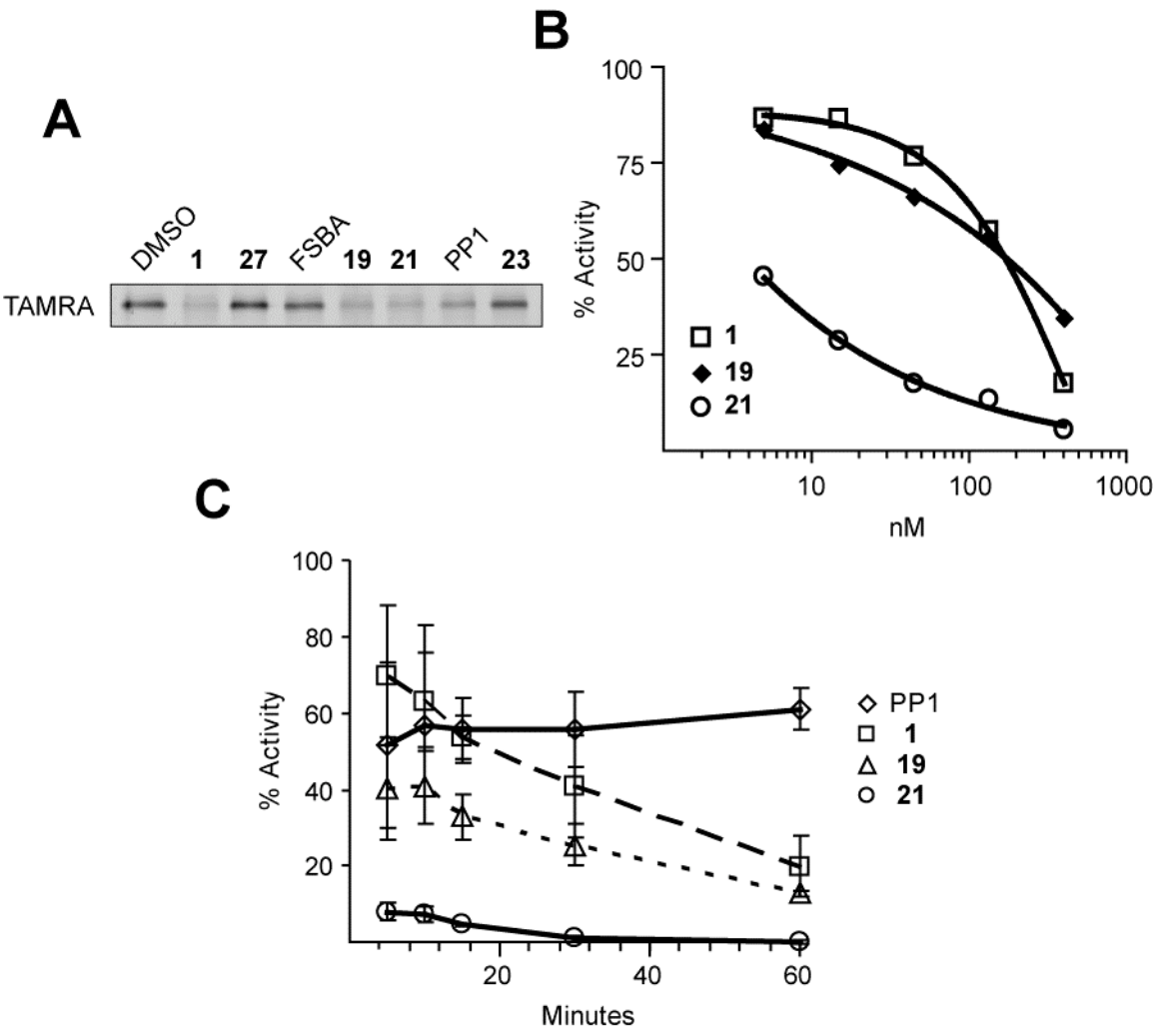


Figure 1-5: Src is inhibited by nucleosides with different 5'-electrophiles. (A) Fluorescent image of recombinant Src labeled with compound **16**. 1 μM enzyme was pretreated with 5 μM untagged inhibitors for 30 min then labeled with 5 μM **16** and resolved by SDS-PAGE. (B) *In vitro* inhibition of Src. Recombinant Src was incubated ± inhibitors for 30 minutes then diluted 1:4 into 125 μM ATP, $^{32}P^{-}$ATP and substrate peptide. After 30 min phosphotrasfer was quantitated. Values are normalized to DMSO controls. (C) Time-dependent inhibition of recombinant Src by new 5' electrophiles. The enzyme was incubated ± 250 nM inhibitors and 100 μM ATP for the indicated time then diluted 1:1 into 100 μM ATP$^{32}P^{-}$ATP and substrate peptide and assayed as in (A). Values are normalized to DMSO controls.

Synthesis of 5-(*p*-phenoxyphenyl)pyrrolopyrimidine nucleosides with 5'-electrophiles

In Src-family kinases a complex network of interdomain interactions mediates auto-inhibition of the kinase domain. Phosphorylation of a C-terminal tyrosine,by Csk allows the Src SH2 domain to bind to this residue at the base of the catalytic domain. This allows the N-terminal SH3 to bind the type II polyproline helix near the catalytic domain's N-terminal lobe[23]. The combination of these interactions causes a change within the active site[14,15,16,24]. ATP can still bind, but catalytic residues are not in the correct position to facilitate phospho-transfer.

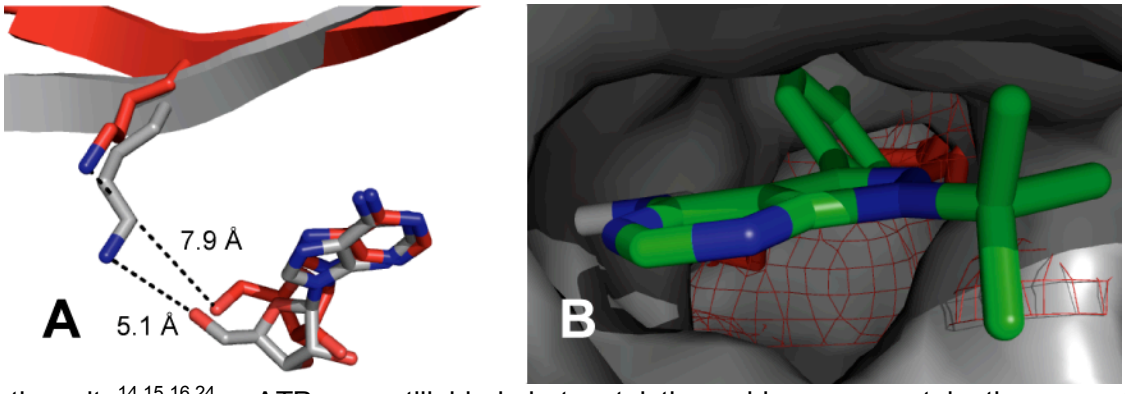
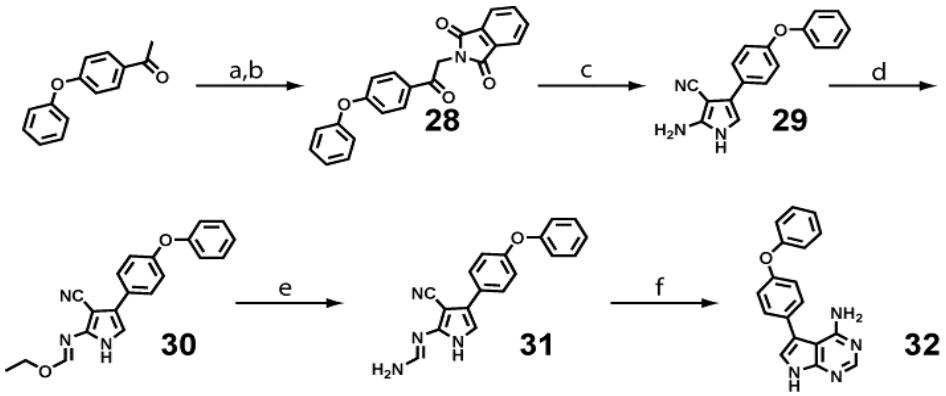


Figure 1-6: Activation induced changes in the ATP binding pocket of Src-family kinases. (A) A superposition of auto-inhibited Src and active Lck. Shown are adenosine portion of the ATP, β-strand 3, and the catalytic lysine (K295-Src). (B) A superposition of active Lck (red mesh) bound to inhibitor PP2 and auto-inhibited Src (gray surface). Activation of Src-family kinases shifts the position of Helix C, contracting the hydrophobic pocket occupied by the *p*-chlorophenyl ring. The figures were constructed with PyMOL from PDB files 2SRC, 1QPC and 1QPE.

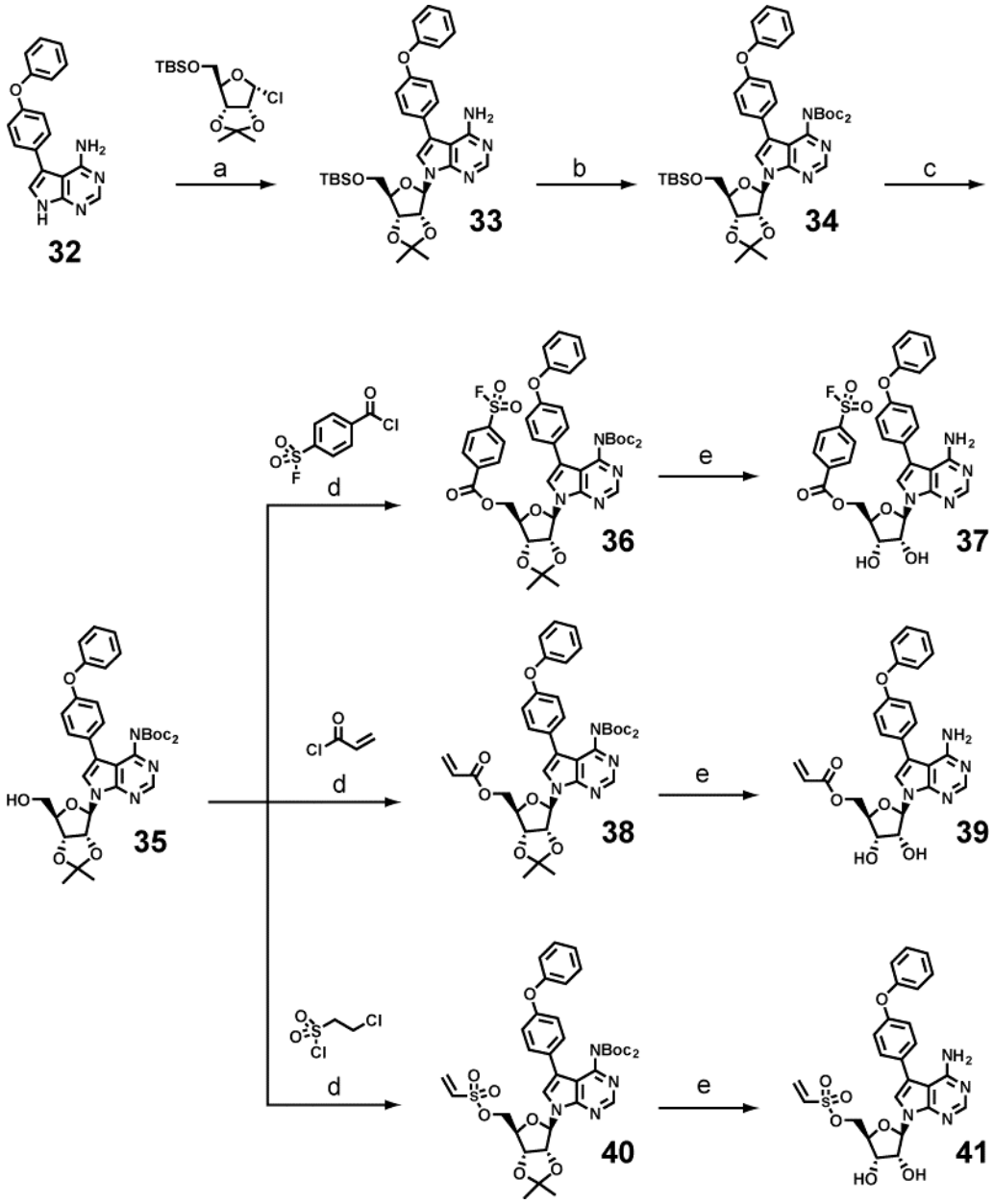
Two activation-induced structural changes in the active site of Src-family members may have implications for binding and bond formation with electrophilic nucleosides like compound **1**. First, the position of the catalytic lysine is slightly shifted (Figure 1-6a). In the inactive structure, K295 switches a hydrogen bonding partners from E310 to the D404[24]. Second, there is a contraction of the hydrophobic pocket accessed by the *p*-tolyl group of PP1 and our nucleoside scaffold (figure 1-6b). The C-helix of the N-terminal lobe forms the back of this cavity and is rolled "out" in the inactive conformation, expanding the pocket. In the Src-family member Lck the hydrophobic

pocket contracts by approximately 20% upon activation. This estimate aided in the design of aryl pyrrolopyrimidine inhibitors that had a 200-fold preference for the down-regulated conformation of Lck[17]. This dramatic result motivated our synthesis of a nucleoside scaffold with and extended 4-phenoxyphenyl substituent at the 5 position of the pyrrolopyrimidine.

We synthesized three new electrophilic nucleosides to test whether the activity state of Src (active vs. auto-inhibited) would influence the binding of these reactive compounds. The *p*-pyrrolopyrimidine **32** was prepared in good yield using similar route to the *p*-tolyl heterocycle with a few minor changes (Scheme 1-5). Omission of acetic anhydride while making **30** increased the purity of the product without decreasing yields. Using ethanol in place of methanol for the pyrimidine annulation had a similar effect, and pure **32** was obtained with an overall yield of 72%. Electrophilic nucleosides **37**, **39**, **41** were prepared in an analogous manner to **1**, **19**, and **21** (Scheme 1-6). Alkylation of the heterocycle was significantly less efficient with the phenoxy species, but all other yields were comparable. TFA at the final deprotection produced an isomer that was difficult to resolve from the desired product. Most likely a 5' to 3' migration of the electrophile, this byproduct was eliminated by using TMS triflate to remove the Boc groups; residual water or workup effected isopropylidene removal.



Scheme 1-5: Synthesis of pyrolo[2,3-d]pyrimidine **32**. Reagents and conditions: (a) $AlCl_3$, Br_2, Et_2O; (b) potassium phthalimide, DMF (90% - 2 steps); (c) *aq*NaOH, malononitrile, MeOH (100%); (d) triethyl orthoformate; (e) NH_3, MeOH; (f) NaOEt, EtOH (80% - 3 steps).



Scheme 1-6: Synthesis of electrophilic *p*-phenoxyphenyl pyrolo[2,3-d]pyrimidine nucleosides. Reagents and conditions: (a) KOH, TDA-1, toluene (18%); (b) Boc_2O, DMAP, THF (82%); (c) TBAF, THF (65%); (d) DIPEA, DCM (**36**: 68%, **38**: 80%, **40**: 82%); (e) TMSOTf, DCM (**37**: 40%, **39**: 56%, **41**: 21%).

5'-lectrophilic nucleosides do not discriminate between active and down-regulated Src

Dose-response curves for compounds **1**, **37**, **39**, and **41** showed that these compounds apparently have no preference for either the active (pY416) or down-regulated (pY527) form of Src (Figure 1-7). It is not surprising that compound **1** inhibits both variants of the kinase equally. Comparing the structures of active Lck and inactive Src, the position of the ε-nitrogen in the catalytic lysine shifts approximately 2.8 Å relative to the 5'-hydroxyl of adenosine (Figure 1-6). Given the apparent low affinity of 5'-benzoyl **23**, it is unlikely that compound **1** initially binds to the active site in its lowest energy conformation. The *p*-phenoxy substituent had no discernible effect on selectivity, but it did substantially improve the potency of **37** relative to **1**. There are many examples of changes to the phenyl ring of PP1-like inhibitors affecting their affinity for Src-family members[17,25,26]. Careful kinetics or lower temperatures may reveal differences in these compounds' selectivity, but this was not investigated further.

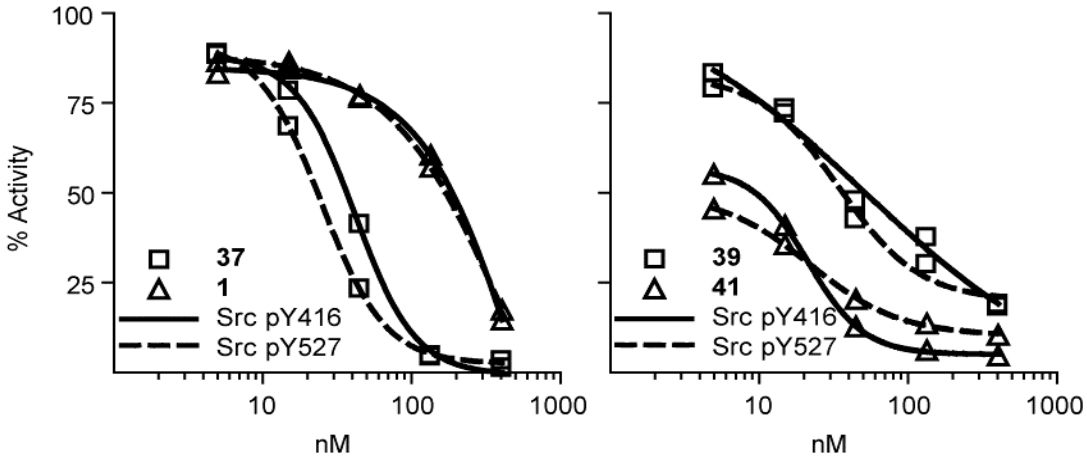


Figure 1-7: 5'-electrophilic nucleosides with a *p*-phenoxyphenyl pyrrolopyrimidine do not distinguish between active (pY416) and down-regulated (pY527) Src. Recombinant Src (5 nM) was incubated ± inhibitors for 30 minutes then diluted 1:4 into 125 μM ATP, $^{32}P^{-}$ATP and substrate peptide. After 30 min phosphotrasfer was quantitated. Values are normalized to DMSO controls. (A) Inhibition of active and down regulated Src by 5'-fluorosulfonyl benzoyl nucleosides with *p*-tolyl (**1**) or *p*-phenoxy phenyl (**37**) substituents. (B) Inhibition of active and down regulated Src by *p*-phenoxyphenyl pyrrolopyrimidine nucleosides with a 5' acrylate (**39**) or vinylsulfonate (**41**).

Compound 21 shows surprising selectivity for Src over Hck

Src-family member, Hck was tested in *in vitro* assays with compounds **1** and **21** under identical conditions to Src (Figure 1-5). There is very little difference in the kinases' sensitivity to compound **1** (Figure 1-8). Literature IC_{50} values for PP1 are 170 nM and 20 nM, respectively, for Src and Hck[13], and their catalytic domain share 70% identity. Given these similarities and the relative levels of inhibition by PP1 and **1**, it was surprising to see such dramatic differences in the enzymes' sensitivity to vinylsulfonate **21**; Values Changed from an IC_{50} of less than 5 nM for Src(equal to the concentration of enzyme in the assay) to ~120 nM for Hck. IC_{50} values for members of this family are usually in the same order of magnitude[27]. Results with 5'-vinylsulfonate **21** strongly suggest that there are important, previously unrecognized differences between Src and Hck. These differences are explored and exploited in Chapter 2.

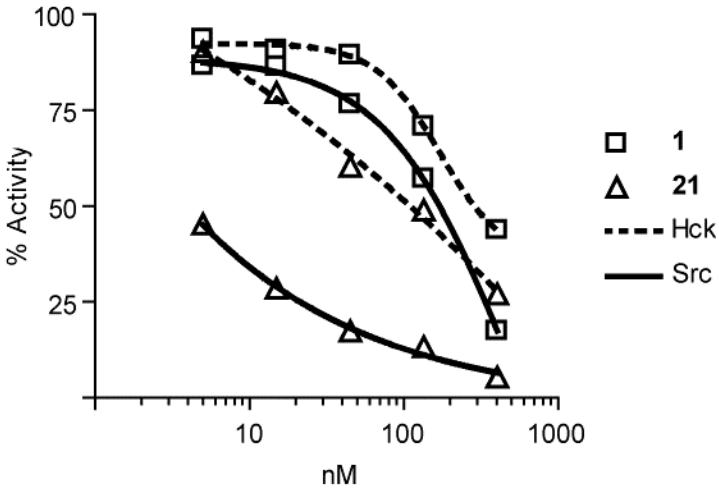


Figure 1-8: Src-family kinases Src and Hck are differentially inhibited by a *p*-tolyl pyrrolopyrimidine nucleoside bearing a 5'-vinylsulfonate (**21**), but not a 5'-fluorosulfonyl benzoyl (**1**) Recombinant Src and Hck (5 nM) were incubated ± inhibitors for 30 minutes then diluted 1:4 into 125 μM ATP, $^{32}P^{-}$ATP and substrate peptide. After 30 min phosphotrasfer was quantitated. Values are normalized to DMSO controls.

Conclusion:

We have developed a new series of electrophilic nucleosides that inhibit Src kinase in a time-dependent, irreversible manner. The core scaffold combines key elements of the Src-family inhibitor PP1 and the covalent inhibitor FSBA. The *p*-tolyl-substituted pyrrolopyrimidine scaffold increases binding affinity relative to the adenosine core of FSBA and effectively delivers the same 5' fluorosulfonyl benzoyl electrophile to the kinase active site. I replaced the *p*-tolyl group of our scaffold with a bulkier p-phenoxy phenyl and the increased the compounds' potency vs. Src, but I saw no preference with regard to Src activation status. I synthesized a relatively unexplored 5'-electrophile, the vinylsulfonate. Interestingly vinylsulfonate **21** inhibits Src ~10 times more potently than its close relative Hck.

Methods:

Synthesis

NMR spectra were recorded on a Varian 400 spectrometer at 400 MHz. Chemical shifts are reported as parts per million from an internal tetramethylsilane standard or solvent references. LC/MS was performed on a Waters AllianceHT LC/MS with a flow rate of 0.2 mL/min using an Xterra MS C18 column (Waters). Dichloromethane, dimethylformamide, methanol, tetrahydrofuran, toluene, and diisopropylethylamine were dried using the solvent purification system manufactured by Glss contour Inc. All other solvents were of ACS chemical grade (Fisher) and used without further purification. Commercially available reagents were used without further purification. Analytical and preparative thin-layer chromatography were performed with silica gel 60 F_{254} glass plates (EM Science). Silica gel chromatography was performed with 230-400 mesh silica gel (Selecto Scientific). HPLC was performed on a Prostar 210 (Varian) with a flow rate of 10 mL/min using COMBI-A C18 preparatory column (Peeke Scientific).

Compound **2**: Potassium phthalimide (4.8 g, 31 mmol) was added to 4-methyl-2-bromo-acetophenone (5 g, 24 mmol) in dimethylformamide (25 mL), The reaction was stirred at room temperature for 2 h, then diluted into methylene chloride and washed sequentially with water, 0.2 N sodium hydroxide, and brine. The organic phase was dried over sodium sulfate and the solvent was removed by rotary evaporation and subsequent exposure to high vacuum to yield pure **2** (5.9 g, 90%). ^{1}HNMR ($CDCl_3$): δ 2.44 (s, 3H), 5.11 (s, 2H), 7.31 (d, J = 7.6, 2H), 7.75 (m, 2H), 7.88 (m, 4H).

Compound **3**: Sodium hydroxide (5.1 g, 130 mmol) was dissolved in water (11 mL) and added to a solution of malononitrile (1.74 mL, 28 mmol) in methanol (34 mL).

Compound **2** was added with vigorous stirring. After one hour, the reaction was diluted into ethyl acetate and washed with water, followed by brine. The organic phase was dried over sodium sulfate then the solvent was removed by rotary evaporation and subsequent exposure to high vacuum to yield pure **3** (3.88 g, 93%). ^{1}HNMR (DMSO): δ 2.28 (s, 3H), 5.69 (s, 2H), 6.46 (s, 1H), 7.14 (d, J = 8.0 Hz, 2H), 7.43 (d, J = 8.0 Hz, 2H), 10.32 (s(br), 1H).

Compound **4**: Compound **3** (3.56 g, 18 mmol) was refluxed in triethyl orthoformate (20 mL) with acetic anhydride (0.17 mL, 1.8 mmol) for 2 h. The reaction was concentrated by rotary evaporation and azeotroped three times with toluene. Residual solvent was removed under high vacuum and the crude product was used directly in subsequent reactions. ^{1}HNMR (DMSO): δ 1.28 (t, J = 6.8 Hz, 3H), 2.26 (s, 3H), 4.35 (q, J = 6.8 Hz, 2H), 6.96 (s, 1H), 7.15 (d, J = 8.0 Hz, 2H0, 7.46 (d, J = 8.0 Hz, 2H), 8.33 (s, 1H), 11.68 (s(br), 1H).

Compound **5**: Compound **4** (~18 mmol) was dissolved in 7N ammonia in methanol at 0º C. The reaction was stirred and allowed to come to room temperature over 2 hours. The ammonia was removed by bubbling argon through the solution, and pure **5** (3.75 g, 93% - 2 steps) was obtained after concentration *in vacuo*. ^{1}HNMR(DMSO): δ 2.29 (s, 2H), 6.77 (s, 1H), 7.17 (d, J = 8.0 Hz, 2H), 7.48 (d, J = 8.0 Hz, 2H), 11.17 (s (br), 1H).

Compound **6**: A 25% (w/v) solution of sodium methoxide (1.9 mL, 8.4 mmol) was added to a suspension of **5** (3.8 g, 16.8 mmol) in methanol (300 mL). The mixture was refluxed one hour, cooled, diluted into methylene chloride, and washed with water. The organic phase was dried over sodium sulfate; the solvent was removed by rotary evaporation

and subsequent exposure to high vacuum to yield pure **6** (2.9 g, 78%) as a gray to brown powder. ^{1}HNMR (DMSO): δ 2.35 (s, 3H),5.98 (s(br),2H), 7.18 (s, 1H), 7.27 (d, J = 8.0 Hz, 2H), 7.36 (d, J = 8.0, 2H), 8.10 (s, 1H), 11.75 (s, H).

(5-*O*-*tert*-butyldimethylsilyl-2,3-*O*-isopropylidene-α-D-ribofuranose): Concentrated sulfuric acid (0.25 mL) was added to a suspension of D(-)Ribofuranose (10 g, 66 mmol) in acetone (100 mL). The reaction was stirred 2 h at room temperature then neutralized with solid calcium hydroxide, filtered through a celite pad, and concentrated by rotary evaporation. The residue was dissolved in dimethylformamide (125 mL) to which was added imidazole (10.9 g, 160 mmol) and *t*-butyl(chloro)dimethylsilane (10.9 g, 70 mmol). The reaction was stirred overnight at room temperature, then quenched with water (200 mL) and extracted three times with ethyl acetate. The combined organic fractions were washed once with water, dried over sodium sulfate and concentrated by rotary evaporation. The 2,3,5 protected sugar was purified by silica gel chromatography (3:1 hexanes:ethyl acetate) to yield a clear oil (12.1 g, 62%). ^{1}HNMR($CDCl_3$): δ 0.13 (s, 6H), 0.92 (s, 9H), 1.31 (s, 3H), 1.47 (s, 3H), 3.75 (m, 2H), 4.34 (s(br), 1H), 4.49 (d, J = 5.6 Hz, 1H), 4.69 (d, J = 5.6, 1H), 4.73 (d, J = 11.6 Hz, 1H), 5.27 (d, H = 11.6 Hz, 1H).

(5-*O*-*tert*-Butyldimethylsilyl-2,3-*O*-isopropylidene-α-D-ribofuranosyl Chloride): (5-*O*-*tert*-Butyldimethylsilyl-2,3-*O*-isopropylidene-α-D-ribofuranose) (12.1 g, 40 mmol) was dissolved in tetrahydrofuran (80 mL) and cooled to -80º C on a dry ice acetone bath. To this solution, carbon tetrachloride (4.4 mL, 46 mmol) was added followed by drop-wise addition of tris(dimethylamino)phosphine (8.2 mL, 45 mmol). After stirring for one hour at -80º C, the reaction was diluted 1:1 with diethyl ether and washed with cold brine. The organic was dried briefly over sodium sulfate and concentrated by rotary evaporation to yield a clear yellowish oil. Based on ^{1}HNMR integration of the anomeric

proton at ~6.15 ppm, the crude product contained a mixture of α and β isomers at a 5:1 ratio. This product was used immediately in subsequent reactions without further purification. ^{1}HNMR (α-isomer, $CDCl_3$): δ 0.07 (s, 6H), 0.88 (s, 9H), 1.39 (s, 3H), 1.66 (s, 3H), 3.82 (m, 2H), 4.39 (m, 1H), 4.77 (m, 2H), 6.16 (d, J = 4.8 Hz, 1H).

Compound **7**: To a heterogeneous solution of **6** (10 g, 45 mmol) in toluene (150 mL), was added powdered KOH (5 g, 90 mmol) and Tris[2-(2methoxyethoxy)ethyl]amine (3.6 mL, 11 mmol). To this mixture was added freshly prepared (5-*O*-*tert*-Butyldimethylsilyl-2,3-*O*-isopropylidene-α-D-ribofuranosyl Chloride) (~90 mmol based on starting material prior to chlorination) in toluene (150 mL), and the reaction was stirred at room temperature overnight. The completed reaction was diluted into ethyl acetate and washed with 0.1 N HCl and brine, dried over sodium sulfate, and concentrated. Compound **7** (9.7 g, 42%) was obtained as an light tan amorphous solid by silica gel chromatography (1:3 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 0.00 (s,3H), 0.02 (s,3H), 0.85 (s,9H), 1.39 (s,3H), 1.64 (s, 3H), 2.40 (s, 3H), 3.82 (dd, J = 11.4, 4.0, 1H), 3.87 (dd, J = 11.4, 4.0, 11.4, 1H), 4.29 (q, J = 4.0, 1H), 4.96 (dd, J = 3.2, 3.2, 1H), 5.16 (dd, J = 3.2, 3.2, 1H), 5.18 (s (b), 2H), 6.38 (d, J = 3.2, 1H), 7.13 (s, 1H), 7.24 (d, J = 8.0, 2H), 7.34 (d, J = 8.0, 2H), 8.32 (s, 1H). HRMS: calcd m/z 511.2741; found m/e 511.2738.

Compound **8**: To a solution of **7** (9.58 g, 19 mmol) in tetrahydrofuran (100 mL) was added Boc anhydride (8.6 g, 39.4 mmol) and 4-(dimethylamino)pyridine (115 mg, 1 mmol). The reaction was sealed and stirred overnight at room temperature. Water (2 mL) was added and the reaction was stirred 40 min prior to dilution into ethyl acetate and sequential washes with 10% citrate (pH 4.0), water, and brine. The organic phase was dried over sodium sulfate and concentrated to yield pure **8** (13.2 g, 99.4%) as a light tan amorphous solid. ^{1}HNMR ($CDCl_3$): δ 0.01 (s, 3H), 0.03 (s, 3H), 0.86 (s, 9H), 1.26 (s,

9H), 1.29 (s, 9H), 1.40 (s, 3H), 1.67 (s, 3H), 2.38 (s, 3H), 3.81 (dd, J = 10.8, 3.6 1H), 3.89 (dd, J = 10.8, 3.6, 1H), 4.35 (q, J = 3.6, 1H), 4.97 (dd, J = 6.4, 3.2, 1H), 5.14 (dd, J = 6.4, 3.2, 1H), 6.48 (d, J = 3.2, 1H), 7.17 (d, J = 8.0, 2H), 7.33 (d, J = 8.0, 2H), 7.47 (s, 1H), 8.84 (s, 1H). HRMS: calcd m/e 711.3789; found m/e 711.3779.

Compound **9**: Tetrabutylammonium fluoride (0.26 mL, 1M in THF, 0.26 mmol) was added to a solution of **8** (92 mg, 0.13 mmol) in tetrahydrofuran (2 mL). Upon completion (~1-2 h), the reaction was concentrated, and the product was purified by silica gel chromatography for a light tan amorphous solid (1:1 ethyl acetate:hexanes, 74 mg, 95%). ^{1}HNMR ($CDCl_3$): δ 1.28 (s, 9H), 1.29 (s, 9H), 1.39 (s, 3H), 1.64 (s, 3H), 2.38 (s, 3H), 3.82 (m, 1H), 3.99 (m, 1H), 4.50 (m, 1H), 5.15 (dd, J = 5.6, 2.0, 1H), 5.32 (dd, J = 5.6, 5.2, 1H), 5.44 (d (b) J = 5.6, 1H), 5.91 (d, J = 5.2, 1H), 7.18 (d, J = 7.6, 2H), 7.28 (s, 1H), 7.33 (d, J = 7.6, 2H), 8.79 (s, 1H).

Compound **10**: 4-(Dimethylamino)pyridine (5.1 mg, 0.042 mmol) was added to a solution of **9** (25 mg, 0.042 mmol) and 4-(fluorosulfonyl)benzoyl chloride (9.4 mg, 0.042 mmol) in dimethylformamide. The reaction was stirred overnight at room temperature. Though incomplete, the reaction was diluted into ethyl acetate and washed with 10 % citrate (pH 4.0) then brine. The organic was dried over sodium sulfate and concentrated by rotary evaporation. Pure **10** (15 mg, 46%) was obtained as a light tan amorphous solid after silica gel chromatography using a step gradient of hexanes in ethyl acetate (2:1, 1:1, 1:3, hexanes:ethyl acetate). A small amount of pure starting material was also recovered (6.7 mg, 27%). ^{1}HNMR ($CDCl_3$): δ 1.30 (s, 18H), 1.42 (s, 3H), 1.66 (s, 3H), 2.40 (s, 3H), 4.58 (m, 2H), 4.71 (dd, J = 4.0, 11.6 Hz, 1H), 5.15 (m, 1H), 5.42 (m, 1H),

6.34 (s, 1H), 7.17 (d, J = 8.0 Hz, 2H), 7.29 (s, 1H), 7.30 (d, J = 8.0 Hz, 2H), 7.98 d, J = 7.6 Hz, 2H), 8.20 (d, J = 7.6 Hz, 2H), 8.82 (s, 1H).

Compound **1**: Aqueous trifluoroacetic acid (90% v/v, 0.9 mL) was added to a solution of **10** (12 mg, 0.015 mmol) in tetrahydrofuran (0.1 mL) and stirred for two hours at room temperature. The reaction was concentrated by rotary evaporation and residual solvent was removed under high vacuum. Pure **1** (6.8 mg, 82%) was obtained as a chalky white solid after high pressure liquid chromatography (20-100% acetonitrile in water). ^{1}HNMR (DMSO): δ 2.35 (s, 3H), 4.21 (m, 1H), 4.39 (q, J = 5.6 Hz, 1H), 4.48 (m, 2H), 4.69 (dd, J = 3.6, 12.0 Hz, 1H), 5.37 (d, J = 5.6, 1H), 5.55 (d, J = 5.6 Hz, 1H), 6.21 (d, J = 4.4, 1H). 7.25 (s, 4H), 7.31 (s, 1H), 8.13 (d, J = 8.0 Hz, 2H), 8.14 (d, J = 8.0, 2H).

5OH: Nucleosides in various states of deprotection at the exocyclic amine and hydroxyls collected as side products from other syntheses were combined and treated with an approximately ten-fold excess of trimethylsilyl trifluoromethylsulfonate (280 uL, 1.5 mmol) in methylene chloride (2.0 mL) at 0 º C. After one hour the reaction was quenched with 2 drops of methanol and then diluted into methylene chloride and washed with water. The organic fraction was concentrated and the residue was washed with diethyl ether to yield the pure *p*-toyl-nucleoside as a white powder (**5OH**, 40 mg). ^{1}HNMR (DMSO): δ 2.30 (s, 3H), 3.53 (ddd, J = 4.0, 6.0, 12.0 Hz, 1H), 3.62 (ddd, J = 4.4, 6.0, 12.0 Hz, 1H), 3.90 (ddd, J = 4.0, 4.4, 8.0 Hz, 1H), 4.09 (ddd, J = 4.0, 4.8, 8.0 Hz, 1H), 4.45 (ddd, J = 4.0, 6.0, 6.4 Hz, 1H), 5.10 (d, J = 4.8 Hz, 1H), 5.17 (t, J = 6.0 Hz, 1H), 5.30 (d, J = 6.4 Hz, 1H), 6.10 (s(br), 2H), 6.11 (d, J = 6.0 Hz ,1H) 7.30m (d, J = 8.0 Hz, 2H), 7.36 (d, J = 8.0 Hz, 2H), 7.49 (s, 1H), 8.14 (s, 1H).

Compound **11**: Aqueous HCl (25 mL, 1 N) was added to a solution of compound **8** (2.1 g, 2.9 mmol) in tetrahydrofuran (25 mL). The reaction was stirred vigorously overnight at RT. The reaction was diluted into dichloromethane and washed sequentially with saturated sodium bicarbonate, water, and brine. The organic fraction was dried over sodium sulfate and concentrated to yield compound **11** (1.4 g, 85%) with 1 equivalent of the resulting *tert*-butyldimethyl silyl alcohol. ^{1}HNMR ($CDCl_3$): 1.26 (s, 9H), 1.29 (s, 9H), 2.37 (s, 3H), 3.39 (s, 1H), 3.67 (d (b), J = 10.8, 1H), 3.88 (d, J = 10.8, 1H), 3.92 (s (b), 1H), 4.18 (s, 1H), 4.29 (d, J = 5.6, 1H), 4.88 (dd, J = 6.4, 5.6, 1H), 5.49 (s, 1H), 5.88 (d, J = 6.4, 1H), 7.17 (d, J = 8.4, 2H), 7.32 (d, J = 8.4, 2H), 7.36 (s, 1H), 8.77 (s, 1H). HRMS: calcd m/e 557.2611; found m/e 557.2618.

Compound **12**: Compound **11** (1.4 g, 2.5 mmol) was azeotropically dried three times with pyridine then dissolved in fresh pyridine (30 mL). To this solution was added 1,3dichloro-1,1,3,3-tetraisopropyldisiloxane (0.84 mL, 2.6 mmol). The reaction was stirred overnight at room temperature then concentrated, dissolved in ethyl acetate, and washed sequentially with 5% citrate (pH 4.0) and brine. The organic fraction was concentrated, and pure **12** (1.1 g, 57%) was obtained as an amorphous white solid by silica gel chromatography (2:1 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 1.08 (m, 28H), 1.26 (s, 9H), 1.30 (s, 9H), 2.38 (s, 3H), 3.11 (d, J = 1.6, 1H), 4.10 (m, 3H), 4.54 (d, J = 5.6, 1H), 4.95 (dd, J = 7.6, 5.6, 1H), 6.24 (d, J = 1.6, 1H), 7.17 (d, J = 8.0, 2H), 7.33 (d, J = 8.0, 2H), 7.39 (s, 1H), 8.79 (s, 1H). HRMS: calcd m/e 799.4134; found 799.4117.

Compound **13**: 1,1'-Carbonyldiimidazol (21 mg, 0.13 mmol) was added to a solution of **12** (61 mg, 0.076 mmol) in methylene chloride (1 mL) and stirred overnight at room temperature. To this reaction was added N,N-diisopropylethylamine (66 uL, 0.38 mmol) and 1,11-diamino-3,6,9-trioxaundecane (30 mg, 0.15 mmol). After one hour the reaction

was diluted and washed with water. The aqueous fraction was extracted twice with additional methylene chloride and all organic fractions were combined, dried over sodium sulfate, and concentrated *in vacuo*. The residue was taken up in methylene chloride (2 mL), added to a flask containing Di-tert-butyl-dicarbamate (165 mg, 0.76 mmol), and stirred for three hours at room temperature. The reaction was diluted into ethyl acetate and washed with 10% citrate (pH 4.0), then brine. The organic was dried over sodium sulfate and concentrated *in vacuo*. Pure **13** (77 mg, 90%) was obtained after silica gel chromatography (1:3 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 1.06 (m, 28H), 1.26 (s, 9H), 1.29 (s, 9H). 1.44 (s, 9H), 2.38 (s, 3H), 3.33 (m, 4H), 3.56 (m, 4H), 3.65 (m, 8H), 4.04 (m, 2H), 4.13 (dd, J = 4.0, 12.8 Hz, 1H), 4.90 (dd, J = 5.6, 8.0 Hz, 1H), 5.10 (s(br), 1H), 5.47 (s(br), 1H), 5.61 (d, J = 5.6 Hz, 1H), 6.37 (s, 1H), 7.17 (d, J = 8.0 Hz, 2H), 7.33 (d, J = 8.0 Hz, 1H), 7.38 (s, 1H), 8.79 (s, 1H). LC/MS [$M+Na^+$] m/z = 1139.25.

Compound **14**: Tetrabutylammonium fluoride (170 uL, 1M in tetrahydrofuran) was added to a solution of **13** (61 mg, 0.057 mmol) in tetrahydrofuran (1mL) and stirred for one hour at room temperature. The reaction was concentrated under a stream of argon and pure **14** (43 mg, 90%) was obtained after silica gel chromatography (7.5% methanol in methylene chloride). ^{1}HNMR ($CDCl_3$): δ 1.28 (s, 9H), 1.29 (s, 9H), 1.44 (s, 9H), 2.39 (s, 3H), 3.33 (m, 4H), 3.60 (m, 12H), 3.85 (d, J = 12.0 Hz, 1H), 4.02 (d, J = 12.0 Hz, 1H), 4.02 (d, J = 12.0 Hz, 1H), 4.30 (m, 1H), 4.77 (m, 1H), 5.78 (m, 1H), 6.17 (m, 1H), 7.19 (d, J = 8.0 Hz, 2H), 7.33 (d, J = 8.0 Hz, 2H), 7.35 (s, 1H), 8.83 (s, 1H). LC/MS [$M+H^+$] m/z = 876.

Compound **15**: 4-(fluorosulfonyl)benzoyl chloride (4.9 mg, 0.02 mmol), *N,N*-diisopropylethylamine (5.2 uL, 0.03 mmol), and 4-(dimethylamino)pyridine (1.2 mg,

0.004 mmol) were added to a solution of **14** (17.3 mg, 0.02 mmol) in dimethylformamide (0.5 mL). The reaction was stirred overnight at room temperature then additional acid chloride (4.4 mg, 0.02 mmol) and 4-(dimethylamino)pyridine (1.2 mg, 0.004 mmol) were added. After 2 h at 50ºC, the reaction was diluted into ethyl acetate and washed sequentially with sodium bicarbonate, 10% citrate (pH 4.0), and brine, then concentrated *in vacuo*. Pure **15** (4.4 mg, 21%) was obtained after high pressure liquid chromatography (60-100% methanol and 0.1% trifluoroacetic acid in water). Major byproducts were **14** (27%) and the diacylated species (12%). LC/MS $[M+H^+]$ m/z = 875.7. ^{1}HNMR ($CDCl_3$): δ 1.26 (s, 9H), 1.31 (s, 9H), 1.43 (s, 9H), 2.39 (s, 3H), 3.56 (m, 16H), 4.44 (m, 1H), 4.64 (dd, J = 4.0, 12.0 Hz, 1H), 4.79 (m, 2H), 5.60 (m, 1H), 6.57 (d, J = 3.6 Hz, 1H), 7.15 (d, J = 8.0 Hz 2H), 7.20 (d, J = 8.0 Hz, 2H), 7.35 (s, 1H), 7.92 (d, J = 7.6 Hz, 2H), 8.22 (d, J = 7.6 Hz, 2H), 8.86 (s, 1H).

Compound **16**: Trifluoroacetic acid (0.15 mL) was added to a solution of **15** (4.4 mg, 4.2 μmol) in methylene chloride (0.15 mL). The reaction was stirred 50 min at room temperature then the volatiles were removed *in vacuo* and the residue was taken up in dimethylformamide (0.1 mL). To this solution was added TAMRA-NHS ester (2.4 mg, 4.6 μmol) in dimethylformamide (0.1 mL) and diisopropylethyl amine (6 uL, 0.034 mmol). The reaction was stirred for 30 minutes then concentrated *in vacuo*. Compound **16** (2.0 mg, 41%) was obtained as a red oil after HPLC purification (30-100% methanol with 0.1% trifluoroacetic acid in water). LC/MS $[M^+]$ m/z = 1173.52.

Compound **17**: Trifluoroacetic acid (0.5 mL) was added to a solution of **15** (4.8 mg, 4.5 nmol) in methylene chloride (0.5 mL) and stirred for one hour at room temperature. Volatiles were removed *in vacuo* and the residue was resuspended in dimethylformamide (0.1 mL). To this solution was added BODIPY-NHS ester (2.5 mg, 5

nmol; Molecular Probes – D6102) and diisopropylethyl amine (6 uL, 0.03 mmol). The reaction was stirred at room temperature for 2 h then concentrated *in vacuo*. Compound **17** (3.2 mg, 62%) was obtained by high pressure liquid chromatography (65-75% methanol and 0.1% trifluoroacetic acid in water). LC/MS [$M+H^+$] m/z = 1148.54.

Compound **18**: Acryloyl chloride (10 uL, 0.25 mmol) and Diisopropylethylamine (15 uL, 0.09 mmol) were added to a solution of **9** (50 mg, 0.08 mmol) in methylene chloride that had been chilled on an ice bath. The bath was allowed to melt as the reaction stirred overnight. The reaction was diluted into ethyl acetate and washed sequentially with 10% citrate (pH 4.0), water, and brine. The organic fraction was concentrated by rotary evaporation, and pure **18** (28.5 mg, 52%) was obtained after silica gel chromatography (2:1 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 1.29 (s, 18H), 1.41 (s, 3H), 1.65 (s, 3H), 2.39 (s, 1H), 4.34 (m, 3H), 4.50 (m, 2H), 5.04 (dd, J = 3.6, 6.4 Hz, 1H), 5.34 (dd, J = 1.2, 6.4 Hz, 1H), 5.79 (dd, J = 1.2, 10.8, Hz, 1H), 6.06 (dd, J = 10.8, 17.2 Hz, 1H), 6.36 (d, J = 2.4 Hz, 1H), 6.38 (dd, J = 1.2, 17.2, 1H), 7.18 (d, J = 8.0, 2H), 7.31 (s, 1H), 7.33 (d, J = 8.0 Hz, 1H), 8.85 (s, 1H).

Compound **19**: Trifluoroacetic acid (0.5 mL) was added to a solution **18** (28.5 mg, 0.044 mmol) in methylene chloride (0.5 mL). The reaction was stirred at RT for 45 min; the solvent was then removed under a stream of argon. The residual oil was dissolved in tetrahydrofuran (0.1 mL), and a mixture of trifluoroacetic acid and water (9:1, 0.9 mL) was added. Solvent was removed under vacuum and **19** (17.2 mg, 99%) was purified by high pressure liquid chromatography (10–80 % acetonitrile in water with 0.1% trifluoroacetic acid). ^{1}HNMR (DMSO): 2.38 (s, 3H), 4.14 (ddd, J = 3.6, 4.0, 5.6 Hz, 1H), 4.20 (dd, J = 4.0, 5.2 Hz, 1H), 4.29 (dd, J = 5.6, 12.0 Hz, 1H), 4.42 (dd, J = 3.6, 12.0, 1H), 4.48 (dd, J = 4.8, 5.2 Hz, 1H), 5.91 (dd, J = 1.6, 10.4 1Hz, 1H), 6.18 (dd, J = 10.4,

17.2 Hz, 1H), 6.20 (d, J = 4.8 Hz, 1H), 6.32 (dd, J = 1.6, 17.2 Hz, 1H), 7.33 (d, J = 8.0 Hz, 2H), 7.39 (d, J = 8.0 Hz, 2H), 7.63 (s, 1H), 8.40 (s, 1H). LC/MS [$M+H^+$] m/z = 411.33.

Compound **20**: Diisopropylethylamine (87 uL, 0.5 mmol) was added to a solution of **9** (50 mg, 0.084 mmol) in methylene chloride (2 mL) and chilled in an ice bath. 2-chloroethanesulfonyl chloride (22 uL, 0.21 mmol) was added, and the solution was stirred 30 min at 0ºC followed by an additional 30 min at room temperature. The reaction was diluted into ethyl acetate and washed with 10% citrate (pH 4.0) then brine. The organic fraction was dried over sodium sulfate and concentrated *in vacuo*. Compound **20** (33.5 mg, 58%) was purified by silica gel chromatography to yield an amorphous tan solid (1:1 ethylacetate:hexanes). ^{1}HNMR ($CDCl_3$): δ 1.30 (s, 18H), 1.40 (s, 3H), 1.64 (s, 3H), 2.39 (s, 3H), 4.31 (dd, J = 10.8, 5.2 , 1H), 4.37 (dd, J = 10.8, 4.0, 1H), 4.49 (ddd, J = 5.2, 4.0, 4.0, 1H), 5.12 (dd, 6.4, 4.0, 1H), 5.38 (dd, J = 6.4, 2.4, 1H), 5.91 (d, J = 8.8, 1H), 6.23 (d, J = 16.4, 1H), 6.29 (dd, J = 16.4, 8.8, 1H), 6.35 (d, J = 2.4, 1H), 7.19 (d, J = 8.0, 2H), 7.34 (s, 1H), 7.36 (d, J = 8.0, 2H), 8.83 (s, 1H).

Compound **21**: Trifluoroacetic acid (0.5 mL) was added to a solution of compound **20** (33 mg, 0.047 mmol) in methylene chloride (0.5 mL). The reaction was stirred at room temperature 45 minutes then the solvent was removed under a stream of argon. The residual oil was dissolved in tetrahydrofuran (0.1 mL) and a mixture of Trifluoroacetic acid and water (9:1, 0.9 mL) was added. After 30 minutes, the solvent was removed *in vacuo* and compound **21** (15.5 mg, 96%) was purified by high pressure liquid chromatography to yield a chalky white solid (20– 100 % acetonitrile in water with 0.1% trifluoroacetic acid). 1HNMR (DMSO): δ 2.38 (s, 3H), 4.15 (m, 2H), 4.32 (dd, J = 11.2, 5.6, 1H), 4.37 (dd, J = 11.2, 3.6, 1H), 4.46 (m, 1H), 6.22 (d, J = 5.6, 1H), 6.26 (d, J =

10.4, 1H), 6.32 (d, J = 16.8, 1H), 6.98 (dd, J = 16.8, 10.4, 1H), 7.33 (d, J = 8.0, 2H), 7.40 (d, J = 8.0, 2H), 7.64 (s, 1H), 8.41 (s, 1H).

Compound **22**: Benzoic acid (11 mg, 0.09 mmol) and **9** (37 mg, 0.06 mmol) were dissolved in methylene chloride (1 mL) and cooled to 0ºC on an ice bath. (N-(3-dimethylaminopropyl)-*N*-ethyl carbodiimide and 4-(dimethylamino)pyridine (7.0 mg, 0.06 mmol) were added and the reaction was stirred for 45 min at 0ºC. The reaction mixture was diluted into methylene chloride and washed with 0.1 N HCl then brine. The organic phase was dried over sodium sulfate and concentrated *in vacuo*. Pure **22** (31 mg, 71%) was obtained as a tan amorphous solid after silica gel chromatography (2:1 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 1.28 (s, 18H), 1.42 (s, 3H), 1.67 (s, 3H), 2.38 (s, 3H), 4.54 (m, 2H), 4.67 (dd, J = 3.6, 11.2 Hz, 1H), 5.12 (dd, J = 3.6, 6.0 Hz, 1H), 5.37 (dd, J = 2.4, 6.0 Hz, 1H), 6.38 (d, J = 2.4 Hz, 1H), 7.14 (d, J = 8.0 Hz, 2H), 7.26 (d, J = 8.0 Hz, 2H), 7.31 (s, 1H), 7.38 (t, J = 7.2 Hz, 2H), 7.54 (t, J = 7.2 Hz, 1H), 8.00 (d, J = 7.2 Hz, 2H), 8.82 (s, 1H).

Compound **23**: Trifluoroacetic acid (0.25 mL) was added to a solution of **22** (31 mg, 0.044 mmol) in methylene chloride (0.25 mL). The reaction was stirred 45 min at room temperature and the solvent was then blown off with a stream of argon. The residue was dissolved in tetrahydrofuran (0.05 mL) and aqueous trifluoroacetic acid (90% v/v, 0.45 mL) was added. The reaction was stirred 45 minutes at room temperature. Volatiles were removed *in vacuo* and pure **23** (19 mg, 95%) was obtained after high pressure liquid chromatography (20 – 100% acetonitrile in water). ^{1}HNMR (DMSO): δ 2.36 (s, 3H), 4.20 (m, 1H), 4.33 (dd, J = 5.2, 6.0 Hz, 1H), 4.43 (dd, J = 5.2, 12.4 Hz, 1H), 4.51 (q, 5.2 Hz, 1H), 4.60 (dd, J = 3.2, 12.4 Hz, 1H), 5.36 (d, J = 5.2 Hz, 1H), 5.50 (d, J =

6.0 Hz, 1H), 6.19 (d, J = 5.2 Hz, 1H), 7.25 (s, 4H), 7.31 (s, 1H), 7.46 (t, J = 8.0 Hz, 2H), 7.65 (tt, J = 1.2, 8.0 Hz, 1H), 7.96 (dd, J = 1.2, 8.0 Hz, 2H), 8.15 (s, 1H).

Compound **24**: Compound **9** (100 mg, 0.17 mmol) and (*N*-(3-dimethylaminopropyl)-*N'*-ethyl carbodiimide (130 mg, 0.67 mmol) were dissolved in dimethylsulfoxide (2.0 mL). Dichloroacetic acid (7 uL, 0.84 mmol) was added and the reaction was stirred at room temperature until judged complete by thin layer chromatography (1:1 hexanes:ethyl acetate, 5 hours). The reaction was diluted into ethyl acetate and washed twice with water and once with brine. The organic phase was dried over sodium sulfate and concentrated *in vacuo*. The crude product was used for subsequent reactions without further purification. ^{1}HNMR showed a mix of the desired aldehyde and **9** (3:1). ^{1}HNMR ($CDCl_3$): δ 1.29 (s, 9H), 1.30 (s, 9H), 1.41 (s, 3H), 1.62 (s, 3H), 2.39 (s, 3H), 4.67 (d, J = 1.6 Hz, 1H), 5.35 (d, J = 6.4 Hz, 1H), 5.56 (dd, J = 1.6, 6.4 Hz, 1H), 6.21 (s, 1H), 7.20 (d, J = 8.0 Hz, 2H), 7.29 (s, 1H), 7.34 (d, J = 8.0 Hz, 2H), 8.68 (s, 1H), 9.31 (s, 1H).

Compound **25**: Tetrahydrofuran (60 mL) was added to ethymethyl sulfone (2.1 mL, 20 mmol) and cooled to -78ºC. To this solution was added *n*-butyl lithium (13.8 mL, 1.6 M in hexanes) The reaction was stirred for 15 minutes then diethyl chlorophosphate was added (1.6 mL, 11 mmol). The temperature was raised to -50ºC and stirred an additional hour. Ammonium chloride (1.2 g) was added and the temperature was slowly raised to room temperature. The reaction was diluted into methylene chloride, washed with water, dried over sodium sulfate, and concentrated *in vacuo*. Pure **25** (1.04 g, 36%) was obtained after silica gel chromatography (1:3 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 1.28 (t, J = 7.2 Hz, 6H), 1.44 (t, J = 6.8 Hz, 3H), 3.73 (d, J = 16.8 Hz, 2H), 4.25 (m, 4H), 4.42 (q, J = 6.8 Hz, 2H).

Compound **26**: To a solution of **25** (30 mg, 0.12 mmol) in tetrahydrofuran (0.75 mL) at -78ºC was added *n*-butyl lithium (50 µL, 2.5 M in hexanes). To this reaction was added a solution of crude **24** (75 mg) in tetrahydrofuran (0.75 mL). The reaction was stirred for 1 h at -78ºC, then overnight at room temperature. The completed reaction was diluted into ethyl acetate and washed with water and brine. The organic phase was dried over sodium sulfate and concentrated *in vacuo*. Pure **26** (54 mg, 61% - 2 steps) was obtained after silica gel chromatography (5:4 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 1.29 (m, 21H), 1.40 (s, 3H), 1.65 (s, 3H), 2.39 (s, 3H), 4.07 (m, 2H), 4.83 (m, 1H), 5.12 (dd, J = 4.8, 5.6 Hz, 1H), 5.39 (dd, J = 1.6, 5.6 Hz, 1H), 6.33 (d, J = 4.8 Hz, 1H), 6.36 (d, J = 15.6 Hz, 1H), 7.07 (dd, J = 4.0, 15.6 Hz, 1H), 7.20 (d, J = 8.0 Hz, 2H), 7.23 (s, 1H), 7.36 (d, J = 8.0 Hz, 1H), 8.82 (s, 1H).

Compound **27**: Compound **26** (54 mg, 0.077 mmol) was refluxed in acetone (5 mL) with tertrabutylammonium iodide (31 mg, 0.085 mmol) until no starting material was visible by thin layer chromatography (10% methanol in methylene chloride). The reaction was concentrated *in vacuo*. The residue was azeotropically dried twice with toluene then dissolved in methylene chloride (1 mL). To this solution was added triphosgene (23 mg, 0.077 mmol) and dimethylformamide (0.6 µL, 0.007 mmol). The reaction was stirred overnight at room temperature. The reaction was diluted into hexanes and ethyl acetate (1:1), flushed through a plug of silica gel and concentrated *in vacuo*. LC/MS (50-100% acetonitrile and 0.1% formic acid in water) showed one major peak corresponding to the desired 5'-vinylsulfonyl chloride (m/z [M+Na^+] = 713.44). The residue was resuspended in acetonitrile (0.5 mL) that had been dried over calcium hydride. A mixture of CaF_2 and KF (4:1) was prepared by heating with a torch under high vacuum. This salt mix (13 mg,

0.23 mol) was added to the above suspension and the mix was stirred overnight at room temperature. The reaction was diluted into ethyl acetate and washed with water and brine. The organic phase was dried over sodium sulfate and concentrated *in vacuo*. LC/MS (50-100% acetonitrile and 0.1% formic acid in water) indicated the principle product to be the fully protected vinylsulfonyl fluoride (m/z $[M+Na^+]$ = 675.31). The HPLC solvents were removed *in vacuo* and the residue was resuspended in methylene chloride (0.2 mL). Trifluoroacetic acid (0.2 mL) was added and the resulting solution was stirred at room temperature for one hour. The solvent was blown off with a stream of argon and the residue was resuspended in tetrahydrofuran (0.04 mL). To this solution was added aqueous trifluoroacetic acid (90%, 0.36 mL) and the reaction was stirred for 20 min at room temperature. The solvent was removed *in vacuo* to yield pure **27** (5.1 mg, 15% - 4 steps). ^{1}HNMR (DMSO): δ 2.38 (s, 1H), 4.32 (t, J = 5.2 Hz, 1H), 4.60 (t, J = 5.2 Hz, 1H), 4.69 (m, 1H), 6.23 (d, J = 5.2 Hz, 1H), 7.26 (dt, J = 1.6, 14.8 Hz, 1H), 7.33 (d, J = 8.0 Hz, 2H), 7.41 (d, J = 8.0 Hz, 2H), 7.52 (dd, J = 5.2, 14.8 Hz, 1H), 7.71 (s, 1H), 8.34 (s, 1H). LC/MS $[M+H^+]$ m/z = 435.24.

Compound **28**: Aluminum chloride (160 mg, 1.2 mmol) was added to a solution of 4'-phenoxyacetophenone (5 g, 24 mmol) in diethyl ether (50 mL) that had been chilled on an ice bath. Bromine (1.3 mL, 25 mmol) was added drop-wise to the reaction, and the ice bath was allowed to melt as the solution stirred overnight. The reaction was poured over ice and extracted with excess ether, washed with brine, dried over magnesium sulfate, and concentrated by rotary evaporation. The crude bromoacetophenone was dissolved in dimethylformamide (50 mL) and potassium phthalate (4.8 g, 26 mmol) was added. After 2 h, the reaction was diluted into ethyl acetate and washed sequentially with water, 0.2 sodium hydroxide, and brine, then dried over sodium sulfate. The organic was concentrated by rotary evaporation and pure **28** (7.6g, 90%) was recrystlized from a

distilling mixture of methanol, methylene chloride, and hexanes (1:49:125) There was a small amount of 4'-bromo-4-phenoxy phenyl contaminant that was carried through all subsequent steps, but was removed in the final HPLC purification of the derived nucelosides ^{1}HNMR ($CDCl_3$): δ 5.09 (s, 2H), 7.04 (d, J = 8.8 Hz, 2H), 7.09 (d, J = 8.8 Hz, 2H), 7.22 (t, J = 8.0 Hz, 1H), 7.42 (t, J = 8.0 Hz, 2H), 7.75 (dd, J = 3.2, 5.6 Hz, 2H), 7.90 (dd, J = 3.2, 5.6, 2H), 7.99 (d, J = 8.0 Hz, 2H).

Compound **29**: Aqueous sodium hydroxide (10 N, 12.5 mL) was added to a solution of malononitrile (1.7 mL, 28 mmol) in methanol (38 mL). Compound **28** (7.6 g, 21 mmol) was added and the reaction was stirred one h at room temperature. The reaction was diluted into ethyl acetate, washed with water and brine, and dried over sodium sulfate. Pure **29** (5.9 g, 100%) was obtained after removal of solvent *in vacuo*. ^{1}HNMR (DMSO): δ 5.75 (s(br), 2H), 6.49 (d, J = 2.0 Hz, 1H), 7.00 (d, J = 8.0 Hz, 2H), 7.01 (d, J = 8.0 Hz, 1H), 7.13 (t, J = 8.0 Hz, 1H), 7.39 (t, J = 8.0 Hz,2H), 7.55 (d, J = 8.0 Hz, 2H), 10.42 (s(br), 1H).

Compound **30**: Triethyl orthoformate (45 mL) was added to **29** (5.9 g, 21 mmol), and the solution was heated to 100ºC for 3 h. The reaction was concentrated by rotary evaporation and the residue was azeotropically dried three times with toluene. Crude **30** was used directly in the subsequent reaction. ^{1}HNMR (DMSO): δ 1.33 (t, J = 7.6 Hz, 3H), 4.30 (q, J = 7.6 Hz, 2H), 7.04 (m, 5H), 7.14 (t, J = 8.4 Hz, 1H), 7.40 (t, J = 8.4 Hz, 2H), 7.61 (d, J = 8.4 Hz, 2H), 8.39 (s, 1H), 11.78 (s(br), 1H).

Compound **31**: Crude **30** (~21 mmol) was dissolved in 7N ammonia in methanol at 0ºC. The reaction was stirred and allowed to come to room temperature over 2 hours. The ammonia was removed by bubbling argon through the solution, and remaining solvent

was removed *in vacuo*. Crude **31** was used directly for the subsequent annulation without further purification. ^{1}HNMR (DMSO): δ 6.79 (s, 1H), 7.02 (d, J = 8.8 Hz, 4H), 7.13 (t, J = 7.6 Hz, 1H), 7.59 (d, J = 8.8 Hz, 2H), 7.97 (dd, J = 7.6, 8.8 Hz, 2H), 8.02 (m, 1H), 11.22 (s(br), 1H).

Compound **32**: Ethanol (135 mL) and freshly prepared sodium ethoxide (15 mL, 0.14 M) were added to crude **31**, and the mixture was refluxed overnight. The complete reaction was diluted into methylene chloride and washed with water and brine. The organic phase was dried over sodium sulfate and concentrated *in vacuo* to yield pure **32** (5.1 g, 80% - 3 steps). ^{1}HNMR (DMSO): δ 6.03 (s(br), 2H), 7.09 (d, J = 8.4 Hz, 2H), 7.09 (d, J = 8.4 Hz, 2H), 7.15 (t, J = 7.6 Hz, 1H), 7.22 (s, 1H), 7.41 (dd, J = 7.6, 8.8 Hz, 2H), 7.47 (d, J = 8.8, 2H), 8.11 (s, 1H), 11.78 (s, 1H).

Compound **33**: To a heterogeneous mixture of **32** (2 g, 6.7 mmol) in toluene (50 mL) was added powdered potassium hydroxide (750 mg, 13.4 mmol) and tris[2-(2methoxyethoxy)ethyl]amine (0.5 mL, 1.7 mmol). Freshly prepared 5-*O*-*tert*-butyldimethylsilyl-2,3-*O*-isopropylidene-α-D-ribofuranosyl chloride (13.4 mmol) was added in toluene (15 mL). The reaction was left to stir overnight at room temperature. The reaction was diluted into ethyl acetate and washed sequentially with 0.1 N HCl, water, and brine, then dried over sodium sulfate and concentrated by rotary evaporation. Pure **33** (680 mg, 18%) was obtained after silica gel chromatography (1:2 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 0.01 (s, 3H), 0.03 (s, 3H), 0.09 (s, 9H), 1.39 (s, 3H), 1.65 (s, 3H), 3.79 (dd, J = 4.0, 11.2 Hz, 1H), 3.88 (dd, J = 4.0, 11.2 Hz, 1H), 4.31 (m, 1H), 4.96 (dd, J = 3.6, 6.0 Hz, 1H), 5.09 (s, 2H), 5.14 (dd, J = 2.8, 6.0 Hz, 1H), 6.40 (d, J = 2.8 Hz, 1H), 7.07 (m, 4H), 7.15 (m, 1H), 7.17 (s, 1H), 7.40 (m, 5H), 8.34 (s, 1H).

Compound **34**: To a solution of **33** (270 mg, 0.46 mmol) in THF (20 mL) was added di-tert-butyl dicarbonate (500mg, 2.3 mmol) and 4-(dimethylamino)pyridine (5mg, 0.046 mmol). The reaction was sealed and stirred overnight. The completed reaction was diluted into ethyl acetate and washed with 10% citrate (pH 4.0), and then brine. The organic phase was dried over sodium sulfate and concentrated by rotary evaporation. Pure **34** (290 mg, 82%) was obtained by silica gel chromatography (3:1 hexanes;ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 0.01 (s, 3H), 0.03 (s, 3H), 0.85 (s, 9H), 1.30 (s(br), 9H), 1.32 (s(br), 9H), 1.41 (s, 3H), 1.67 (s, 3H), 3.81 (dd, J = 3.6, 11.2 Hz, 1H), 3.91 (dd, J = 3.2, 11.2 Hz, 1H), 4.36 (m, 1H), 4.96 (dd, J = 3.2, 6.4 Hz, 1H), 5.12 (dd, J = 2.8, 6.4 Hz, 1H), 6.51 (d, J = 2.8, 1H), 7.04 (m, 2H), 7.12 (m, 1H), 7.36 (m, 2H), 7.40 (m, 2H), 7.52 (s, 1H), 8.85 (s, 1H).

Compound **35**: Tetrabutylammonium fluoride (0.26 mL 1M, in THF, 0.26 mmol) was added to a solution of compound **34** (138 mg, 0.18 mmol) in tetrahydrofuran (3.5 mL) and left to stir overnight at room temperature. The reaction was diluted into ethyl acetate and washed with water, followed by brine. The organic phase was dried over sodium sulfate and concentrated *in vacuo.* Compound **35** was purified by column chromatography (1:1 ethyl acetate:hexanes, 77 mg, 65%). ^{1}HNMR ($CDCl_3$): δ 1.32 (s, 18H), 1.39 (s, 3H), 1.65 (s, 3H), 3.83 (m, 1H), 4.01 (m, 1H), 4.51 (m, 1H), 5.15 (dd, J = 1.6, 6.4 Hz, 1H), 5.33 (dd, J = 4.8, 6.4 Hz, 1H), 5.42 (dd, J = 1.6, 10.8 Hz, 1H), 5.92 (d, J = 4.8 Hz, 1H), 7.02 (d, J = 8.8 Hz, 2H), 7.06 (m, 2H), 7.14 (m, 1H), 7.30 (s, 1H), 7.37 (m, 2H), 7.40 (m, 2H), 8.81 (s, 1H).

Compound **36**: Diisopropylethylamine (47 µL, 0.27 mmol) was added to a heterogeneous mixture of 4-(fluorosulfonyl)benzoyl chloride (30 mg, 0.13 mmol) and **35** (40 mg, 0.059 mmol) in methylene chloride (0.5 mL). The reaction was stirred overnight

at room temperature, diluted into methylene chloride, and washed sequentially with sodium bicarbonate, 0.1 N HCl, and brine. The organic was concentrated *in vacuo,* and pure **36** (31mg, 68%) was isolated by preparative silica gel thin layer chromatography (1:1 hexanes:ethyl acetate) ^{1}HNMR ($CDCl_3$): δ 1.33 (s, 18H), 1.42 (s, 3H), 1.66 (s, 3H), 4.80 (m, 2H), 4.71 (m, 1H), 5.17 (dd, J = 4.0, 6.4 Hz, 1H), 5.48 (dd, J = 2.0, 6.4 Hz, 1H), 6.27 (d, J = 2.0 Hz, 1H), 7.01 (d, J = 8.8 Hz, 2H), 7.06 (d, J = 8.8 Hz, 2H), 7.15 (t, J = 3.6 Hz, 1H), 7.29 (s, 1H), 7.38 (m, 4H), 8.04 (d, J = 8.4 Hz, 2H), 8.21 (d, J = 8.4 Hz, 2H), 8.83 (s, 1H).

Compound **37**: Trimethylsilyl trifluoromethanesulfonate (46 µL, 0.26 mmol) was added to a solution of **36** (20 mg, 0.023 mmol) in methylene chloride (0.5 mL) at 0º. After one hour, the reaction was concentrated under a stream of argon and spotted directly onto a preparative silica gel thin layer chromatography plate. Pure **37** (5.8 mg, 40%) was obtained after elution (10% methanol in methylene chloride) and further resolution from a *p*-bromophenoxy nucleoside by HPLC (50-90% acetonitrile in water). ^{1}HNMR (DMSO): δ 4.22 (q, J = 4.8 Hz, 1H), 4.37 (m, 1H), 4.50 (dd, J = 5.6, 12.0, 1H), 4.52 (m, 1H), 4.69 (dd, J = 3.6, 12.0 Hz, 1H), 5.39 (d, J = 5.2 Hz, 1H), 5.53 (d, J = 6.0 Hz, 1H), 6.17 (s(br), 2H), 6.19 (d, J = 5.2 Hz, 1H), 7.06 (d, J = 8.8 Hz, 2H), 7.08 (d, J = 8.8 Hz, 2H), 7.17 (t, J = 7.6 Hz, 1H), 7.40 (m, 5H), 8.14 (s, 1H), 8.21 (d, J = 8.4 Hz, 2H), 8.27 (d, J = 8.4 Hz, 2H).

Compound **38**: Diisopropylethylamine (47 uL, 0.27 mmol) was added to a solution of acyloyl chloride (11 uL, 0.13 mmol) and **35** (40 mg, 0.059 mmol) in methylene chloride (0.5 mL). The reaction was stirred overnight then diluted into methylene chloride and sequentially washed with sodium bicarbonate, 0.1 N HCl, and brine. The organic phase was concentrated by rotary evaporation and pure **38** (35 mg, 80%) was isolated by

preparative silica gel thin layer chromatography (1:1 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 1.32 (s, 18H), 1.41 (s, 3H), 1.66 (s, 3H), 4.34 (dd, J = 6.8, 13.2 Hz, 1H), 4.51 (m, 2H), 5.05 (dd, J = 3.2, 6.4 Hz, 1H), 5.36 (dd, J = 2.4, 6.4 Hz, 1H), 5.80 (d, J = 10.4 Hz, 1H), 6.06 (dd, J = 10.4, 18.0 Hz, 1H), 6.35 (d, J = 2.4 Hz, 1H), 6.38 (d, J = 18.0 Hz, 1H), 7.02 (d, J = 6.8 Hz, 2H), 7.06 (d, J = 6.8 Hz, 2H), 7.14 (t, J = 3.6 Hz, 1H), 7.32 (s, 1H), 7.37 (t, J = 3.6 Hz, 2H), 7.40 (d, J = 3.6 Hz, 2H), 8.86 (s,1H).

Compound **39**: Trimethylsilyl trifluoromethanesulfonate (46 µL, 0.26 mmol) was added to a solution of **38** (20 mg, 0.027 mmol) in methylene chloride (0.5 mL) at 0º C. After one hour the reaction was concentrated under a stream of argon and spotted directly onto a preparative silica gel thin layer chromatography plate. Pure **39** (7.5 mg, 56%) was obtained after elution (10% methanol in methylene chloride) and further resolution from a *p*-bromophenoxy-containing nucleoside by HPLC (30-70% acetonitrile in water). ^{1}HNMR (DMSO): δ 4.04 (m, 1H), 4.14 (q, J = 5.2 Hz, 1H), 4.21 (dd, J = 6.4, 12.0 Hz, 1H), 4.37 (dd, J = 3.6, 12.4 Hz, 1H), 4.42 (q, J = 5.6 Hz, 1H), 5.28 (d, J = 5.2 Hz, 1H), 5.42 (d, J = 5.6 Hz, 1H), 5.86 (dd, J = 1.6, 10.4 Hz, 1H), 6.13 (d, J = 5.6 Hz, 1H), 6.14 (dd, J = 10.4, 17.2 Hz, 1H), 6.28 (dd, J = 1.6, 17.2 Hz, 1H), 7.05 (d, J = 8.0 Hz, 2H), 7.07 (d, J = 8.0 Hz, 2H), 7.11 (t, J = 7.6 Hz, 1H), 7.37 (t, J = 7.6 Hz, 2H), 7.37 (s, 1H), 7.42 (d, J = 7.6 Hz, 2H), 8.11 (s, 1H).

Compound **40**: Diisopropylethylamine (47 µL, 0.27 mmol) was added to a solution of 2-chloroethanesulfonyl chloride (14 µL, 0.13mmol) and **35** (40 mg, 0.059 mmol) in methylene chloride (0.5 mL). The reaction was stirred overnight then diluted into methylene chloride and sequentially washed with sodium bicarbonate, 0.1 N HCl, and brine. The organic was concentrated by rotary evaporation and pure **40** (42mg, 82%) was isolated by preparative silica gel thin layer chromatography (5% methanol in

methylene chloride). ^{1}HNMR ($CDCl_3$): δ 1.30 (s, 18H), 1.38 (s, 3H), 1.61 (s, 3H), 4.28 (dd, J = 5.2, 10.8 Hz, 1H), 4.35 (dd, J = 4.0, 10.8 Hz, 1H), 4.47 (m, 1H), 5.09 (dd, J = 3.6, 6.4 Hz, 1H), 5.35 (dd, J = 2.4, 6.4 Hz, 1H), 5.91 (d, J = 9.2 Hz, 1H), 6.27 (m, 3H), 7.00 (d, J = 8.4 Hz, 2H), 7.03 (d, J = 8.4 Hz, 2H), 7.11 (t, J = 7.6 Hz, 1H), 7.32 (s, 1H), 7.34 (t, J = 7.6 Hz, 2H), 7.40 (d, J = 7.6 Hz, 2H), 8.81 (s, 1H).

Compound **41**: Trimethylsilyl trifluoromethanesulfonate (46 µL, 0.26 mmol) was added to a solution of **40** (20 mg, 0.026 mmol) in methylene chloride (0.5 mL) at 0ºC. After one hour the reaction was concentrated under a stream of argon and spotted directly onto a preparative silica gel thin layer chromatography plate. Pure **41** (3.0 mg, 22%) was obtained as a chalky white solid after elution (10% methanol in methylene chloride) and further resolution from a *p*-bromophenoxy nucleoside by HPLC (40-80% acetonitrile in water). ^{1}HNMR (DMSO): δ 4.11 (m, 1H), 4.16 (q, J = 5.2 Hz, 1H), 4.30 (dd, J = 6.0, 11.2 Hz, 1H), 4.37 (dd, J = 3.6, 11.2 Hz, 1H), 4.46 (q, J = 5.6 Hz, 1H), 5.41 (d, J = 5.2 Hz, 1H), 5.52 (d, J = 5.6 Hz, 1H), 6.20 (d, J = 5.6 Hz, 1H), 6.25 (d, J = 10.0 Hz, 1H), 6.32 (d, J = 16.8 Hz, 1H), 6.97 (dd, J = 10.0, 16.8 Hz, 1H), 7.09 (d, J = 8.4 Hz, 2H), 7.12 (d, J = 8.4 Hz, 2H), 7.16 (t, J = 7.6 Hz, 1H), 7.42 (t, J = 7.6 Hz, 2H), 7.44 (s, 1H), 7.48 (d, J = 7.6 Hz, 2H), 8.16 (s, 1H).

Kinase assays:

Active Src and Hck kinases were obtained from Upstate. Down-regulated pY528 Src was a generous gift from Michael Eck (Harvard Medical School). Substrate peptide (LEEIYGEFKK) was synthesized in-house by standard solid phase peptide synthesis using Wang resin, and γ^{32}P-ATP (6000 Ci/mmol) was purchased from Perkins Elmer.

Inhibition of Src by compound 1, PP1, and FSBA (Figure 1-2): Active Src (0.5 nM) was incubated with inhibitors in reaction buffer (20 mM HEPES (pH 7.4), 10 mM $MgCl_2$, 0.2 mM EDTA, and 1 mg/mL BSA) and 250 μM ATP, for 30 min at room temperature. The kinase reaction was initiated by a 1:1 dilution into reaction buffer containing 250 μM ATP, 200 μM substrate peptide and 0.8 mCi/mL γ^{32}P-ATP. After 30 min, reactions were spotted onto a phosphocellulose sheet, briefly dried, and washed with 10% acetic acid, twice with 1% phosphoric acid, and finally with acetone before drying. Kinase activity was quantified using a Typhoon Imaging System (Molecular Dynamics) and ImageJ software (NIH).

Time-dependent inhibition of Src by compound 1 (Figure 1-3): Active Src (0.2 μM) was incubated in the presence or absence of 1 μM compound **1** in reaction buffer with 100 μM ATP. At selected times, 2 μL samples were withdrawn and diluted into 23 μL of reaction buffer with 100 μM ATP, 100 μM substrate peptide and 0.1 mCi/mL γ^{32}P-ATP. The kianse reactions were run for 30 min, and then spotted onto phosphocellulose disks. The disks were washed once with 10% acetic acid, twice with 10% phosphoric acid, and once with acetone before drying. Kinase activity was quantified by scintillation counting.

Dialysis of inhibited Src (Figure 1-3): Phospho-Y527 Src (5 nM) was incubated in reaction buffer with compound **1** or **5OH** for one hour at room temperature. An aliquot of each reaction was removed and stored at 4°C and the remainder was dialyzed over 5 h against 2 changes of reaction buffer at 4ºC. Following dialysis, the kinase was diluted 1:5 with reaction buffer containing 125 μM ATP, 125 substrate peptide, and 0.2 mCi/mL γ^{32}P-ATP and incubated for 30 min RT. Reactions were spotted onto phosphocellulase disks and treated as described above.

Time dependent inhibition of Src by compounds 1, 19, and 21 (figure 1-5): Phospho-Y527 Src (10 nM) was incubated with 250 nM inhibitors or DMSO control in reaction buffer with 100 μM ATP at room temperature. At selected times, the reactions were diluted 1:1 with reaction buffer containing 100 μM ATP, 200 μM substrate peptide and 0.2 mCi/mL γ^{32}P-ATP. After 30 min at room temperature, reactions were spotted onto phosphocellulose disks and treated as described above.

Inhibition of Src and Hck by compounds 1, 19, 21, 37, 39, and 41 (Figures 1-5, 1-7, and 1-8): Kinases (5 nM) were incubated at room temperature with inhibitors in reaction buffer for 30 min, and then diluted 1:4 with reaction buffer containing 125 μM ATP, 125 μM substrate, and 0.2 mCi/mL γ^{32}P-ATP. After 30 min the reactions were spotted onto phosphocellulose disks and treated as described above.

Competitive labeling of recombinant Src:

Phospho-Y527 Src (1 μM) was incubated in kinase reaction buffer with 5 μM inhibitors at room temperature for 20 min. Src was labeled with 5 μM **16** for 5 minutes. The reaction was quenched with 2X Laemmli sample buffer and the proteins were resolved by SDS-PAGE. Fluorescent species were detected with the Typhoon Imaging System using standard TAMRA settings (Molecular Dynamics).

In vivo kinase studies:

NIH-3T3 cells stably expressing I338T v-Src (CZ9 cells) were a generous gift from Chao Zang and Kevan Shokat (University of California San Francisco). Cells were cultured in Dulbecco's modified Eagle's Media supplemented with 10% fetal bovine serum.

***In vivo* inhibition of v-Src**: CZ9 cells were treated with PP1 or **1** in serum-free media for 1.5 hours at 37°C. Cells were harvested and lysed in 1% NP-40, 50 mM Tris (pH 7.4), 2 mM EDTA, 150 mM NaCl, 1X Complete Protease Inhibitor Cocktail (Roche), and 1X Phosphatase Inhibitor Cocktail 2 (Sigma). Lysates were resolved by SDS-PAGE, blotted, and probed with a 1:2500 dilution of the monoclonal anti-phosphotyrosine antibody 4G10 (Upstate), followed by a 1:5000 dilution of a HRP conjugated secondary (Santa Cruz Biotechnology).

***In vivo* labeling of v-Src**: CZ9 cells at 80% confluency in serum-free media were incubated in the presence or absence of 5 μM **1** for 30 minutes at room temperature. The media was aspirated and replaced with serum-free media containing 5 μM **17** for an additional 30 minutes. The media was aspirated and the cells were collected and lysed by two freeze-thaw cycles in 10 mM HEPES (pH 7.5), 10 mM KCl, 150 mM NaCl, and 1X Complete Protease Inhibitor Cocktail (Roche). The lysates were spun at top speed in a benchtop centrifuge and the pellets were resusupended in 1% NP-40, 50 mM Tris (pH 7.4), 2 mM EDTA, 150 mM NaCl, 1X Complete Protease Inhibitor Cocktail (Roche). Proteins were resolved by SDS-PAGE and fluorescent species were detected with the Typhoon Imaging System (Molecular Dynamics).

References:

1) DeCamp DL, Colman RF. Identification of tyrosine and lysine peptides labeled by 5'-p-fluorosulfonylbenzoyl adenosein in the active site of pyruvate kinase. J Biol Chem 1986, 261: 4499-4503.

2) Hanoulle X, Van Damme J, Staes A, Martens L, Goethals M, Vandekerckhove J, Gevaert K. A new functional, chemical proteomics technology to identify purine nucleotide binding sites in complex proteomes. J Proteome Res 2006, 5: 3428-3445.

3) Oudot C, Jault J, Jaquinod M, Negre D, Prost J, Cozzone AJ, Cortay J. 1998. Inactivation of isocitrate dehydrogenase kinase/phosphatase by 5'-[*p*-(fluorosulfonyl)benzoyl]adenosine is not due to the labeling of the invariant lysine residue found in the protein kinase family. Eur J Biochem 258: 579-585.

4) Zoller MJ, Nelson, NC, Taylor SS. Affinity labeling of cAMP-dependent protein kinase with p-fluorosulfonylbenzoyl adenosine. Covalent modification of lysine 71 J Biol Chem 1981, 256: 10837-10842.

5) Kamps MP, Taylor SS, Sefton BM. Direct evidence that oncogenic tyrosine kinases and cyclic AMP-dependent protein kinase have homologous ATP-binding sites. 1984. Nature 1984, 310: 589-592.

6) Ratcliffe SJ, Yi T, Khandekar SS. Synthesis and characterization of 5'-p-fluorosulfonylbenzoyl-2'(or 3')-(biotinyl)adenosine as an activity-base probe for protein kinases. J Biomolecular Screening 2007, 12: 126-132.

7) Bell IM, Stirdivant SM, Ahern J, Culberson JC, Darke PL, Dinsmore CJ, Drakas RA, Gallicchio SN, Graham SL, Heimbrook DC, Hall DL, Hua J, Kett NR, Kim AS, Kornienko M, Kuo LC, Munshi SK, Quigley AG, Reid JC, Trotter RB, Waxman LH, Williams TM, Zartman CB. Biochemical and structural characterization of a novel class of inhibitors of the type I insulin-like growth factor and insulin receptor kinases. Biochemistry 2005, 44: 9430-9440.

8) Yee M, Fas, SC, Stohlmeyer MM, Wandles TJ, Cimprich KA. A cell-permeable, activity based probe for protein and lipid kinases. J Biol Chem 2005, 280: 29053-29059.

9) Hanks SK, Hunter T. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. FASEB J. 1995, 9: 576-596

10) Knight ZA, Shokat KM. Features of selective kinase inhibitors. Chem Biol 2005, 12: 621-637

11) Moore LL, Fulton AM, Harrison ML, Geahlen RL. Anti-sulfonylbenzoate antibodies as a tool for the detection of nucleotide-binding proteins for functional proteomics. J Proteome Res 2004, 3: 1184-1190.

12) Liu Yi, Bishop A, Witucki L, Kraybill B, Shimizu E, Tsien J, Ubersax J, Blethrow J, Morgan DO, Shokat KM. Structural basis for selective inhibition of Src family kinases by PP1. Chemistry & Biology 1999, 6: 671-678.

13) Hanke JH, Gardner JP, Dow RL, Changelian PS, Brissette WH, Weringer EJ, Pollok BA, Connelly PA. Discovery of a novel, potent, and Src family-selective tyrosine kinase inhbitor. J Biol Chem 1996, 271: 695-701.

14) Zhu X, Kim, JL, Newcomb JR, Rose PE, Stover DR, Toledo LM, Zhao H, Morgenstern KA. Structural analysis of the lymphocyte-specific kinase Lck in complex with Src family selective kinase inhibitors. Structure 1999, 7: 651-661.

15) Schindler T, Sicheri F, Pico A, Gazit A, Levitski A, Kuriyan J. Crystal structure of Hck in complex with a Src family-selective tyrosine kinase inhibitor. Mol Cell 1999, 3: 639-648.

16) Sicheri F, Moarefi I, Kuriyan J. Crystal structure of the Src family tyrosine kinase Hck. Nature 1997, 385: 602-609.

17) Arnold LD, Calderwood DJ, Dixon RW, Johnston DN, Kamens JS, Munschauer R, Rafferty P, Ratnofsky SE. Pyrrolo[2,3-d]pyrimidines containing an extended 5-substituent as potent and selective inhibitors of lck I. Biorg Med Chem Let. 200, 10: 2167-2170.

18) Ugarkar BG, Castellino AJ, DaRe JM, Kopcho JJ, Wiesner JB, Schanzer JM, Erion MD. Adenosine kinase inhibitors. 2. Synthesis, enzyme inhibition, and antiseizure activity of diaryltubercidin analogues. J Med Chem. 2000, 43(15): 2894-2905.

19) Cohen MS, Zhang C, Shokat KM, Taunton J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. Science 2005, 308: 1318-1321.

20) Parker PJ. Antibodies to fluorylsulfonylbenzoyladenosine permit identification of protein kinases. FEBS 1993, 334: 347-350.

21) Carretero JC, Demillequand M, Ghosez L. Synthesis of α, β unsaturated sulphonates via the wittig-horner reaction. Tetrahedron 1987, 43: 5125-5134.

22) Roskoski R. Src protein-tyrosine kinase structure and regulation. Biochem Biophys Res Com 2004, 324: 1155-1164.

23) Ichihara J, Matsuo T, Hanafusa T, Ando T. The combination of potassium fluoride and calcium fluoride: a useful heterogeneous fluorinating reagent. J Chem Soc, Chem Commun. 1986, 10: 793-794

24) Xu W, Doshi A, Lei M, Eck MJ, Harrison SC. Crystal structure of c-Src reveal features of it autoinhibitory mechanism. Mol Cell 1999, 3: 629-638.

25) Missbach M, Jeschke M, Feyen J, Müller K, Glatt M, Green J, Susa M. A novel inhibitor of the tyrosine kinase Src suppresses phosphorylation of its major cellular substrates and reduces bone resorption in vitro and in rodent models in vivo. Bone 1999, 24: 437-449.

26) Burchat A, Borhani DW, Calderwood DJ, Hirst GC, Li B, Stachlewitz RF. Discovery of A-770041, a src-family selective orally active lck inhibitor that prevents organ allograft rejection. Biorg Med Chem Lett. 2006, 16: 118-122.

27) Bain J, Plater L, Elliott M, Shpiro N, Hastie CJ, Mclauchlan H, Klevernic I, Arthur JSC, Alessi DR, Cohen P. The selectivity of protein kinase inhibitors: a further update. Biochem J 2007, 408: 297-315.

Chapter 2: Electrophilic Kinase Inhibitors:

A Chemical Switch for Nucleophile Selectivity

Introduction:

Protein phosphorylation is a post-translational modification that dominates signal transduction and enzymatic regulation in eukaryotic cells. Transfer of the γ-phosphate of ATP to specific serine, theonine, or tyrosine residues is mediated by protein kinases. Missregulation of these important enzymes has dire consequences, manifesting in a number of human disorders, especially in many cancer. Considerable effort has been invested to develop inhibitors for specific members of this enzyme super-family[1]. The challenge of developing selective inhibitors is rooted in the high structural conservation of the ATP binding pocket within the active site of the 510 human protein kinases[2,3]. This hydrophobic cleft lies between the two lobes of the enzymes' catalytic core and is the target of the vast majority of small-molecule inhibitors[1,4]. Specific structural elements that determine compound selectivity among specific kinases are important lessons in inhibitor design, as revealed by x-ray crystallographic and SAR studies can provide important clues that guide the design of inhibitors for other distantly related kinases.

Src is the prototypical member of the Src family of tyrosine kinases. The human genome encodes eleven family members, all sharing a highly similar catalytic domain, along with SH2 and SH3 domains. These enzymes' modular design and regulation allow them to interact with and regulate a large number of different proteins and complexes. Src itself is a ubiquitous enzyme tied to cell migration, growth, and division[5,6,7]. Due in part to these functions, Src has recently become a target for cancer therapy. It is a potential therapeutic target in breast[8], colorectal[9], non-small-cell lung[10] cancers, and chonic myeloid leukemia, both alone and in conjunction with other mitogenic kinases[11]. Selective inhibitors of this kinase could be useful in treating a number of debilitating diseases. However, for all of these diseases, it is not yet clear whether Src alone, or Src in addition to other Src-family kinases must be inhibited to attain efficacy.

Previous work described in Chapter 1 produced an inhibitor with a surprising difference in potency against Src and close family member Hck. Compound **21** is a 5-(*p*-tolyl)pyrrolo[2,3-d]pyrimidine nucleoside with an electrophilic 5'-vinylsulfonate ester. The same nucleoside scaffold bearing a lysine reactive electrophile (compound **1**) showed little discrimination between these kinases (Figure 1-8). Both compounds demonstrated time-dependent inhibition of Src (Figure 1-5), and though it is possible Lys295 adds to the α,β unsaturated system, a sequence alignment of the two kinases suggests another intriguing possibility for selectivity of compound **21** (Figure 2-1a) [12].

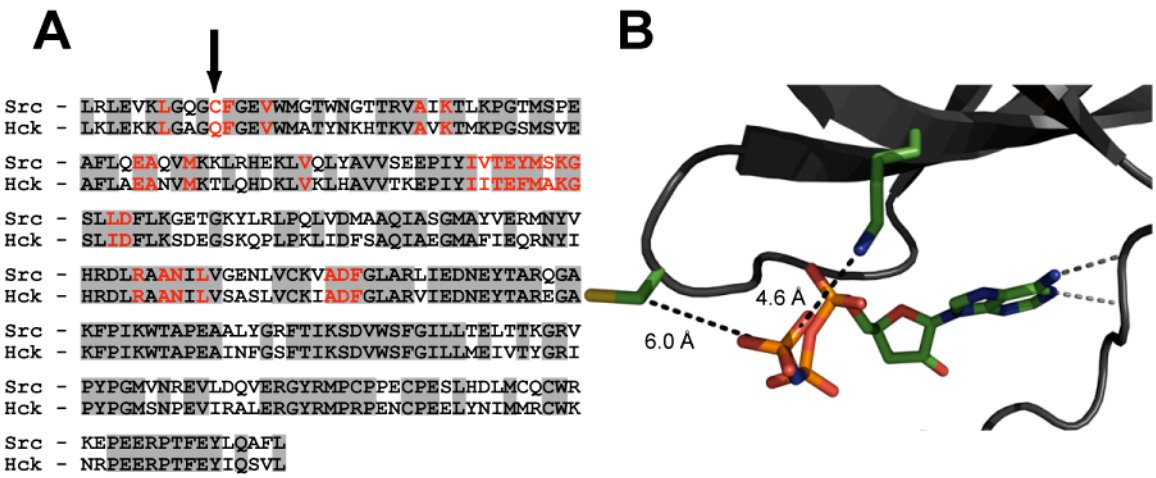


Figure 2-1: Src has a non-conserved cysteine in its ATP binding site. (A) Sequence alignment of the catalytic domains of Src related kianse Hck. Shaded areas denote regions of identity. ATP contacting residues are in red text. The arrow indicates C277 in the "glycine-rich" loop of Src. (B) The structure of AMP-PNP bound to Src reveals the close proximity of C277 to the gamma-phosphate of the nucleotide. Figure created with PyMOL from PDB 2SRC.

The catalytic domain of Src and Hck share 70% sequence identity and 25 of the 28 residues that contact ATP are the same. One very noTable exception is C277 in the "glycine-rich" loop of kinase subdomain I (Figure 2-1a) [13]. Several cysteine-reactive kinase inhibitors have been developed, including at least four that are in clinical trials. Cysteine-reactive electrophiles have been incorporated into advanced inhibitors of EGFR[14] and Btk[15] kinases to further boost their *in vivo* potency. Kinase inhibition by the

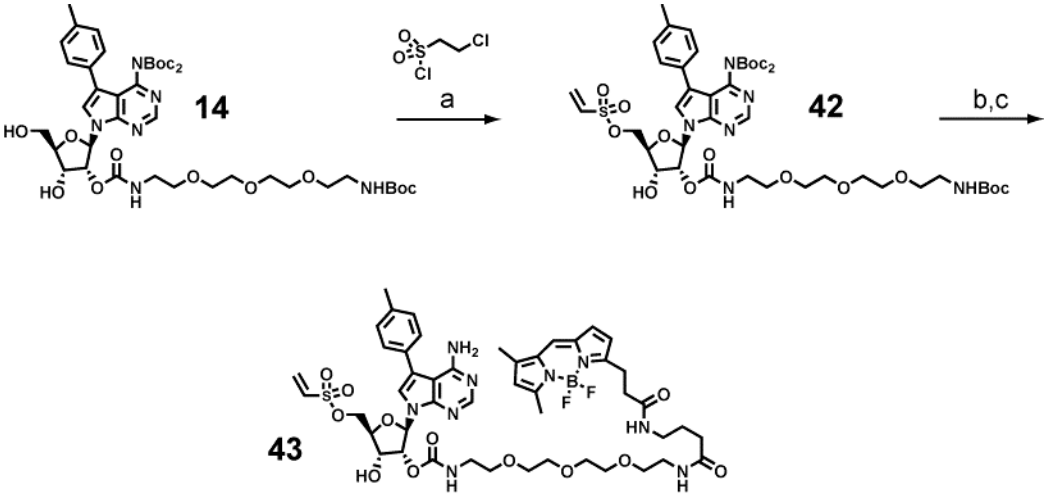
natural product hypothemycin relies on an α,β unsaturated ketone and its attack on a semi-conserved cysteine in ~10% of kinases at the base of the active site[16]. The Rsk-1/2 inhibitor, fmk, and the Btk inhibitor 32765 deserve special mention. The fluoromethylketone and acrylamide electrophiles, respectively, increase both the potency *and* selectivity of their semi-promiscuous scaffold[15,17].

We hypothesized that potent inhibition of Src by compound **21** derives from reaction with cysteine-277 in the "glycine-rich" loop, and the ability to target this non-conserved nucleophile can be exploited as a selectivity filter for kinase inhibition. C277 and the catalytic lysine K295 are within 6 Å of the γ-phosphate of ATP (Figure 2-1b). The lysine is the proposed nucleophilic target of compound **1** but both residues are candidates for reaction with the 5'-electrophiles of **1** and **21**. We present our investigation of inhibition and covalent binding by derivatives of compounds **1** and **21**. We delineate the nucleophile-electrophile pairings in the active site of Src, and demonstrate that these have implications beyond the Src-family kinases.

Results and Discussion:

Synthesis and testing of fluorescent electrophilic nucleosides

We synthesized a BODIPY tagged version of vinylsulfonate **21** to directly monitor covalent bond formation with Src. Compound **43** was made in thee steps from a previously synthesized intermediate **14** (Schemes 1-3 and 2-1). The product contained a mixture of 2' and 3' linkages in a ratio of approximately 1:1. Prior studies with the 5' electrophilic nucleoside FSBA showed little difference in the effective labeling of kinases by 2' or 3' conjugated biotin, so no further effort was spent on resolving the isomers of **43**[18]. A BODIPY conjugate of fluorosulfone **1** had been previously prepared (BODIPY-fluorosulfone **17**, Scheme 1-3).



Scheme 2-1: Synthesis of fluorescent 5'-vinylsulfonate nucleoside **43**. Reagents and conditions: (a) DIPEA, DCM (31%); (b) TFA, DCM; (c) BODIPY-NHS, DIPEA, DMF (48%).

BODIPY conjugates **43** and **17** and their 2' hydroxyl analogues **21** and **1** were tested for their ability to inhibit Src kinase. *In vitro* kinase assays showed that the tagged compounds are severely crippled in their ability to inhibit Src (Figure 2-2). Relative to vinylsulfonate **21**, BODIPY-vinylsulfonate **43** was greater than 100-fold less potent, and little to no inhibition was detected for the fluorosulfone **17**. These failures led us to try a new approach to compound labeling.

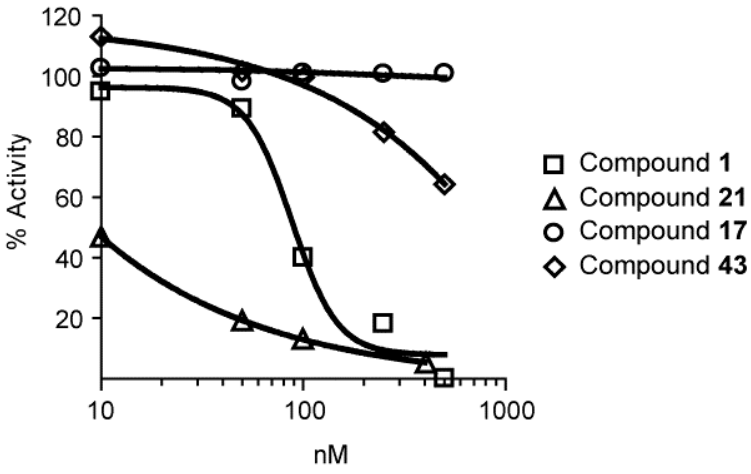


Figure 2-2: Large tags have a severe effect on Src inhibition by electrophilic nucleosides. Src was incubated with 2'-OH (**1** and **21**) or 2'-BODIPY (**17** and **43**) nucleosides for 30 min then diluted 1:4 into 125 μM ATP, ^{32}P-ATP and substrate peptide. After 30 min phosphotrasfer was quantitated. Values are expressed as a % of DMSO control.

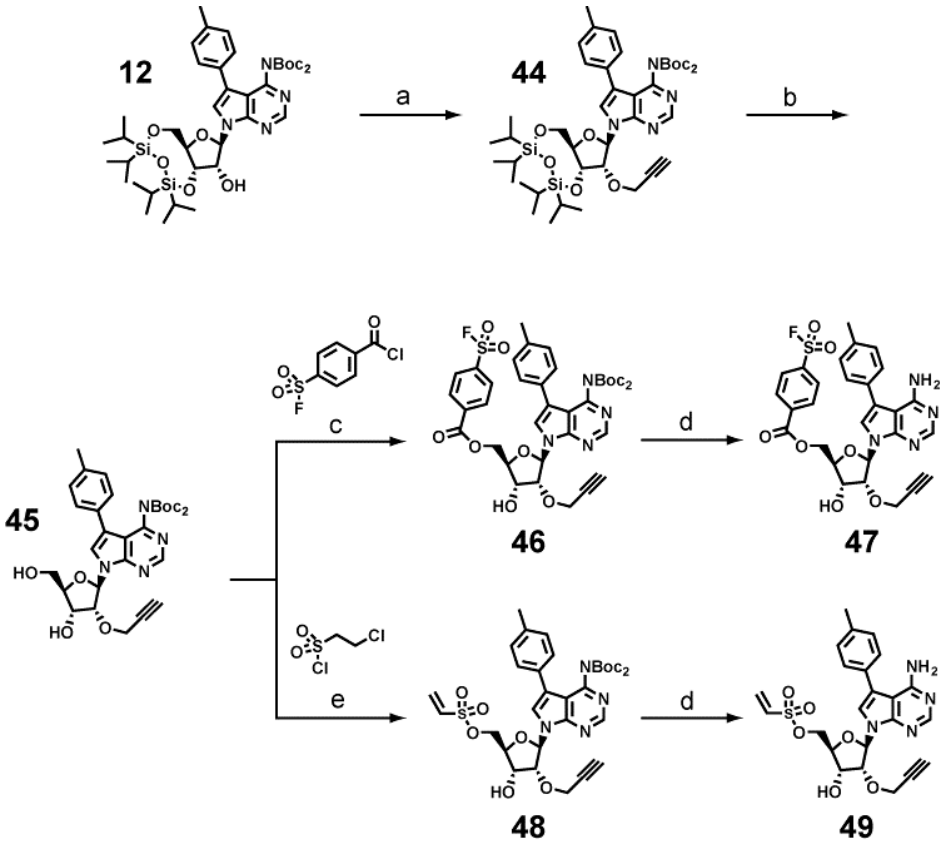
Synthesis of 2' propargylated nucleosides for click chemistry

Propargylated compounds **47** and **49** were synthesized from the 2',3'-protected intermediate **12** (Schemes 2-2 and 1-3). The 2'-propargyl ether is a relatively small structural change to the scaffold, and the copper-catalyzed Huisgen cyclization ("click" chemistry) would give us access to a variety of different tags[19]. A similar strategy has been used to detect in-cell occupancy of Rsk by a covalent inhibitor[20]. Alkylation of **12** after deprotection with stiochiometric sodium hydride produced a variety of mono, di, and tri-Boc species with varying degrees of propargylation. A biphasic reaction with aqueous sodium hydroxide and excess propargyl bromide gave **44** in good yields with no intermolecular Boc exchange. Synthesis of final electrophilic nucleosides **47** and **49** was accomplished in thee easy steps (Scheme 2-2).

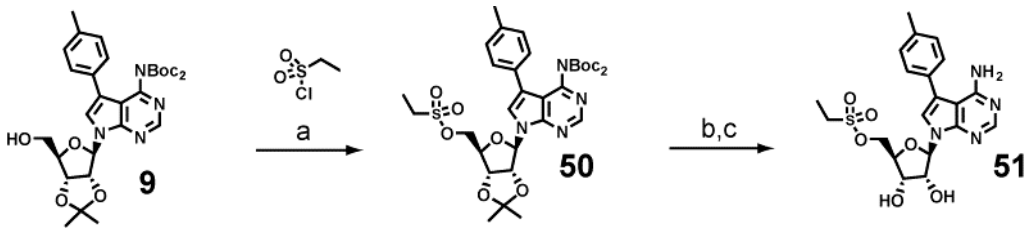
We synthesized ethyl sulfonate **51** as a non-electrophilic analogue of **21**. The compound was easily prepared from intermediate **9** using established chemistry (Scheme 2-3). It is nearly isosteric to **21** and maintains the phospho-mimetic sulfonate without the distal electrophilic center.

The 2'-propargyl tag increases the potency of the electrophilic nucleosides for Src and Hck while maintaining SAR trends

The new propargylated nucleosides were tested along with related compounds for their ability to inhibit Src and Hck kinase activity *in vitro*. All of the compounds, with the exception of the adenosine-derived FSBA, had IC_{50} values of ≤ 1 μM in the presence of 250 μM ATP (Table 1). Addition of a propargyl ether to the 5'-fluorosulfone nucleoside increased its potency five-fold for vs. Src and an ten-fold vs. Hck. The propargyl ether effect was less dramatic for the vinylsulfonate series, increasing potency ~3-fold for Src and Hck



Scheme 2-2: Synthesis of propargylated electrophilic nucleosides **47** and **48**. Reagents and conditions: (a) NaOH, propargyl bromide, toluene (77%); (b) TBAF, THF (88%); (c) HMPA (76%); (d) TMSOTf, DCM (**47**: 100%, **49**: 100%); (e) DIPEA, DCM (41%).



Scheme 2-3: Synthesis a non-electrophilic control ethyl 5'-sulfonate nucleoside **51**. Reagents and conditions: (a) DIPEA, DCM (97%); (b) TFA, DCM; (c) TFA, H_2O, THF (96%).

The 5'-vinylsulfonate nucleosides potently inhibited both Src and Hck, but the overall trend in structure-activity-relationships points to a role for Src C277 which is a glutamine in Hck. The nucleoside scaffold of our electrophilic inhibitors borrows from the structure of Src-family inhibitor PP1 (Figure 1-1). Literature shows PP1 is a more potent inhibitor of Hck than Src[21], and this is recapitulated here with PP1 and reversible inhibitor **51** (Table 1). This trend is continued with the 5'-fluorosulfone inhibitors **1** and **47** which were designed to target the catalytic lysine present in both enzymes. Selectivity between the kinases is inverted for the vinylsulfonates **21** and **49**. The ratio of Src and Hck IC_{50} values for non-electrophilic ethyl sulfonate **51** is 1:2 but inverts to 3:1 for the isosteric vinylsulfonate **21**. It could be argued that this ratio does not constitute significant selectivity, but its *in vitro* magnitude is a function of assay conditions. The two-stage binding of these covalent inhibitors is influenced by factors including reversible binding affinity, time of incubation, and ATP content[3,22]. The selectivity of electrophile **21** for Src over Hck increased dramatically when the kinases were pre-incubated with the compound in the absence of competing ATP (Figure 1-8).

	Src		Hck	
Compound	EC_{50} nM	SD	EC_{50} nM	SD
FSBA	>10000	NA	>10000	NA
PP1	555	174	146	68
1	1060	200	521	349
47	200	21	42	12
21	26	6	72	29
51	423	85	178	36
49	9	5	28	6

Table 2-1: Inhibition of Src and Hck by electrophilic nucleosides and related compounds. Kinases were incubated with compounds and 250 μM ATP. After 30 min the reactions were diluted 1:1 into a solution containing ^{32}P-ATP and substrate peptide and incubated an additional 30 min. ^{32}P incorporation was measured. Values are the mean and standard deviation (SD) of 3 experiments.

Compounds 47 and 49 show different rates of covalent reaction with Src and Hck

To test for covalent binding recombinant Src and Hck were incubated with propargyl ethers **47** and **49** followed by click chemistry with TAMRA azide. No ATP was included and inhibitors were used at saturating concentrations to approximate a pseudo-first-order reaction. Figure 2-3 shows the results after attaching a rhodamine tag with "click" chemistry and SDS-PAGE. Both compounds rapidly form a stable covalent bond with Src; binding saturates before the first time point of 40 sec. No binding was seen when the kinase was incubated with the compounds in the presence of 1% SDS (data not shown). As with Src, compound **47** formed a stable covalent bond with Hck within 40 sec. There were striking differences in labeling of the two kinases by the vinylsulfonate **49**. Even after 5 min labeling of Hck by vinylsulfonate **49** was almost undetectable Lacking C277, the slow increase in labeling by **49** is most likely due to reaction with the catalytic lysine (or other nucleophile) in the ATP binding site.

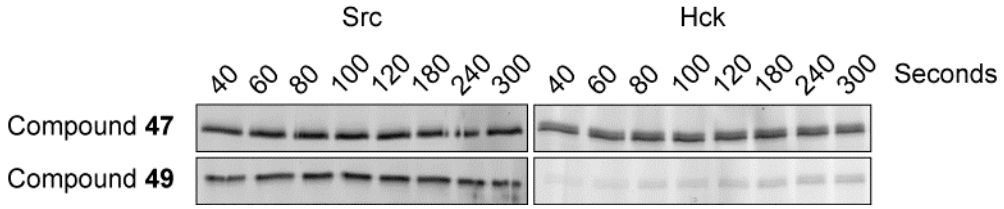


Figure 2-3: Labeling of recombinant Src and Hck by 5'-fluorosulfone **47** or vinylsulfonate **49**. Kinases were incubated with 5 μM compound for the indicated time denatured with 1% SDS, submitted to the click reaction with a rhodamine azide, resolved by SDS-PAGE, and scanned for fluorescence.

Fluorosulfone 47 irreversibly inhibits Src-family kinases *in vivo* in a K295-dependent manner

COS-7 cells expressing cytoplasmic FLAG-tagged Src were treated with compound **47** and assayed for covalent binding and functional inhibition. Src was the most intensely labeled band in the lysates, with little increase in signal between samples treated with 1-3 μM of **47** (Figure 2-4a). The same lysates were assayed for Src auto-phosphorylation.

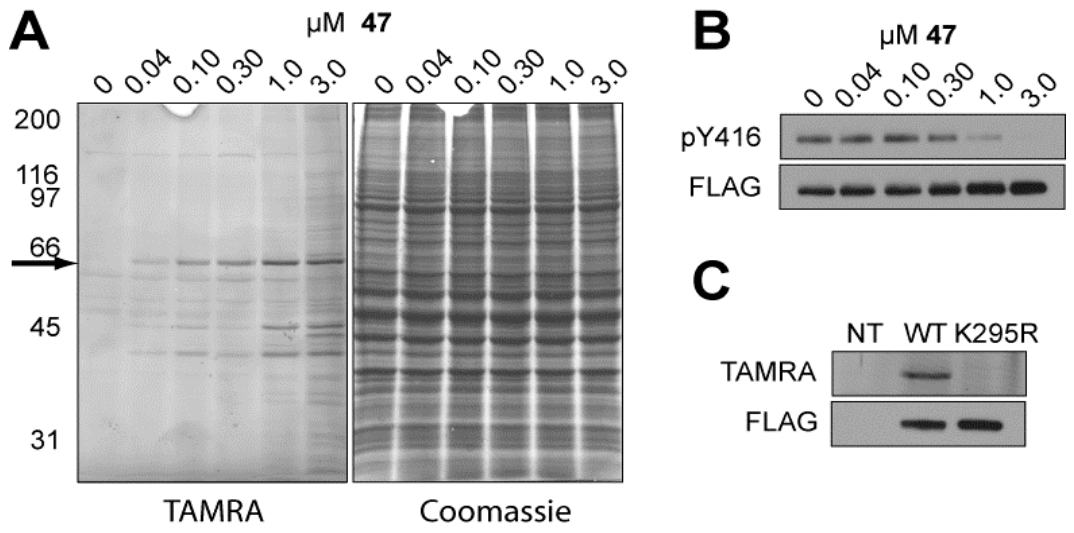


Figure 2-4: *In vivo* covalent inactivation of Src by fluorosulfone **47**. COS-7 cells expressing N-terminal FLAG-tagged Src were treated with the indicated concentrations of **47** for 1 h. Cellular lysates were submitted to the click reaction with rhodamine-azide, resolved by SDS-PAGE, and either scanned for fluorescence and Coomassie stained (A), or blotted for Src autophosphorylation at Y416 (B). The arrow indicates the MW of the FLAG-Src construct. (C) Untransfected COS-7 cells (NT) or those expressing wild type (WT) or mutant (K295R) Src were treated with 1 uM **47** for 1 h and analyzed as in (A), then blotted for FLAG-Src.

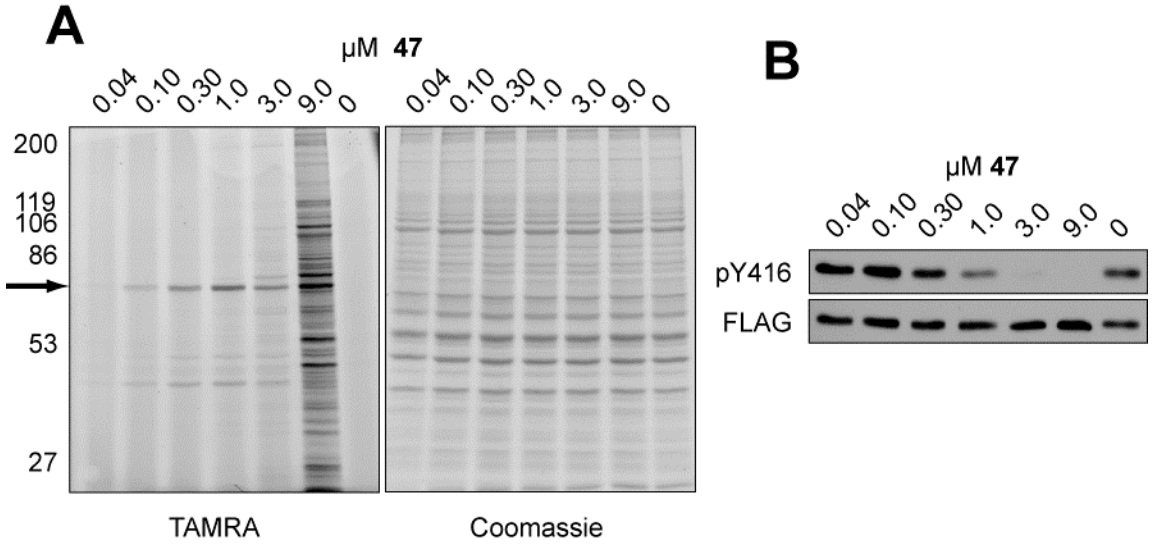


Figure 2-5: *In vivo* covalent inactivation of Hck by fluorosulfone **47**. COS-7 cells expressing N-terminal FLAG-tagged Hck were treated with the indicated concentrations of **47** for 1 h. Cellular lysates were submitted to the click reaction with rhodamine-azide, resolved by SDS-PAGE, and either scanned for fluorescence and Coomassie stained (A), or blotted for Hck autophosphorylation (B). The arrow indicates the MW of the FLAG-Hck construct.

The intensity of labeled Src correlated with the loss of pY416, suggesting that Src inhibition involves covalent bond formation between the fluorosulfone and K295 (Figure 2-4b). Consistent with this model mutation of K295 to a non-nucleophilic arginine eliminated all covalent binding (Figure 2-4c). This is a conservative mutation, and though detrimental to the enzyme's catalytic activity, does not significantly affect nucleotide binding[23].

Hck was also labeled and inhibited *in vivo* by the lysine-selective fluorosulfone **47**. Labeling and inhibition of Hck saturated at similar concentrations of **47** as Src (Figure 2-5). The *in vivo* IC_{50} values for both kinases are surprisingly close to the *in vitro* values (Table 2-1). There is a larger difference in the *in vitro* and *in vivo* values for Hck, which binds with a greater reversible affinity. This difference along with the strong correlation between labeling and pY416 inhibition for both enzymes highlights the covalent mechanism of vinylsulfonate **47**.

Vinylsulfonate 49 selectively reacts with C277 in Src kinase's active site

Experiments comparing the action of vinylsulfonate **49** toward Src and Hck (Table 2-1 and Figure 2-3) suggested that the vinylsulfonate electrophile may derive its Src selectivity by forming a covalent bond with C277, a position occupied by Gln in Hck. Unlike fluorosulfone **47**, the 5'-vinylsulfonate **49** labeled both wild-type and K295R Src (Figure 2.-6a), thus ruling out K295 as the relevant nucleophile. To directly assess the role of C277, I treated cells expressing wild-type (WT) and C277Q Src with vinylsulfonate **49** and quantified pY416 levels after a brief "washout" period in which cells were incubated in media lacking the inhibitor. Lysates from cells expressing the C277Q mutant showed a full recovery of kinase activity while WT Src was potently inhibited by vinylsulfonate **49** in a dose-dependant manner (Figure 2-6b). Once again, Src inhibition correlated with covalent bond formation as revealed by "click" chemistry. WT Src was robustly labeled and there was little to no labeling of the mutant. Labeling

by fluorosulfone **47** and robust auto-phosphorylation show that the C277Q mutation affects neither binding to the nucleoside scaffold of **49** nor catalytic activity.

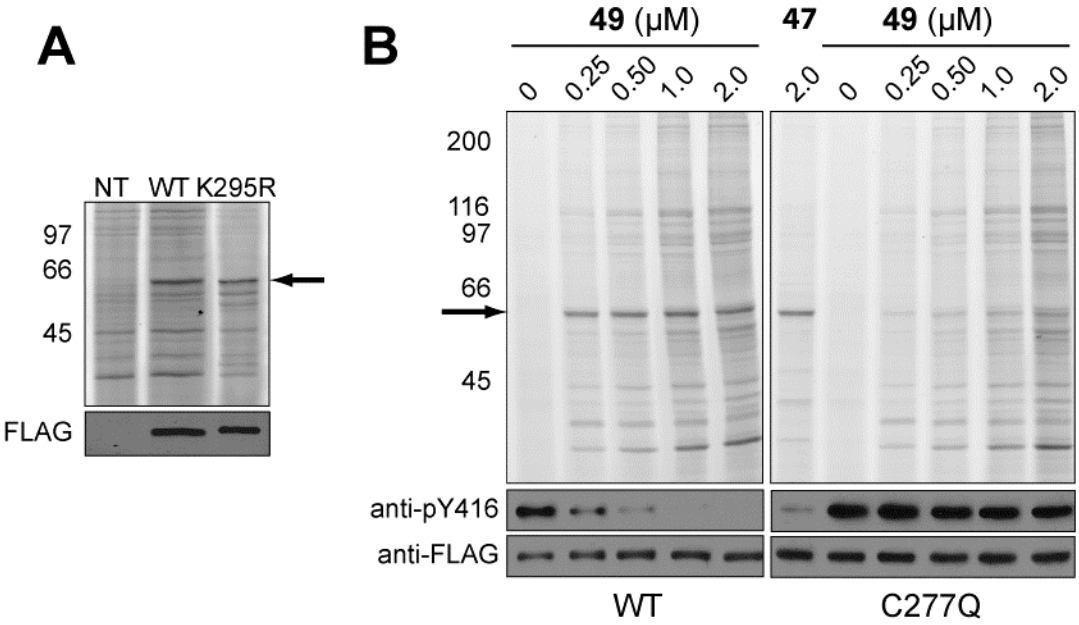


Figure 2-6. Cysteine-dependent inhibition and labeling of Src. (A) Untransfected COS-7 cells (NT) or those expressing wild type (WT) or mutant (K295R) FLAG-Src were treated with 1.0 □M **49** for 20 min. Lysates were submitted to the click reaction with a rhodamine azide, resolved by SDS-PAGE, scanned for fluorescence, and blotted for protein expression. (B) COS-7 cells expressing wild type (WT) or mutant (C277Q) FLAG-Src were treated with the indicated concentrations of **49** or 2.0 μM **47** for 30 min, washed, and allowed to recover for 15 min in compound-free media. After rhodamine conjugation, the lysates were resolved by SDS-PAGE and either scanned for fluorescence, or blotted to detect Src autophosphorylation at Y416 and total FLAG-Src.

We made the reciprocal Q277C mutant in Hck and assayed its ability to form a covalent bond with vinylsulfonate **47**. As expected from *in vitro* assays (Figure 2-3), wild-type Hck was not detectably labeled in cells (Figure 2-7). By contrast Q277C Hck was efficiently labeled. These experiments confirm our hypothesis that the 5'-vinylsulfonate electrophile mediates potent and selective inhibition of Src kinase by targeting the non-conserved cysteine 277 in the “glycine-rich” loop.

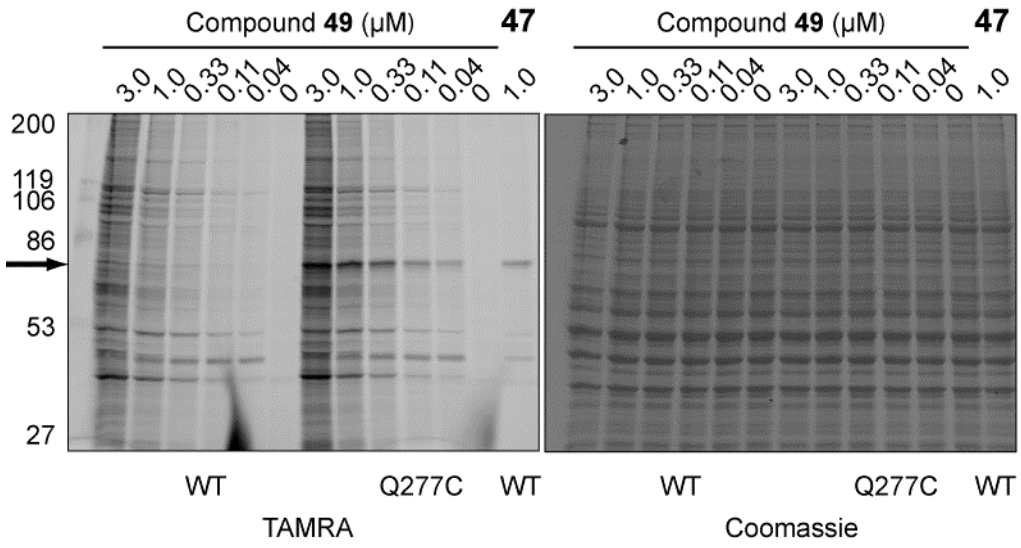


Figure 2-7: Q277C mutation renders Hck susceptible to labeling by vinylsulfonate **49**. COS-7 cells expressing wild type (WT) or mutant (Q277C) FLAG-Hck were treated with the indicated concentrations of **49** or **47** for 30 min, washed, and lysed. Lysates were submitted to the click reaction with a rhodamine-azide, resolved by SDS-PAGE, scanned for fluorescence and Coomassie stained.

A crystal structure of Src reveals how nucleoside 49 binds in the active site

We collected data from crystals of the Src kinase domain treated with compound **49** (collaboration with Haralambos Hadjivassiliou and Ville Paavilainen, UCSF) . The crystals diffracted to 2.2 Å, and the majority the catalytic domain could be modeled with the exception of residues 300-304 between β3 and αC, and 406-424 in the activation loop. The compound can be clearly seen in the ATP binding site (Figure 2-8a). Like PP1, the pyrrolopyrimidine of **49** superimposes with the adenine of ATP, and the *p*-toly group projects back into the hydrophobic pocket. The one major difference between Src-bound AMP-PNP and compound **49** is an inward rotation of the ribose of **49**. This inward rotation allows the 2' propargyl group to access a hydrophobic surface below the furanose ring. This rotation and increased hydrophobic contact likely explain the increased potency of the propargylated inhibitors (Table 2-1).

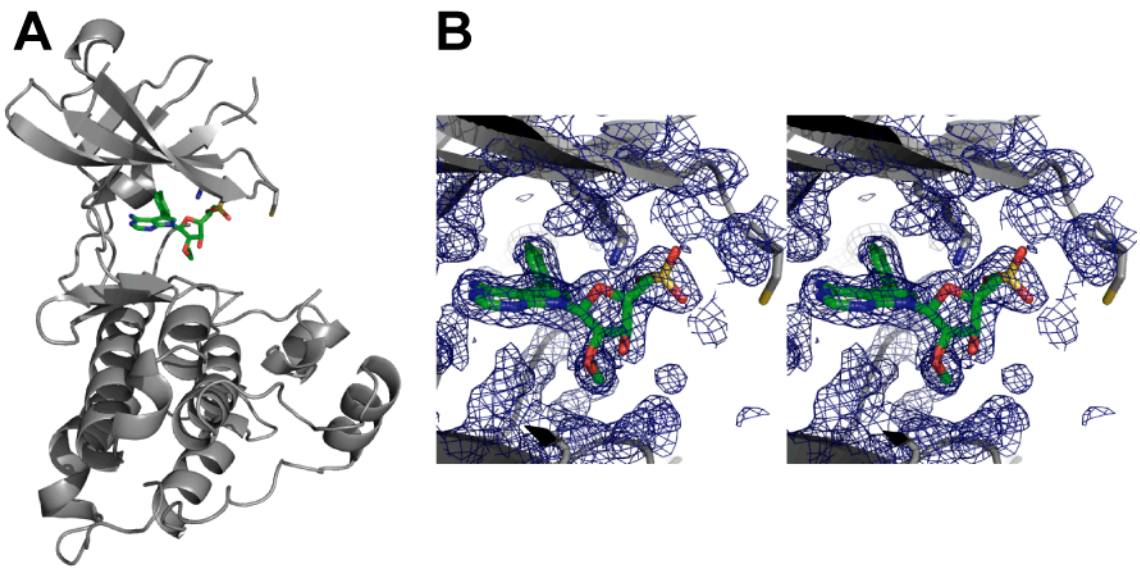


Figure 2-8: Crystal structure of compound **49** bound to the catalytic domain of Src kinase. (A) Ribbon diagram of the Src kinase domain showing compound **49** bound in the active site. (B) Stereo diagram of the electron density at the site of **49** binding. No electron density is seen for the C277. Data for this Figure were produced by Haralambos Hadjivassiliou and Ville Paavilainen (Taunton Laboratory, UCSF)

Unfortunately, multiple independent rounds of refinement failed to show electron density for a covalent bond between the 5'-vinlysulfonate and C277 (Figure 2-8b). Mass spectrometry with crystals used for data collection clearly demonstrated the covalent and stoichiometric addition **49**, but there was not enough density to model the sidechain of C277. We tentatively conclude that the vinylsulfonate reacts with C277, but the bond is unstable in the crystal or resides in multiple conformations.

C277 of Src is also found in the FGFR family kinases

Selective targeting of C277 in the active site of Src has implications beyond the Src family. The "glycine-rich" loop cysteine is found in a small number of other human kinases, most of which are within the Src and fibroblast growth factor receptor (FGFR) kinase families (Figure 2-9). Interestingly, this cysteine has been implicated in redox-mediated inactivation of both Src and FGFR1[24]. The FGFR family has a valine in the

gatekeeper position, similar in size to the threonine that allows the *p*-tolyl group of PP1 to bind to Src. This combination of "selectivity filters," shared between Src and FGFR1-4 led us to test our cysteine-directed vinylsulfonate inhibitors against FGFR3, the mutation of which has been implicated in 15% of multiple myelomas, an incurable plasma cell cancer[25].

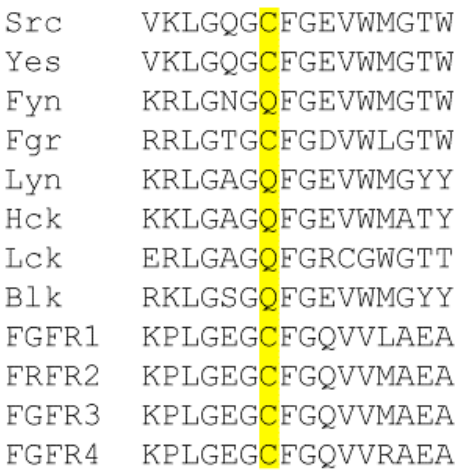


Figure 2-9: Sequence alignment of the "glycine-rich" loop in kinase subdomain I of the Src and FGFR families of tyrosine kinases. Src family members Frk, Brk and Srm have been omitted. The cysteine targeted by vinylsulfonate **49** in Src is highlighted in yellow.

Our collaborators Sumin Kang and Jing Chen (Emory University) found that vinylsulfonate **21** abolishes FGFR3 mediated signaling in a cysteine-dependent manner. Ba/F3 cells expressing the constitutively active TEL-FGFR3 fusion protein showed a dose-dependent decrease in cell viability when treated with **21** (Figure 2-10). In the absence of IL-3, these cells rely on continuous FGFR kinase activity for survival which was blocked by low minro-molar concentrations of vinylsulfonate **21** which had no effect on the viability of the parental cell line[25]. Mutation of the "glycine-rich" loop cysteine of FGFR3 to alanine (C482A) conferred complete resistance to compound **21**. Consistent with these results, TEL-FGFR3 cells were unaffected by ethylsulfonate **51** with lacks the electrophile.

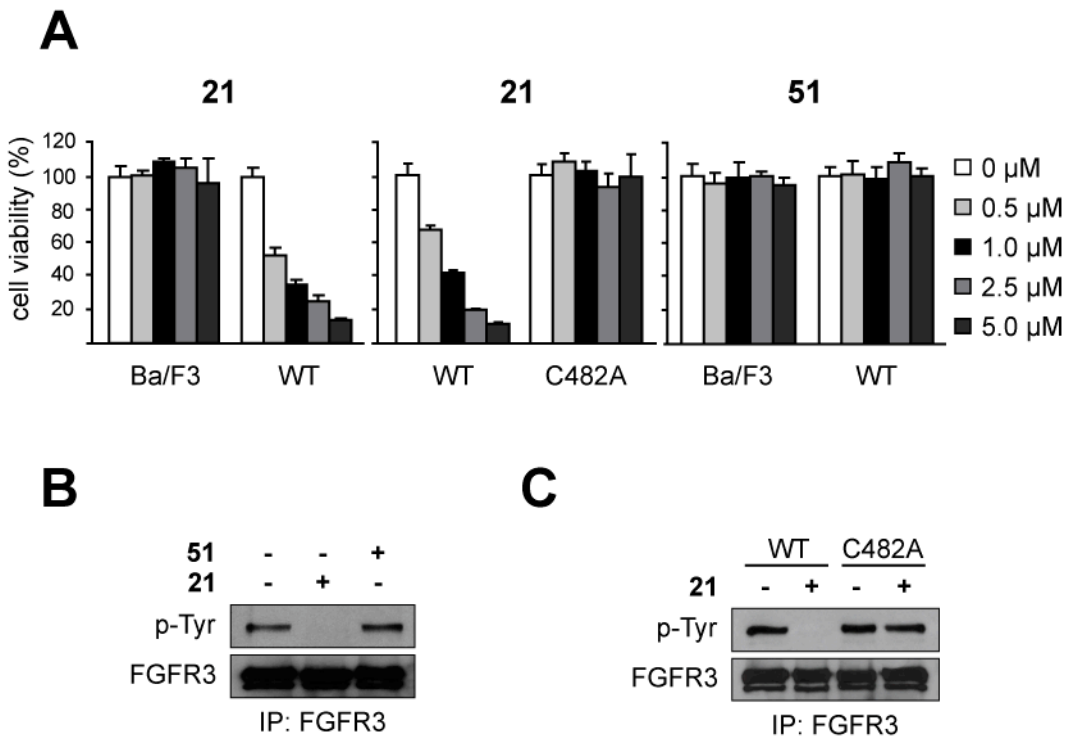


Figure 2-10: Cysteine-dependent inhibition of FGFR3. (A) Ba/F3 cells stably expressing wild type (WT) or mutant (C482A) TEL-FGFR3, along with the parental Ba/F3 cells, were treated with vinylsulfonate **49** or ethyl sulfonate **51** and assayed for cell viability after 72 h. (B) TEL-FGFR3 expressing cells were treated with 1.0 μM **49** or **51** for 4 h. TEL-FGFR3 was immunoprecipitated from cell lysates and blotted for phospho-tyrosine. (C) Cells expressing wild type (WT) or mutant (C482A) TEL-FGFR3 were treated with 1.0 uM **49** and analyzed as in (B). All data for this Figure was produced by Dr.Sumin Kang in the Chen Laboratory at Emory University.

Consistent with its effects on TEL-FGFR3 Ba/F3 cell viability, vinylsulfonate **21** specifically inhibited FGFR3 auto-phosphorylation. TEL-FGFR3 was immunoprecipitated after treatment of cells with **21** or **51** and probed for auto-phosphorylation. The electrophilic inhibitor **21** completely inhibited auto-phosphorylation while the ethylsulfonate had no effect (Figure 2-10b). As with cell viability, the C482A mutation conferred resistance to **21** in the auto-phosphorylation assays (Figure 2-10c). These data show the power of precise nucleophile-electrophile pairing. Though specific inhibition by **21** requires that the compound first bind non-covalently, the related but non-electrophilic compound **51** is completely inactive (Figures 2-10a & b). Electrophilic

compounds like **21** can reversibly occupy the active site of a target (or off-target) kinase below a biologically relevant theshold, and rely on the time-dependent action of their chemical warhead to deplete targets containing a complimentary nucleophile.

Conclusion:

We have developed two novel 5'-electrophilic nucleosides that kinetically discriminate between two nucleophilic residues, K295 and C277, within the active site of Src kinase. We developed a “click” chemistry tagging system to show that 5'-vinylsulfonate **49** selectively reacts with a non-conserved cysteine (C277) in the “glycine-rich” loop. This electrophile selectively inhibits and labels Src whereas the related kinase Hck is resistant. Introducing a cysteine into Hck confers sensitivity to **49.** We demonstrate the portability of this nucleophile-electrophile pairing showing that FGFR3 kinase is also inhibited by **49** in a cysteine dependent manner.

Methods:

Synthesis

Compound **42**: Diisopropylethylamine (21 uL, 0.12 mmol) was added to a solution of **14** (21 mg, 0.024 mmol) and 2-chloroethanesulfonyl chloride (2.9 µL, 0.028 mmol) in methylene chloride (0.5 mL) and stirred at room temperature overnight. The reaction was concentrated under a stream of argon and spotted onto a preparative silica gel thin layer chomatography plate. Pure **42** (7.2 mg, 31%) was eluted with methanol in methylene chloride (7.5%). 1HNMR (DMSO): δ 1.28 (s, 9H), 1.31 (s, 9H), 1.45 (s, 9H), 2.39 (s, 3H), 3.35 (m, 4H), 3.62 (m, 12H), 4.21 (m, 1H), 4.37 (dd, J = 4.0, 11.2 Hz, 1H), 4.46 (dd, J = 2.8, 11.2 Hz, 1H), 4.73 (m, 1H), 5.30 (s, 1H), 5.35 (s(br), 1H), 5.52 (t, J = 4.8 Hz, 1H), 5.81 (s(br), 1H), 5.98 (d, J = 9.6 Hz, 1H), 6.34 (d, J = 16.8 Hz, 1H), 6.42 (dd, J = 9.6, 16.8 Hz, 1H), 6.55 (d, J = 4.8 Hz, 1H), 7.19 (d, J = 8.0 Hz, 2H), 7.37 (d, J = 8.0 Hz, 2H), 7.42 (s, 1H), 8.82 (s, 1H). LC/MS [$M+H^+$,$M+Na^+$] m/z = 965, 987.

Compound **43**: Trifluoroacetic acid (0.5 mL) was added to a solution of **42** (4.4 mg, 4.5 nmol) in methylene chloride (0.5 mL) and stirred for one hour at room temperature. Volatiles were removed *in vacuo* and the residue was resuspended in dimethylformamide (0.1 mL). To this solution was added Bodipy-NHS ester (2.5 mg, 5 nmol; Molecular Probes – D6102) and Diisopropylethylamine (6 uL, 0.03 mmol). The reaction was stirred at room temperature for 2 h then concentrated *in vacuo*. Compound **43** (2.3 mg, 48%) was purified as a mixture of 2' and 3' linked isomers (2:1) by high pressure liquid chomatography (55-65% methanol and 0.1% trifluoroacetic acid in water). LC/MS [$M+H^+$] m/z = 1052.48.

Compound **44**: Compound **12** (100 mg, 0.13 mmol) and tetrabutylammonium iodide were dissolved in a solution of propargyl bromide in toluene (1 mL, 80%). Aqueous

sodium hydroxide (0.6 mL, 60% w/v) was added drop-wise while stirring vigorously. Upon completion, the reaction was diluted into ethyl acetate and washed with water and brine. The organic was dried over sodium sulfate and concentrated *in vacuo.* Compound **44** (80 mg, 77%) was purified by silica gel chomatography (1:4 ethyl acetate:hexanes). ^{1}HNMR ($CDCl_3$): δ 0.93 (m, 3H), 0.98 (m, 3H), 1.07 (m, 22H), 1.25 (s, 9H), 1.33 (s, 9H), 2.38 (s, 3H), 2.38 (s, 1H), 4.02 (dd, J = 13.2, 2.4, 1H), 4.16 (m, 1H), 4.22 (d, J = 13.2, 1H), 4.44 (d, J = 4.8, 1H), 4.59 (m, 1H), 4.66 (m, 1H), 4.73 (m, 1H), 6.33 (s, 1H), 7.16 (d, J = 8.0, 2H), 7.53 (d, J = 8.0, 2H), 7.53 (s, 1H), 8.81 (s, 1H). HMS calcd [$M+H^+$] m/e 837.4290; found 837.4305.

Compound **45**: Tetrabutylamoniumfluoride (0.29 mL 1M solution, 0.3 mmol) was added to a solution of compound **44** (80 mg, 0.1 mmol) in tetrahydrofuran (1 mL) and stirred at room temperature for 15 min. The reaction was diluted into methylene chloride and washed with water. The organic was dried over sodium sulfate, concentrated, and submitted to silica gel chomatography (2.5:1.5 ethyl acetate:hexanes) to yield pure compound **45** (50 mg, 88%). ^{1}HNMR ($CDCl_3$): δ 1.27 (s, 9H), 1.34 (s, 9H), 2.25 (t, J = 2.4, 1H), 2.39 (s, 3H), 2.68 (s, 1H), 3.81 (m, 1H), 4.02 (m, 1H), 4.11 (dd, J = 16.0, 2.4, 1H), 4.20 (dd, J = 16.0, 2.4, 1H), 4.38 (s, 1H), 4.65 (d, J = 4.8, 1H), 5.09 (dd, 4.8, 7.2, 1H), 5.89 (d, J = 7.2, 1H), 6.09 (dd, J = 11.6, 1.6, 1H), 7.19 (d, J = 8.0, 2H), 7.27 (s, 1H), 7.34 (d, J = 8.0, 2H), 8.82 (s, 1H). HMS calcd [$M+H^+$] m/e 595.2768; found 595.2755.

Compound **46**: Hexamethylphosphoramide (1 mL) was added to **45** (150 mg, 0.25 mmol) and 4-(fluorosulfonyl)benzoyl chloride (84 mg, 0.38 mmol) and stirred at room temperature under argon overnight. The reaction was diluted into methylene chloride and washed with aqueous sodium bicarbonate then 0.1 N HCl. The organic was dried over sodium sulfate, concentrated and submitted to silica gel chomatography (1:1 ethyl

acetate:hexanes) to yield pure **46** (150mg, 76%). ^{1}HNMR ($CDCl_3$): δ 1.27 (s, 9H), 1.33 (s, 9H), 2.40 (s, 3H), 2.43 (dd, J = 2.8, 1H), 2.75 (d, J = 8.0, 1H), 4.36 (m, 1H), 4.49 (dd, J = 16.0, 2.8, 1H), 4.55 (dd, J = 16.0, 2.8, 1H), 4.58 (m, 1H), 4.66 (dd, J = 5.6, 2.8, 1H), 4.73 (dd, J = 12.4, 4.0, 1H), 4.82 (dd, J = 12.4, 2.8, 1H), 6.45 (d, J = 2.8, 1H), 7.16 (d, J = 8.0, 2H), 7.26 (d, J = 8.0, 2H), 7.36 (s, 1H), 7.84 (d, J = 9.2, 2H), 8.17 (d, J = 9.2, 2H), 8.81 (s, 1H).

Compound **47**: Trimethylsilyltrifluoromethanesulfonate (0.26 mL, 1.4 mmol) was added to a solution of **46** (150 mg, 0.19 mmol) in methylene chloride (1.7 mL). The reaction was stirred at room temperature overnight then diluted into methylene chloride and washed with aqueous sodium bicarbonate. The organic was concentrated *in vacuo* and **47** (112 mg, 100%) was purified by silica gel chomatography to yield a amorphous white solid (2.5 % methanol in methylene chloride). ^{1}HNMR (DMSO): δ 2.43 (s, 3H), 2.44 (t, J = 2.4, 1H), 3.36 (s(b), 1H), 4.35 (m, 1H), 4.46 (dd, J = 16.0, 2.4, 1H), 4.53 (dd, J = 16.0, 2.4, 1H), 4.63 (m, 2H), 4.68 (dd, J = 12.4, 4.0, 1H), 4.81 (dd, J = 12.4, 2.8, 1H), 5.47 (s(b), 2H), 6.37 (d, J = 2.8, 1H), 7.05 (s, 1H), 7.25 (d, J = 2.8, 2H), 7.25 (d, J = 2.8, 2H), 7.85 (d, J = 8.8, 2H), 8.17 (d, J = 8.8, 2H), 8.25 (s, 1H).

Compound **48**: Diisopropylethylamine (9 uL, 0.05 mmol) was added to a solution of **45** (10 mg, 0.017 mmol) in methylene chloride (1 mL) at -50º C. 2-chloroethanesulfonyl chloride (2 uL, 0.02 mmol) was added over five min and the solution was allowed to come to room temperature over five h then stirred overnight. The reaction was diluted into ethyl acetate, washed with 10 % citrate (pH 4.0) then brine. The organic was dried over sodium sulfate, concentrated, and **48** (4.6 mg, 41%) was purified by preparative silica gel layer chomatography (3:1, ethyl acetate:hexanes). ^{1}HNMR ($CDCl_3$): δ 1.29 (s, 9H), 1.30 (s, 9H), 2.39 (s, 1H), 2.42 (t, J = 2.4, 1H), 2.76 (d, J = 6.0, 1H), 4.29 (dt, J = 5.2,

3.2, 1H), 4.35 (dd, J = 16.0, 2.4, 1H), 4.39 (dd, J = 11.2, 3.2, 1H), 4.44 (dd, J = 16.0, 2.4, 1H), 4.48 (dd, J = 11.2, 2.8, 1H), 4.56, (m, 1H), 4.61 (dd, J = 4.4, 5.2, 1H), 6.03 (d, J = 9.6, 1H), 6.38 (d, J = 16.8, 1H), 6.47 (dd, J = 16.8, 9.6, 1H), 6.49 (d, J = 4.4, , 1H), 7.20 (d, J = 8.0, 1H), 7.37 (d, J = 8.0, 1H), 7.45 (s, 1H), 8.83 (s, 1H).

Compound **49**: Trifluoroacetic acid (0.25 mL) was added to **48** (4.6 mg, 0.007 mmol) in methylene chloride (0.25 mL). After 45 min, the solvent blown off with a stream of argon and residual volatiles were removed under high vacuum to yield pure **49** which was dissolved directly in deuterated dimethylsulfoxide for analysis and biological assays. ^{1}HNMR (DMSO): δ 2.38 (s, 3H), 3.42 (t, J = 2.0, 1H), 4.19 (m, 1H), 4.22 (dd, J = 15.6, 2.0, 1H), 4.32 (dd, J = 15.6, 2.0, 1H), 4.36 (m, 3H), 4.54 (m, 1H), 6.27 (d, J = 10.0, 1H), 6.33 (d, J = 16.4, 1H), 6.34 (d, J = 6.0, 1H), 6.98 (dd, J = 16.4, 10.0, 1H), 7.34 (d, J = 8.4, 1H), 7.41 (d, J = 8.4, 1H), 7.73 (s, 1H), 8.47 (s, 1H).

Compound **50**: Diisopropylethylamine (40 µL, 0.22 mmol) was added to a solution of **9** (22 mg, 0.037 mmol) and ethanesulfonyl chloride (11 uL, 0.11 mmol), on an ice bath in methylene chloride (0.5 mL). The ice was allowed to melt as the reaction stirred overnight. The reaction was diluted into methylene chloride and washed with 0.1 N HCl then brine. The organic phase was concentrated *in vacuo* and **50** (25 mg, 97%) was obtained by preparative silica gel thin-layer chomatography (2% methanol in methylene chloride). ^{1}HNMR ($CDCl_3$): δ 1.26 (t, J = 7.6, 3H), 1.29 (s, 18H), 1.41 (s, 3H), 1.65 (s, 3H), 2.39 (s, 3H), 2.99 (q, J = 7.6, 2H), 4.41 (m, 1H), 4.48 (dd, J = 13.2, 3.6, 1H), 4.49 (m, 1H), 5.12 (dd, J = 6.4, 3.6, 1H), 5.38 (dd, J = 6.4, 2.4, 1H), 6.35 (d, J = 2.4, 1H), 7.19 (d, J = 8.0, 2H), 7.35 (s, 1H), 7.36 (d, J = 8.0, 2H), 8.85 (s, 1H).

Compound **51**: Trifluoroacetic acid (0.5 mL) was added to a solution of compound **50**

(25 mg, 0.036 mmol) in methylene chloride (0.5 mL). The reaction was stirred at room temperature for 45 min, and the solvent was removed under a stream of argon. The residual oil was dissolved in tetrahydrofuran (0.1 mL) and a mixture of trifluoroacetic acid and water (9:1, 0.9 mL) was added. After 30 min, the solvent was removed *in vacuo* and compound **51** (15.5 mg, 96%) was purified by high pressure liquid chomatography (10 – 80 % acetonitrile in water with 0.1% trifluoroacetic acid). 1HNMR (DMSO): δ 1.17 (t, J = 7.6, 3H), 2.38 (s, 3H), 3.34 (q, J = 7.6, 2H), 4.15 (m, 1H), 4.41 (m, 2H), 4.47 (m, 1H), 6.23 (d, J = 6.0, 1H), 7.32 (d, J = 8.0, 1H), 7.39 (d, J = 8.0, 2H), 7.67 (s, 1H), 8.39 (s, 1H).

Kinase assays

Src and Hck were obtained from Upstate. Substrate peptide (LEEIYGEFKK) was synthesized in house, and γ^{32}P-ATP (6000 Ci/mmol) was purchased from Perkin Elmer.

Inhibition of Src by compounds 1, 17, 21, and 43 (Figure 2-3): Src (5 nM) was incubated at room temperature with inhibitors and 100 μM ATP in reaction buffer (20 mM HEPES (pH 7.4), 10 mM $MgCl_2$, 0.2 mM EDTA, and 1 mg/mL BSA) for 30 min then diluted 1:4 with reaction buffer containing 125 μM ATP, 125 μM substrate, and 0.2 mCi/mL γ^{32}P-ATP. After 30 min, the reactions were spotted onto phosphocellulose disks. The disks were washed once with 10% acetic acid, twice with 10% phosphoric acid, and once with acetone before drying. Kinase activity was quantitated by scintillation counting.

Inhibition of Src and Hck: Src (0.5 nM) was incubated with inhibitors in reaction buffer and 250 μM ATP, for 30 min at room temperature. The kinase reaction was initiated by a 1:1 dilution into reaction buffer containing 250 μM ATP, 200 μM substrate peptide and 0.8 mCi/mL γ^{32}P-ATP. After 30 min, the reactions were spotted onto a phosphocellulose sheet, briefly dried, and washed with 10% acetic acid, twice with 1% phosphoric acid,

and finally with acetone before drying. Kinase activity was quantitated using a Typhoon Imaging System (Molecular Dynamics) and ImageJ software (NIH). Plots for IC_{50} values were made with Prism GraphPad 4 (GraphPad Software).

Click Chemistry

Reactions were performed on a 25 µL scale. To a solution of protein and inhibitor (20.75 µL) was added 1.25 µL 20% SDS, 0.5 µL 5 mM azide tag, 0.5 µL 50 mM TCEP (~pH 7.0), 1.5 uL 1 mM TBTA ligand in 1:4 DMSO:*tert*-butyl alcohol, and 0.5 µL 5 mM $CuSO_4$. After one hour at room temperature, 6 µL 5 X Laemmli sample buffer was added and samples were either analyzed immediately or snap frozen in $_lN_2$ and stored at -80ºC.

***In vitro* labeling**

50 nM recombinant Src (pY527, generous gift from Michael Eck, Harvard Medical School) and Hck (pY527, generous gift from Todd Miller, Stoney Brook University) were incubated for the given times in PBS and 1 mg/mL BSA with 5 µM compounds. Samples were denatured with 1% SDS then submitted as a group to the click reaction with rhodamine azide. Samples were resolved by SDS-PAGE and scanned with the Typhoon Imaging System (Molecular Dynamics).

***In vivo* labeling**

COS-7 cells were cultured in Dulbecco's Modified Eagle's Media supplemented with 10% fetal bovine serum. p3XFLAG-CMV-7.1 with Y527F Src (mouse) and Hck (human) inserted between EcoRI and XbaI restriction sites were a gift from Todd Miller (SUNY, Stony Brook University). Mutations were introduced by QuickChange Site-Directed mutagenesis (Strategene). Transfections were performed using Lipofectamine 2000

Reagent according to manufacturer's recommendations (Invitrogen). Cells were grown for 48 h pos- transfection before labeling.

Labeling and inhibition of Src and Hck with 47: Media was aspirated from the cells and **47** added in serum-free media. Cells were incubated for one hour at room temperature. Plates were washed once with cold PBS and frozen. PBS with 1X Complete Protease Inhibitor Cocktail (Roche), and 1X Phosphatase Inhibitor Cocktail 2 (Sigma) was added while thawing as the plates were scraped. Cellular debris was pelleted and the supernatant was submitted to the click reaction with a TAMRA azide. Labeled lysates were resolved by SDS-PAGE and either scanned for fluorescence (Typhoon Imaging System, Molecular Dynamics) and coomassie stained or blotted for protein expression (F3165, Sigma) and phospho-tyrosine (2101s, Cell Signaling).

Labeling and inhibition of Src and Hck with 49: Media was aspirated from the cells and **49** added in serum-free media. Cells were incubated for thirty min at room temperature. Plates for Hck were washed once with cold PBS and frozen. Plates for Src were washed once with regular media, returned to the incubator for 15 min, washed once with cold PBS, and frozen. Lysates were collected and treated as above.

Crystalization of Src with 49 – Experiments performed by Haralambos Hadjivassiliou and Ville Paavilainen (Taunton Laboratory, UCSF)

The catalytic domain of Src (chicken c-Src residues 251-533) was expressed and purified as previously described[26]. 20 mg/mL protein was incubated with 1.5 equivalents of **49** in 50 mM Tris (pH 8.0) 100 mM NaCl, 1 mM DTT, 5% glycerol for two h at room temperature. Precipitated protein was removed and the protein crystallized by the hanging drop vapor diffusion method with mother liquor of 50 mM Tris (pH 8.0) 100 mM

NaCl, 1 mM DTT, and 20% ethylene glycol. Crystals belonged to the space group P1 with unit cell parameters a = 41.9; b = 63.4; and c = 74.1; α = 78.9º; β = 88.6º; and γ = 90.0º. Crystals were snap frozen with 20% ethylene glycol cryoprotectant and data collected on the 8.3.1 beamline of the Advanced Light Source at Lawrence Berkeley National Laboratory. Diffraction data were integrated and scaled with the program iMosflm (Supplementary Table 1). The structure of c-Src-**49** complex was solved by molecular replacement using data to 2.2 Å and structure 2hwo as a search model in the program MOLREP, followed by several rounds of manual rebuilding and restrained refinement with programs COOT and REFMAC5 (supplementary Table 1)[26,27,28,29]. The structure was validated with the MolProbity server[30].

Inhibition of FGFR3 – all experiments preformed by Sumin Kang (Chen Laboratory, Emory University):

Ba/F3 cells were cultured in RPMI 1640 medium supplemented with 10% fetal bovine serum and 1.0 ng/mL interleukin-3 (R & D Systems). Mutation C482A was introduced into TEL-FGFR3 by using QuikChange-XL site-directed mutagenesis kit (Stratagene). Ba/F3 cell lines stably expressing TEL-FGFR3 and TEL-FGFR3 C482A were generated by retroviral transduction as described using pMSCV-neo^{R} plasmids encoding TEL-FGFR3 variants, followed by antibiotic selection with 1 mg/ml neomycin[25].

Ba/F3 IL-3-independent proliferation assay: For cell viability assays, 1×10^{5} Ba/F3 cells stably expressing TEL-FGFR3 variants were cultured in 96-well plates with media containing increasing inhibitor concentrations in the absence of IL-3. Parental Ba/F3 cells were cultured with IL-3. The relative cell viability was determined using the Celltiter96AQ$_{ueous}$ One solution proliferation kit (Promega) after 72 h.

Drug treatment and immunoprecitation: Ba/F3 cells stably expressing TEL-FGFR3 variants were treated with 1 μM **21** or **51** for 4 h in the absence of serum and IL-3. Cells were lysed and the TEL-FGFR3 variants were immunoprecipitated (anti-FGFR3, Santa Cruz Biotechnology) and auto-phosphorylation was quantified by western blotting (4G10, Millipore).

References:

1) Noble MEM, Endicott JA, Johnson LN. Protein kinase inhibitors: insights into drug design from structure. Science 2004, 303: 1800-1805.

2) Manning G, Whyte DB, Martinez R, Hunter T, Sudarsanam S. The protein kinase complement of the human genome. Science 2002, 298: 1912-1934.

3) Knight ZA, Shokat KM. Features of selective kinase inhibitors. Chem Biol 2005, 12: 621-637.

4) Zhang J, Yang PL, Gray NS. Targeting cancer with small molecule kinase inhibitors. Nature Reviews Cancer 2009, 9: 28-39.

5) Roskoski R. Src protein-tyrosine kinase structure and regulation. Biochem Biophys Res Com 2004, 324: 1155-1164.

6) Ingley E. Src family kinases: regulation of their activities, levels and identification of new pathways. Biochim Biophys Acta 2008, 1784: 56-65.

7) Brown MT, Cooper JA. Regulation, substrates and functions of src. Biochim Biophys Acta 1996, 1287: 121-149.

8) Trevino JG, Summy JM, Gallick GE. Src inhibitors as potential therapeutic agents for human cancer. Mini Rev Med Chem 2006, 6: 681-687.

9) Giaccone G, Zucali PA. Src as a potential therapeutic target in non-small-cell lung cancer. Ann Oncol 2008, 19:1219-1223.

10) Hiscox S, Nicholson RI. Src inhibitors in breast cancer therapy. Expert Opin Ther Targets 2008, 12: 757-67.

11) Martinelli G, Soverini S, Rosti G, Baccarani M. Dual tyrosine kinase inhibitors in chonic myeloid leukemia. Leukemia 2005, 19: 1872-9.

12) Yee M, Fas, SC, Stohlmeyer MM, Wandles TJ, Cimprich KA. A cell-permeable, activity based probe for protein and lipid kinases. J Biol Chem 2005, 280: 29053-29059.

13) Hanks SK, Hunter T. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. FASEB J. 1995, 9: 576-596.

14) Fry DW, Bridges AJ, Denny WA, Doherty A, Greis KD, Hicks JL, Hook KE, Keller PR, Leopold WR, Loo, JA, McNamara DJ, Nelson JM, Sherwood V, Smaill JB, Trumpp-Kallmeyer S, Dorbrusin EM. Specific, irreversible inactivation of the epidermal growth factor receptor and erB2, by a new class of tyrosine kinase inhibitor. Proc Natl Acad Sci USA 1998, 95: 12022-12027.

15) Pan Z, Scheerens H, Li S, Schultz BE, Sprengerler PA, Burrill LC, Mendonca RV, Sweeney MD, Scott KCK, Grothaus PG, Jeffery DA, Spoerke JM, Honigberg LA, Young PR, Dalrymple SA, Palmer JT. Discovery of selective irreversibly inhibitors for burton's tyrosine kinase. ChemMedChem 2007, 2: 58-61.

16) Schirmer A, Kennedy J, Murli S, Reid R, Santi DV. Targeted covalent inactivation of protein kinases by resorcylic acid lactone polyketides. PNAS 2006, 103: 4234-4239.

17) Cohen MS, Zhang C, Shokat KM, Taunton J. Structural bioinformatics-based design of selective, irreversible kinase inhibitors. Science 2005, 308: 1318-1321.

18) Ratcliffe SJ, Yi T, Khandekar SS. Synthesis and characterization of 5'-p-fluorosulfonylbenzoyl-2'(or 3')-(biotinyl)adenosine as an activity-base probe for protein kinases. J Biomolecular Screening 2007, 12: 126-132.

19) Speers AE, Cravatt BF. Profiling enzyme activities in vivo using click chemistry methods. Chem Biol 2004, 11: 535-546.

20) Cohen MS, Hadjivassiliou H, Taunton J. A clickable inhibitor reveals context-dependent autoactivation of p90 RSK. Nature Chem Bio 2007, 3: 156-160.

21) Hanke JH, Gardner JP, Dow RL, Changelian PS, Brissette WH, Weringer EJ, Pollok BA, Connelly PA. Discovery of a novel, potent, and Src family-selective tyrosine kinase inhbitor. J Biol Chem 1996, 271: 695-701.

22) Maurer TS, Fung H. Comparison of methods for analyzing kinetic data from mechanism-based enzyme inactivation: application to nitric oxide synthase. AAPS Pharmsci 2000, 2: E8.

23) Adams JA. Kinetic and catalytic mechanisms of protein kinases. Chem Rev. 2001, 101: 2271-2290.

24) Kemble DJ, Sun G. Direct and specific inactivation of protein tyrosine kinases in Src and FGFR families by reversible cysteine oxidation. PNAS 2009, 106: 5070-5075.

25) Chen J, Williams IR, Lee BH, Duclos N, Huntly BJP, Donoghue DJ, Gilliland DG. Constitutively activated FGFR3 mutants signal though PLCγ-dependent and –independent pathways for hematopoietic transformation. Blood 2005, 106: 328-337.

26) Blair JA, Rauh D, Kung C, Yun C, Fan Q, Rode H, Zhang C, Eck MJ, Weiss WA, Shokat KM. Structure-guided development of affinity probes for tyrosine kinases using chemical genetics. Nature Chemical Biology 2007, 3: 229-238.

27) Vagin A, Teplyakov A. MOLREP: an automated program for molecular replacement. J. Appl. Cryst. 1997, 30: 1022-1025.

28) Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004, 60:2126-32.

29) Murshudov GN, Vagin AA, Dodson EJ. Refinement of Macromolecular Structures by the Maximum-Likelihood method. Acta Cryst. 1997, D53: 240-255.

30) Davis et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Research 2007, 35:W375-W383.

Chapter 3: Chemical Proteomic Identification of Essential Protein Kinases in *Trypanosoma brucei*

Introduction:

Trypanosoma brucei is a member of an ancient group of eukaryotic parasites and is the causative agent of African sleeping sickness[1]. *T. brucei* and related kinetoplastids cause fatal and debilitating diseases that threaten millions of the world's poorest people[2,3]. Current treatments rely on highly toxic chemotherapeutics, some of which were developed over fifty years ago[4]. Although they are highly divergent, humans and kinetoplastids share many of the same cellular regulatory components, including an array of protein kinases[5]. Protein kinases have shown great promise as drug targets and have received much attention in recent drug discovery efforts for cancer therapeutics. Due to their essential cellular roles in all eukaryotic organisms, protein kinases are a rich source of potential therapeutic targets for lesser known diseases caused by tropical parasites[6,7].

Encoded within the *T. brucei* genome are more than 150 protein kinases, many of which, in principle, could be targeted with drugs to treat this deadly disease[7]. The functions of the majority of trypanosomal protein kinases are currently assigned by sequence homology alone. A small number of kinases have been shown to be essential to the clinically relevant forms of the parasite, but the vast majority remain unexplored[8].

Affinity and activity-based probes are chemical tools for investigating enzymatic families that require no prior knowledge of a particular protein's specific cellular function. Activity-based probes covalently react with enzymes, which can be targeted individually or as a family[9]. Affinity probes reversibly bind their targets (usually with sub-micromolar affinity) and, when immobilized to a solid support, can be used to purify individual proteins or entire families of related proteins. When activity-based or affinity probes are applied to cell lysates (or, in rare cases, intact cells or whole animals) in the presence of an unlabeled drug candidate (or any small molecule of interest, which may bind to targets shared by the probe), the drug candidate will act as a competitor and prevent the

binding of a subset of protein targets to the probe. Using quantitative mass spectrometry and a variety of differential isotopic labeling schemes, it is possible to quantitatively elucidate the full spectrum of probe targets that are competitively bound by any given drug candidate (or small molecule of interest). Activity-based and affinity probes thus offer powerful tools for quantifying interactions between drugs or drug leads and their desired therapeutic targets, as well as their undesired off-targets[10,11,12,13]. Moreover, they do so in the context of native proteins expressed at their endogenous levels. Two large-scale studies using affinity and activity-based probes have focused on protein kinases[11,14]. Both methods enable the quantitation of binding interactions between drug leads and a large swath of the human kinome under quasi-native conditions in whole-cell lysates. One limitation of both of these methods is that, because the probes are not cell permeable, they cannot readily report on interactions between drug leads and their targets in intact cells and animals.

We propose to use the electrophilic 5'-fluorosulfone-containing nucleoside **47** from Chapter 2 as an activity-based probe to investigate protein kinases in *T. brucei*. This compound forms a covalent bond with a conserved lysine in the kinase active site, and "click" chemistry permits tagging of labeled proteins with a variety of fluorophores and affinity labels[15]. Fluorosulfone **47** is cell permeable, and in preliminary experiments, it inhibited the viability of the bloodstream form of *T. brucei*. This reagent's reversible binding affinity relies on a *p*-tolyl pyrrolopyrimidine heterocycle that accesses a hydrophobic pocket found in a subset of protein kinases. Typically, these kinases contain a threonine or smaller residue at the "gatekeeper" position[16]. Requirement for a small gatekeeper is not absolute, and similar compounds bind many more kinases than predicted by this "gatekeeper" hypothesis[10,17]. Additionally, the ability to form an irreversible covalent bond facilitates saturable binding of targets that are only modestly bound by non-covalent analogues (Figure 2-10)[18].

Based on its ability to kill parasites growing in culture, we hypothesized that fluorosulfone **47** worked by irreversibly inactivating a subset of protein kinases in *T. brucei*. Such a covalent mechanism of inhibition would greatly facilitate target identification and would permit further investigation of their inhibition by other candidate inhibitors. In this scheme, progargyl-tagged fluorosulfone **47** would serve as an activity-based probe and would reveal inhibition of untagged drug leads in competitive binding experiments. Because the activity-based probe **47** is cell permeable, such competitive binding experiments could be conducted, in principle, in intact *T. brucei* cells.

In this chapter, I demonstrate that fluorosulfone **47** selectively labels a significant subset (~15%) of *T. brucei* protein kinases and that labeling of individual kinases is selectively blocked by co-incubation with a variety of known kinase inhibitors. In collaboration with Juan Oses-Prieto and Al Burlingame at UCSF, I developed a quantitative mass spectrometry (MS) method to identify proteins that are covalently bound to **47**. Most importantly, we used this novel chemical proteomics method to quantify the extent of occupancy of individual kinases by a variety of trypanocidal kinase inhibitors, thus providing a short list of likely *in vivo*-relevant targets. Finally, I have characterized a new electrophilic nucleoside that targets an atypical protein kinase (among other kinase and non-kinase targets), the *T. brucei* ortholog of the budding yeast Bud32p. Based on its known functions in budding yeast, *T. brucei* Bud32 may play a role in the parasite's antigen variation and immune evasion.

Results and Discussion:

Electrophilic nucleosides 1 and 47 kill *Trypanosoma brucei*

The 5'-fluorosulfonyl benzoyl nucleosides **1** and **47**, along with non-electrophilic analogs, were tested for effects on the bloodstream form of *T. brucei*. All of the compounds blocked cell proliferation with IC_{50}'s < 10 μM (Figure 3-1). Reversible kinase inhibitors PP1 and the related nucleoside **5OH** were nearly equipotent to one another, but addition of the 5'-fluorosulfone significantly increased potency. As demonstrated in Chapter 2 and consistent with previous studies with FSBA,[19] the 5'-fluorosulfone **47** specifically targets the catalytic lysine for covalent modification when bound to the human protein kinase, Src (Figure 2-4). Thus, the increased *in vivo* potency of the electrophilic compounds is likely due to similar covalent modification of endogenous *T. brucei* kinases. Propargylated fluorosulfone **47** killed *T. brucei* at nanomolar concentrations (IC_{50}~200 nM), whereas non-propargylated fluorosulfone **1** was significantly less potent. With human Src, the 2'-propargyl ether increased the potency of **47** compared to **1**, presumably by making additional hydrophobic contacts in the "ribose pocket". An analogous trend in Figure 3-1 supports the hypothesis that the cytotoxicity of **1** and **47** toward *T. brucei* is due to kinase inhibition.

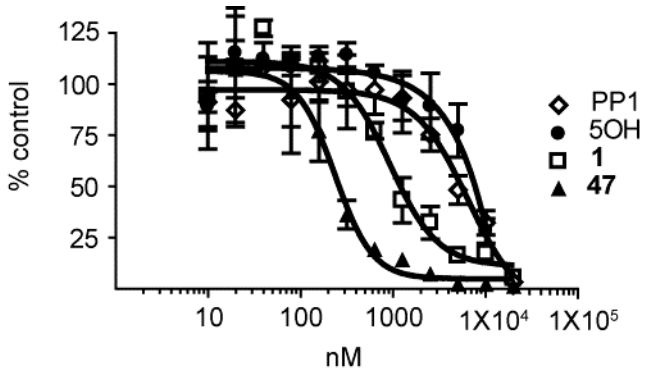


Figure 3-1: 5-electrophilic nucleosides significantly inhibit cell viability of *Trypanosoma brucei*. Cultures of the bloodstream form of *T. brucei* were treated with compounds for 24 h, and living cells were quantified. Values are normalized to DMSO control, and error bars show standard deviation.

Fluorosulfone 47 selectively reacts with a subset of *T. brucei* proteins

Labeling of select proteins in *T. brucei* lysates by propargylated fluorosulfone **47** was diminished by pretreatment with **1** or a variety of known kinase inhibitors. Several intensely labeled bands were visible in lysates treated with **47** after attaching a rhodamine tag via "click" chemistry (Figure 3-2)[15]. The majority of these bands disappeared when the lysates were pretreated with fluorosulfone **1**, which lacks the click-reactive propargyl ether. This indicates that labeling by **47** is not a non-specific reaction with surface residues. Additionally, labeling of several proteins decreased when lysates were pretreated with a mixture of known "semi-promiscuous" kinase inhibitors, PD0166326 (PD) and SB203580 (SB). Like the nucleoside scaffold of **47**, these compounds bind to the ATP binding site of protein kinases and have a preference for a small residue at the "gatekeeper" position[20,21].

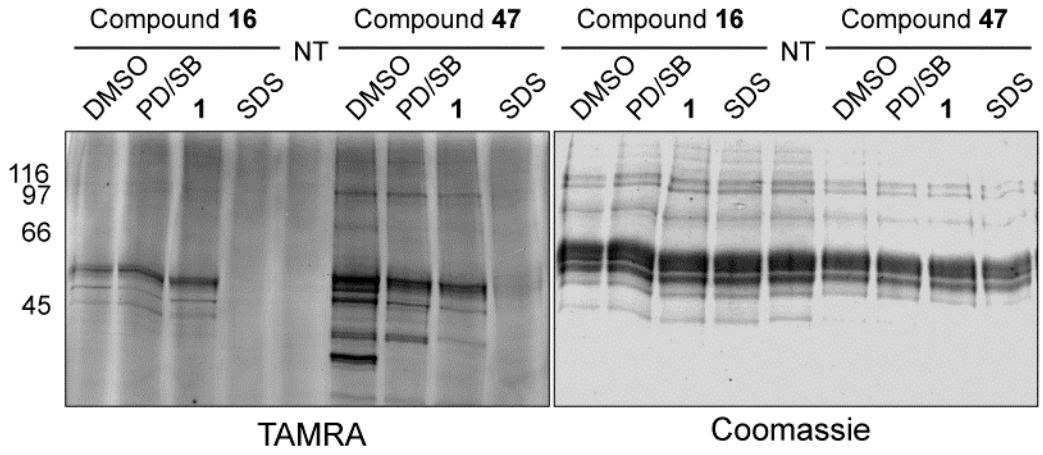


Figure 3-2: Electrophilic nucleoside **47** forms specific covalent bonds to endogenous proteins of *Trypanosoma brucei*. Whole cell lysates were incubated with 10 μM PD0166326 (PD) and SB203580 (SB), compound **1**, or DMSO, then treated with 2 μM **47** or the TAMRA conjugated nucleoside **16**. Lysates labeled with **47** were submitted to Cu catalyzed "click" chemistry with TAMRA azide. All lysates were resolved by SDS-PAGE, scanned for fluorescence, and Coomassie stained.

Unfortunately, direct conjugation of a fluorophore to compound **1** abolished specific labeling of lysate proteins (Figure 3-2), as none of the proteins labeled by **16** in *T. brucei* lysates were competed by **1**. Compound **16** is connected to TAMRA by a 2'-carbamate/PEG linker (Scheme 1-3), which may impede binding to the kinase active site (as shown in Figure 2-2 with BODIPY-labeled compound vs. Src) without impeding non-specific reactions with surface lysines on abundant proteins (note that the intense TAMRA-labeled bands in Figure 3-2 correspond to the most abundant Coomassie-stained bands). Together, these data indicate that propargylated fluorosulfone **47** is superior TAMRA fluorosulfone **16** and that chemical proteomics experiments with this class of probes require that the affinity or fluorescent tag be attached via click chemistry *after* irreversible inhibition of cellular kinases.

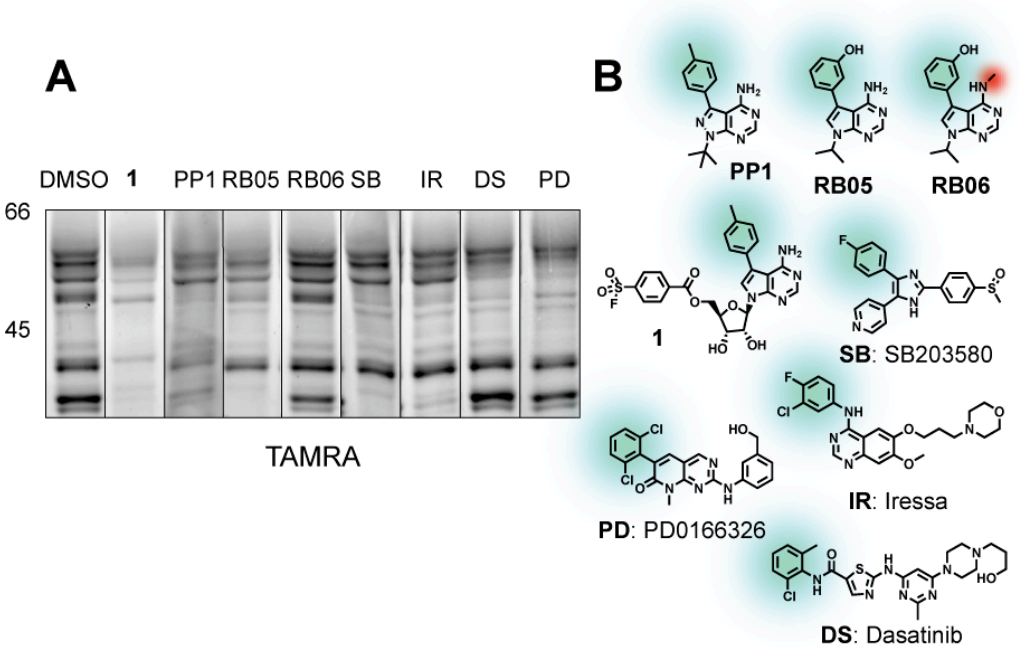


Figure 3-3: Labeling of *Trypanosoma brucei* proteins by **47** is differentially inhibited by ATP competitive protein kinase inhibitors. (A) Whole-cell lysates were incubated with 10 μM **1** or 20 μM of other kinase inhibitors, then labeled with **47**. Lysates were submitted to the "click" reaction with TAMRA-azide and resolved by SDS-PAGE. Lane images were selected from different gels that were processed in parallel and contrasted identically. (B) Structures of competing compounds. In blue are regions that bind to the "hydrophobic pocket," and red indicates a methylation that disrupts a critical hydrogen bond.

ATP-competitive kinase inhibitors from several different structural classes were tested for their ability to block labeling by **47** (Figure 3-3). Like the nucleoside scaffold, all of the inhibitors share a preference for a small gatekeeper and thus likely bind to a subset of kinases targeted by **47** (Figure 3-3b). Compound RB06 is an important control. It differs from the related RB05 by a single methyl group that disrupts an important hydrogen bond formed by the exocyclic amine of the inhibitor and a backbone carbonyl in the hinge region of the kinase active site[22]. RB06 did not inhibit labeling of any of the bands blocked by RB05, strongly suggesting that these bands are protein kinase targets of **47**. Moreover, each of the kinase inhibitors showed a unique pattern with respect to competed protein bands. For example, co-incubation with Iressa (IR) predominantly abolished labeling of an intense ~38 kDa band and competed other bands to a lesser extent. By contrast, dasatinib (DS) had no effect on the ~38 kDa band, whereas it potently competed for labeling of a ~48 kDa band that was unaffected by Iressa. In conclusion, competitive labeling experiments with **47** in the presence of structurally diverse kinase inhibitors, including two approved drugs (Iressa and dasatinib), revealed >10 specifically labeled bands. Moreover, different kinase inhibitors had varying abilities to bind competitively to different proteins. Thus, a semi-promiscuous activity-based probe, such as fluorosulfone **47**, can reveal a "kinomic fingerprint" for clinically relevant kinase inhibitors in a relatively understudied kinome from a disease-causing pathogen.

Trypanosome kinase TbPK50 as a model for kinase labeling and affinity purification

Gene Tb10.70.2260 encodes a Ser/Thr protein kinase with a threonine residue in the gatekeeper position. This gene's protein product is TbPK50, which may be a functional homologue of the essential AGC family kinase, Orb6, from *S. pombe*[23,24]. Recombinant TbPK50 showed a modest ability to auto-phosphorylate, and was robustly labeled by compound **47** *in vitro* (data not shown, Figure 3-4a). When spiked into lystates from

COS-7 cells, recombinant TbPK50 was the most heavily labeled band, even when not distinguishable by Coomassie staining. Labeled TbPK50 was observed when it represented less than 0.05% of the total protein in a lane. Labeling of TbPK50 can be blocked by co-incubation with PD. This localizes the site of covalent modification to the ATP binding site of this kinase (Figure 3-4b).

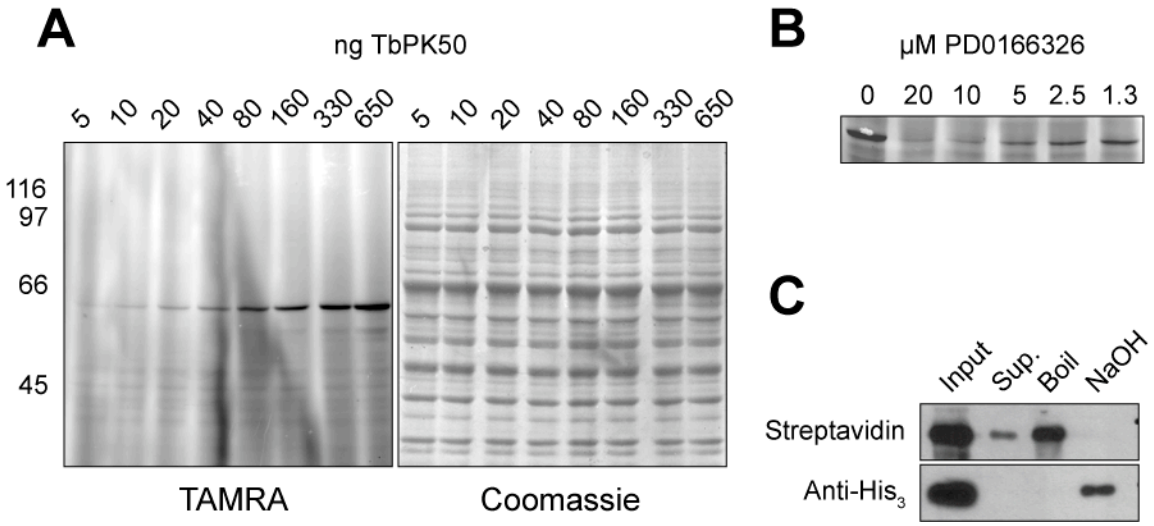
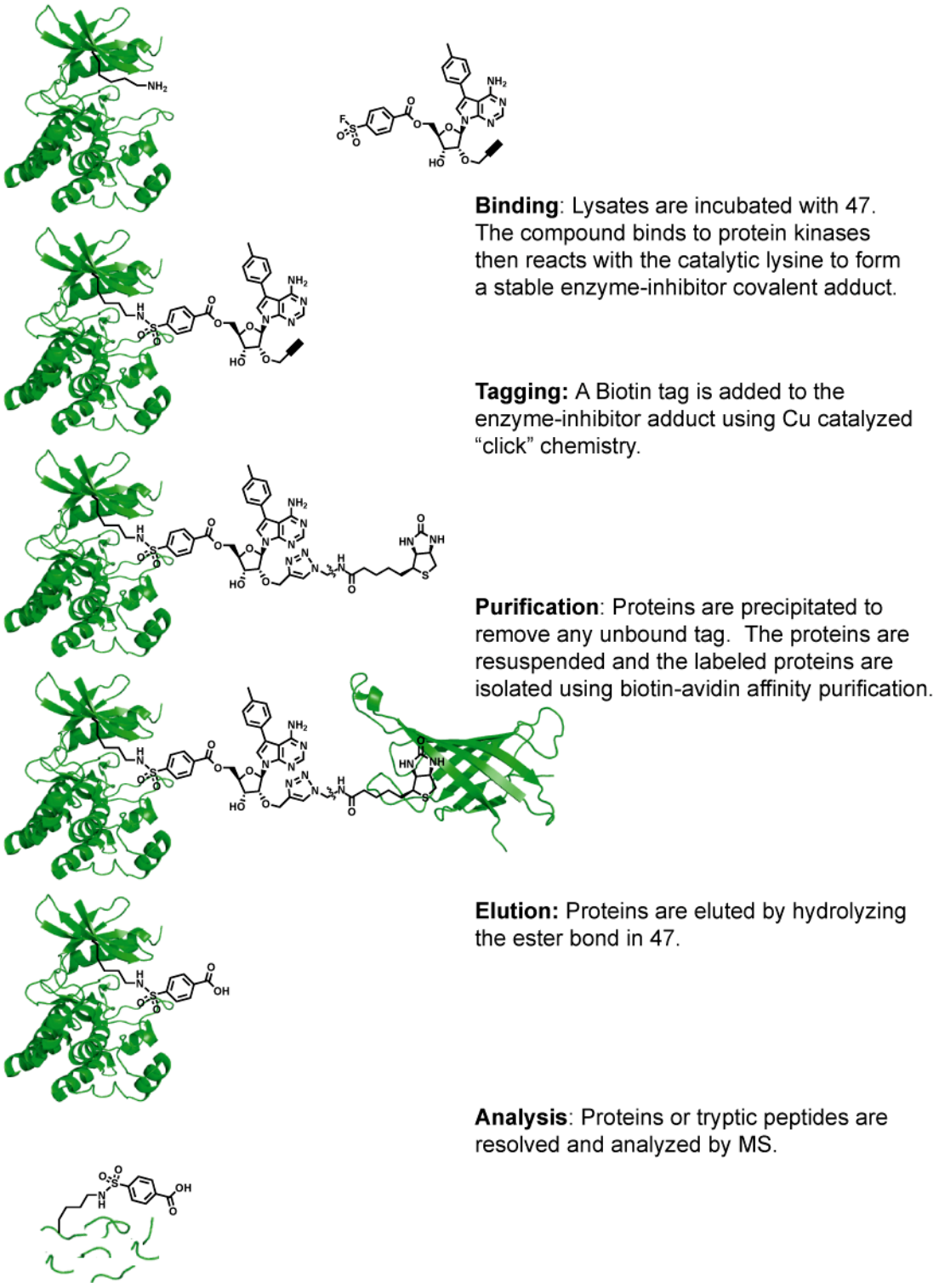


Figure 3-4: Specific labeling and affinity purification of kinase TbPK50. (A) Recombinant TbPK50 was spiked into COS-7 cell lysate and labeled with **47**. Labeled lysates were submitted to a "click" reaction with TAMRA azide and resolved by SDS-PAGE. (B) TbPK50 in COS-7 cell lysate was incubated with PD0166326 then processed as in A. (C) TbPK50 was labeled with **47** and conjugated to a biotin azide with "click" chemistry. Labeled protein was purified with avidin-agarose and eluted by either boiling in 2X Laemmli sample buffer or treating with 0.2 N NaOH. Samples were analyzed by Western blotting.

We used TbPK50 as a test case for affinity purification of proteins labeled by **47**. The kinase was incubated with **47** and tagged using a biotin azide in place of the previous TAMRA-derived label. Following "click" chemistry with biotin-azide the labeled protein was efficiently removed from solution by incubation with avidin-conjugated agarose beads (Figure 3-4c). Rather than attempt to disrupt the avidin-biotin interaction, which often results in low yields of affity purified biotinylated proteins, we used 0.2 N NaOH to specifically elute **47**-modified proteins via hydrolysis of the 5'-benzoate ester[25].

Western blot detection of the His-tag showed that significantly more His-TbPK50 was recovered using ester hydrolysis than by heating the beads in Laemmli sample buffer.



Scheme 3-1: Strategy for purification and analysis of proteins labeled by **47**.

Kinases are highly enriched after affinity purification of lysates labeled with 47

Scheme 3-1 diagrams our strategy for identification of proteins covalently labeled by compound **47**. Lysates were treated with an optimal concentration of **47** (2 mM) that maximized the intensity of specifically labeled bands with a minimal non-specific background signal. After treatment with fluorosulfone **47**, lysates were subjected to click chemistry with a biotin azide under conditions that ensured protein denaturation (1% SDS, see Methods for details) to maximize accessibility of the 2'-propargyl tag. Excess compound **47** and biotin-azide was removed via acetone-mediated protein precipitation. Precipitated proteins were re-solubilized and affinity-purified with avidin-agarose beads. After extensive washing, labeled proteins were specifically released via NaOH-mediated ester hydrolysis. The eluates were resolved by SDS-PAGE and submitted for analysis by mass spectrometry. This process includes two technically important points: (1) use of click chemistry avoids the decreased affinity of inhibitors that contain directly conjugated biotin (Figure 3-2), and (2) the hydrolysis step avoids the need to disrupt the biotin-avidin linkage (Figure 3-4).

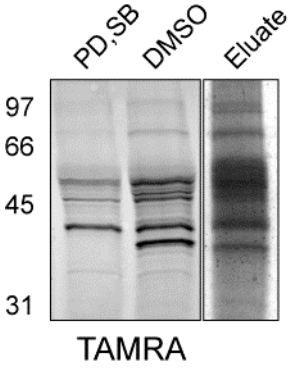
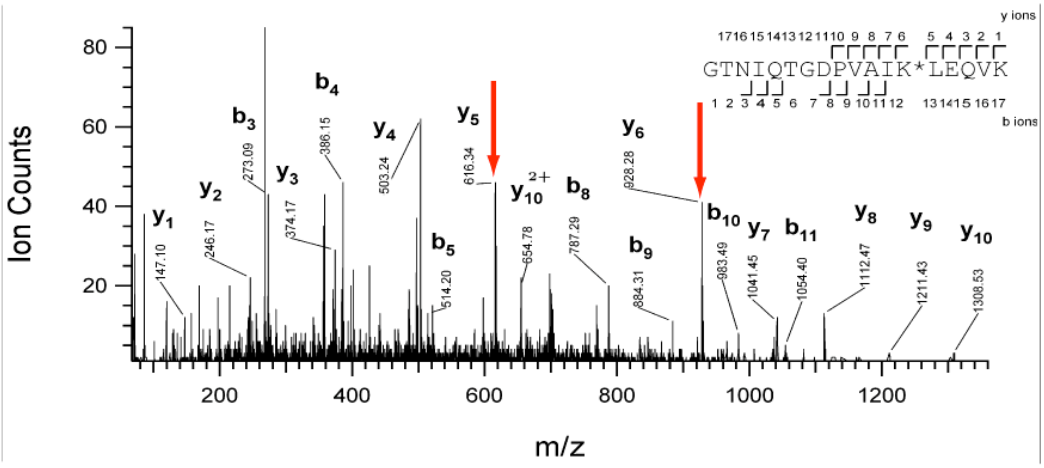


Figure 3-5: Purification of *Trypanosoma brucei* proteins labeled by **47**. Whole-cell lysate from the bloodstream form of *T. brucei* was treated with **47**. Covalently bound proteins were affinity purified after click reaction with biotin-azide. Small aliquots were co-incubated with DMSO or kinase inhibitors PD0166326 (PD) and SB203580 (SB), then treated with **47** and submitted to click reaction with TAMRA-azide. Samples were resolved by SDS-PAGE and either scanned for fluorescence ("PD,SB", "DMSO") or Coomassie stained ("Eluate").

Eluted proteins, stained with Coomassie, had identical molecular weights to those whose labeling was competed by reversible kinase inhibitors. Figure 3-5 shows three lanes from the same gel. Coomassie stained bands in the eluate from an affinity purification correspond to bands competed by a mixture of kinase inhibitors PD and SB.

This demonstrates the effectiveness of our affinity purification strategy and prompted subsequent experiments using MS.

Trypsin digestion and LCMS/MS analysis of the gel in Figure 3-5 identified one or more unique peptides for 25 *T. brucei* protein kinases (Table 3-1). Kinases represent 16% of all the proteins identified, and 10 out of 20 proteins with the greatest sequence coverage (Supplementary Table 2). Protein kinases represent roughly 2% of the *T. brucei* genome. As signaling molecules, kinases are generally present at much lower copy number than other identified proteins such as tubulin isoforms and variant surface glycoprotein[8]. The collection of protein kinases affinity purified by fluorosulfone **47** is enriched in kinases with threonine or other small residues in the gatekeeper position (Table 3-1): 32% of the eluted kinases versus only 9% in the genome. The lysine-sulfonamide benzoate conjugate, resulting from hydrolysis of lysine-adducted **47**, was detected on only 9 peptides, derived from 9 protein kinases (Figure 3-6). Alignments of these peptides reveal that, in every case, this modification occurred at the catalytic lysine. There was no trace of compound **47** (or the sulfonylbenzoate fragment) found at



any other amino acid .

Figure 3-6: LCMS/MS of a covalently modified peptide. Shown is the mass fragmentation spectrum of a peptide derived from *T. brucei* casein kinase (Tb927.5.800) after treatment of

lysates with **47** and affinity purification. Arrows indicate the loss of the catalytic lysine bearing a sulfamide benzoate after hydrolysis of **47**.

Only one of the bands in Figure 3-5 was unambiguously matched to a unique kinase based on molecular weight and on MS-identification of the excised band. The ~38 kDa band competed by SB in labeling experiments contains a homolog of casein kinase 1 (Tb927.5.800) and was identified by over 36 unique peptides (table 3-1). This kinase has previously been identified as the target of a similar class of triarylimidazole compounds in the kinetoplastid *Leishmania major*[26]. The remaining bands in Figure 3-5 contained multiple protein kinases, along with abundant non-specifically bound proteins.

Five kinases targeted by fluorosulfone 47 appear to be essential for cell growth

Several kinases in Table 3-1 have been previously investigated. RNAi-mediated knockdown of Crk6 in conjunction with Crk3 (but not RNAi vs. Crk6 alone) produced a severe growth phenotype[27]. Knockdown of the related kinase, Crk10, had no effect on cell growth[28]. Homozygous knockout of either the 46 kDa SAPK or TbMAPK5 (see Table 3-1) had effect on viability, but conferred resistance to temperature stress and affected differentiation, respectively[29,30]. In contrast to these kinases, RNAi against the Gsk3 homologue led to growth arrest in the bloodstream form of the parasite, and it has been highlighted as a potential drug target[31].

We used RNAi to screen the remaining kinases for essential functions. Segments from each gene were cloned into the pZJM tetracycline-inducible vector and transfected into the 90-13 cell line (Table 3-1)[32]. Induction of RNAi produced a growth defect for four kinases (Figure 3-7). The results for TbPK50 are not unexpected, given the importance of Orb6 in cell polarity and division in fission yeast, as well as the importance of its binding partner, MOB1, in trypanosome cytokinesis[24,33,34,35]. There are rapid and severe effects after induction of RNAi for casein kinase 1 (Table 3-1, Tb927.5.800). This result agrees with those published recently by another group[36]. The

50 kDa NIMA kinase (Tb927.4.5310) and Clk homologs (Tb11.01.4230) represent novel, previously uncharacterized targets in the parasite. Both proteins have one nearly

Genome ID #	Homology	MW	Peptides	GK	FSB-K	RNAi construct
TB10.70.2070	Erk 8	50	47	F	X	ACAAGGATATTTACCTTG/CATGTCAACACCCTTCGTG
Tb10.70.2260	Orb 6	50	44	T		CTATTCGCCTTGTACTGTTAC/CATGGCATAGACAATGTGAGG
Tb927.5.800	Casein Kinase 1	38	36	I	X	GGGCCCTCATTGGAGGATC/AAACCCTGCCACGGCAGTG
Tb927.3.2440	PKC1 / Akt	46	30	M	X	GAGGGATGGCGAAAGCTCC/TACAACTTGGATGTGGGTAAG
Tb927.7.6220	RAD53	60	24	M	X	GACCATCATTACGAACTTTTGG/AAGTCACCCTCCGAAAGAAG
Tb927.8.3770	MAPK 3	47	20	T	X	GGTGCACCGTTTTCTCAGC/ACAAGTACATTTCCTGGGTG
Tb10.329.0030	MAPK 3	64	23	M	X	TCCCTGTAACC/GACGAGCTTTCCAAATACC
Tb11.47.0031	Crk 6	37	21	F	X	
Tb927.7.5770	STK 38L	53	23	M	X	TGTTTCTCATCACAAGGCTTCC/TGTTTGATGCACAACCATG
Tb927.6.1780	SAPK	46	17	T		
Tb05.5K5.80	Stk36	39	18	T	X	TGTGGGGCGAGGAACGTATG/AAGTGGAATCCACCATCTTC
Tb927.4.5310	NIMA 1	50	15	M		ATGACGGAACATAAGCGTCCCG/CTGTCTGTAGAGATCACC
Tb927.8.7110	NIMA 1	50	4	A		GACCGCGCACAAACGTGCAAC/TTTGCCTGCAGAGGTCGCC
Tb927.3.4670	Crk 10	48	11	T		
Tb927.8.6810	SAPK	47	10	M		CGAAACCCAG/AACGTTCCAGTTTTATTACC
Tb10.70.0960	NIMA 1	55	7	L		CGACTACGCAAATGCTGG/AGAGGACACCAAGAGAGAAC
Tb10.70.0970	NIMA 1	55	5	L		As Above
Tb927.3.4560	SAPK	71	10	M		TGGTAAAG/CCATCCTTTGCTATGTTGG
Tb10.61.1880	Ysk4 / SAPK	31	7	L		GTTCATTATATGCATCAC/TTTGTCAC
Tb11.01.4230	Clk3	53	6	M		CCACGTAAACAACAATCC/TATTTTGGCACATTACGTAC
Tb927.3.1570	EIF2AK2	69	3	M		AGTCCATCACTTGAAGC/AATACGC
Tb927.6.4220	TbMAPK5	44	4	T		
Tb09.160.3480	PDPK1	43	1	T		
Tb10.70.4250	SAPK	42	2	M	X	
Tb10.61.3140	Gsk3	40	1	M		

Table 3-1: Identity of protein kinases bound by **47** after MS analysis of affinity purified proteins. Genome ID numbers are from the *T brucei* GeneDB. The number of unique peptides for each identification is listed. Gatekeeper (GK) residue assignment is based on sequence alignments with Lck and p38. FSB-K indicates a peptide containing a covalently modified catalytic lysine. Listed are the 5' and 3' sequences of gene fragments used to make RNAi constructs. MS analysis performed by Juan Oses-Prieto, UCSF.

identical paralog (Tb927.8.7110 for NIMA and Tb11.01.4250 for Clk), and our RNAi construct for Clk shares significant sequence overlap with both paralogs. In yeast and humans, Clks phosphorylate SR splicing factors and influence mRNA splicing[37,38]. Given the unique modes of mRNA splicing and regulation in *T brucei*, the Clk homologues represent an exciting target for future research[39,40].

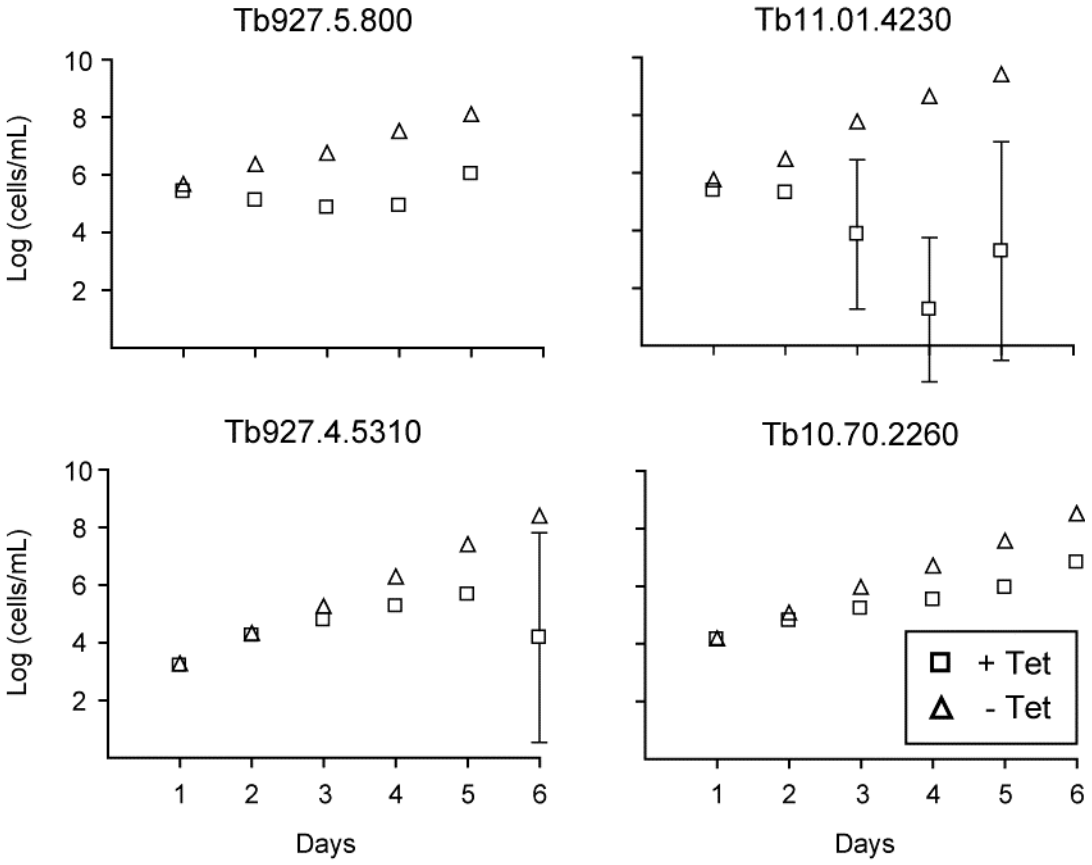


Figure 3-7: RNAi of *Trypanosoma brucei* kinases. RNAi was induced for different kinases and cell densities were monitored daily. Values are the average of three separate populations and error bars represent standard deviations.

We used RNAi to help screen the remaining kinases for essential function. Segments from each gene were cloned into the pZJM inducible vector and transfected into the 90-13 cell line (table 3-1)[32]. Induction of RNAi gives a growth phenotype for four kinases (figure 3-7). The results for TBPK50 are not unexpected given its homologue's importance in cell polarity and division in yeast as well as the importance of its binding

partner, MOB1, in trypanosome cytokinesis[24,33,34,35]. There are rapid and sever effects after induction of RNAi for casein kinase 1. This parallels results in Leishmania, and recently our results in *T. brucei* have been duplicated[36]. The 50 kDa NIMA kinase (Tb927.4.5310) and Clk homologues represent novel targets in the parasite. Both proteins have one nearly identical paralogue (Tb927.8.7110 and Tb11.01.4250), and our RNAi construct for Clk shares significant sequence overlap with both versions of the gene. In yeast and humans, Clks phosphorylate SR splicing factors and influence mRNA splicing[37,38]. Given the unique modes of mRNA splicing and regulation in *T brucei*, the Clk homologues represent an exciting target for future research[39,40].

Competition vs. fluorosulfone 47 via quantitative mass spectrometry yields binding profiles for other kinase inhibitors

We used affinity purification in conjunction with iTRAQ isobaric tags and mass spectrometry to quantitate decreased labeling of proteins in lysates co-incubated with **47** and compound **1** or PD[41]. This allowed us to quantitate the extent of competitive binding of a covalent and non-covalent inhibitor for specific kinases labeled by **47**.

Competition with the non-propargylated fluorosulfone **1** demonstrated the potential of compound **47** as an activity-based probe for protein kinases. Comparison of recovered kinases and non-kinases at different concentrations of compound **1** shows that the majority of specifically labeled proteins are kinases (Figure 3-8a). Little competition by **1** for non-kinases indicates that such proteins reacted non-specifically with **47** and/or were non-specifically bound by avidin-agarose purification. With the exception of two kinases, all of the kinases identified in this experiment were effectively competed by compound **1** (Figure 3-8b). These enzymes represent more than 10% of the *T. brucei* kinome and include four of the five essential gene products discussed above. Contrary to expectations, those kinases that were competed most effectively by compound **1** did not have a small gatekeeper residue. The 50 kDa Erk8 homolog

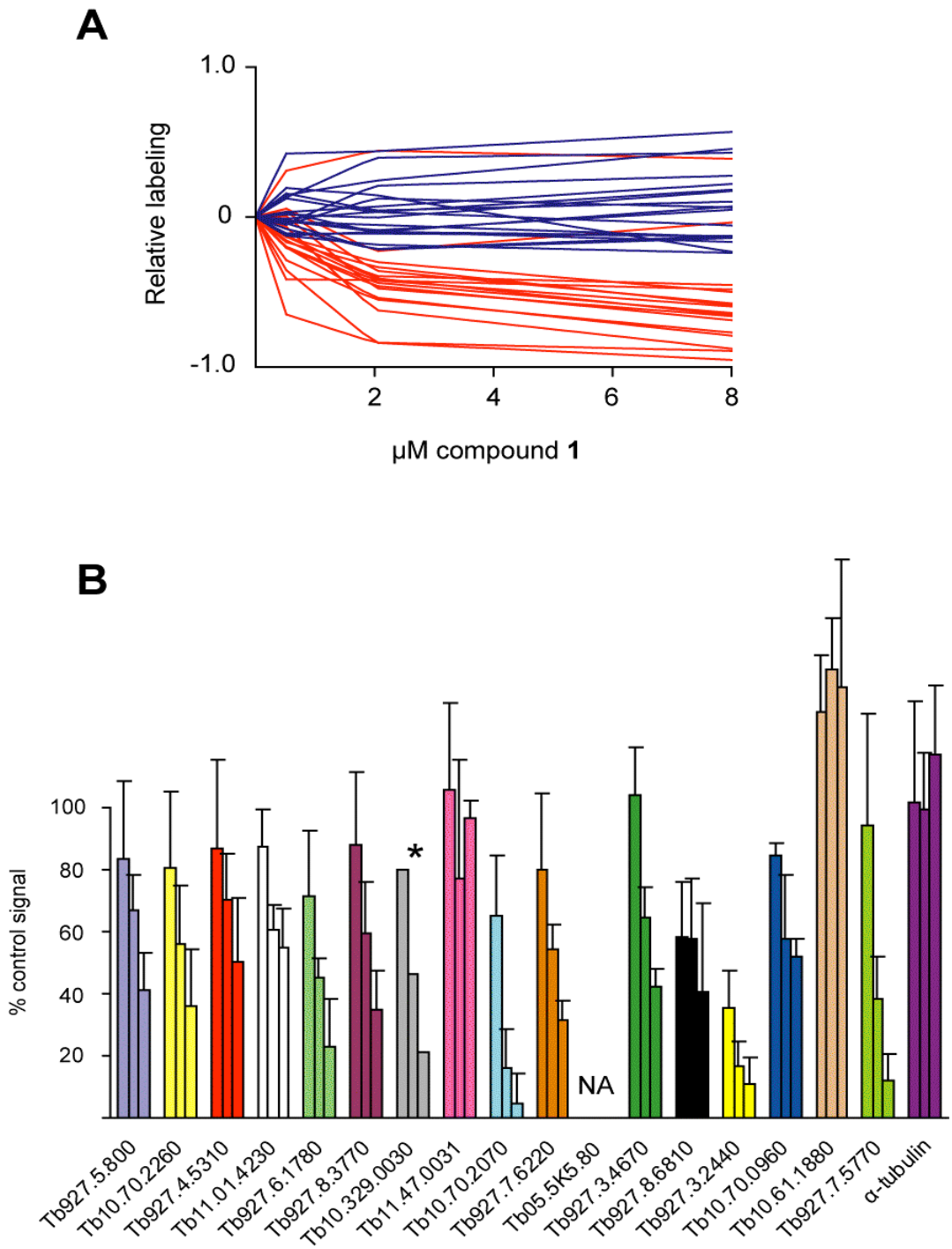


Figure 3-8: Quantitative analysis of competitive binding by non-propargylated fluorosulfone **1**. Whole cell lysates from the bloodstream form of *T. brucei* were incubated with **1**, then labeled with **47**. After click chemistry and affinity purification, samples were analyzed using iTRAQ/MS. (A) The relative abundance of proteins labeled by **47** was normalized to a DMSO control and plotted. Red trajectories are protein kinases, blue are all other identified proteins. (B) Relative abundance of labeled protein kinases after pretreatment with propargylated fluorosulfone **1** at 0.5, 2.0, or 8.0 μM. Data from a single peptide is indicated with an asterisk; all other data are normalized mean intensities for all identified peptides with error bars for standard deviations. iTRAQ analysis performed by Juan Oses-Prieto, UCSF.

(Tb10.70.2070) and the 46 kDa PKC/Akt homolog (Tb927.3.2440) have Phe and Met gatekeepers, respectively. This unexpected result shows that at least certain kinases with large gatekeepers can accomodate the bulky *p*-tolyl group shared by fluorosulfones **1** and **47**.

Competition experiments with the reversible inhibitor PD show a much more selective profile than **1**. PD is a tyrosine kinase inhibitor that inhibits BCR-ABL and Src in human cell lines at nanomolar concentrations, but PD may have several other cellular targets[10,20]. Though lacking tyrosine kinases, *T. brucei* show a significant decrease in cell viability when treated with micromolar concentrations of PD (Figure 3-9a). Analysis of fluorescent labeling experiments show that PD blocks labeling of several bands ranging in molecular weight from 45 to 50 kDa (Figures 3-3 and 3-5). The *p*-tolyl group of **47** and the dichlorophenyl ring of PD both require access to a large hydrophobic pocket adjacent to the ATP binding site for efficient binding. This requirement, and a shared affinity for Src suggest these compounds may share a number of kinase targets[16,20].

The ability of PD to compete with fluorosulfone **47** for labeling of specific kinases was assayed using iTRAQ/MS. As expected, this sterically more demanding inhibitor showed greater kinase selectivity compared to compound **1**. Nearly half of the labeled kinases show little to no decrease in labeling upon co-incubation with 8.0 μM PD, the highest concentration tested. Labeling of five kinases decreased >40% in the presence of 0.5 μM PD (Figure 3-9b). Consistent with a more stringent steric requirement for occupancy of the hydrophobic pocket, three of these kinases have a Thr gatekeeper. PD did not strongly compete with compound **47** for binding to endogenous TbPK50, despite this kinase's Thr gatekeeper. PD potently blocked labeling of the essential 50 kDa NIMA kinase (Tb927.4.5310, Figures 3-9b and 3-7). This essential kinase has a Met gatekeeper and may have otherwise been overlooked as a potential target for PD.

Inhibition of the 50 kDa NIMA kinase, Tb927.4.5310, may explain, in part, the trypanocidal effects of PD.

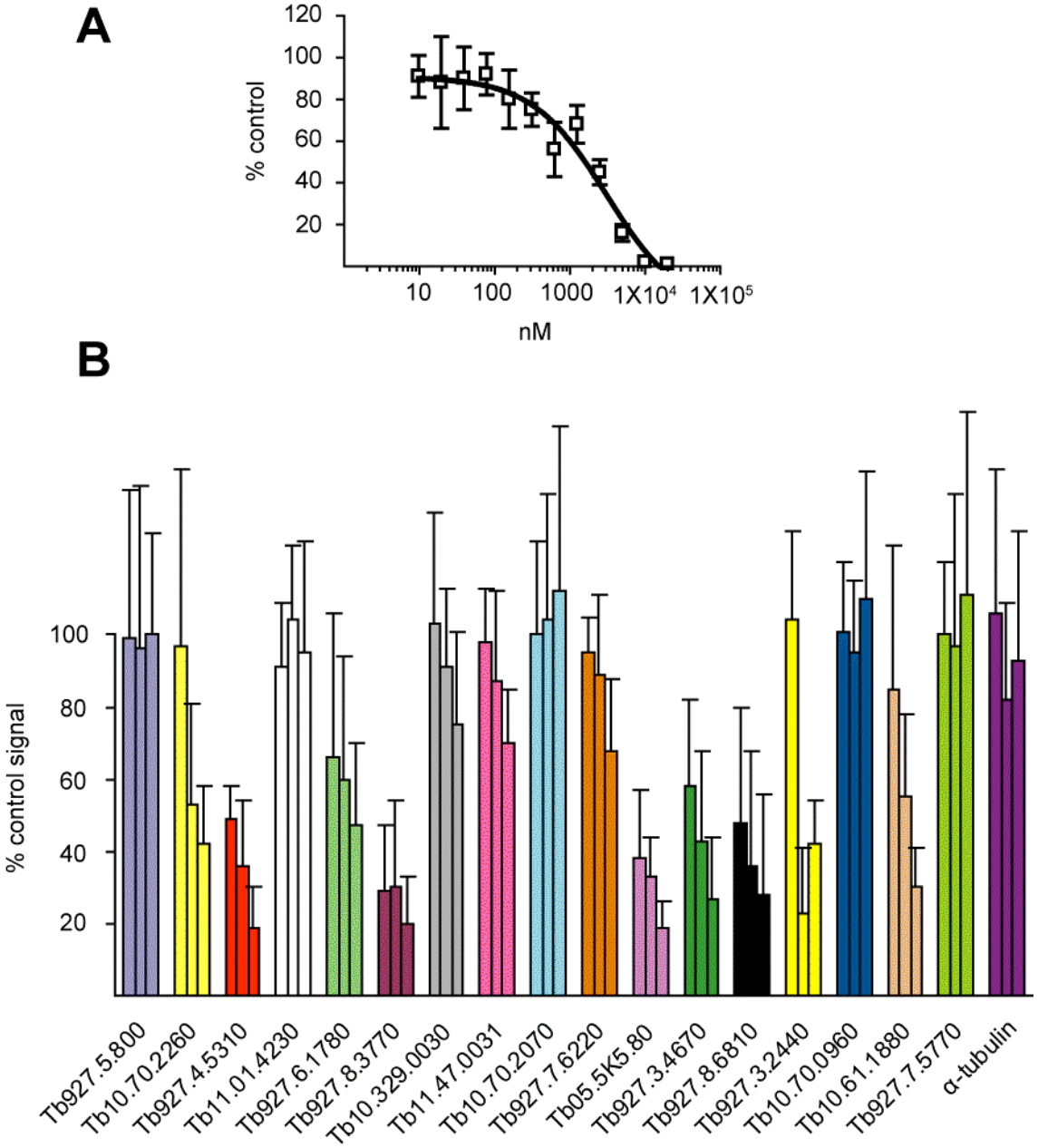
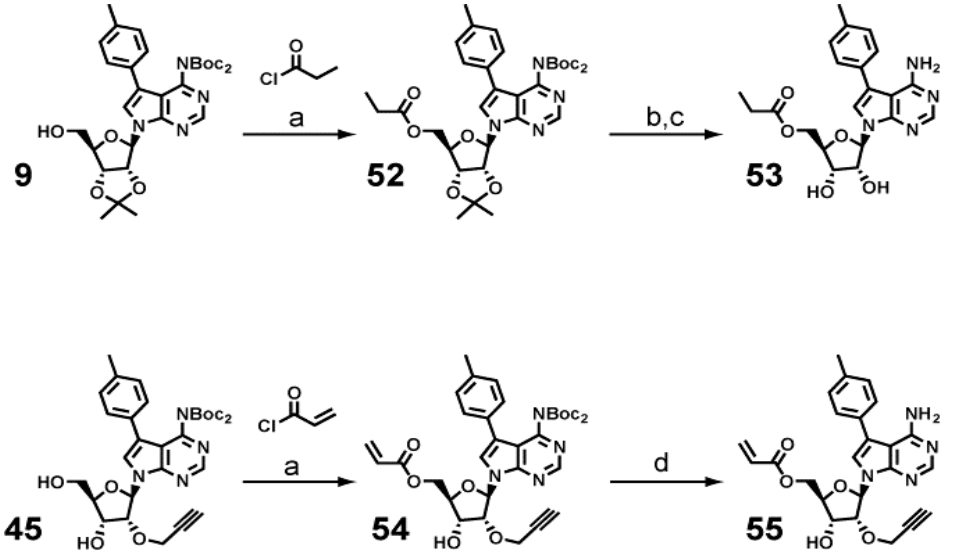


Figure 3-9: Effects of kinase inhibitor PD0166326 (PD). (A) Cultures of the bloodstream form of *T. brucei* were treated with PD for 24 h then assayed for living cells. Values are normalized to a DMSO control, and error bars represent standard deviation. (B) Relative recovery of labeled protein kinases after pretreatment with 0, 0.5, 2.0, or 8.0 μM PD. Whole cell lysates from the bloodstream form of *T. brucei* were incubated with PD, then labeled with **47**. After click chemistry and affinity purification, samples were analyzed using iTRAQ/MS. Values are the mean signal intensities for all identified peptides normalized to the DMSO control with error bars for standard deviations. iTRAQ analysis was performed by Juan Oses-Prieto, UCSF.

Changing the 5' electrophile yields a novel kinase target

Cell viability assays with *T. brucei* revealed that the 5'-acrylate nucleoside **19** was nearly equipotent to compound **1**. Compound **19** shares the same nucleoside scaffold as **1** and **47**, as well as an ability to inhibit Src kinase in a time-dependent manner (Figure 1-5). Given the utility of compound **47** as an activity-based probe, we synthesized the propargylated nucleoside **55** to identify proteins covalently bound by this 5'-acrylate (Scheme 3-2). We synthesized both the labeling reagent **55,** and a non-electrophilic control compound **53** in two steps from late-stage intermediates of the synthetic route described in Chapter 1.



Scheme 3-2: Synthesis of non-electrophilic nucleoside **53** and propargylated acrylate **55**. Reagents and conditions: (a) DIPEA, DCM (**52**: 87%, **54**: 25%); (b) TFA, DCM; (c) TFA, H_2O, THF (79% - 2 steps); (d) TMSOTf, DCM (100%).

Compound **55** showed remarkable toxicity when tested in *T. brucei* viability assays. As expected, the non-electrophilic compound **53** had an effect similar to other reversible inhibitors bearing the *p*-tolyl heterocycle (Figures 3-1, 3-10a). The 2'-propargyl ether had a remarkable effect on cell viability, dropping the IC_{50} to below 10 nM.

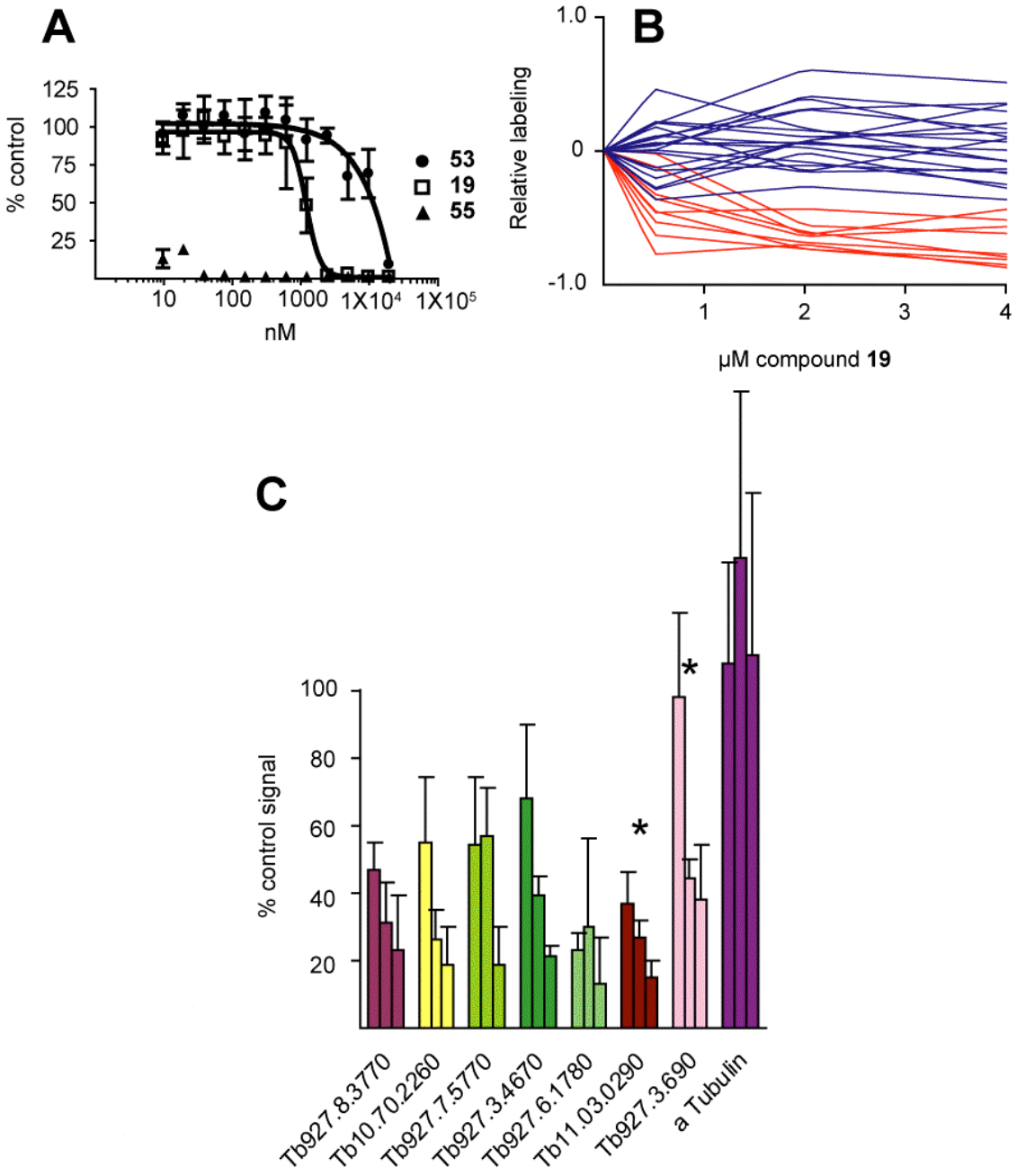


Figure 3-10: Effects of 5'-acrylate nucleosides. (A) Cultures of the bloodstream form of *T. brucei* were treated with compounds for 24 h, then assayed for viable cells. Values are normalized to DMSO control, and error bars represent standard deviations. (B) Quantitative analysis of competitive binding by acrylate **19**. Whole-cell lysates from the bloodstream form of *T. brucei* were incubated with **19**, then labeled with **55**. After click chemistry and affinity purification, samples were analyzed using iTRAQ/MS. The relative abundance of proteins labeled by **55** was normalized to a DMSO control and plotted. Red trajectories are protein kinases, blue are all other identified proteins. (C). Relative abundance of labeled protein kinases after pretreatment with 0.5, 2.0, or 4.0 μM **19**. Values are the mean signal intensity for all identified peptides normalized to the DMSO control with error bars for standard deviations. Asterisks indicate kinases labeled by **55** but not **47**. iTRAQ analysis was performed by Juan Oses-Prieto, UCSF.

A competitive binding assay with **55** showed that **19** preferentially binds a small group of protein kinases. Analysis of affinity purified proteins with iTRAQ showed that, like the fluorosulfone nucleosides, the 5'-acrylates preferentially target protein kinases (Figure 3-10b). Protein kinases represented ~12% of all identified proteins, and in all cases, their labeling was >50% blocked by 4.0 μM **19**. Among the identified kinases is TbPK50. Labeling of this essential kinase was >80% blocked by the highest concentration of **19**, which was more efficient than either fluorosulfone **1** or PD toward TbPK50. Acrylate **55** also labeled two kinases that were not detected in MS-experiments with fluorosulfone **47**. One, Tb11.03.0290, is a homologue of Bud32 in *S. cerevisiae*, and represents a fascinating kinase with ancient roots.

Bud32 is an atypical kinase conserved in all eukaryotes with homologues in archaea[42,43,44]. It phosphorylates a Grxp4p glutaredoxin in yeast and p53 in humans in an Akt-dependant manner[45,46,47]. Perhaps much more importantly, Bud32 is a component of the KEOPS complex whose core member Kae1 is conserved in all kingdoms of life. This complex is essential for normal cell growth, helping to maintain genome stability and telomere maintenance in higher organisms[42,48,49].

There are several reasons why this target of acrylate **55** is of particular interest. One of three *T. brucei* homologues of the Grxp4p glutaredoxin targeted by Bud32 (in yeast) is an essential gene[50]. Potentially more important is this enzyme's role within the KEOPS complex and telomere maintenance[51]. Double-stranded breaks at or near the telomeres of *T. brucei* play a critical role in antigen variation and avoidance of the immune response[52,53]. A third point of interest is the potential for selective inhibition of the parasite's homolog over host kinases. Bud32 substantially differs from the structure of other protein kinases (Figure 3-11a). Additionally, the sequence of the kinetoplastid homologs of Bud32 reveals a non-conserved cysteine within the putative "glycine-rich" loop. A similar cysteine in Src and FGFR kinases mediated the selective covalent

inhibition by compound **49** (Figures 2-5 and 2-10). Unfortunately, our MS analysis showed no trace of covalent modification by **55**, but the "glycine-rich" loop cysteine may hold the key to the selective binding of these acrylate compounds.

Given the striking difference in cytotoxicity between non-propargylated acrylate **19** and propargylated acrylate **55**, the target responsible for mediating the extreme toxicity of **55** may not be highly competed by **19**. One such non-kinase target, identified by MS, is inosine monophosphate dehydrogenase (IMPDH). This protein is among the non-kinase targets bound by **55**, but not by **47** (Supplementary Tables 3 - 5). Its inhibition by mycophenolic acid kills trypanosomes in culture and has been studied as a model for acquired drug resistance in *T. brucei*[54]. The fact that that IMPDH or some other enzyme may mediate the cytotoxic effects of acrylate **55** does not detract from the novelty of Bud32 or the other kinases that are competed by **19.** The non-propargylated compound **19** retains a modest toxicity, which is potentially caused by kinase inhibition. Additionally, the increased kinase selectivity displayed by acrylate **55** provides structure-activity relationships that will be useful in the future design of compounds that selectively inhibit *T. brucei* Bud32 kinase.

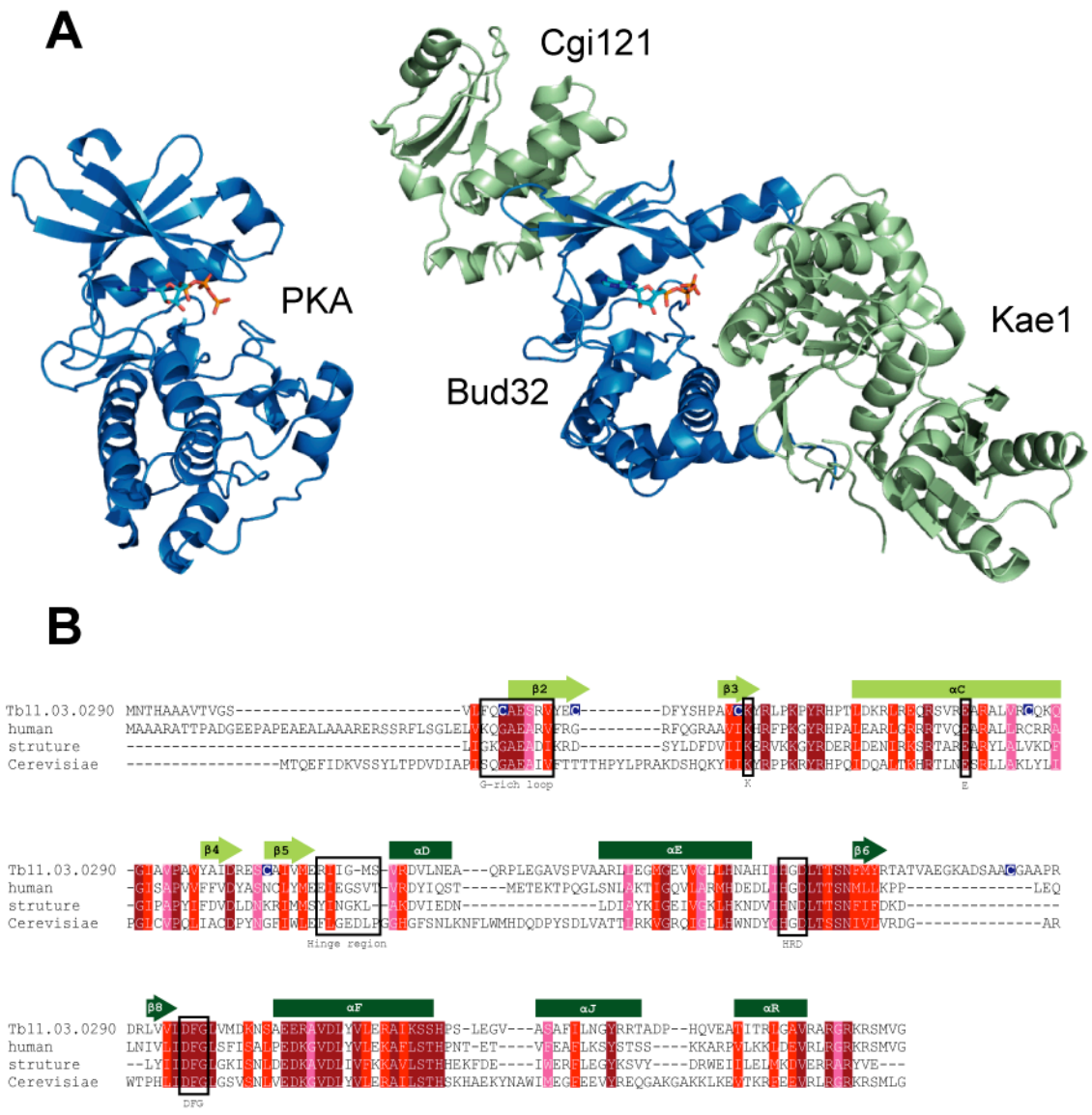


Figure 3-11: Tb11.03.0290 encodes a homologue of the atypical protein kinase Bud32. (A) Comparison of structures of archetype protein kinase PKA and Bud32 homologue from *Methanococcus jannaschii*. Bud32 is shown with other KEOPS complex members Cgi121 and Kae1. ATP was placed in the Bud32 structure based on a superposition with related kinase Rio1. Image was created in PyMOL using PDB files 1ATP, 3ENH, and 1ZP9. (B) Sequence alignment of Bud32 homologues from *T. brucei*, humans, *M jannaschii*, and *S. cereisiae* highlighting important residues and structural features. Marroon denotes regions of identity, red high conservation, and pink moderate conservation. Figure is based on results from Mao et al[42].

Conclusion and future directions:

We have explored the use of 5'-electrophilic nucleosides as activity-based probes for protein kinases in the eukaryotic parasite *Trypanosoma brucei*. The 5'-fluorosulfone **47** kills trypanosomes and specifically binds a subset of *T. brucei* protein kinases in cellular lysates. MS analysis of covalently modified proteins revealed a set of protein kinases, at least five of which are likely to be essential. Competitive profiling experiments with the reversible kinase inhibitor PD, demonstrated that fluorosulfone **47** can reveal at least a subset of the targets of reversible kinase inhibitors. Experiments with the 5'-acrylate **55** revealed a smaller subset of kinases bound by this electrophile compared to fluorosulfone **47**. Included in this subset is the essential gene product TbPK50 and an atypical kinase related to a budding yeast kinase, Bud32. Inferred functions of the trypanosomal Bud32, as well as the potential for its selective covalent inhibition, make this kinase of particular interest to kinetoplastid biology.

Use of these compounds has provided a glimpse into the trypanosome kinome, but much work remains to cement their utility as effective probes and inhibitors. The list of bound kinases must be further parsed for essential proteins and biological function. Our initial RNAi studies highlighted a few essential genes, but additional kinases in our set may prove to be essential or play roles in morphogenesis and/or virulence. Many more compounds like PD remain to be screened, and development of a medium throughput method using fluorescent tags and 2D gels could potentially speed this process. Our experiments used lysates for labeling; this was necessitated by the large amount of protein needed for the initial MS analysis. The cell permeability of these nucleosides and a list of potential targets clears the way for *in vivo* competition experiments. This will provide a more physiologically relevant analysis of competitive binding by potential kinase inhibitors. We have presented one hypothesis regarding the interaction of compound **55** with Bud32. Targeting the non-conserved cysteine in the

glycine-rich loop needs to be verified, along with a more thorough investigation of the biological function of this kinase. The electrophilic nucleosides presented here have opened many doors to future research opportunities.

Methods:

Synthesis

Compound **52**: Diisopropylethylamine (40 uL, 0.22 mmol) was added to a solution of **9** (22 mg, 0.037 mmol) and propionyl chloride (9.7 uL, 0.11 mmol), on an ice bath in methylene chloride (0.5 mL). The ice was allowed to melt as the reaction stirred overnight. The reaction was diluted into methylene chloride and washed with 0.1 N HCl then brine. The organic phase was concentrated and pure **52** (21 mg, 87%) was obtained by silica gel thin layer chromatography (1:1 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 1.08 (t, J = 7.6 Hz, 3H), 1.29 (s, 18H), 1.40 (s, 3H), 1.65 (s, 3H), 2.28 (dd, J = 2.0, 7.6 Hz, 2H), 2.39 (s, 3H), 4.25 (dd, J = 5.2, 11.2 Hz, 1H), 4.44 (m, 2H), 5.01 (dd, J = 3.6, 6.4 Hz, 1H),5.33 (dd, J = 2.4, 6.4 Hz, 1H), 6.35 (d, J = 2.4 Hz, 1H), 7.19 (d, J = 8.0 Hz, 2H), 7.32 (s, 1H), 7.33 (d, J = 8.0 Hz, 2H), 8.85 (s, 1H).

Compound **53**: Trifluoroacetic acid (0.5 mL) was added to a solution of compound **52** (21 mg, 0.032 mmol) in methylene chloride (0.5 mL). The reaction was stirred at room temperature 45 min, and the solvent was removed under a stream of argon. The residual oil was dissolved in tetrahydrofuran (0.1 mL) and aqueous trifluoroacetic acid (90%, 0.9 mL) was added. Solvent was removed under vacuum and compound **53** (10.5 mg, 79%) was obtained by HPLC purification (10 – 80 % acetonitrile in water with 0.1% trifluoroacetic acid). After removal of the residual solvent, 1HNMR revealed a mixture of 2 isomers in a ratio of 1:2.7 favoring the desired 5'-acylated α-isomer. The compound was used for assays without further purification. ^{1}HNMR (DMSO): δ 0.97 (t, J = 7.2 Hz, 3H), 2.30 (q, J = 7.2 Hz, 2H), 2.38 (s, 3H), 4.09 (m, 1H), 4.17 (m, 2H), 4.31 (m, 1H), 4.45 (t, J = 5.2 Hz, 1H), 6.18 (d, J = 4.8 Hz, 1H), 4.45 (t, J = 5.2 Hz, 1H), 6.18 (d, J = 4.8 Hz, 1H), 7.32 (d, J = 7.6 Hz, 2H), 7.40 (d, J = 7.6 Hz, 2H), 7.61 (s, 1H), 8.36 (s, 1H).

Compound **54**: Diisopropylethylamine (3 uL, 0.017 mmol) was added to a solution of **45** (10 mg, 0.017 mmol) in methylene chloride (1.0 mL) at -50 ºC. Acryloyl chloride (1.5 uL, 0.018 mmol) was added and the reaction was allowed to slowly warm to -10ºC over 30 min. The reaction was transferred to an ice bath and stirred overnight as the bath melted. The following day, the temperature was again lowered to -50ºC and additional base (3 uL, 0.017 mmol) and acryloyl chloride (1.0 uL, 0.012 mmol) were added. The temperature was slowly raised to 0ºC over the next two hours, and then the reaction was diluted into ethyl acetate and washed with 10% citrate (pH 4.0) and brine. The organic phase was dried over sodium sulfate and concentrated *in vacuo*. Compound **54** (2.7 mg, 25%) was obatined after purification by silca gel thin layer chromatography (1:2 hexanes:ethyl acetate). ^{1}HNMR ($CDCl_3$): δ 1.28 (s, 9H), 1.31 (s, 9H), 2.39 (s, 3H), 2.41 (t, J = 2.4 Hz, 1H), 2.71 (s(br), 1H), 4.31 (ddd, J = 3.2, 4.4, 6.0 Hz, 1H), 4.39 (dd, J = 2.4, 16.0 Hz, 1H), 4.45 (dd, J = 4.4, 16.0 Hz, 1H), 4.48 (dd, J = 2.4, 16.0 Hz, 1H), 4.49 (m, 1H), 4.57 (dd, J = 3.6, 5.2 Hz, 1H), 4.61 (dd, J = 3.2, 12.0 Hz, 1H), 5.79 (dd, J = 1.2, 10.4 Hz, 1H), 6.11 (dd, J = 10.4, 17.2 Hz, 1H), 6.42 (dd, J = 1.2, 17.2 Hz, 1H), 6.47 (d, J = 3.6 Hz, 1H), 7.18 (d, J = 8.0 Hz, 2H), 7.33 (d, J = 8.0 Hz, 2H), 7.40 (s, 1H), 8.83 (s, 1H).

Compound **55**: Trifluoroacetic acid (0.25 mL) was added to a solution of compound **54** (2.7 mg, 0.004 mmol) in methylene chloride (0.25 mL). After 45 min, the solvent was blown off with a stream of argon and residual volatiles were removed *in vacuo* to yield pure **55**, which was dissolved directly in deuterated dimethylsulfoxide for analysis and biological assays. ^{1}HNMR (DMSO): δ 2.38 (s, 3H), 3.43 (t, J = 2.4 Hz, 1H), 4.19 (m, 1H), 4.25 (dd, J = 2.4, 16.0 Hz, 1H), 4.31 (dd, J = 6.0, 12.0 Hz, 1H), 4.33 (dd, J = 2.4, 16.0 Hz, 1H), 4.41 (m, 1H), 4.42 (dd, J = 4.4, 16.0 Hz, 1H), 4.57 (t, J = 6.0 Hz, 1H), 5.94 (d, J = 10.4 Hz, 1H), 6.20 (dd, J = 10.4, 17.2 Hz, 1H), 6.33 (d, J = 6.0 Hz, 1H), 6.35 (d, J = 17.2

Hz, 1H), 7.34 (d, J = 8.0 Hz, 2H), 7.39 (d, J = 8.0 Hz, 2H), 7.69 (s, 1H), 8.48 (s, 1H). LC/MS [$M+H^+$] m/z = 449.4

Biotin azide: A biotinylated azide (39 mg, 51%) was prepared in four steps from 6-bromohexanoic acid as previously described[55].

Cell culture

Strains of the bloodstream form of *T brucei* were cultured in HMI 9 media supplemented with 10 % fetal bovine serum and 10% Serum Plus (Sigma). The 90-13 strain was maintained in the presence of 5 µg/mL Hygromycin and 2.5 µg/mL G418.

Growth inhibition

Compounds were added in 1 µL of DMSO to 99 µL of 221 strain cells at a density of 1 X 10^5 cells/mL. Cells were cultured for 24 hours and cell number was assayed using the CellTiterGlo cell viability assay (Promega).

RNAi

Gene fragments were cloned from genomic DNA with gene-specific primers (Table 3-1) flanked by XhoI and HindIII cut sites. Fragments were ligated into the pZJM vector and propagated in DH5α cells. Vectors were linearized with NotI and transfected into the 90-13 strain as previously described[56]. Transformants were selected with 2.5 µg phleomycin and clonal populations were isolated from agar plates. RNAi was induced by addition of 1 µg/mL tetracycline and cell number was assayed daily using the CellTiterGlo cell viability assay (Promega).

Competitive labeling

Cells were pelleted and lysed in PBS with 1X Complete Protease Inhibitor Cocktail (Roche), and 1X Phosphatase Inhibitor Cocktail 2 (Sigma) by brief sonication (*T. brucei*) or freeze-thaw cycle (COS-7). Lysates were spun at top speed in a bench-top centrifuge to pellet cellular debris. Clarified lysates were adjusted to 1 mg/mL (*T. brucei*) and 10 mg/mL (COS-7). Competitor compounds were added and incubated for 30 min at room temperature before adding 2 μM compound **47** for an additional 30 min. Click reactions were preformed on a 25 μL scale. To 20.75 μL lysate was added 1.25 μL 20% SDS, 0.5 μL 5 mM azide tag, 0.5 μL 50 mM TCEP (~pH 7.0), 1.5 uL 1 mM TBTA ligand in 1:4 DMSO, *tert*-butyl alcohol, and 0.5 μL 5 mM $CuSO_4$. After one hour at room temperature, 6 μL 5 X Laemmli sample buffer was added and samples were either used immediately or snap frozen in liquid N_2 and stored at -80ºC.

TbPK50 expression

BL21-CodonPlus(DE3)-RIL cells (Stratagene) were transfected with a pRSET vector containing an 5' truncation of Tb10.70.2260 (Generous gift from Jose Garcia-Salcedo, Laboratory of Molecular Parasitology, ULB)[23]. Expression was induced overnight with 0.1 mM IPTG at 18º C. Cells pellets were sonicated in 50 mM Tris 8.0, 500 mM NaCl, 5 mM Imidazole, and 1 mM DTT. Clarified lysates were incubated with Ni-NTA agarose and TbPK50 was eluted with 50 mM Tris 8.0, 150 mM NaCl, 300 mM Imidazole, 1 mM DTT. The buffer was exchanged for 50 mM HEPES 7.5, 0.1 mM EGTA, 150 mM NaCl, 0.05% BME, and 0.03% Brij-35 to yield pure TbPK50 at 2.6 mg/mL.

Affinity purification of labeled TbPK50

TbPK50 in PBS (1 μM) was incubated with 5 uM **47** for 30 min at room temperature and the following reagents were added 1% SDS, 0.1 mM biotin azide, 1 mM TCEP, 0.1 mM TBTA ligand, and 1 mM $CuSO_4$. After one hour at room temperature, NP-40 was added

to 1% and proteins were precipitated with TCA. Pellets were resuspended in 0.1 M Tris (pH 8.8), 1% SDS the diluted 1:10 in 1% NP-40 in PBS. Avidin-agarose (Sigma) was added, and samples were incubated on a rotator overnight at 4ºC. Beads were washed five times with 1% NP-40, 0.1% SDS in PBS then divided and either boiled in 1.5 X Laemmli sample buffer with 1 mM biotin or hydrolyzed with 0.2 N NaOH, neutralized and brought to 1X with concentrated Laemmli sample buffer. Equivalent fractions of all samples were resolved by SDS-PAGE and blotted with anti-His antibody (Santa Cruz Biotechnology) and HRP conjugated streptavidin (Jackson Immunoresearch).

Affinity purification of *T. brucei* proteins bound to compound 47 and 55

Lysate from the bloodstream form of *T. brucei* (3 mg/mL) in PBS were prepared as above, then passed through a PD-10 desalting column (GE Healthcare Sciences), eluting with additional lysis buffer. Lysates (18 mg for gel analysis, 3 mg/treatment for iTRAQ) were labeled with either 2 μM compound **47** or 0.5 μM compound **55** for 30 minutes at room temperature. 20% SDS was added to a final concentration of 1% and the click reagents were added to the following concentrations: 0.1 mM biotin azide, 1 mM TCEP, 0.1 mM TBTA ligand, 1 mM $CuSO_4$ and incubated for one hour at room temperature. Proteins were precipitated with cold acetone (80% final concentration) and pellets were washed three times with additional acetone and dried. Pellets were resuspended (with bath sonication) in a minimal volume of 50 mM Tris (pH 8.0), 1% SDS, then diluted with 1% NP-40 in PBS and passed through a PD-10 column, eluting with 1% NP-40 and 0.1% SDS in PBS. Avidin-agarose (Sigma) was added (50 μL for gel analysis, 30 μL for iTRAQ) and the samples were rotated overnight at 4ºC. Supernatant was removed and the beads were washed twice for one hour at room temperature with 1% NP-40, 0.1% SDS in PBS, twice for one hour at 4ºC with 6M urea in PBS, one hour with PBS at 4ºC, and one additional hour with PBS at room temperature. The

supernatant was removed, 20 µL of 0.4 N NaOH were added, and the beads were incubated for 20 minutes at room temperature before neutralization with 10 µL of 0.8 N HCl. 20% SDS was added to a final concentration of ~1%, and the beads were heated to 90ºC for 3 minutes before collecting the supernatant for analysis.

Gel analysis of affinity purified proteins

Eluted proteins were diluted 1:1 with 2X Laemmli sample buffer and loaded onto a 12% precast polyacrylamide gel (Readygel, BioRad Laboratories) for SDS-PAGE analysis. The gel was stained, cut into 16 slices, and submitted to the UCSF Mass Spectrometry Facilty for in-gel digestion and MS analysis using a QSTAR Pulsar mass spectrometer (Applied Biosystems/MDS Sciex)[57]. Digest and MS analysis were performed by Juan Oses-Prieto, Burlingame Laboratory, UCSF.

iTRAQ analysis of competitive labeling – All methods and analysis after affinity purification were performed by Juan Oses-Prieto, Burlingame Laboratory, UCSF.
Lysates were prepared as described above then treated with competitor compounds for 30 minutes before labeling and affinity purification as previously described.

Eluted proteins were acetone precipitated and pellets were washed twice with additional cold acetone. Dried pellets were resuspended in 8 M guanidinium HCl then reduced (5 mM TCEP, 50 mM ammonium bicarbonate, 6 M guanidinium HCl) and alkylated (10 mM IAA). The solution was adjusted to 1 M guanidinium HCl, 100 mM ammonium bicarbonate, 10% acetonitrile and trypsinized overnight. Volatiles were removed and peptides resuspended in 0.1% formic acid, extracted with C18 OMIXtips (Varian) and eluted with 50% acetonitrile. Volatiles were removed and peptides were resuspended in 0.5 M triethylammonium bicarbonate (pH 8.5) and labeled with iTRAQ Reagents (Applied Biosystems). Samples were mixed, dried, and resuspended in 30%

acetonitrile 5 mM potassium phosphate (pH 2.7), and fractionated on a 50.0 X 1.0 mm 5 uM 200Å Polysulfoethyl A column (PolyLC) with a 1-40% gradient of 350 mM potassium chloride in 30% acetonitrile 5 mM potassium phosphate (pH 2.7). Fractions were dried, resuspended in 0.1% formic acid, extracted using C18 ZipTips (Millipore), eluted with 50 % acetonitrile, dried, and resuspended in 0.1% formic acid.

The digests were separated by nano-flow liquid chromatography using a 75-µm x 150-mm reverse phase C18 PepMap column (Dionex-LC-Packings) at a flow rate of 350 nL/min in a NanoLC-1D Proteomics high-performance liquid chromatography system (Eksigent Technologies) equipped with a FAMOS autosampler (Dionex-LC-Packings). Peptides were eluted using a 30 minute 5-50% gradient of acetonitrile with 0.1% formic acid. The eluate was coupled to a microionspray source attached to a QSTAR Pulsar mass spectrometer (Applied Biosystems/MDS Sciex).

Peptides were analyzed in positive ion mode. MS spectra were acquired for 1s in the m/z range between 310 and 1400. MS acquisitions were followed by four 2.5-s collision-induced dissociation (CID) experiments in information-dependent acquisition mode. For each MS spectrum, the 2 most intense multiple charged peaks over a threshold of 30 counts were selected for generation of CID mass spectra. 2 MSMS spectra of each one of them were acquired, first on the m/z range 180-1500, and then on the m/z range 112-120.The CID collision energy was automatically set according to mass to charge (m/z) ratio and charge state of the precursor ion. A dynamic exclusion window was applied which prevented the same m/z from being selected for 1 min after its acquisition.

Peak lists were generated using the mascot.dll script. The peak list was searched against the Trypanosoma subset of the NCBInr database as of December 16, 2008 (containing 30522 entries) using in-house ProteinProspector version 5.2.2 A minimal ProteinProspector protein score of 15, a peptide score of 15, a maximum expectation

value of 0.1 and a minimal discriminant score threshold of 0.0 were used for initial identification criteria. Modification of the amino terminus or the epsilon-amino group of lysines, carbamidomethylation of cysteine; acetylation of the N terminus of the protein and oxidation of methionine were allowed as variable modifications. Peptide tolerance in searches was 100 ppm for precursor and 0.2 Da for product ions, respectively. Peptides containing two miscleavages were allowed. The number of modification was limited to two per peptide.

Quantization was based on the relative areas of the reporter ions generated by the isobaric iTRAQ reagents at m/z 114, 115, 116 and 117 during CID experiments. Abundance ratios of individual peptides between the different samples were calculated taking as reference the vehicle-treated sample, by dividing the areas of their respective iTRAQ reporter ions. Peptides with peak areas lower than 30 for the most intense reporter ion were discarded. For changes in relative abundance at the protein level, all MSMS taken for the different peptides belonging to a particular protein were used to calculate the average and SD of the abundance ratios.

References:

1) Simpson AG, Stevens JR, Lukes J. The evolution and diversity of kinetoplastid flagellates. Trends in parasitol 2006, 22: 168-174.

2) Barrett MP, Burchmore RJS, Stich A, Lazzari JO, Frasch AC, Cassulo JJ, Krishna S. The trypanosomiases. Lancet 2003, 362:1469-1480.

3) Kedzierski L, Sakthianandesvaren A, Curtis JM, Andrews PC, Junk PC, Kedzierska K. Leishmaiasis: current treatment and prospects for new drugs and vaccines. Curr Med Chem. 2009, 16: 599-614.

4) Docampo R, Moreno SNJ. Current chemotherapy of human African trypanosomiasis. Parasitol Res. 2003, 90: S10-S13

5) Berriman et. al. The genome of the African trypanosome, Trypanosoma brucei. Science 2005, 309: 416-422.

6) Zhang J, Yang PL, Gray NS. Targeting cancer with small molecule kinase inhibitors. Nature Reviews Cancer 2009, 9: 28-39.

7) Naula C, Parsons M, Mottram JC. Protein kinases as drug targets in trypanosomes and Leishmania. Biochimica et Biophysica Acta 2005, 1754: 151-159.

8) Parsons M, Worthey EA, Ward PN, Mottram JC. Comparative analysis of the kinomes of three pathogenic trypanosomatids: Leishmania major, Trypanosoma brucei and Trypanosoma cruzi. BMC Genomics 2005, 6:127.

9) Campbell DA, Szardenings AK. Functional profiling of the proteome with affinity labels. Curr Opin Chem Biol. 2003, 7: 296-303.

10) Wissing J, Godl K, Brehmer D, Blencke S, Weber M, Habenberger P, Stein-Gerlach M, Missio A, Cotton M, Muller S, Daub H. Chemical proteomic analysis reveals alternative modes of action for the pyrido[2,3-d]pryimidine kinase inhibitors. Molecular & Cellular Proteomics 2004, 3: 1181-1193.

11) Bantscheff M, Eberhard D, Abrahma Y, Basteck S, Boesche M, Hobson S, Mathieson T, Perrin J, Raida M, Rau C, Reader V, Sweetman N, Baur A, Bouwmeester T, Hopf C, Kruse U, Neubauer G, Ramsden N, Rick J, Kuster B, Drewes G. Quantitative chemical proteomics reveals mechanisms of action of clinical ABL kinase inhibitors. Nature Biotechnology 2007, 25: 1035-1044.

12) Alexander JP, Cravatt BF. Mechanism of carbamate inactivation of FAAH: implications for the design of covalent inhibitors and in vivo functional probes for enzymes. Chemistry & Biology 2005, 12: 1179-1187.

13) Saghatelian A, Jessani N, Joseph A, Humphrey M, Cravatt BF. Activity-based probes for the proteomic profiling of metalloproteases. PNAS 2004, 101: 10000-10005.

14) Patricelli MP, Szrdenings AK, Liyangage M, Nomanbhoy TK, Wu M, Weissig H, Aban A, Chun D, Tanner S, Kozarich JW. Functional interrogation of the kinome using nucleotide acyl phosphates. Biochemistry 2007, 46: 350-358.

15) Speers AE, Cravatt BF. Profiling enzyme activities in vivo using click chemistry methods. Chem Biol 2004, 11: 535-546.

16) Liu Yi, Bishop A, Witucki L, Kraybill B, Shimizu E, Tsien J, Ubersax J, Blethrow J, Morgan DO, Shokat KM. Structural basis for selective inhibition of Src family kinases by PP1. Chemistry & Biology 1999, 6: 671-678.

17) Bain J, Plater L, Elliott M, Shpiro N, Hastie CJ, Mclauchlan H, Klevernic I, Arthur JSC, Alessi DR, Cohen P. The selectivity of protein kinase inhibitors: a further update. Biochem J 2007, 408: 297-315.

18) Michalczhk A, Kluter S, Rode HB, Simard JR, Grutter C, Rabiller M, Rauh D. Sturctural insights into how irreversible inhibitors can overcome drug resistance in EGFR. Nat Chem Biol. 2007, 3: 229-238.

19) Zoller MJ, Nelson, NC, Taylor SS. Affinity labeling of cAMP-dependent protein kinase with p-fluorosulfonylbenzoyl adenosine. Covalent modification of lysine 71 J Biol Chem 1981, 256: 10837-10842.

20) Kraker AJ, Hartl BG, Amar AM, Barvian MR, Showlter HDH, Moore CW. Biochemical and cellular effects of c-Src kinase-selective pyrido[2,3-d]pyrimidine tyrosine kinase inhibitors. Biochemical Pharmacology 2000, 60: 885-898.

21) Lisnock JM, Tebben A, Frantz B, O'Neill EA, Croft G, O'Keefe SJ, Li B, Hacker C, de Laslo S, Smith A, Libby V, Liverton N, Hermes J, LoGrasso P. Molecular basis for p38 protein kinase inhibitor specificity. Biochemistry 1998, 37: 16573-16581.

22) Zhu X, Kim, JL, Newcomb JR, Rose PE, Stover DR, Toledo LM, Zhao H, Morgenstern KA. Structural analysis of the lymphocyte-specific kinase Lck in complex with Src family selective kinase inhibitors. Structure 1999, 7: 651-661.

23) Garcia-Salcedo JA, Nolan DP, Gijon P, Gomez-Rodriguez J, Pays E. A protein specifically associated with proliferative forms of Trypanosoma brucei is functionally related to a yeast kinase involved in the co-ordination of cell shape and division. Molecular Microbiology 2002, 45: 307-319.

24) Hammarton TC, Lillico SG, Welburn SC, Mottram JC. Trypanosoma brucei MOB1 is required for accurate and efficient cytokinesis but not exit from mitosis. Molecular Microbiology 2005, 56: 104-116.

25) Moore LL, Fulton AM, Harrison ML, Geahlen RL. Anti-sulfonylbenzoate antibodies as a tool for the detection of nucleotide-binding proteins for functional proteomics. J Proteome Res 2004, 3: 1184-1190.

26) Allocco JJ, Donald R, Zhong T, Lee A, Tang YS, Hendrickson RC, Liberator P, Nare B. Inhibitors of casein kinase 1 block the growth of Leishmania major promastigotes in vitro. International Journal for Parasitology 2006, 36: 1249-1259.

27) Tu X, Wang CC. Pairwise knockdowns of cdc-2-related kinases (CRKs) in Trypanosoma brucei identified the CRKs for G1/S and G2/M transitions and demonstrated distinctive cytokinetic regulations between two developmental stages of the organism. Eukaryot Cell 2005, 4: 755-764.

28) Gourguechon S, Savich JM, Wang CC. The multiple roles of cyclin E1 in controlling cell cycle progression and cellular morphology of Trypanosoma brucei. J Mol Biol 2007, 368: 939-950.

29) Guttinger A, Schwab C, Morand S, Roditi I, Vassella E. A mitogen-activated protein kinase of Trypanosoma brucei confers resistance to temperature stress. Molecular & Biochemical Parasitology 2007, 153: 203-206.

30) Pfister DD, Burkard G, Morand S, Renggli CK, Roditi I, Vassella E. A mitogen-activated protein kinase controls differentiation of bloodstream forms of Trypanosoma brucei. Eukaryotic Cell 2006, 5: 1126-1135.

31) Ojo, KK, Gillespie JR, Riechers AJ, Napuli AJ, Verlinde CL, Buckner FS, Gelb MH, Domostij MM, Wells SJ, Scheer A, Wells TN, Van Voorhis WC. Glycogen synthase kinase 3 is a potential drug target for African trypanosomiasis therapy.

32) Wang Z, Morris JC, Drew ME, Englund PT. Inhibition of Trypanosoma brucei gene expression by RNA interference using an integratable vector with opposing T7 promoters. J Biol Chem. 2000, 275: 40174-40179.

33) Racki WJ, Becam AM, Nasr F, Hervert CJ. Cbk1p, a protein similar to the human myotonic dystrophy kinase, is essential for normal morphogenesis in Saccharomyces cerevisiae. EMBO J. 2000, 19: 4524-4532.

34) Bidlingmaier S, Weiss EL, Seidel C, Drubin DG, Snyder M. The Cbk1 pathway is important for polarized cell growth and cell separation in Saccharomyces cerevisiae. Mol Cell Biol 2001, 21: 2449-2462.

35) Verde F, Wiley DJ, Nurse P. Fission yeast orb6, a ser/thr protein kinase related to mammalian rho kinase and myotonic dystrophy kinase, is required for maintenance of cell polarity and coordinates cell morphogenesis with the cell cycle. Proc Natl Acad Sci USA 1998, 95: 7526-7531.

36) Urbaniak MD. Casein kinase 1 isoform 2 is essential for bloodstream form Trypanosoma brucei. Mol Biochem Parasitol. 2009, 166: 183-185.

37) Colwill K, Pawson T, Andrews B, Prasad J, Manley JL, Bell JC, Duncan PI. The Clk/Sty protein kinase phosphorylates SR splicing factors and regulates their intranuclear distribution. EMBO J. 1996, 15: 265-275.

38) Muraki M, Ohkawara B, Hosoya T, Onogi H, Koizumi J, Koizumi T, Sumi K, Yomoda J, Murray MV, Kimura H, Furuichi K, Shibuya H, Krainer AR, Suzuki M, Hagiwara M. Manipulation of alternative splicing by a newly developed inhibitor of Clks. J Biol Chem. 2004, 279: 24246-24254.

39) Liang X, Haritan A, Uliel S, Michaeli S. Trans and cis splicing in trypanosomatids: mechanism, factors and regulation. Eukaryotic Cell 2003, 2: 830-840.

40) Stuart KD, Schnaufer A, Ernst NL, Panigrahi AK. Complex management: RNA editing in trypanosomes. Trends Biochem Sci. 30: 97-105.

41) Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He Fm Jacobson A, Pappin DJ. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. Molecular & Cellular Proteomics 2004, 3: 1154-1169.

42) Mao DL, Neculai D, Downey M, Orlicky S, Haffani YZ, Ceccarelli DF, Ho JSL, Szilard RK, Zhang W, Ho CS, Wan L, Fares C, Rumpel S, Kurinov I, Arrowsmith CH, Durocher D, Sicheri F. Atomic structure of the KEOPS complex: an ancient protein kinase-containing molecular machine. Molecular Cell 2008, 24:259-275.

43) Lower BH, Potters MB, Kennelly PJ. A phosphorprotein from the archaeon Sulfolobus solfataricus with protein-serine/threonine kinase activity. J Bacteriology 2004, 186: 463-472.

44) Facchin S, Lopreiato R, Ruzzene M, Marin O, Sartori G, Cotz C, Montenarh M, Carignani G, Pinna LA. Functional homology between yeast piD261/Bud32 and human PRPK: both phosphrylate p53 and PRPK partially complements piD261/Bud32 deficiency. FEBS Letters 2003, 549: 63-66 .

45) Lopreiato R, Facchin S, Sartori G, Arrigoni G, Casonato S, Ruzzene M, Pinna LA, Carignani G. Analysis of the interation between piD261/Bud32, and evolutionarily conserved protein kinase of Saccharomyces cerevisiae, and the Grx4 glutaredoxin. Biochem J. 2004, 377: 395-405.

46) Peggion C, Lopreiato R, Casanova E, Ruzzene M, Facchin S, Pinna LA, Carignani G, Sartori G. Phosphorylation of the Saccharomyces cerevisiae Grx4p glutaredoxin by the Bud32p kinase unveils a novel signaling pathway involving Sch9p, a yeast member of the Akt/PKB subfamily. FEBS Journal 2008, 275: 5919-5933.

47) Facchin S, Ruzzene M, Peggion C, Sartori G, Carignani G, Marin O, Brustolon F, Lopreiato R, Pinna LA. Phosphorylation and activation of the atupical kinase p53-related protein kinase (PRPK) by Akt/PKB. Cell Mol Life Sci. 2007, 64: 2680-2689.

48) Hecker A, Graille M, Madec E, Gadelle D, Le Cams E, van Tilbergh H, Forterre P. The universal Kae1 protein and the associated Bud32 kinase (PRPK), a mysterious protein couple probably essential for gemome maintenance in Archaea and Eukarya. Biochem Soc Trans. 2009, 37: 29-35.

49) Bianchi A, Shore D. The KEOPS complex: a rosetta stone for telomere regulation? Cell 2006, 124: 1125-1128.

50) Comini MA, Rettig J, Dirdjaja N, Hanschmann E, Berndt C, Krauth-Siegl RL. Monothil glutaredoxin-1 is an essential iron-sulfur protein in the mitochondrion of African trypanosomes. J Biol Chem. 2008, 41: 27785-27798.

51) Downey M, Houlsworth R, Maringele J, Rollie A, Brehme M, Galicia S, Guillard S, Partington M, Zubko MK, Krogan NJ, Emili A, Greenblatt JF, Harrington L, Lydall D, Durocher D. A genome-wide screen identifies the evolutionarily conserved KEOPS complex as a telomere regulator. Cell 2006, 124: 1155-1168.

52) Boothroyd DE, Dreesen O, Leonova T, Li KI, Fiqueiredo LM, Cross GA, Papavasiliou FN. A yeast-endonuclease-generated DNA break induces antigenic switching in Trypanosoma brucei. Nature 2009, 459: 278-281.

53) Dreesen O, Li B, Cross GAM. Telomere structure and function in trypanosomes: a proposal. Nature Reviews Microbiology 2007, 5: 70-75.

54) Wilson K, Berens RL, Sifri CD, Ullman B. Amplification of the inosinate dehydrogenase gene in Trypanosoma brucei gambiense due to an increase in chromosome copy number. J Biol Chem. 1994, 269: 28979-28987.

55) Grandjean C, Boutonnier A, Guerreiro C, Fournier JM, Mulard LA. On the preparation of carbohydrate-protein conjugates using the traceless Staudinger ligation. J Org Chem. 2005, 70: 7123-7132.

56) Tu X, Wang CC. The involvement of tow cdc2-related kinases (CRKs) in Trypanosoma brucei cell cycle regulation ad the distinctive stage-specific phenotypes caused by CRK3 depletion. J Biol Chem. 2004, 279: 20519-20528.

57) Rosenfeld J, Capdevielle J, Guillemot JC, Ferrara P. In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. Anal Biochem. 1992, 203: 173-179.

Appendix:

Supplementary Table 1: X-ray data collection and refinement statistics

space group	P1
unit-cell parameters (Å)	a=41.9; b=63.4; c=74.1; alpha=78.9; ,beta=88.6 ,gamma=90.0 deg.
resolution range (Å)	50.0 - 2.2
highest resolution shell (Å)	2.32 - 2.20
measured reflections	72746 (10624)
unique reflections	36599 (5287)
redundancy	2.0 (2.0)
completeness (%)	96.4 (95.4)
average I/σ	9.6 (2.3)
Rcryst	20.7 (29.7)
Rfree	26.3 (35.1)
Wilson B-value (Å2)	31.5
Ramachandran (%)	
favored	93.7
allowed	4.3
outliers	0.4
rms deviations	
bond lengths (Å)	0.019
bond angles (o)	1.9

Values in parentheses are for the highest resolution bin

Supplementary Table 2: Compound **47** binding proteins found by MS analysis after in gel digestion of affinity purified lysates from *T. brucei* (figure 3-5). Gray indicates a protein kinase.

NCBI #	Peptides	MW	NCBI Definition
71747410	47	50	protein kinase
13926114	44	50	serine/threonine protein kinase TbPK6
74025264	39	75	heat shock protein 70
72390395	14	71	heat shock 70 kDa protein, mitochondrial precursor, putative
72388918	36	38	casein kinase, putative
139611	32	51	Variant surface glycoprotein MITAT 1.2 precursor (VSG 221)
115504281	28	50	beta tubulin
115504283	27	50	alpha tubulin
72386827	30	46	serine/threonine-protein kinase, putative
71748504	26	81	heat shock protein 83
72393479	29	243	acetyl-CoA carboxylase, putative
72391902	24	61	protein kinase, putative
71755837	26	71	glucose-regulated protein 78
15626363	19	51	hexokinase
72392815	20	47	mitogen-activated protein kinase, putative
71748500	23	64	protein kinase
71746820	21	49	elongation factor 1-alpha
71754589	21	37	cell-division control protein 2 homolog 6
72391812	23	53	protein kinase, putative
72390009	17	46	mitogen-activated protein kinase, putative
71746974	18	47	enolase
62176656	18	39	protein kinase, putative
74025600	17	46	elongation factor 1 gamma
15485709	17	56	glycerolkinase
72393453	16	74	3-methylcrotonyl-CoA carboxylase, putative
71747296	17	94	elongation factor 2
53849855	15	35	glycosomal glyceraldehyde phosphate dehydrogenase
11245496	17	91	C-terminal kinesin KIFC1
72388510	15	50	serine/threonine-protein kinase A, putative
72393481	4	50	serine/threonine-protein kinase A, putative
71747946	11	41	methyltransferase
72387269	11	48	cell division control protein, putative
71747502	14	41	fructose-bisphosphate aldolase glycosomal
115141	11	56	Bloodstream-specific protein 2 precursor
71747654	12	60	chaperonin Hsp60
71748842	12	91	heat shock protein
72393421	10	47	protein kinase, putative
71744742	11	40	arginine kinase
72392947	9	37	nucleoside phosphorylase, putative
71747566	7	55	protein kinase
71747564	5	55	protein kinase
71755425	7	27	triosephosphate isomerase
266429	10	54	Pyruvate kinase 2 (PK 2)
70800580	10	71	serine/threonine protein kinase, putative
71749480	9	48	peptidylprolyl isomerase-like protein
71755063	8	67	adenylosuccinate synthetase
72392767	8	38	glycerol-3-phosphate dehydrogenase [NAD+], glycosomal
71744678	8	22	tryparedoxin peroxidase
71749234	7	31	protein kinase
72390201	7	33	pyridoxal kinase
115503961	5	47	phosphoglycerate kinase
71756217	6	48	S-adenosylhomocysteine hydrolase
72393337	5	30	lysophospholipase, putative
71745150	6	42	actin A
71744488	6	45	eukaryotic initiation factor 4a
15077032	7	35	cytosolic malate dehydrogenase

NCBI #	Peptides	MW	NCBI Definition
33327788	5	37	L-threonine 3-dehydrogenase
70908166	5	561	microtubule-associated protein 2
74025474	6	53	protein kinase
72387712	4	42	hypothetical protein, conserved
71747184	4	46	ATP-dependent DEAD-box RNA helicase
18765891	5	71	RHS1 protein
71747898	4	75	ABC transporter
72390599	6	83	hypothetical protein, conserved
71755637	4	55	aminopeptidase
71755691	5	29	hypothetical protein Tb11.02.4700
74025270	3	35	guanine nucleotide-binding protein subunit beta - like protein
71747944	4	61	2,3-bisphosphoglycerate-independent phosphoglycerate mutase
71747996	2	58	t-complex protein 1 subunit theta
72387668	4	68	V-type ATPase, A subunit, putative
24474936	3	72	trypanothione synthetase
71744572	3	32	proliferative cell nuclear antigen
71749030	4	38	hypothetical protein Tb10.61.3210
71754749	4	35	60S acidic ribosomal subunit protein
71746514	4	49	ATP-dependent DEAD/H RNA helicase
18413539	3	98	H25N7.12
72386991	4	54	ATP-dependent phosphofructokinase
71747120	4	45	aspartate aminotransferase
71746626	3	22	60S ribosomal protein L9
84043716	3	40	branched-chain amino acid aminotransferase
463287	4	70	paraflagellar rod protein
72389050	3	37	hypothetical protein, conserved
5532470	3	68	transmembrane glycoprotein
72386563	2	24	GTP-binding nuclear protein rtb2, putative
72386653	3	69	protein kinase, putative
71744362	2	24	ribosomal protein S7
71754983	2	60	t-complex protein 1 subunit zeta
71747174	5	29	40S ribosomal protein S3a
72390491	4	44	mitogen-activated protein kinase 5, putative
71744950	1	38	hypothetical protein Tb09.160.5530
71747540	2	48	hypothetical protein Tb10.70.1130
18700061	5	95	RHS6a
72390920	2	42	protein disulfide isomerase, putative
72388452	2	45	calreticulin, putative
74025332	1	56	vacuolar ATP synthase subunit B
72387193	4	69	73 kDa paraflagellar rod protein
2645495	2	36	inosine-adenosine-guanosine-nucleoside hydrolase; IAG-nucleoside hydrolase
71746208	4	35	60S ribosomal protein L5
72389969	2	32	2-hydroxy-3-oxopropionate reductase, putative
74026380	3	110	hypothetical protein Tb11.01.8770
71746618	5	58	t-complex protein 1 subunit delta
71745040	2	25	60S ribosomal protein L10
72391572	2	39	nucleoside hydrolase, putative
71755051	4	31	40s ribosomal protein S4
72392463	2	26	tryparedoxin peroxidase
14029752	1	22	iron-containing superoxide dismutase
71745456	2	62	poly(A)-binding protein 1
84043798	1	45	chaperone protein
71745578	1	26	Gim5A protein
71755409	2	62	malic enzyme
71409156	1	19	RNA-binding protein
71754537	1	47	isocitrate dehydrogenase
72390113	2	38	adenosine kinase, putative
72392307	1	36	hypothetical protein, conserved
71746704	2	87	methionyl-tRNA synthetase
71747584	3	191	clathrin heavy chain
72387942	2	44	thioredoxin, putative
18700053	1	94	RHS2a
115504593	1	63	alanine aminotransferase, putative
71747778	1	38	alternative oxidase
72392691	2	61	t-complex protein 1 gamma subunit, putative

NCBI #	Peptides	MW	NCBI Definition
74025154	1	28	40S ribosomal protein SA
71744526	1	43	protein kinase
71756205	2	30	hypothetical protein Tb11.01.1290
72391516	1	34	hypothetical protein, conserved
72387858	1	23	hypothetical protein, conserved
71747040	2	42	protein kinase
71749044	1	40	protein kinase
72386759	1	34	hypothetical protein, conserved
74025798	1	52	hypothetical protein Tb11.01.5680
71749046	2	75	hypothetical protein Tb10.61.3130
71744988	2	33	spermidine synthase
71756227	1	70	glycyl-tRNA synthetase
71409597	1	69	nucleolar RNA helicase II
72388176	2	34	serine/threonine-protein phosphatase PP1, putative
74025330	2	41	2-oxoglutarate dehydrogenase E2 component dihydrolipoamide succinyltransferase
72387047	1	69	hypothetical protein, conserved
71755597	2	33	hypothetical protein Tb11.02.4250
72393281	1	55	hypothetical protein, conserved
72389000	1	34	short-chain dehydrogenase, putative
18033030	2	53	chaperonin subunit alpha
71755753	1	69	hypothetical protein Tb11.02.5040
71659517	1	41	developmentally regulated GTP-binding protein
71744718	1	30	40S ribosomal protein S3
71748584	3	21	60S ribosomal protein L6
71749200	1	29	40S ribosomal protein S2
72391006	1	28	60S ribosomal protein L7
72388162	1	26	60S ribosomal protein L13a, putative
71747966	1	76	adaptin complex 1 subunit
72393711	1	98	calpain, putative
72392060	1	33	methylthioadenosine phosphorylase, putative
71755047	1	98	aminopeptidase

Supplementary Table 3: iTRAQ analysis of compound **47** binding proteins in *T brucei* lysates after pretreatment with compound **1.** Gray indicates a protein kinase.

				% DMSO control						
NCBI #	Peptides	MW	# MS/MS	0.5 uM	Std Dev	2.0 uM	Std Dev	8.0 uM	Std Dev	NCBI Definition
115504281	30	50	38	95	38	112	30	106	31	beta tubulin
71747410	25	50	30	65	20	16	13	5	10	protein kinase
115504283	20	50	17	102	33	100	18	117	22	alpha tubulin
72393479	20	243	10	115	44	124	31	145	41	acetyl-CoA carboxylase, putative
74025264	18	75	17	97	22	90	23	83	20	heat shock protein 70
71746820	16	49	26	113	32	103	23	118	34	elongation factor 1-alpha
13926114	16	50	18	80	25	56	19	36	18	serine/threonine protein kinase TbPK6
139611	14	51	22	97	44	95	23	87	27	Variant surface glycoprotein MITAT 1.2 precursor (VSG 221)
15485709	14	56	14	116	27	104	19	110	21	glycerolkinase
72392815	13	47	17	88	23	59	17	35	13	mitogen-activated protein kinase, putative
72388918	12	38	13	83	25	67	11	41	12	casein kinase, putative
11245496	12	91	5	103	11	107	23	122	32	C-terminal kinesin KIFC1
72388510	11	50	7	87	29	70	15	50	21	serine/threonine-protein kinase A, putative
71747654	11	60	4	98	14	91	18	105	44	chaperonin Hsp60
266429	11	54	7	103	36	90	20	87	41	Pyruvate kinase 2 (PK 2)
71755837	10	71	5	88	14	90	38	107	15	glucose-regulated protein 78
24474936	10	72	8	113	39	140	29	143	38	trypanothione synthetase
71746974	9	47	10	92	30	79	14	87	28	enolase
74025600	9	46	10	95	37	121	32	128	31	elongation factor 1 gamma
71747296	9	94	10	99	18	103	28	94	17	elongation factor 2
15626363	8	51	7	96	17	78	15	86	14	hexokinase
71747502	8	41	10	89	22	89	25	87	18	fructose-bisphosphate aldolase glycosomal
115141	8	56	4	142	42	144	22	157	42	protein disulfide isomerase 71748004
5726483	8	17	5	119	37	115	40	77	15	nucleoside diphosphate kinase (74026208)
72391812	7	53	9	94	36	38	14	12	8	protein kinase, putative
71744742	7	40	7	88	25	82	25	76	20	arginine kinase
72391612	7	19	6	104	30	98	17	82	19	cyclophilin-type peptidyl-prolyl cis-trans isomerase, putative
72387269	7	48	5	104	16	64	10	42	6	cell division control protein, putative
72387810	7	54	3	84	39	91	13	81	28	ribosomal protein L3, putative
71744678	6	22	5	83	29	76	36	80	22	tryparedoxin peroxidase
71748504	6	81	4	92	21	84	13	98	16	heat shock protein 83
72390201	6	33	7	103	25	108	18	122	36	pyridoxal kinase
71748842	6	91	4	90	18	93	23	77	6	heat shock protein
71756217	6	48	5	98	12	113	6	84	7	S-adenosylhomocysteine hydrolase
71747584	6	191	3	98	15	90	12	124	25	clathrin heavy chain
71754589	6	37	2	106	28	77	38	96	6	cell-division control protein 2 homolog 6
53849855	5	35	3	102	15	70	25	102	3	glycosomal glyceraldehyde phosphate dehydrogenase
72390009	5	46	4	71	21	45	6	23	15	mitogen-activated protein kinase, putative
71747946	5	41	4	107	12	111	14	103	17	methyltransferase
71749234	5	31	3	131	18	144	17	139	41	protein kinase
84043716	5	40	2	116	17	137	54	99	8	branched-chain amino acid aminotransferase
71745150	4	42	5	94	36	101	21	101	23	actin A
71755051	4	31							0	40s ribosomal protein S4
72391902	4	61	3	80	25	54	8	31	6	protein kinase, putative
72393421	4	47	4	58	18	58	19	40	29	protein kinase, putative
71747566	4	55	4	85	4	58	20	52	6	protein kinase
72386563	4	24	3	123	32	103	29	117	23	GTP-binding nuclear protein rtb2, putative
72387942	4	44	2	92	5	87	14	86	25	thioredoxin, putative
71748388	4	255	2	117	8	81	7	136	5	microtubule-associated protein
72388980	4	28	3	117	33	94	11	101	30	60S ribosomal protein L2, putative
74025474	4	53	2	87	12	61	8	55	13	protein kinase
72386827	3	46	3	35	12	16	8	11	9	serine/threonine-protein kinase, putative
71755425	3	27	2	120	34	111	37	94	8	triosephosphate isomerase
72392767	3	38	1	80		83		86	0	glycerol-3-phosphate dehydrogenase [NAD+], glycosomal
71755691	3	29							0	hypothetical protein Tb11.02.4700
71744488	3	45	3	82	10	84	5	74	18	eukaryotic initiation factor 4a
71755063	3	67	2	112	9	102	40	130	48	adenylosuccinate synthetase
115503961	3	47	1	93		111		88	0	phosphoglycerate kinase
72387712	3	42	2	132	46	109		141	28	hypothetical protein, conserved
74026380	3	110	4						0	hypothetical protein Tb11.01.8770
71755047	3	98	3	80	43	94	4	96	35	aminopeptidase
72387087	3	16							0	tryparedoxin
71754533	3	19	2	77	5	101	25	78	25	cyclophilin A
71748126	3	45							0	hypothetical protein Tb10.6k15.1500
115504447	3	20							0	40S ribosomal protein S11, putative
71745744	3	135							0	ubiquitin-activating enzyme E1
71754491	3	18	5						0	ribosomal protein L21E (60S), putative
71746622	3	16							0	40S ribosomal protein S23, putative
71747500	3	22							0	40S ribosomal protein S9
84043378	3	98							0	RHS4a
71749196	3	19	3						0	60S ribosomal protein L17, putative
72392947	2	37	1						0	nucleoside phosphorylase, putative
71747174	2	29	1	85		81		113	0	40S ribosomal protein S3a
71748584	2	21	2						0	60S ribosomal protein L6
71749030	2	38							0	hypothetical protein Tb10.61.3210
71745040	2	25							0	60S ribosomal protein L10

				% DMSO control						
NCBI #	Peptides	MW	# MS/MS	0.5 uM	Std Dev	2.0 uM	Std Dev	8.0 uM	Std Dev	NCBI Definition
71747944	2	61							0	2,3-bisphosphoglycerate-independent phosphoglycerate mutase
72390113	2	38	1	105		131		75	0	adenosine kinase, putative
71747540	2	48							0	hypothetical protein Tb10.70.1130
71745456	2	62							0	poly(A)-binding protein 1
72391516	2	34							0	hypothetical protein, conserved
71754537	2	47	1						0	isocitrate dehydrogenase
72390706	2	11							0	40S ribosomal protein S33, putative
71749292	2	17	2	102	33	82	10	68	32	40S ribosomal protein S13
71748046	2	56							0	hexose transporter
72386999	2	27	1						0	60S ribosomal protein L13, putative
18378042	2	67	3	113	25	86	21	112	15	retrotransposon hot spot protein, RHS3
71745370	2	19	1	92		101		97	0	hypothetical protein Tb09.211.1690
71746212	2	26	2						0	ribosomal protein L15, putative
71747708	2	14							0	dynein light chain, putative
72387824	2	19	2	94	13	99	1	103	23	ribosomal protein S19, putative
72393289	2	25	2							40S ribosomal protein S8, putative
72390884	2	19								glutathione peroxidase-like protein
62176656	2	39								protein kinase, putative
71747826	2	16	1							40S ribosomal protein S24E, putative
72387888	2	17	1							60S ribosomal protein L35A, putative
71748500	1	64	1	80		46		21		protein kinase
72386653	1	69	1	101		53		34		protein kinase, putative
71749480	1	48								peptidylprolyl isomerase-like protein
71746208	1	35								60S ribosomal protein L5
72386975	1	56								serine/threonine-protein kinase, putative
72387952	1	47								protein kinase
74025270	1	35								guanine nucleotide-binding protein subunit beta - like protein
71747748	1	36								glyceraldehyde 3-phosphate dehydrogenase cytosolic
71746514	1	49								ATP-dependent DEAD/H RNA helicase
71744718	1	30								40S ribosomal protein S3
72392691	1	61								t-complex protein 1 gamma subunit, putative
72389000	1	34								short-chain dehydrogenase, putative
72391554	1	16								adenosine 5'-monophosphoramidase, putative
84043668	1	59								glycosomal phosphoenolpyruvate carboxykinase
71744100	1	86								hypothetical protein Tb09.160.1160
71747446	1	18								40S ribosomal protein S18
72390641	1	16								40S ribosomal protein S14
71744668	1	11								60S acidic ribosomalprotein
71755547	1	15								40S ribosomal protein S15a
71411145	1	12								hypothetical protein Tc00.1047053506893.50
71419733	1	61								CDC16, putative
71413905	1	123								hypothetical protein Tc00.1047053510735.90
72392337	1	31								60S ribosomal protein L7a, putative
72386707	1	33								3-oxo-5-alpha-steroid 4-dehydrogenase, putative
71748836	1	19	1							60S ribosomal protein L34, putative

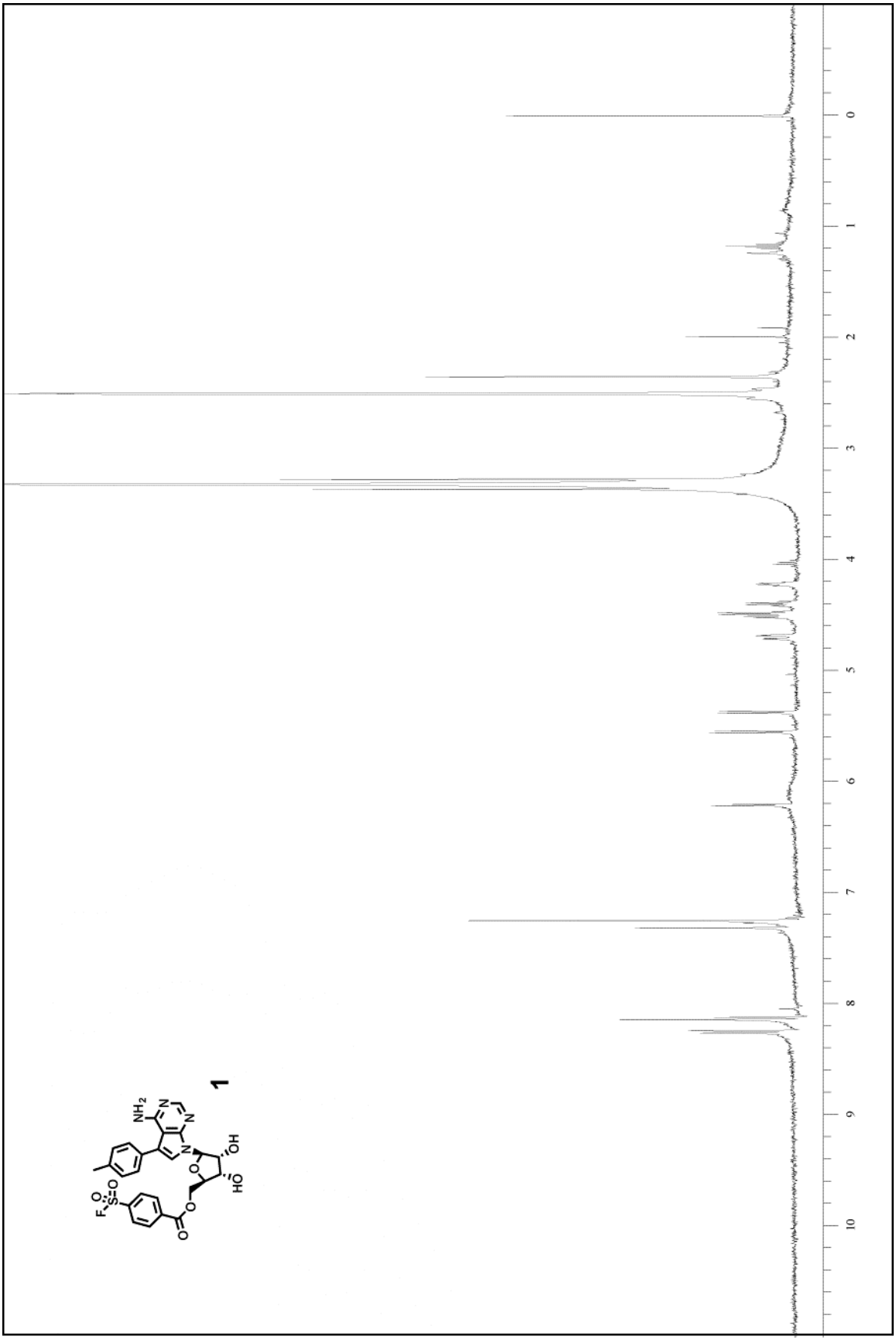
Supplementary Table 4: iTRAQ analysis of compound **47** binding proteins in *T brucei* lysates after pretreatment with PD0166326**.** Gray indicates a protein kinase

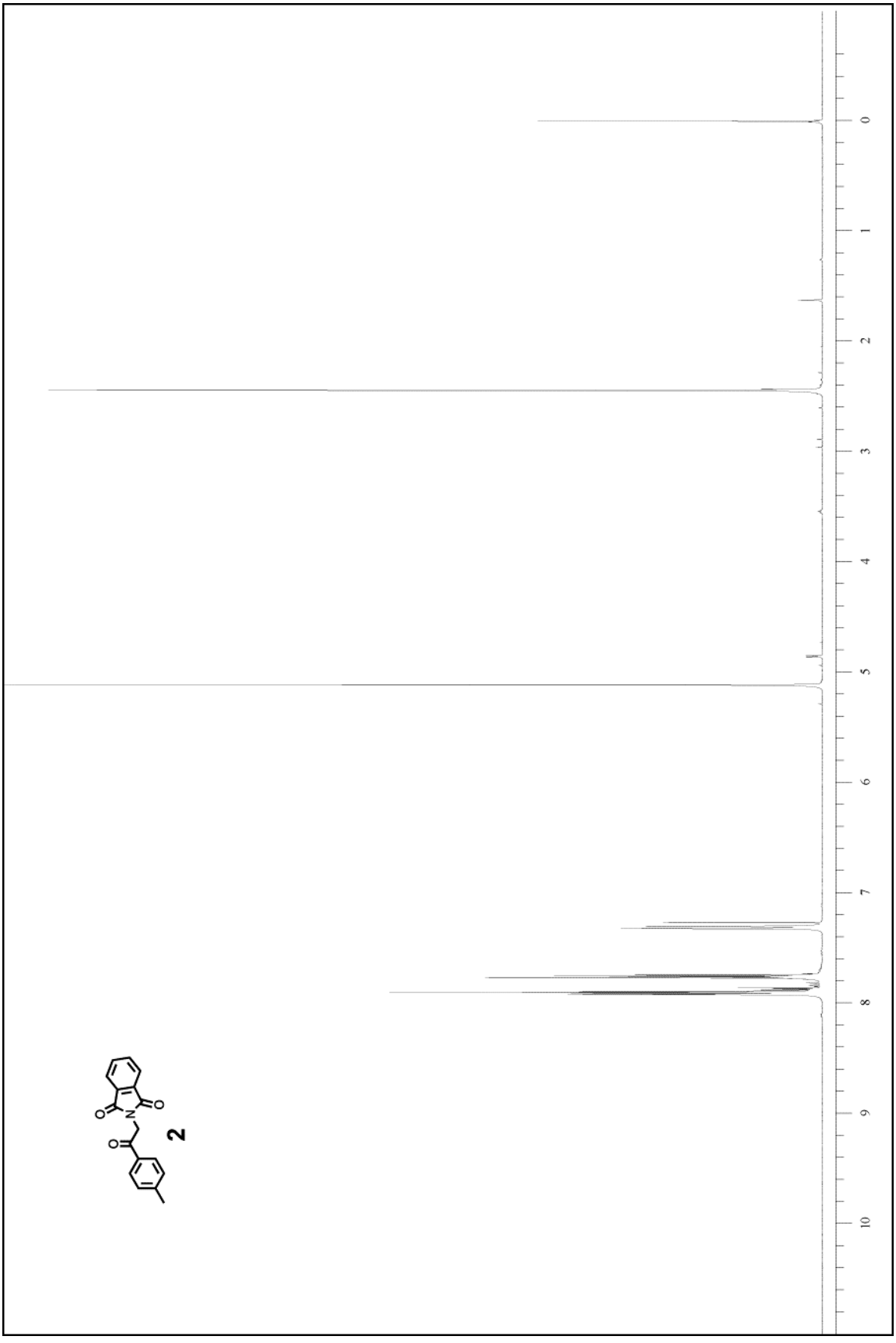
				% DMSO control						
NCBI #	Peptides	MW	# MS/MS	0.5 uM	Std Dev	2.0 uM	Std Dev	8.0 uM	Std Dev	NCBI Definition
72393479	30	243	46	101	24	91	28	107	22	acetyl-CoA carboxylase, putative
15485709	28	56	56	89	25	99	32	111	28	glycerolkinase
71747410	27	50	74	100	26	104	35	112	46	protein kinase
13926114	24	50	66	97	49	65	28	42	16	serine/threonine protein kinase TbPK6
115504281	23	50	50	97	29	101	39	121	29	beta tubulin
115504283	23	50	46	106	40	82	27	93	36	alpha tubulin
72391812	20	53	46	100	20	97	42	111	51	protein kinase, putative
24474936	19	72	25	114	44	121	58	115	51	trypanothione synthetase
72388918	17	38	27	99	41	96	45	100	28	casein kinase, putative
72386827	16	46	25	104	25	77	18	58	12	serine/threonine-protein kinase, putative
139611	15	51	18	92	35	106	22	84	30	Variant surface glycoprotein MITAT 1.2 precursor (VSG 221)
72392815	14	47	22	29	18	31	24	20	13	mitogen-activated protein kinase, putative
71746820	13	49	27	83	22	81	25	92	20	elongation factor 1-alpha
72386991	13	54	18	105	33	99	22	93	20	ATP-dependent phosphofructokinase
71747946	10	41	16	105	15	100	25	105	12	methyltransferase
71747502	10	41	15	93	25	89	21	94	28	fructose-bisphosphate aldolase glycosomal
15626363	9	51	16	92	24	80	17	87	15	hexokinase
72387269	8	48	17	58	24	43	25	27	17	cell division control protein, putative
5726483	8	17	15	93	25	127	29	109	20	nucleoside diphosphate kinase
74025264	8	75	13	100	36	99	32	93	22	heat shock protein 70
71749234	8	31	9	85	40	55	23	30	11	protein kinase
71748500	7	64	14	103	31	91	22	75	26	protein kinase
71748504	7	81	9	79	35	89	23	91	18	heat shock protein 83
71754589	6	37	12	98	15	87	25	70	15	cell-division control protein 2 homolog 6
53849855	6	35	10	114	34	83	22	100	23	glycosomal glyceraldehyde phosphate dehydrogenase
71746974	6	47	9	80	18	84	16	81	17	enolase
74025600	6	46	8	103	15	102	24	115	28	elongation factor 1 gamma
72392767	5	38	10	110	50	109	28	96	29	glycerol-3-phosphate dehydrogenase [NAD+], glycosomal
71747296	5	94	8	110	39	99	30	101	10	elongation factor 2
72390009	5	46	7	66	40	60	34	47	23	mitogen-activated protein kinase, putative
71745150	5	42	7	90	20	94	19	94	23	actin A
266429	5	54	6	92	34	109	31	104	22	Pyruvate kinase 2 (PK 2)
71755047	5	98	6	91	20	98	32	119	34	aminopeptidase
71747566	5	55	5	101	19	95	20	110	35	protein kinase
71744742	5	40	5	78	25	89	8	85	13	arginine kinase
71747584	5	191	5	96	9	90	14	99	10	clathrin heavy chain
72391902	4	61	8	95	10	89	22	68	20	protein kinase, putative
72388510	4	50	6	49	9	34	18	19	11	serine/threonine-protein kinase A, putative
72390201	4	33	5	127	35	135	20	93	23	pyridoxal kinase
71747654	4	60	4	114	35	113	39	91	19	chaperonin Hsp60
71748842	4	91	3	131	51	126	19	111	16	heat shock protein
115504663	4	184	3	109	26	86	46	89	40	hypothetical protein, conserved
72391554	3	16	5	87	12	92	13	83	16	adenosine 5'-monophosphoramidase, putative
72391612	3	19	5	83	5	82	14	96	18	cyclophilin-type peptidyl-prolyl cis-trans isomerase, putative
74025474	3	53	4	91	18	104	21	95	31	protein kinase
71756217	3	48	4	94	19	96	14	95	13	S-adenosylhomocysteine hydrolase
71755425	3	27	3	105	10	102	20	112	23	triosephosphate isomerase
72393453	3	74	3	101	34	89	10	90	9	3-methylcrotonyl-CoA carboxylase, putative
115503961	3	47	3	84	14	104	41	77	19	phosphoglycerate kinase
72387087	3	16	3	102	14	82	13	90	17	tryparedoxin
71754533	3	19	3	127	80	124	46	121	44	cyclophilin A
71747806	3	285	3	100	3	102	5	122	28	hypothetical protein Tb10.6k15.3460
72391324	3	503	3	116	32	138	81	100	35	hypothetical protein, conserved
463287	3	70	1	95		140		99		paraflagellar rod protein
62176656	2	39	3	38	19	33	11	20	7	protein kinase, putative
72393421	2	47	3	48	32	36	32	28	28	protein kinase, putative
72387712	2	42	3	87	13	87	11	106	14	hypothetical protein, conserved
71748388	2	255	3	90	34	99	42	121	7	microtubule-associated protein
71748568	2		3	86	49	124	41	114	8	
71755837	2	71	2	121	61	153	23	105	28	glucose-regulated protein 78
11245496	2	91	2	89	17	96	15	103	18	C-terminal kinesin KIFC1
2645495	2	36	2	121	41	117	9	117	47	inosine-adenosine-guanosine-nucleoside hydrolase
115504569	2		2	123	30	115	26	116	30	
72391254	2		1	79		99		101		
71747944	2	61								2,3-bisphosphoglycerate-independent phosphoglycerate mutase
71754537	2	47								isocitrate dehydrogenase
71744678	1	22	11	86	18	99	18	90	16	tryparedoxin peroxidase
71748046	1		3	62	7	67	2	70	19	
72387193	1	69	2	93	6	71	9	92	12	73 kDa paraflagellar rod protein
71749238	1	43	2	105	31	79	21	73	1	protein kinase
71406041	1		2	89	3	123	9	83	1	
72386653	1	69	1	115		82		55		protein kinase, putative
71746808	1	94	1	53		87		46		protein kinase
72392947	1	37								nucleoside phosphorylase, putative
72393337	1	30								lysophospholipase, putative

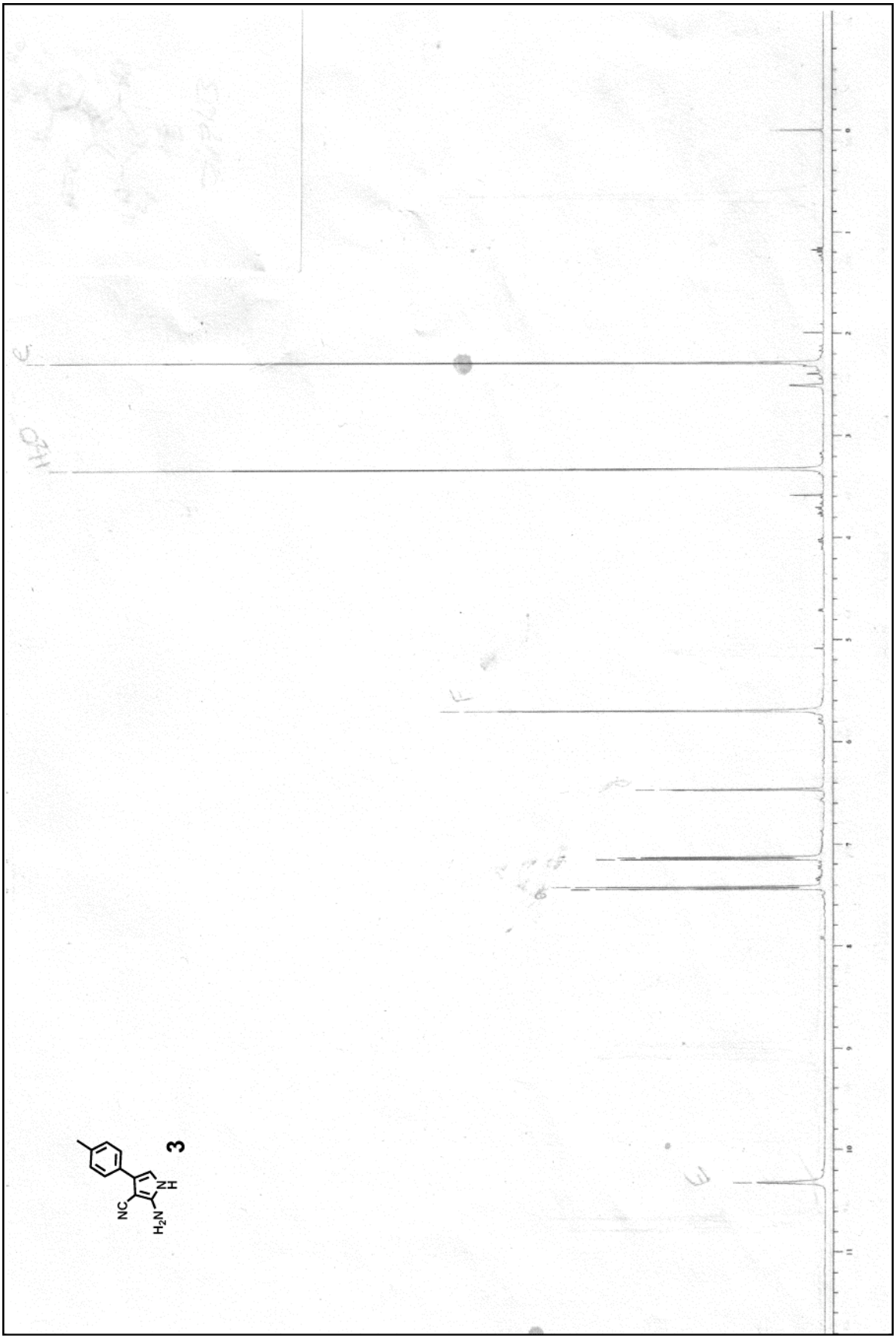
				% DMSO control						
NCBI #	Peptides	MW	# MS/MS	0.5 uM	Std Dev	2.0 uM	Std Dev	8.0 uM	Std Dev	NCBI Definition
71755063	1	67								adenylosuccinate synthetase
71744362	1	24								ribosomal protein S7
72386563	1	24								GTP-binding nuclear protein rtb2, putative
5532470	1	68								transmembrane glycoprotein
71747540	1	48								hypothetical protein Tb10.70.1130
72393281	1	55								hypothetical protein, conserved
72393711	1	98								calpain, putative
71749414	1	36								hypothetical protein Tb10.61.0540
71748126	1	45								hypothetical protein Tb10.6k15.1500
72387025	1	20								ADP-ribosylation factor-like protein 3A, putative
72390696	1	50								hypothetical protein Tb927.7.180
72390013	1									
70908226	1									
72392267	1									
74025638	1									
71749216	1									
72388804	1									
71755061	1									
72392251	1									
71748136	1									
115503949	1									
71744654	1									
72388894	1									
71749396	1									
71744404	1									
15140404	1									
155199173	1									
71746808	1									
72391124	1									
71754499	1									
72387626	1									
72393573	1									
72393015	1									
72393411	1									
74026152	1									
71744334	1									
71747592	1									
71745620	1									
74026374	1									
71748312	1									
72389412	1									
72390343	1									
72390529	1									
72388028	1									

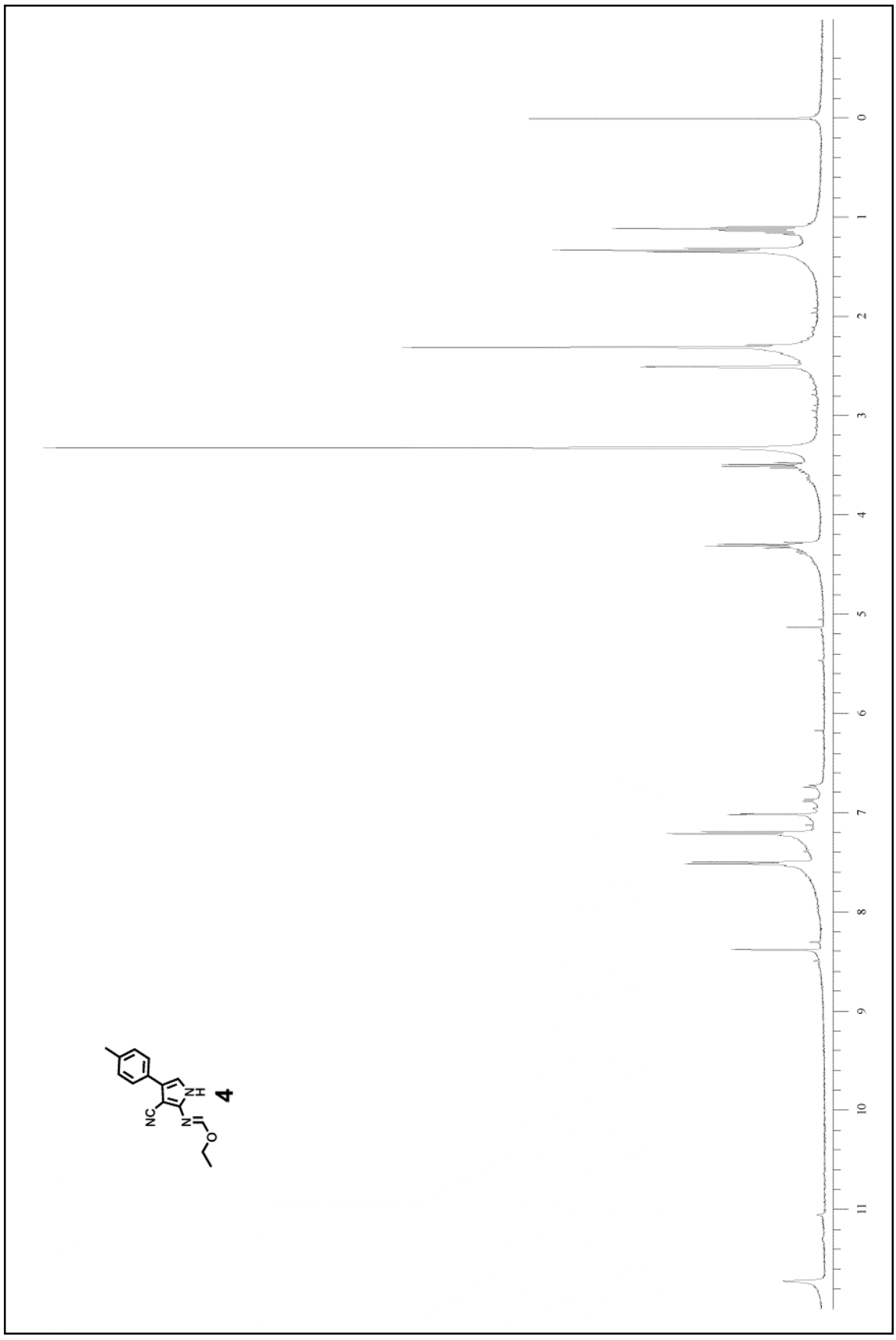
Supplementary Table 5: iTRAQ analysis of compound **55** binding proteins in *T brucei* lysates after pretreatment with compound **19.** Gray indicates a protein kinase

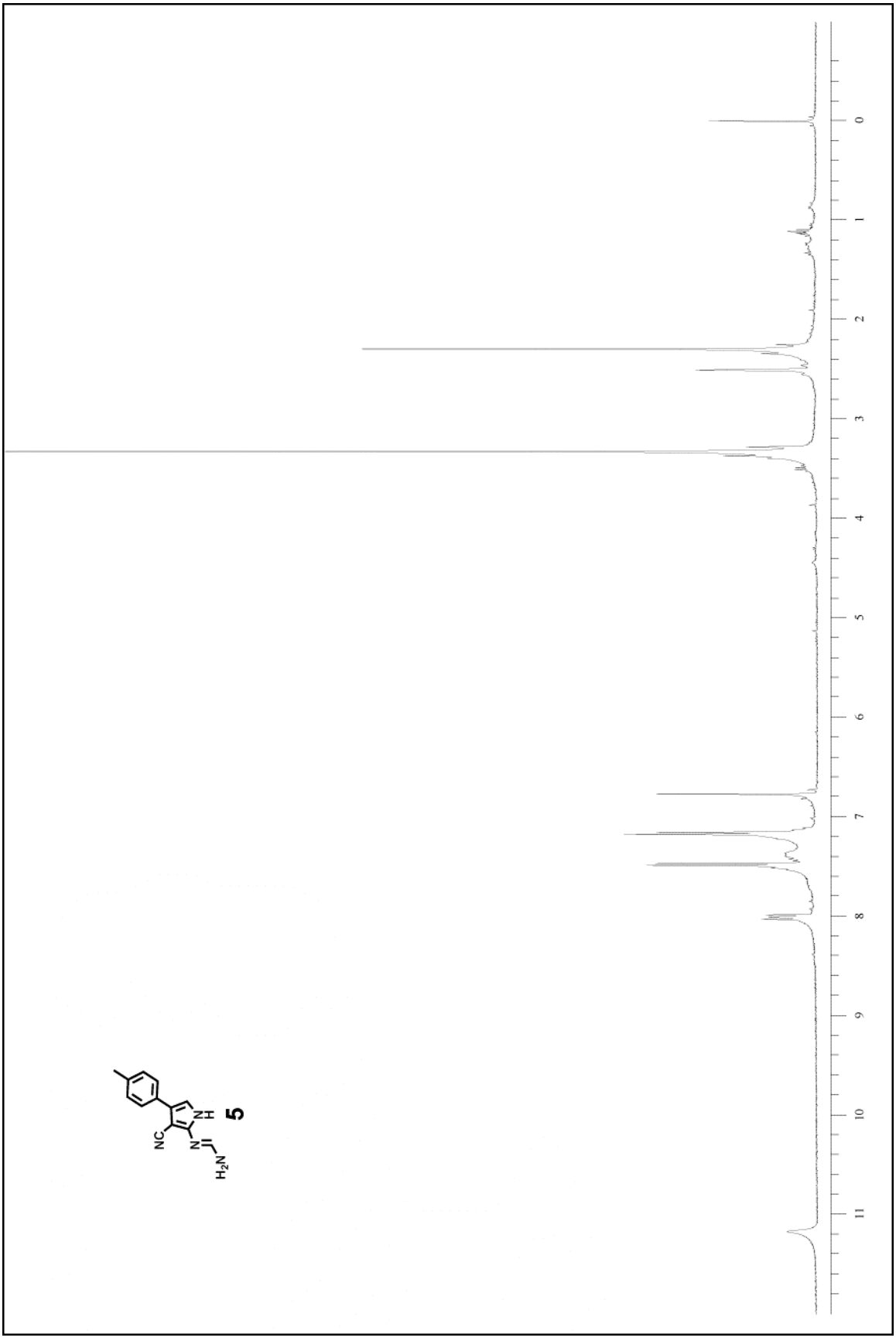
				% DMSO control						
NCBI #	Peptides	MW	# MS/MS	0.5 uM	Std Dev	2.0 uM	Std Dev	8.0 uM	Std Dev	NCBI Definition
115504281	27	50	28	111	34	105	29	100	37	beta tubulin
115504283	25	50	24	109	30	139	49	110	48	alpha tubulin
15485709	21	56	10	106	25	102	18	83	27	glycerolkinase
71746820	20	49	21	103	23	131	32	116	32	elongation factor 1-alpha
72392815	20	47	16	47	8	31	12	23	16	mitogen-activated protein kinase, putative
71747584	19	191	11	122	35	115	29	107	38	clathrin heavy chain
13926114	17	50	24	55	19	26	9	19	11	serine/threonine protein kinase TbPK6
72393479	17	243	11	121	35	110	27	113	38	acetyl-CoA carboxylase, putative
11245496	14	91	10	100	26	141	30	129	33	C-terminal kinesin KIFC1
139611	11	51	7	72	23	107	40	120	38	Variant surface glycoprotein MITAT 1.2 precursor (VSG 221)
71747502	11	41	5	79	19	107	22	93	28	fructose-bisphosphate aldolase glycosomal
115141	11	56	6	120	33	160	70	151	48	Bloodstream-specific protein 2 precursor
71745150	11	42	7	110	26	131	47	135	24	actin A
266429	10	54	4	87	37	107	33	134	34	Pyruvate kinase 2 (PK 2)
1708476	10	56	6	100	20	92	25	75	25	Inosine-5'-monophosphate dehydrogenase
15626363	9	51	8	85	12	89	21	84	24	hexokinase
71754537	9	47	4	105	26	106	50	93	24	isocitrate dehydrogenase
72391812	9	53	3	54	20	57	14	48	5	protein kinase, putative
53849855	8	35	4	71	7	99	31	72	28	glycosomal glyceraldehyde phosphate dehydrogenase
74025264	7	75	2	146	105	117	64	92	21	heat shock protein 70
71744678	7	22	5	118	32	84	26	109	30	tryparedoxin peroxidase
71748504	7	81	1	63		73		63		heat shock protein 83
72392767	7	38	4	99	18	84	21	86	51	glycerol-3-phosphate dehydrogenase [NAD+], glycosomal
71747296	7	94	5	112	51	100	28	103	59	elongation factor 2
72386991	7	54	5	100	29	139	23	112	19	ATP-dependent phosphofructokinase
71744742	6	40	3	86	4	92	18	85	28	arginine kinase
5726483	6	17	2	72	14	61	4	58	7	nucleoside diphosphate kinase
71747654	5	60								chaperonin Hsp60
71746974	4	47	5	78	29	82	15	81	32	enolase
72387269	4	48	3	68	22	39	6	21	3	cell division control protein, putative
71754527	4	27	2	37	9	27	5	15	5	protein kinase
71747540	4	48	2	110	18	129	54	139	32	hypothetical protein Tb10.70.1130
72390009	3	46	2	23	5	30	26	13	15	mitogen-activated protein kinase, putative
72386481	3	64	4	98	25	44	6	38	16	protein kinase, putative
74025600	3	46	2	132	28	100	2	105	3	elongation factor 1 gamma
71755425	3	27								triosephosphate isomerase
72387712	3	42								hypothetical protein, conserved
72386563	3	24	2	77	2	140	43	81	23	GTP-binding nuclear protein rtb2, putative
72387193	3	69								73 kDa paraflagellar rod protein
72391612	3	19	2	125	22	96	31	97	15	cyclophilin-type peptidyl-prolyl cis-trans isomerase, putative
22750416	3	73								dynamin-related protein
71748842	2	91	2	84	3	76	26	50	14	heat shock protein
72388510	2	50	1	64		36		43		serine/threonine-protein kinase A, putative
71756217	2	48	1	123	2	110	31	59	1	S-adenosylhomocysteine hydrolase
18765891	2	71								RHS1 protein
72393711	2	98	2	121	33	116	37	64	8	calpain, putative
72387087	2	16	1	102	0	81	0	84		tryparedoxin
71748126	2	45								hypothetical protein Tb10.6k15.1500
71755547	2	15	2	148	6	78	34	131	73	40S ribosomal protein S15a
2853003	2	86	1	111		93		174		valosin-containing protein homolog
72393421	1	47	1	86		59		75		protein kinase, putative
71747040	1	42	1	88		37		56		protein kinase
71756205	1	30								hypothetical protein Tb11.01.1290
5532470	1	68								transmembrane glycoprotein
72393281	1	55								hypothetical protein, conserved
72391554	1	16								adenosine 5'-monophosphoramidase, putative
72387025	1	20								ADP-ribosylation factor-like protein 3A, putative
71748568	1	14								hypothetical protein Tb10.26.0680
71748046	1	56								hexose transporter
72390377	1	20								ADP-ribosylation factor, putative
72386883	1	187								hypothetical protein, conserved
70908110	1	58								hypothetical protein
74025736	1	16								profilin

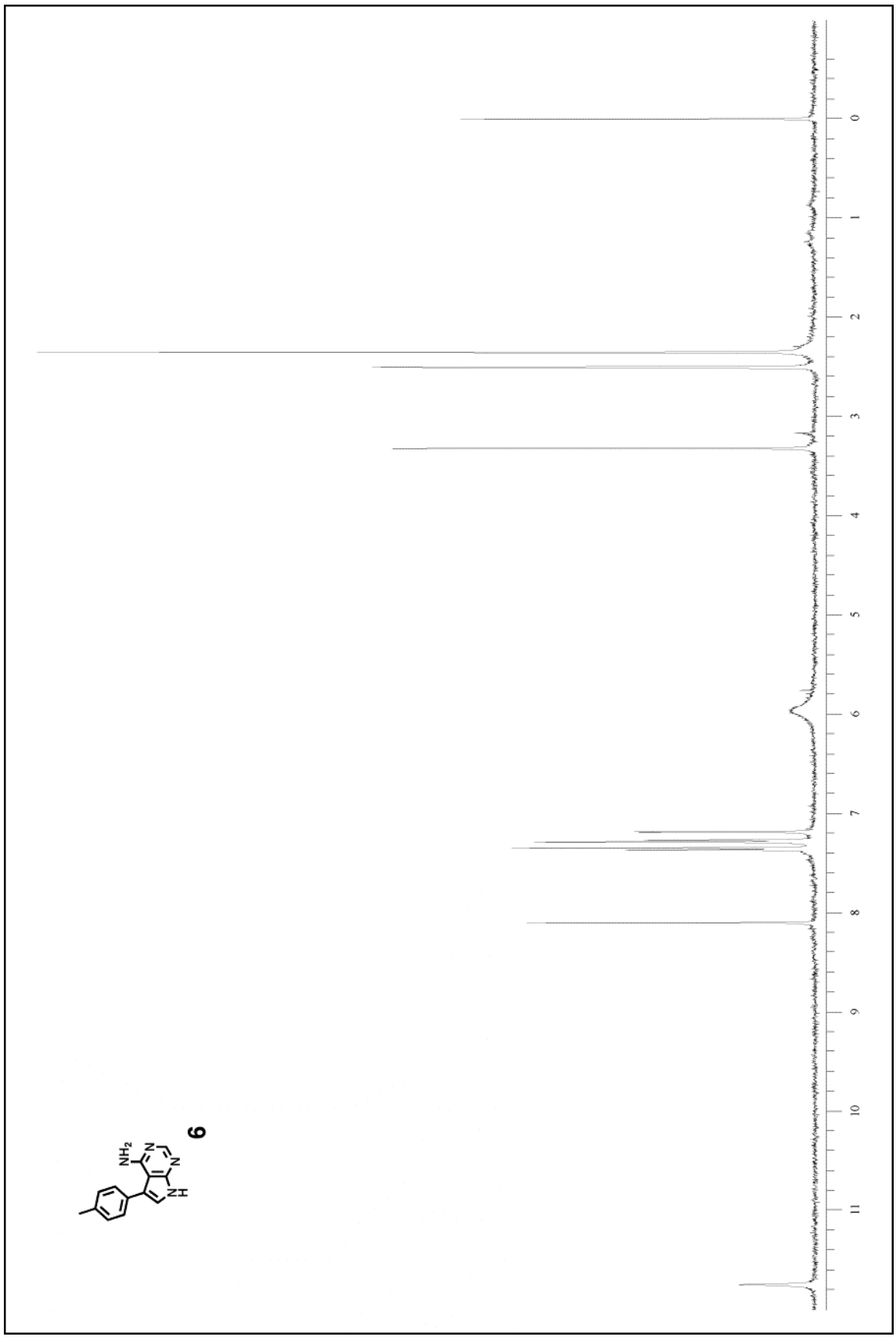

1
NH2
N
N
N
O
OH
HO
F
S
O
O
O
O
0
1
2
3
4
5
6
7
8
9
10

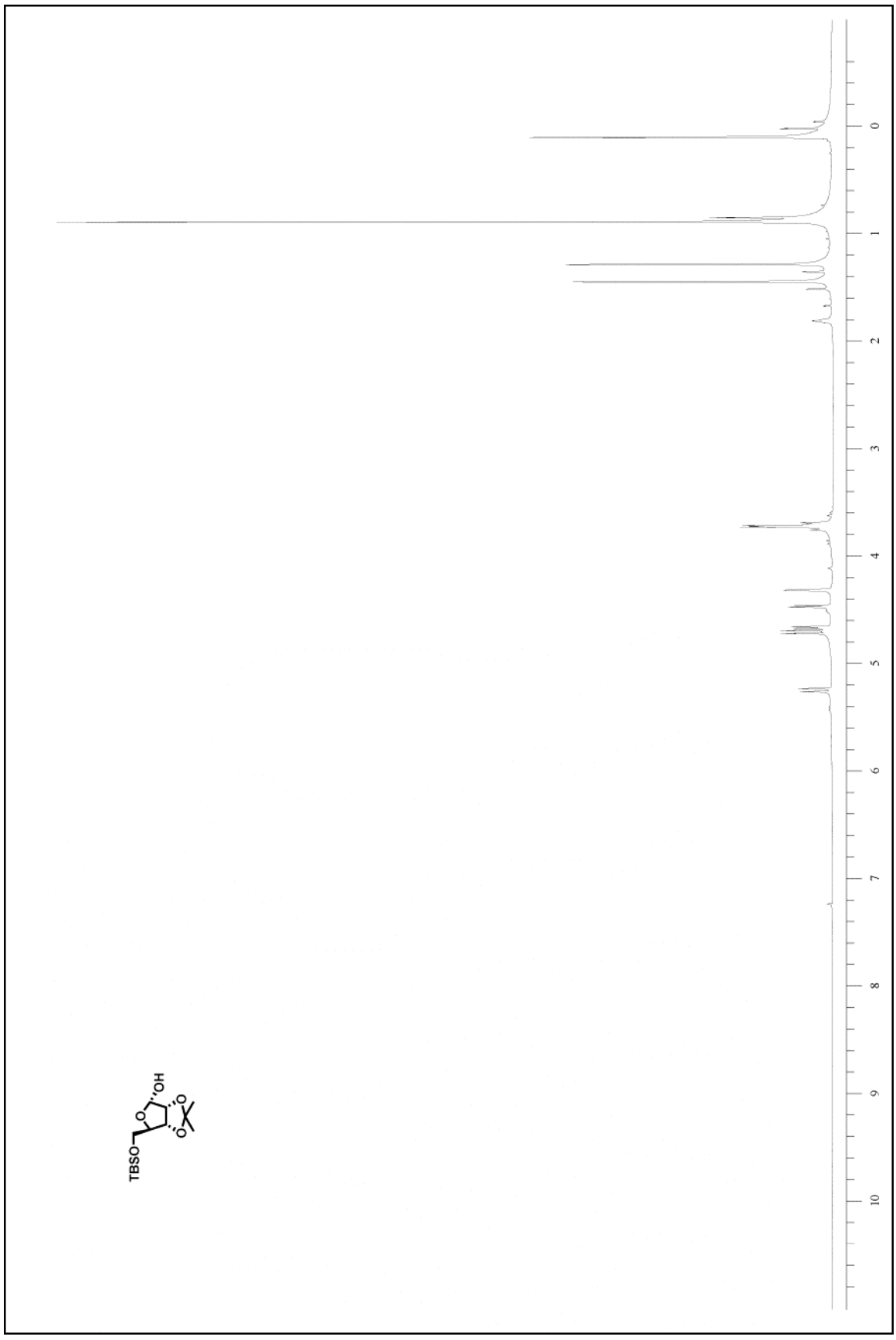

2

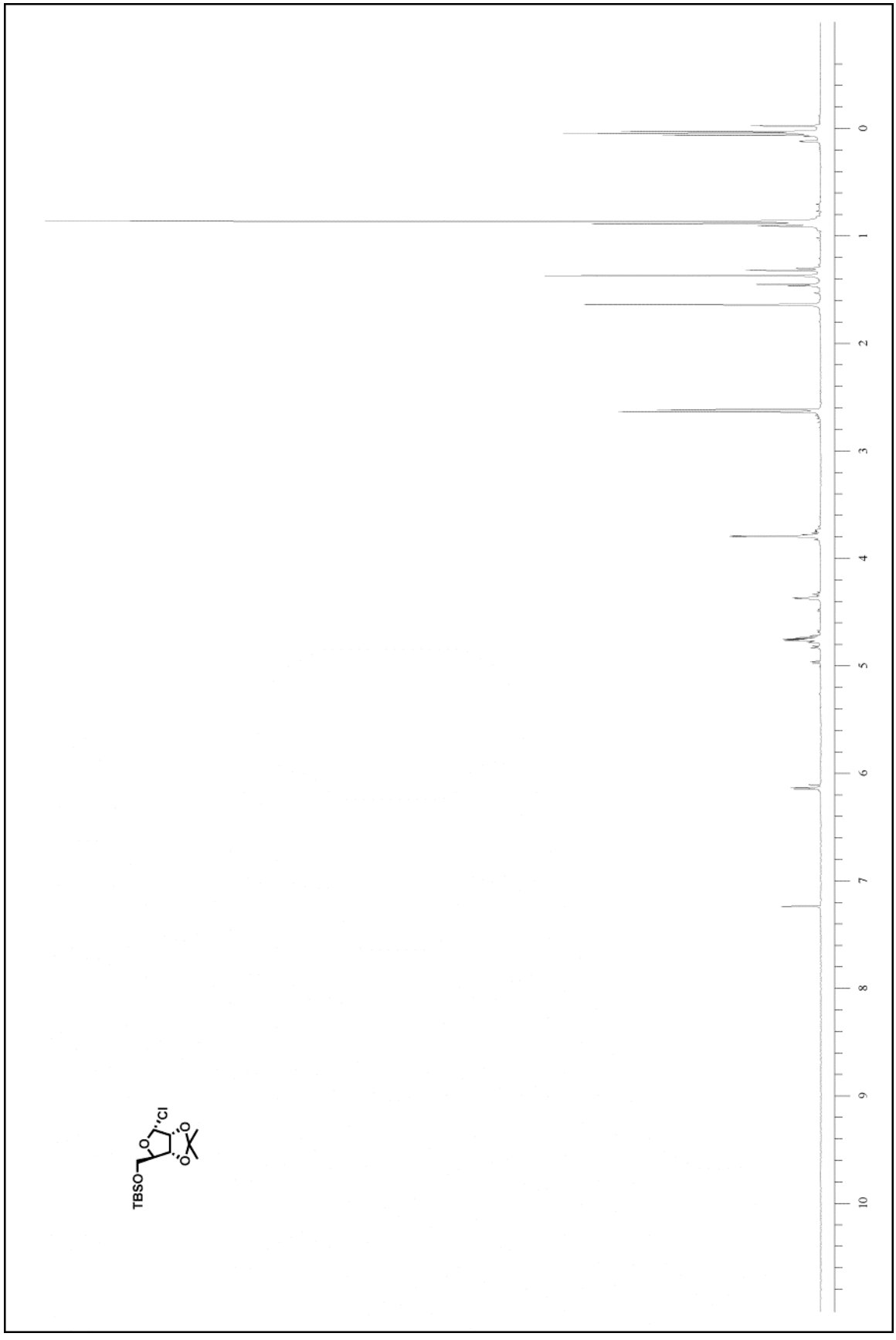

3
NC
H
N
H2N

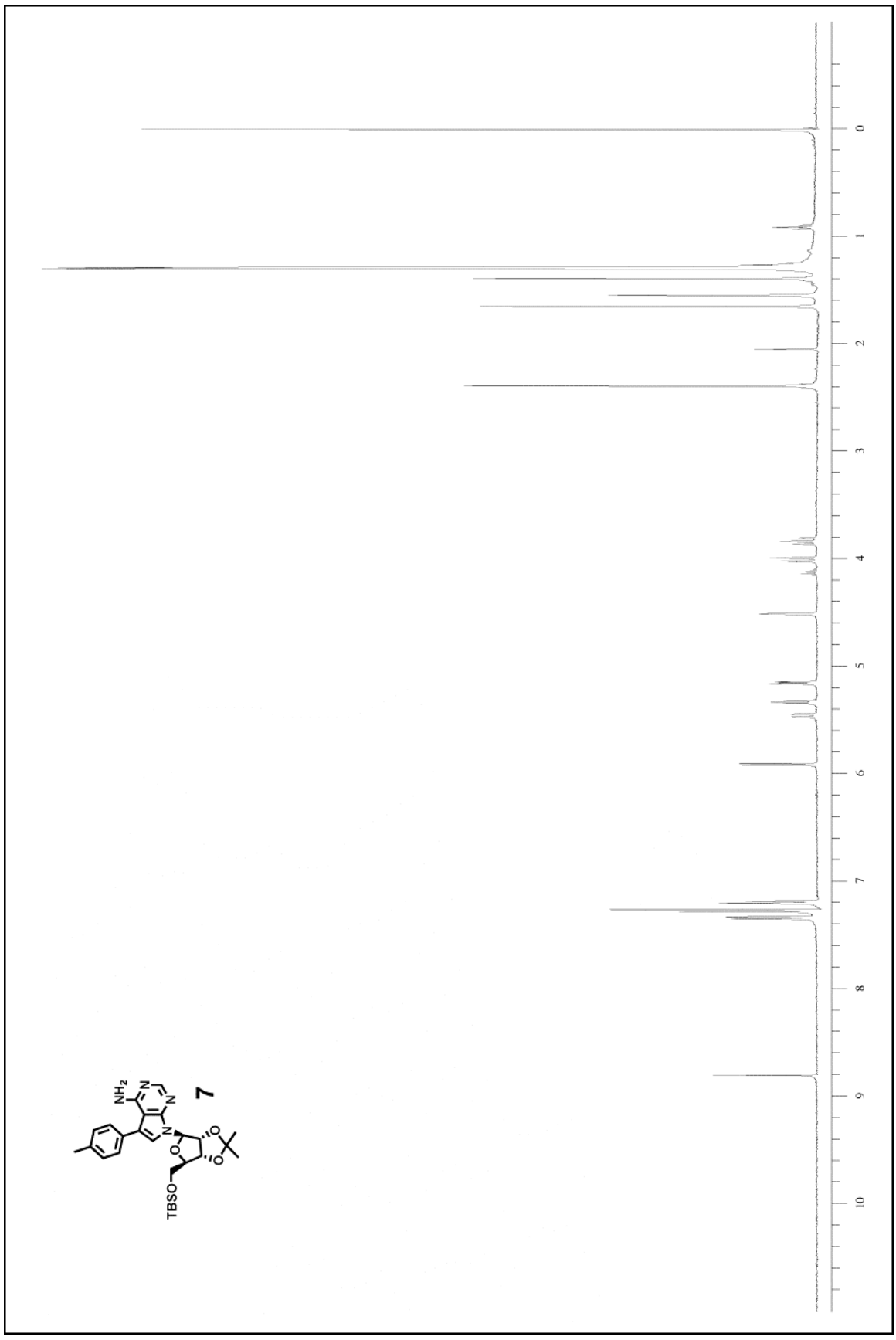

4

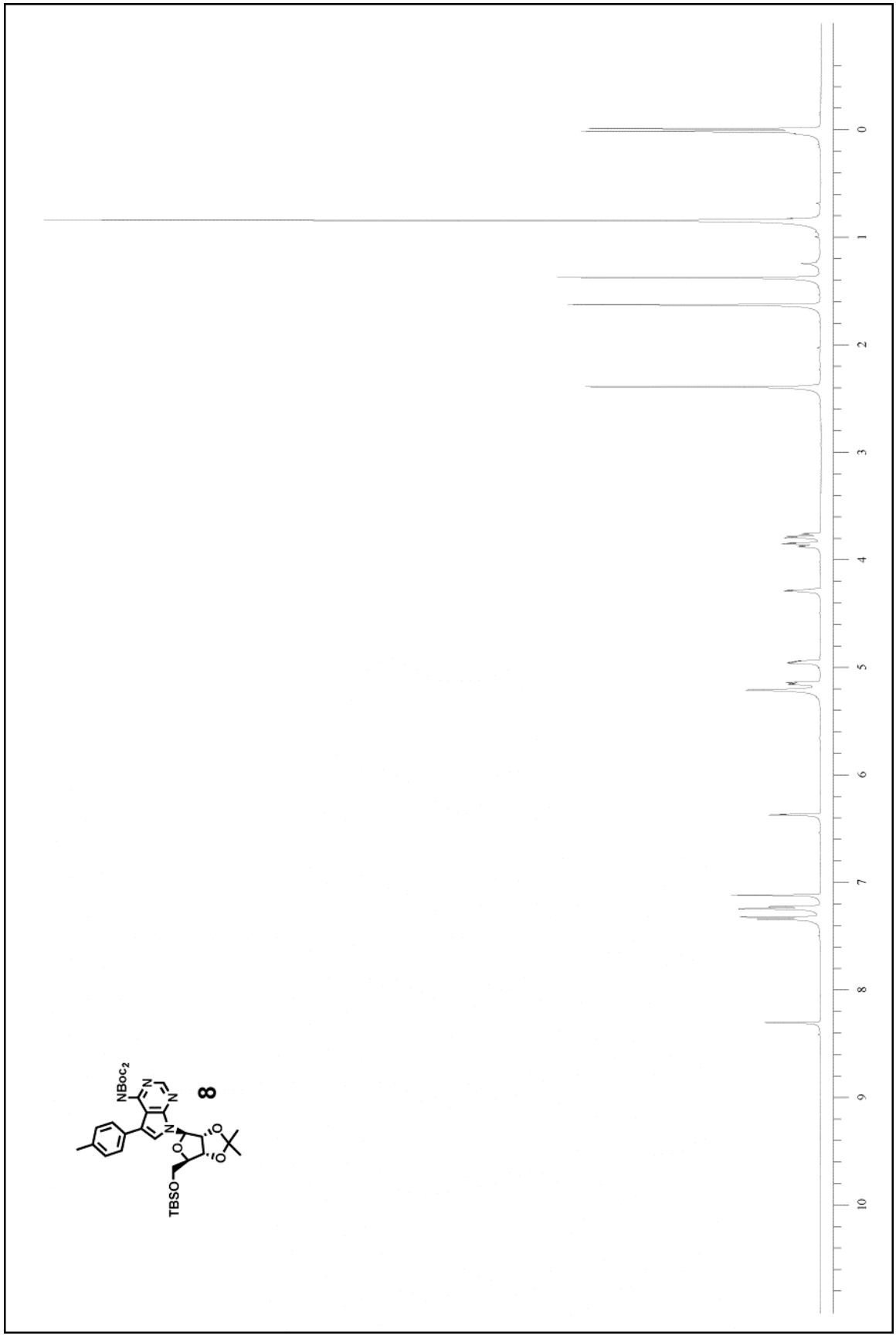

5

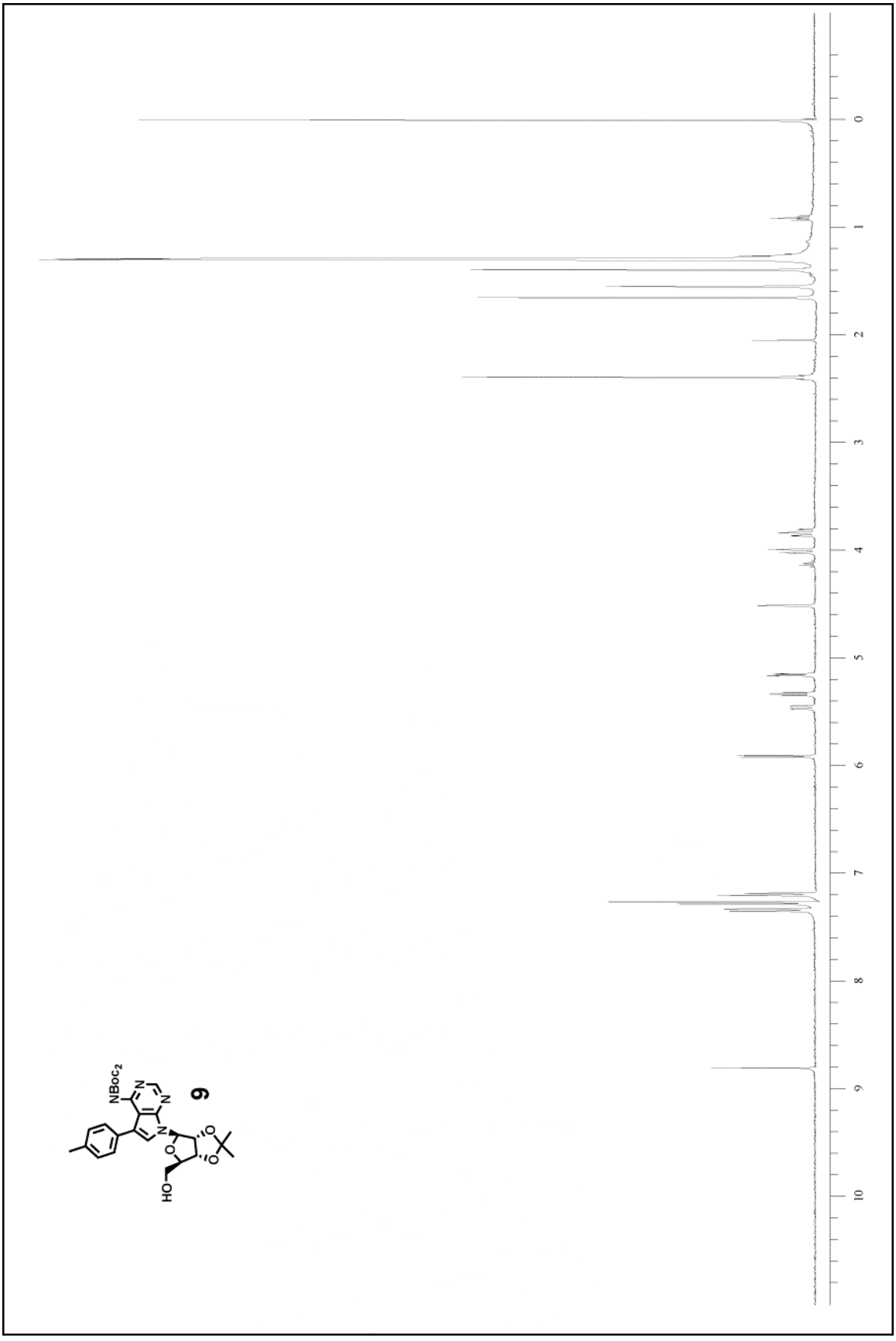

NH2
N
N
NH
6
0
1
2
3
4
5
6
7
8
9
10
11

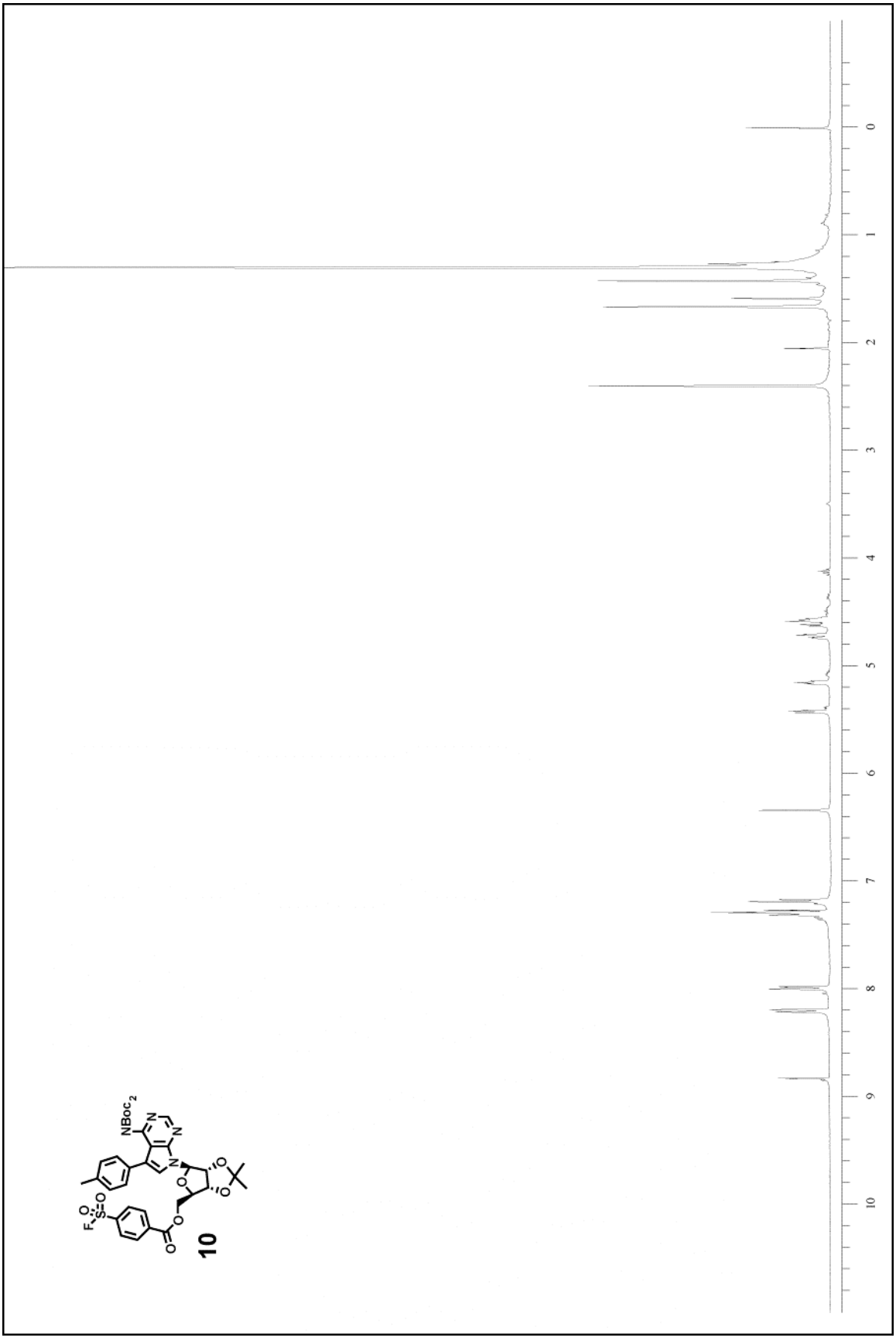

TBSO
OH
O
O
O
0
1
2
3
4
5
6
7
8
9
10

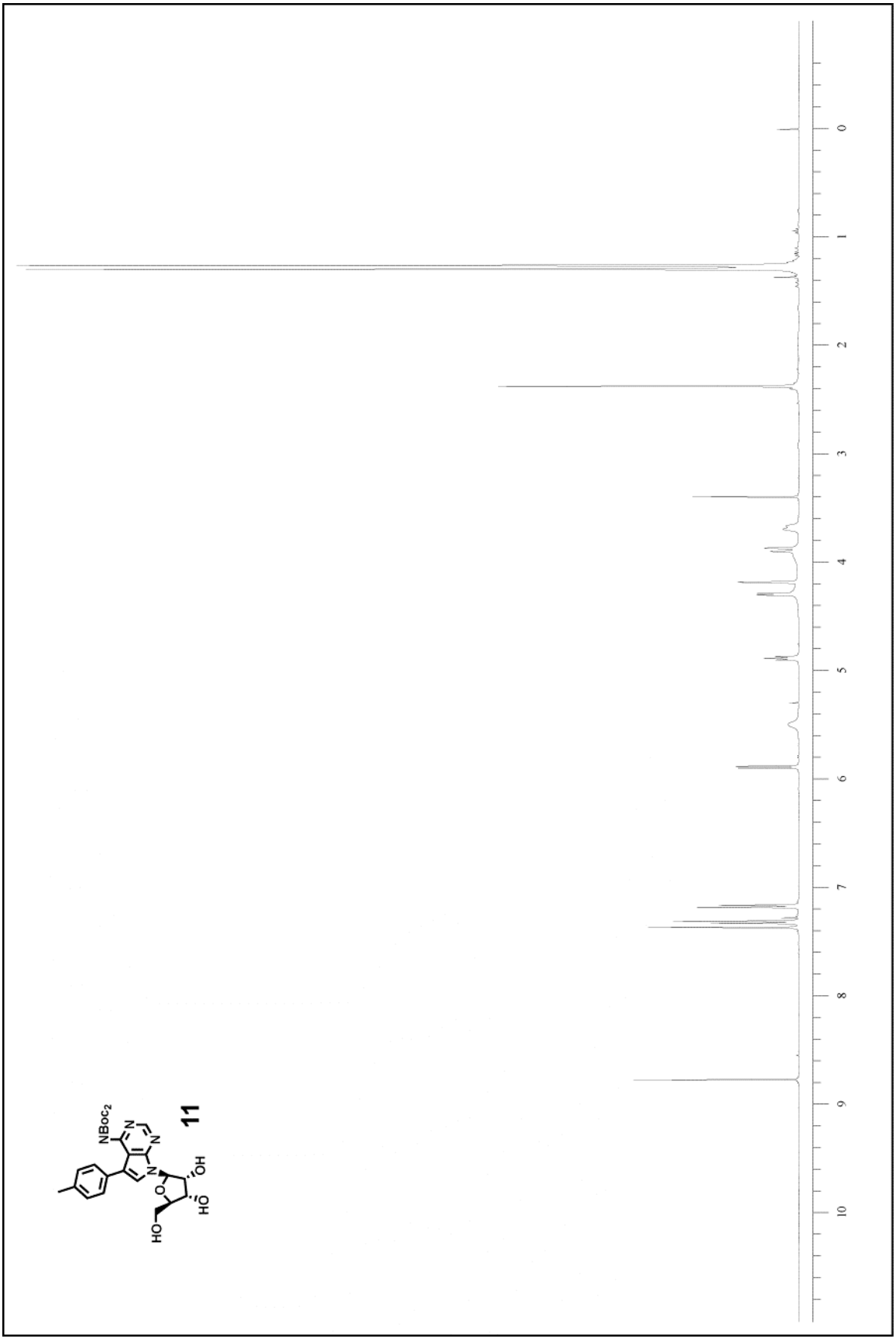

TBSO
Cl
O
O
O
O
0
1
2
3
4
5
6
7
8
9
10

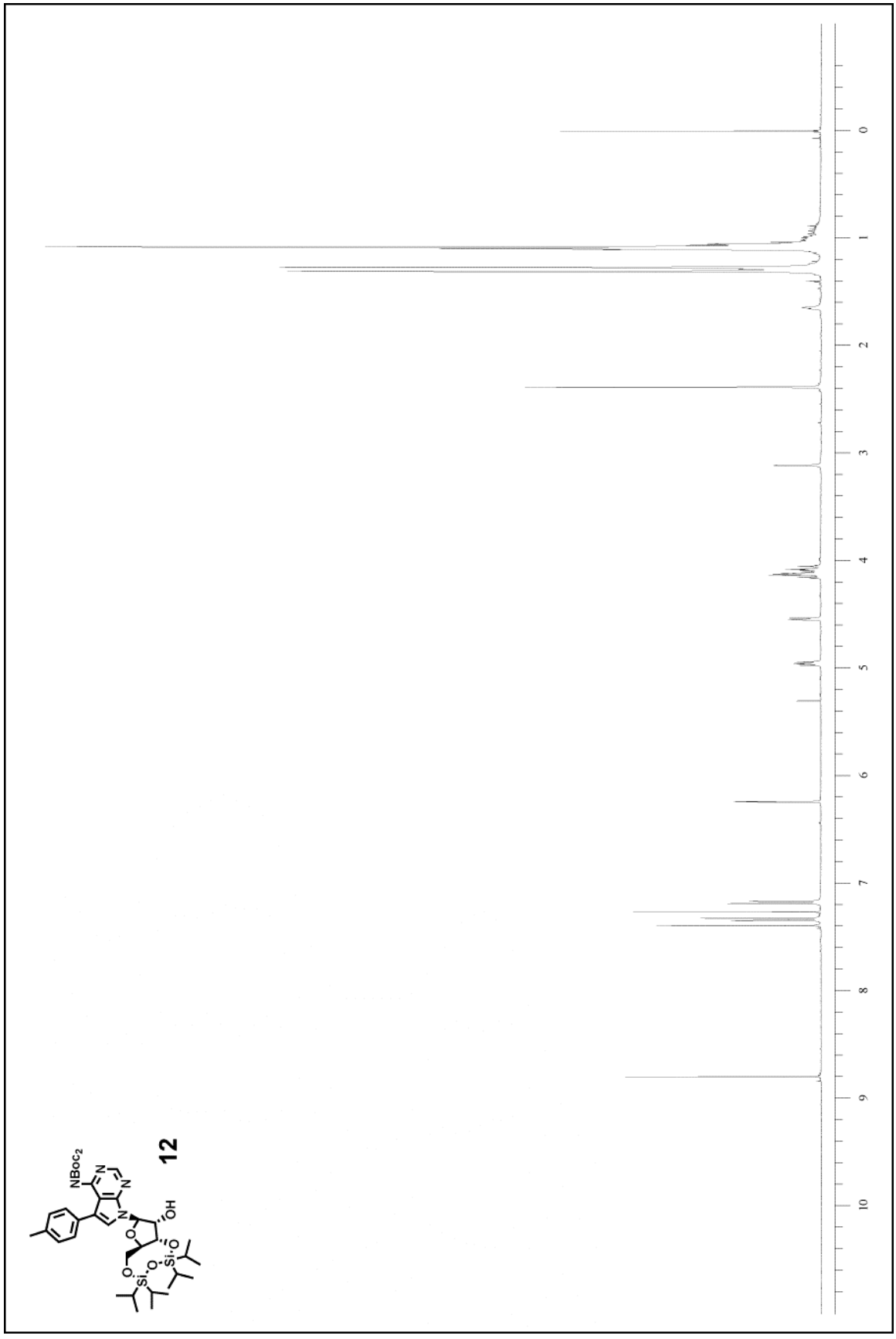

7
NH2
N
N
O
O
TBSO
0
1
2
3
4
5
6
7
8
9
10

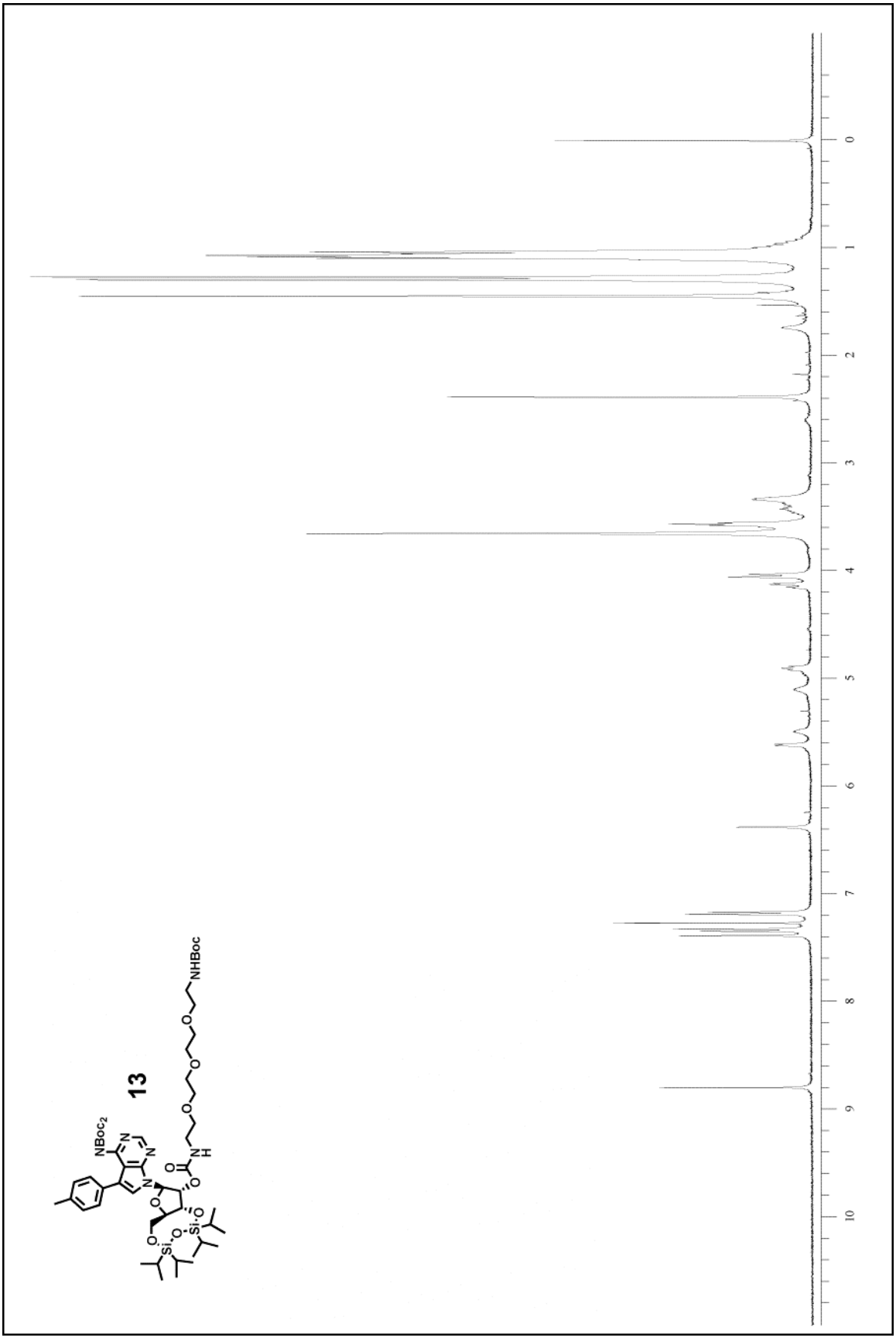

NBoc2
N
N
O
O
TBSO
8
0
1
2
3
4
5
6
7
8
9
10

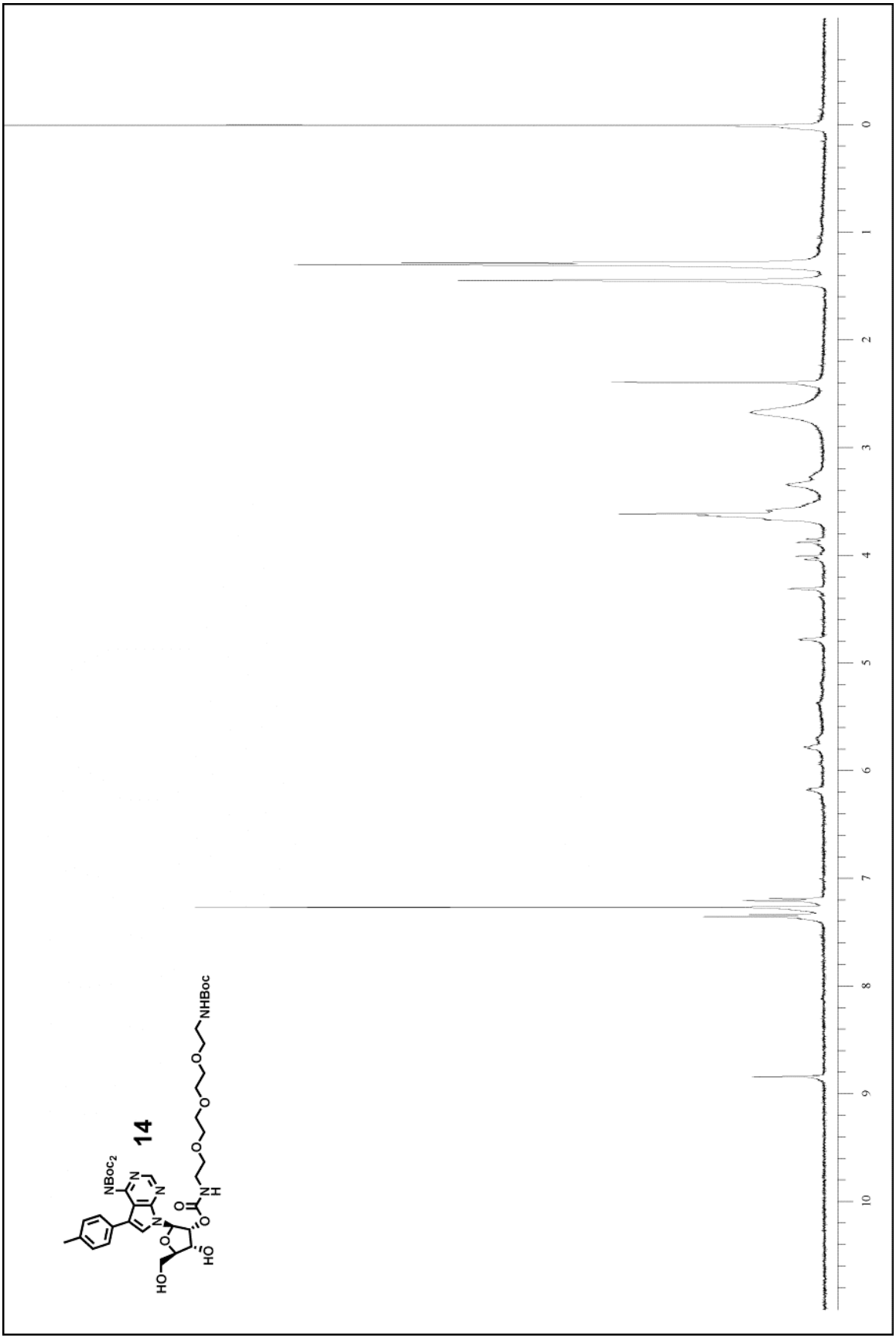

NBoc2
N
N
N
O
O
O
HO
9

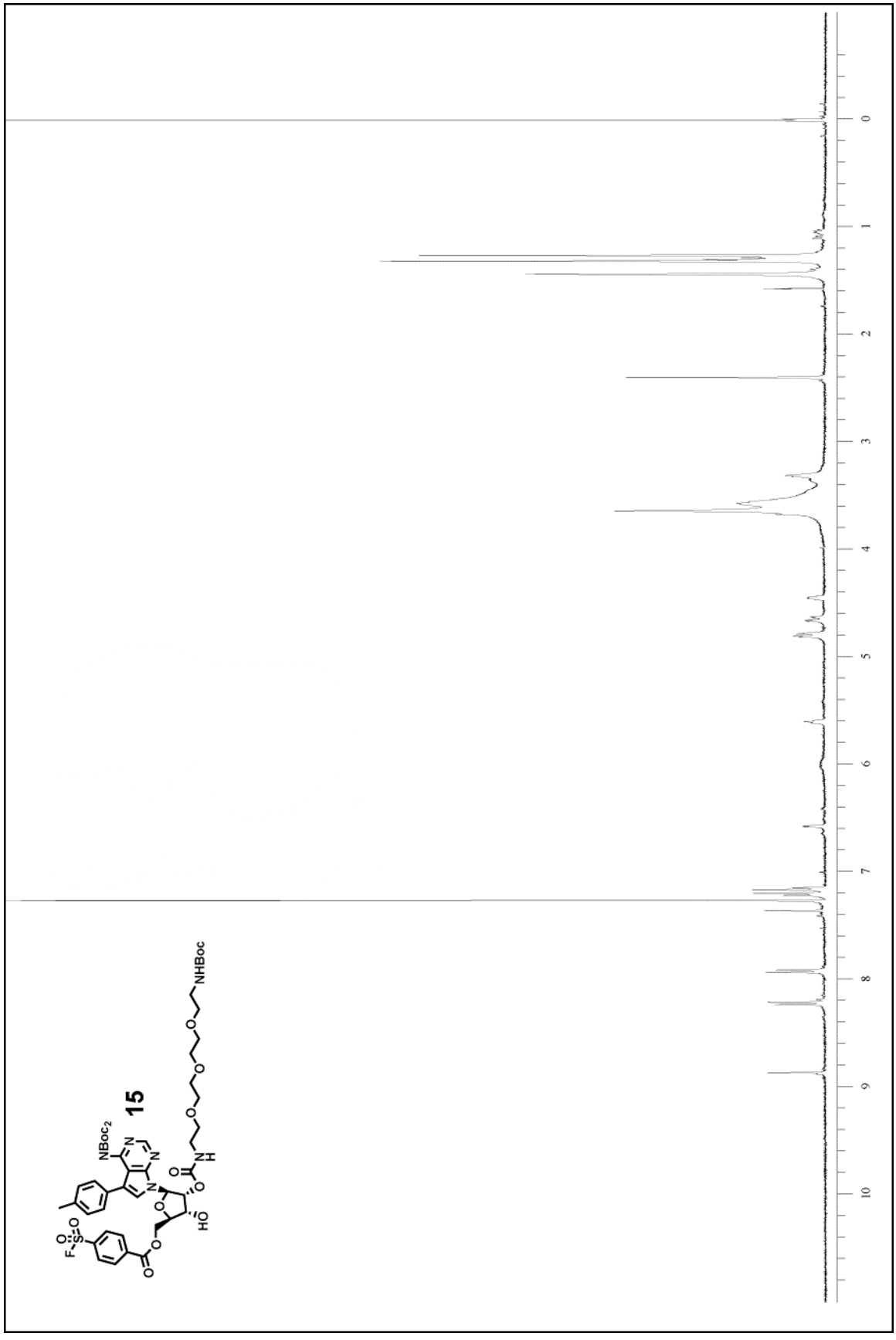

NBoc2
F
S
O
O
O
O
O
O
O
N
N
N
10

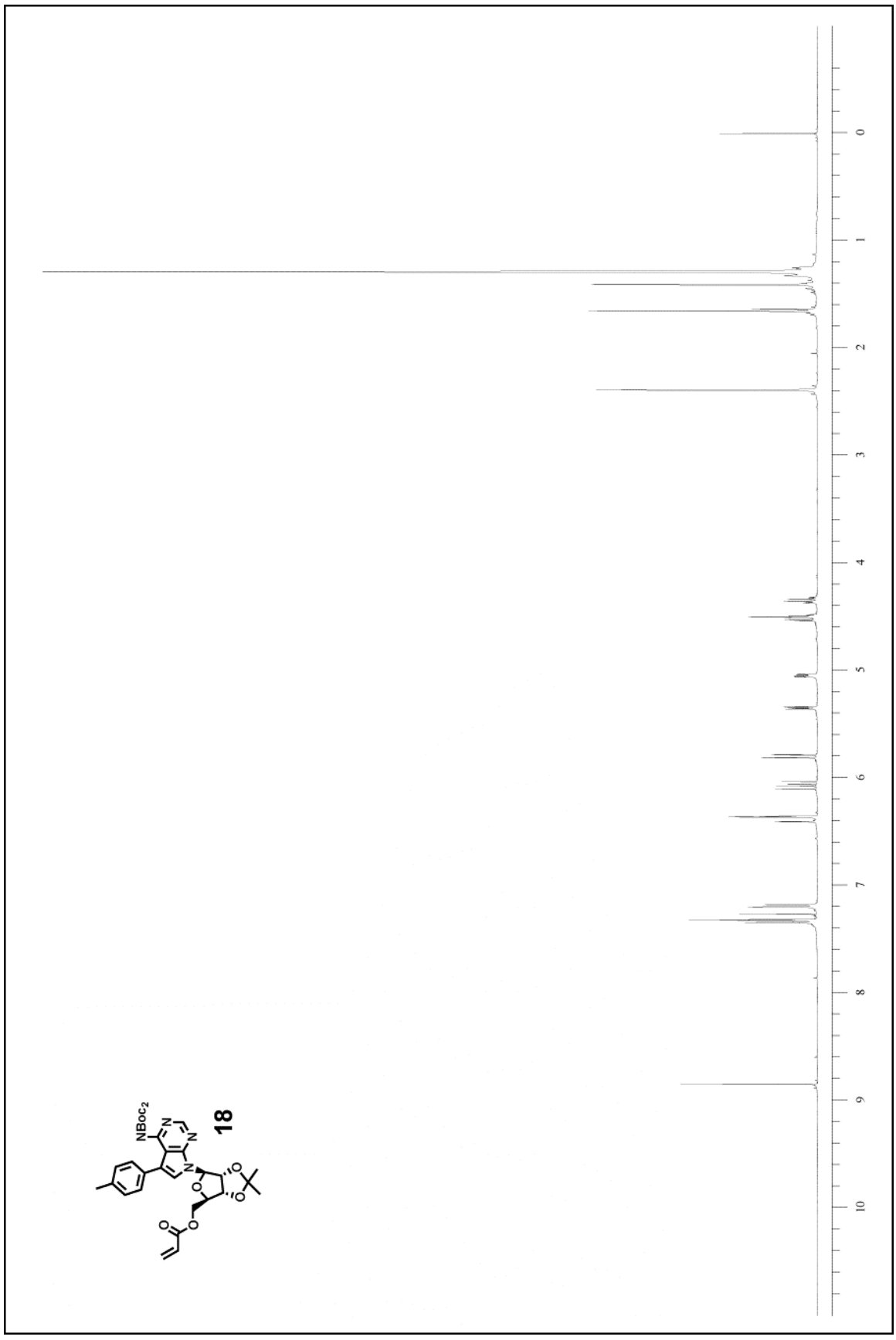

NBoc2
N
N
N
O
OH
HO
HO
11

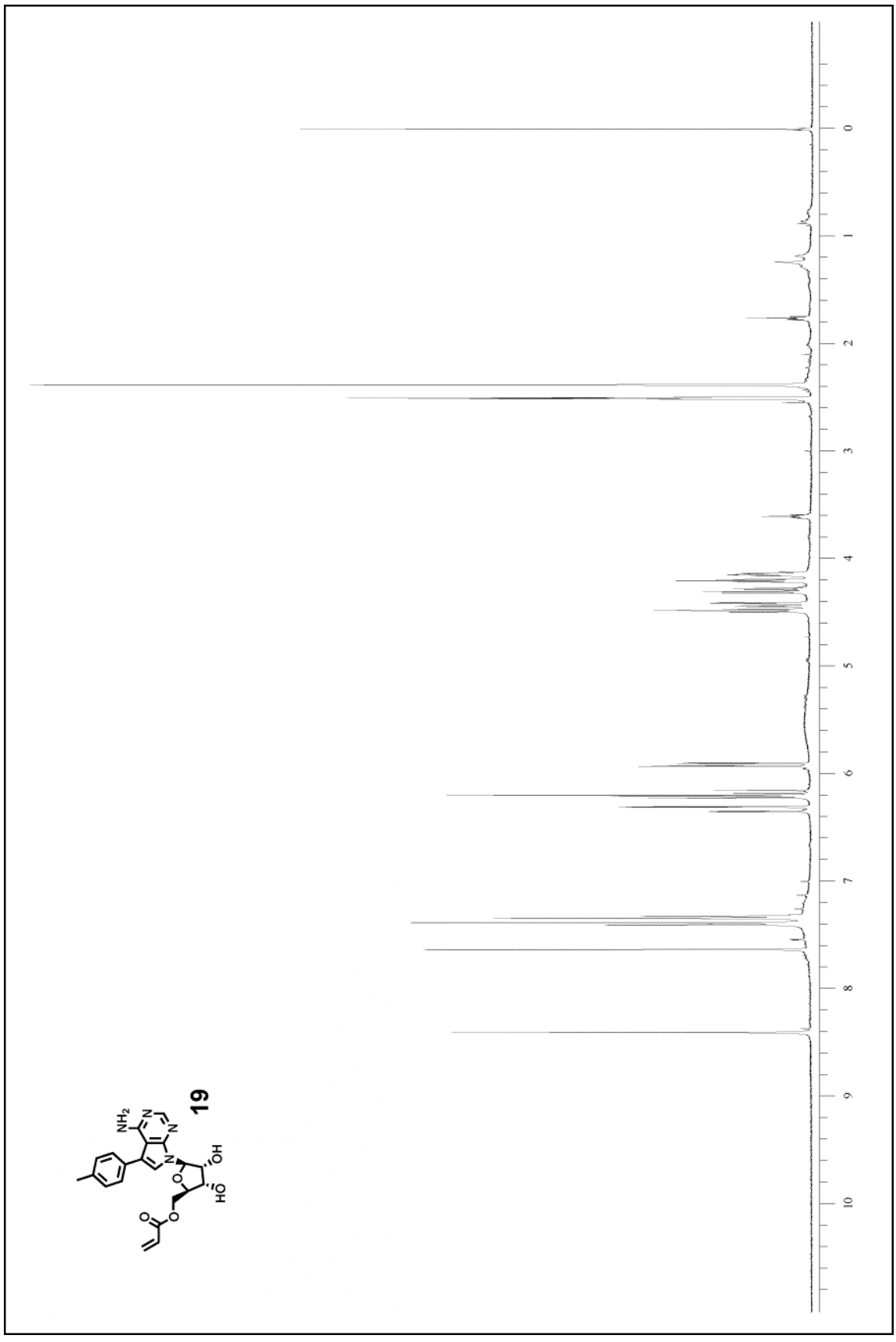

NBoc2
N
N
N
O
OH
O
Si
O
Si
O
12
0
1
2
3
4
5
6
7
8
9
10

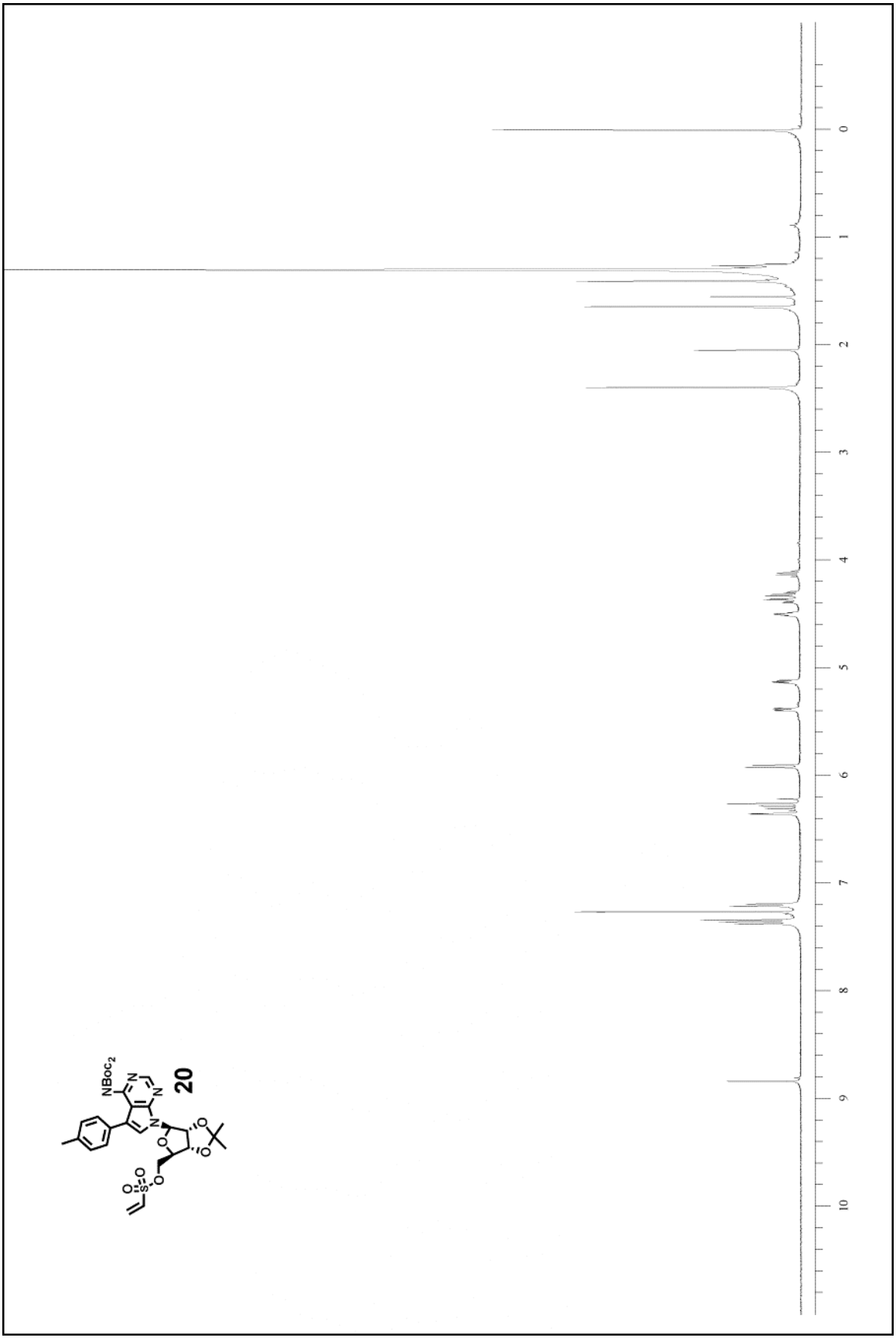

13
NBoc2
NHBoc
Si
Si
O
N
H
10
9
8
7
6
5
4
3
2
1
0

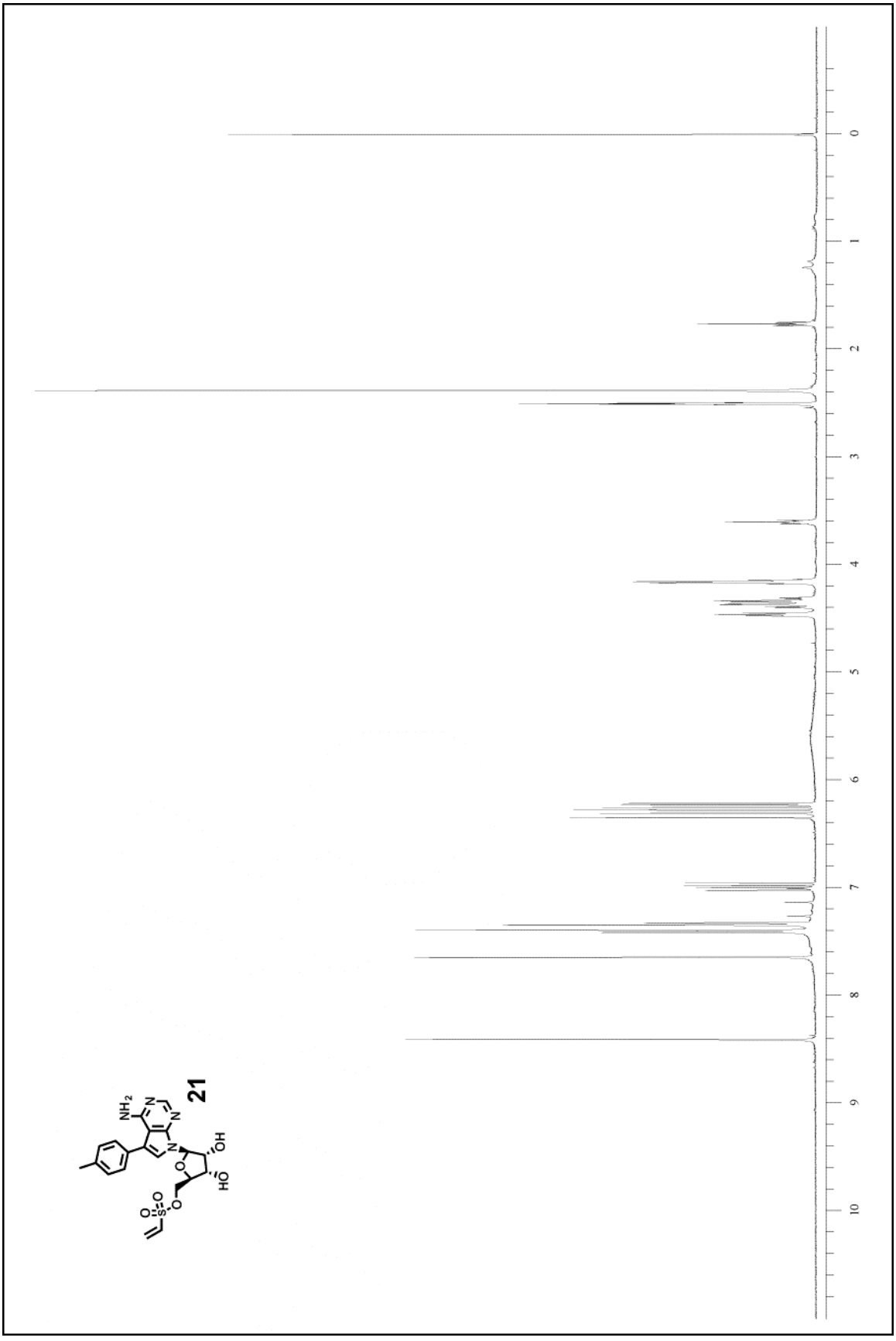

14
NBoc2
NHBoc
HO
HO
0
1
2
3
4
5
6
7
8
9
10

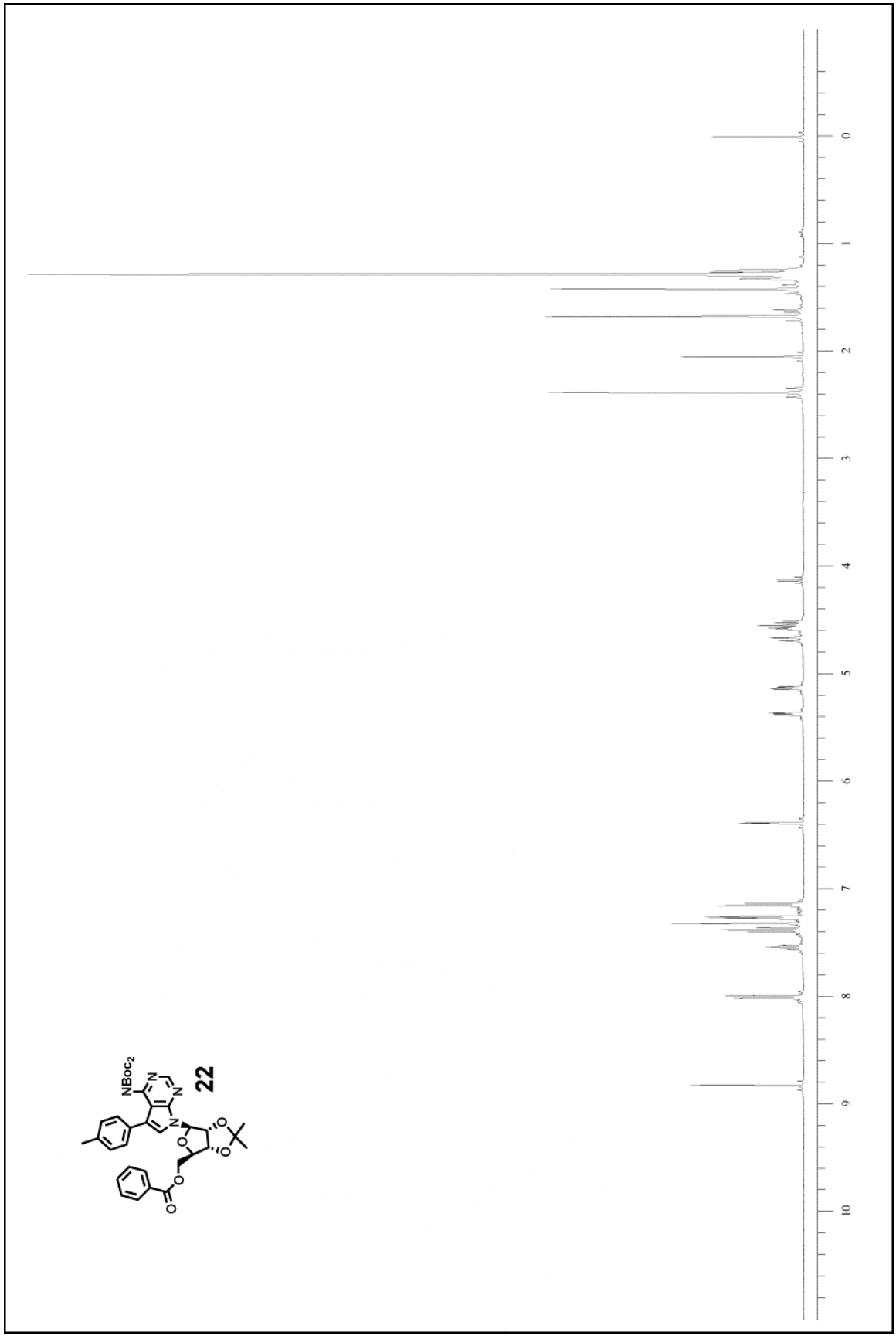

15
NBoc2
NHBoc
HO
F
S
O
N
H
10
9
8
7
6
5
4
3
2
1
0

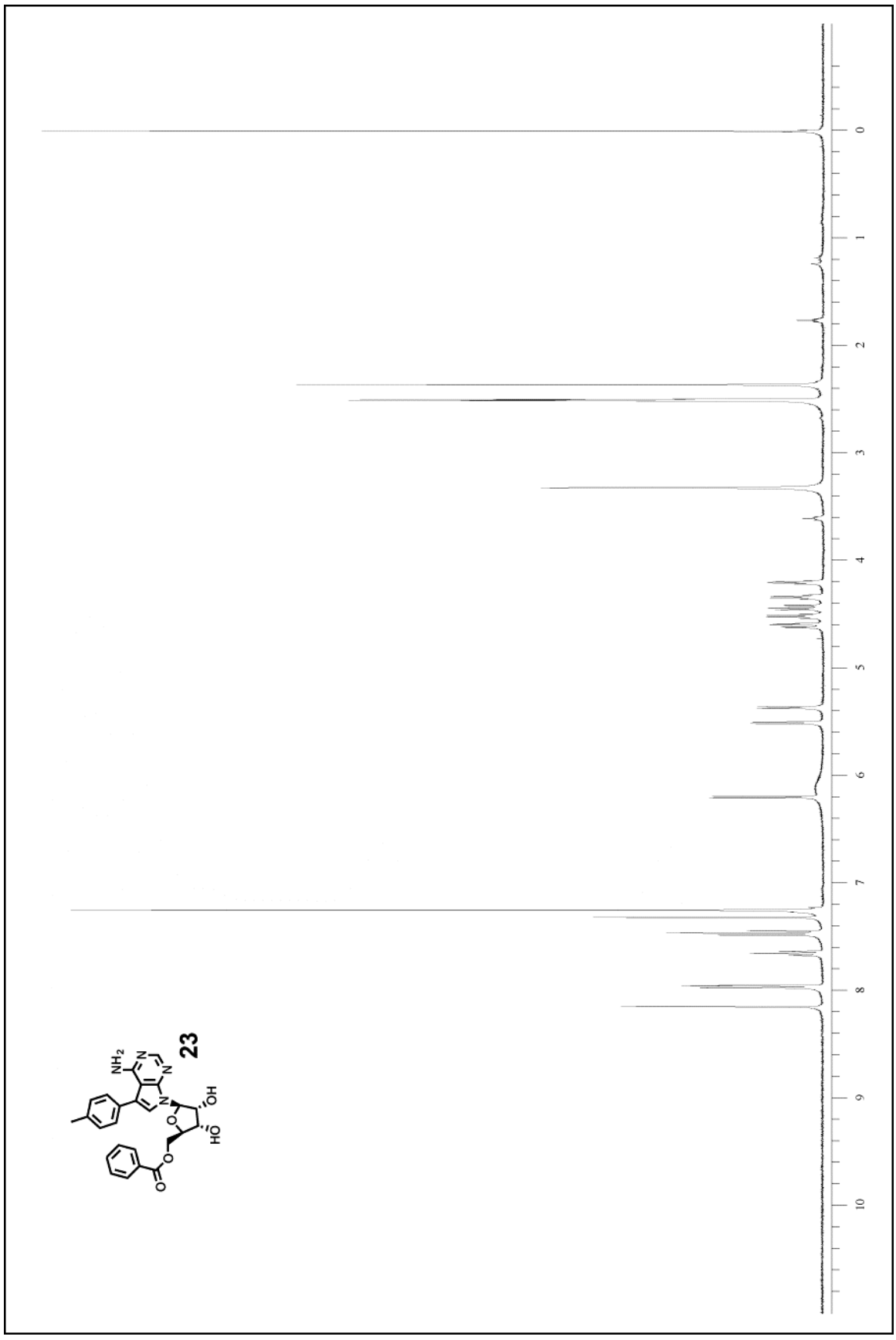

NBoc2
18

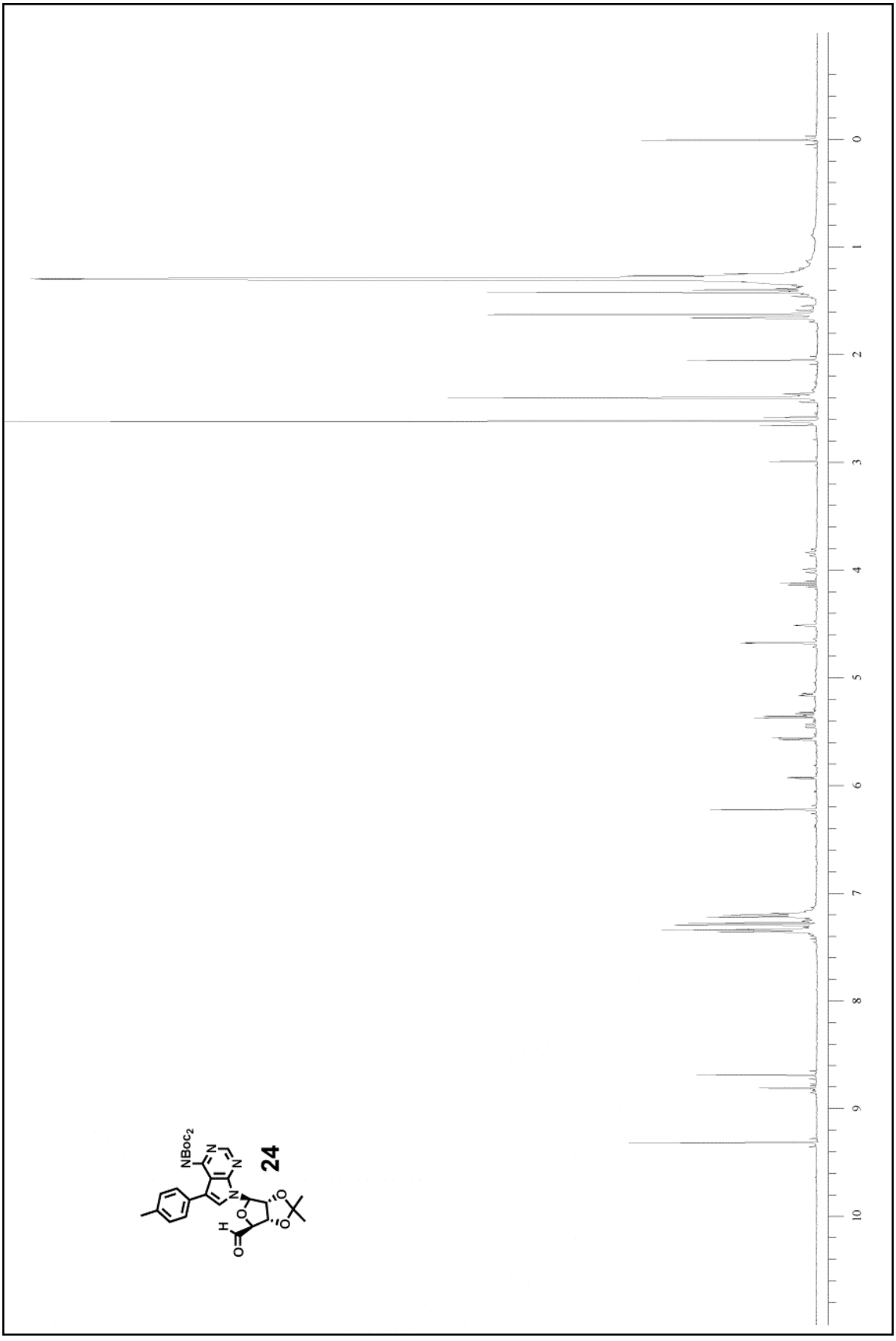

19
NH2
N
N
N
HO
OH
O
O
O
0
1
2
3
4
5
6
7
8
9
10

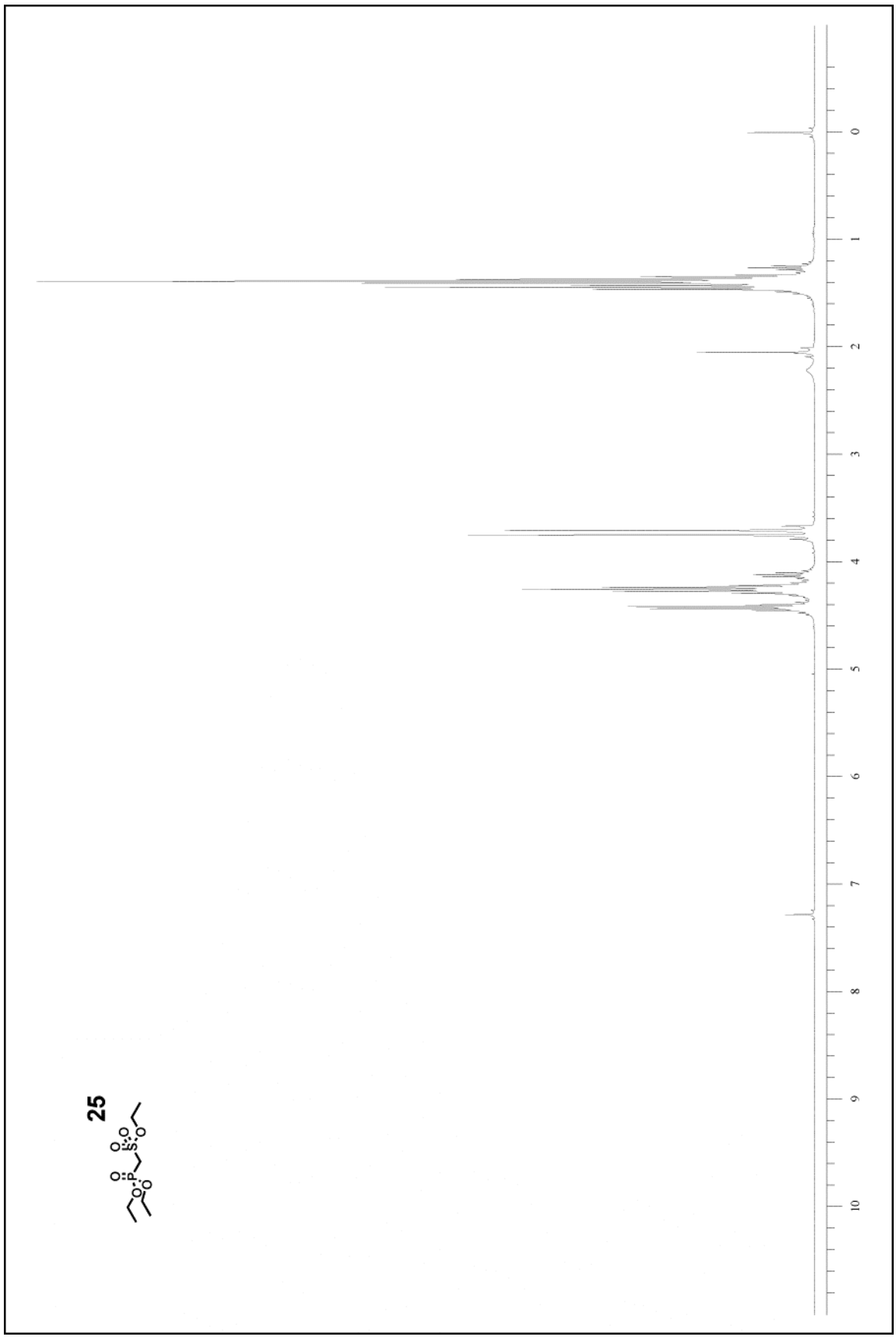

NBoc2
20

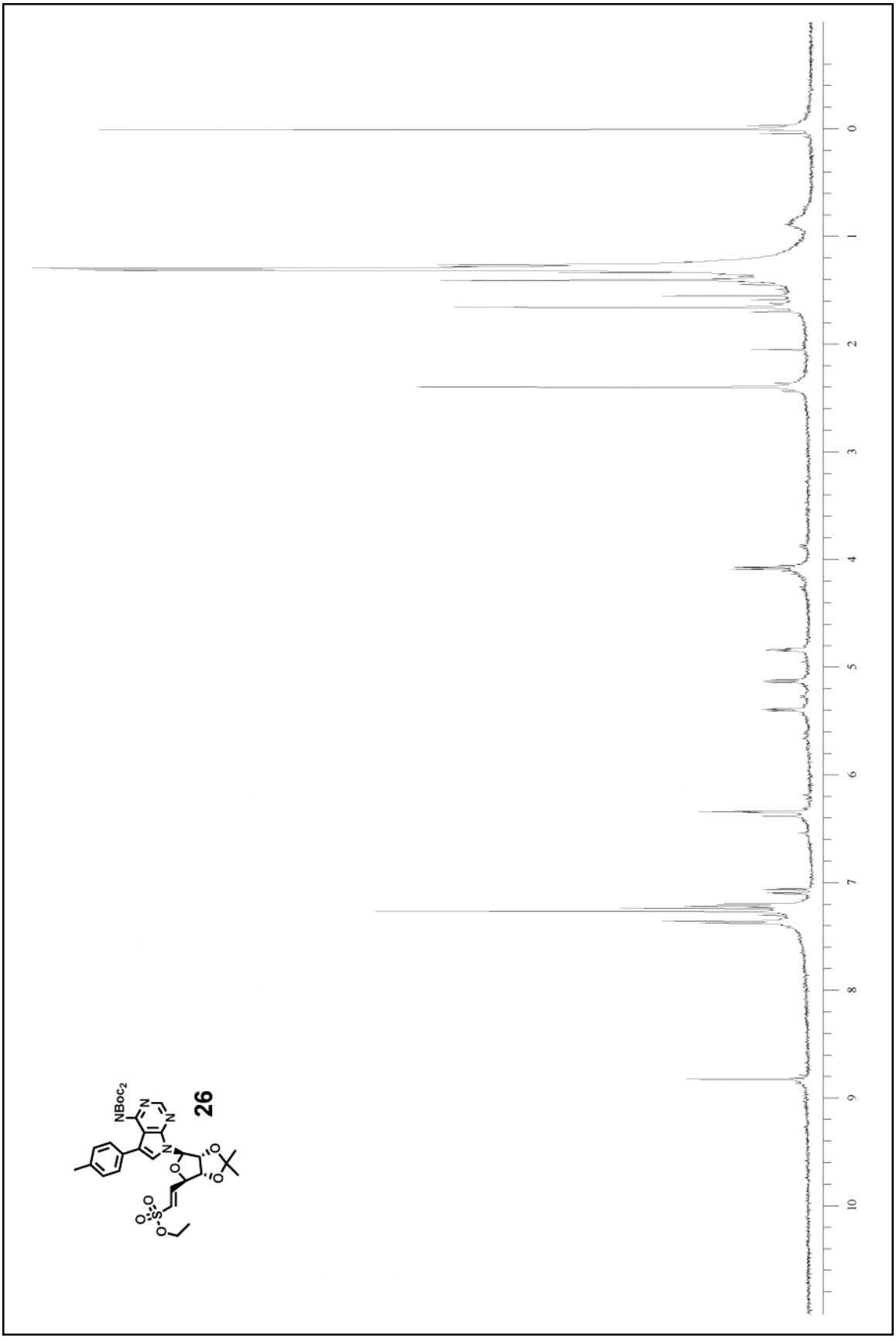

21
NH2
N
N
N
O
OH
HO
O
O
S
O
O
0
1
2
3
4
5
6
7
8
9
10

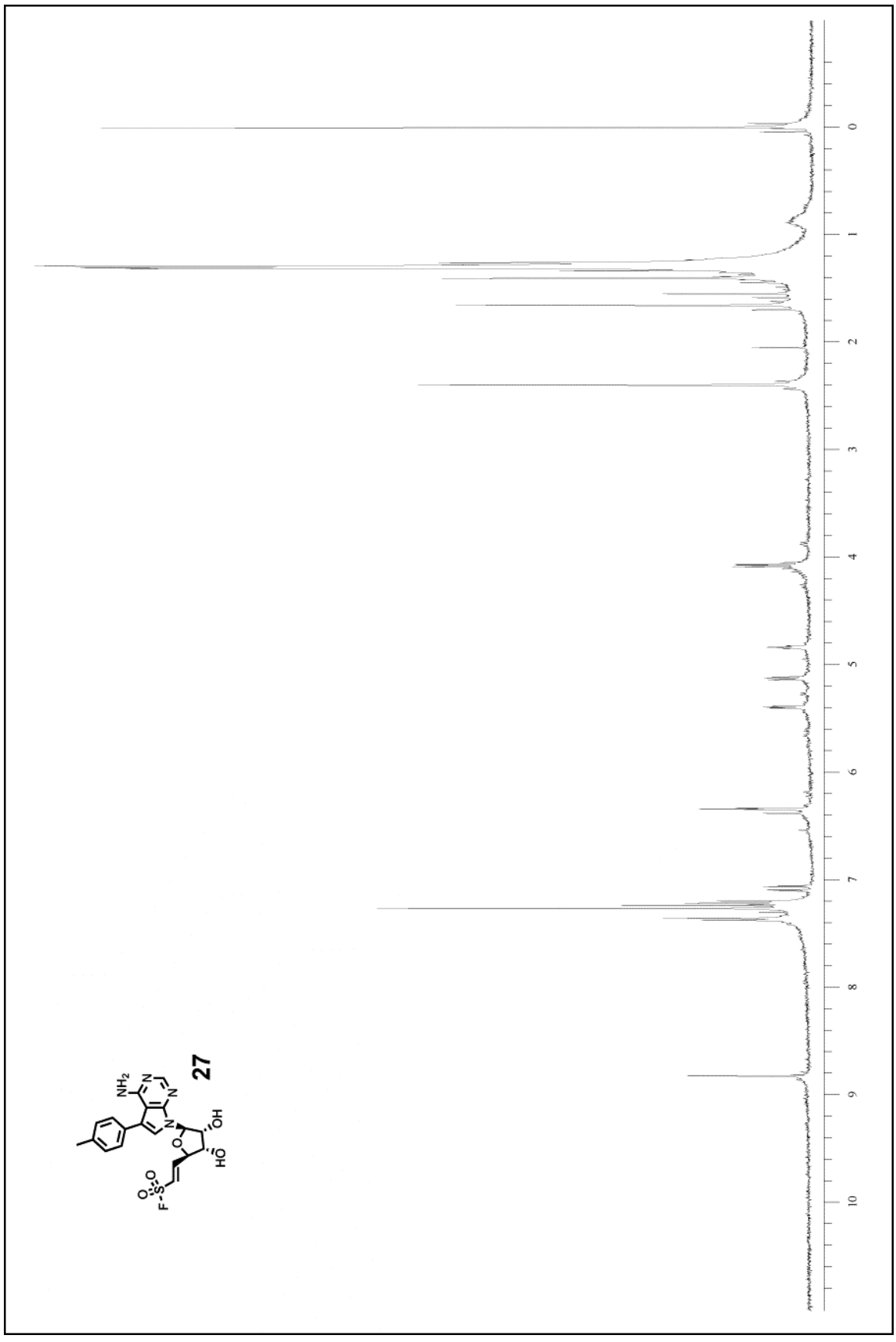

NBoc2
22

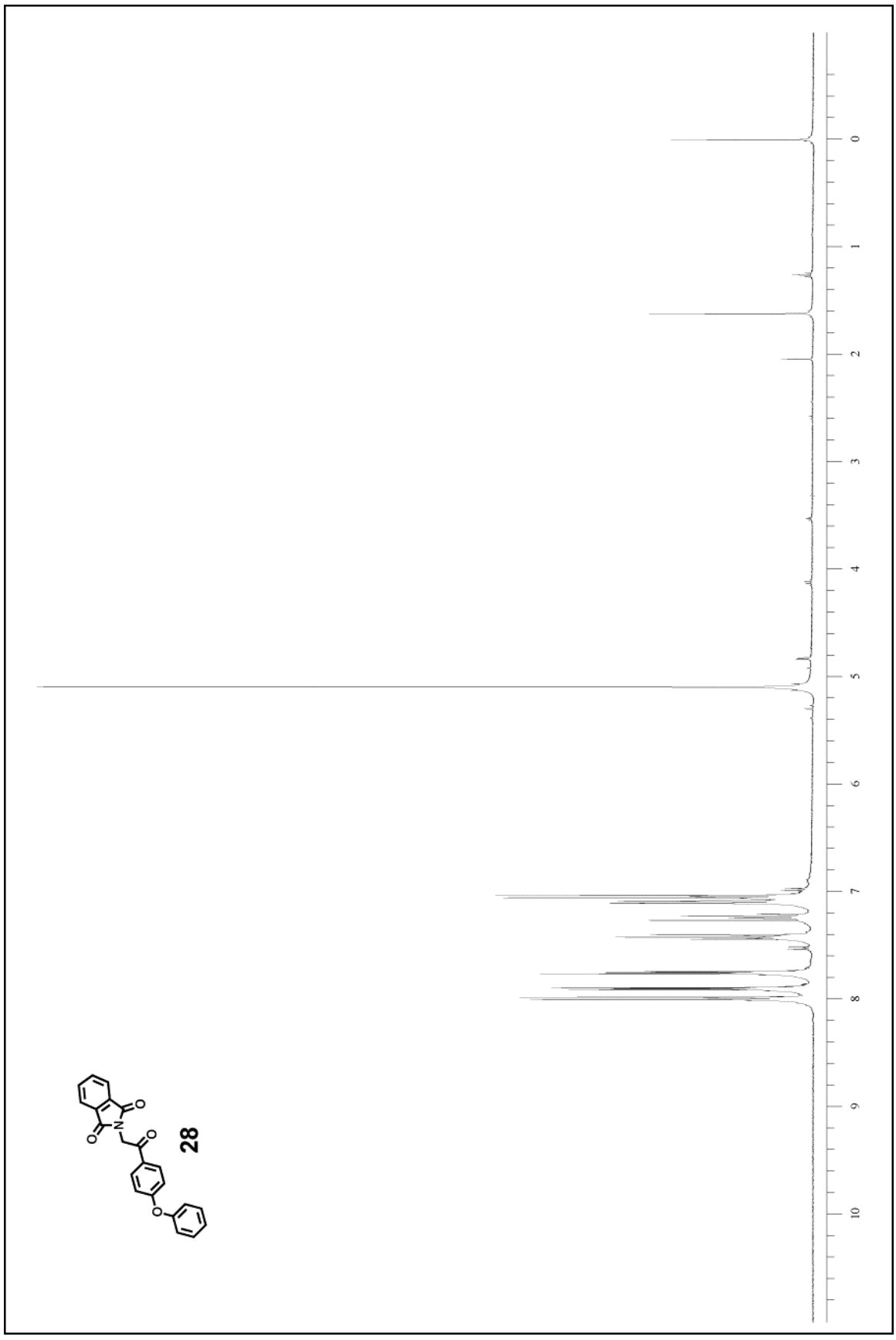

23
NH2
N
N
N
O
HO
HO
O
O
0
1
2
3
4
5
6
7
8
9
10

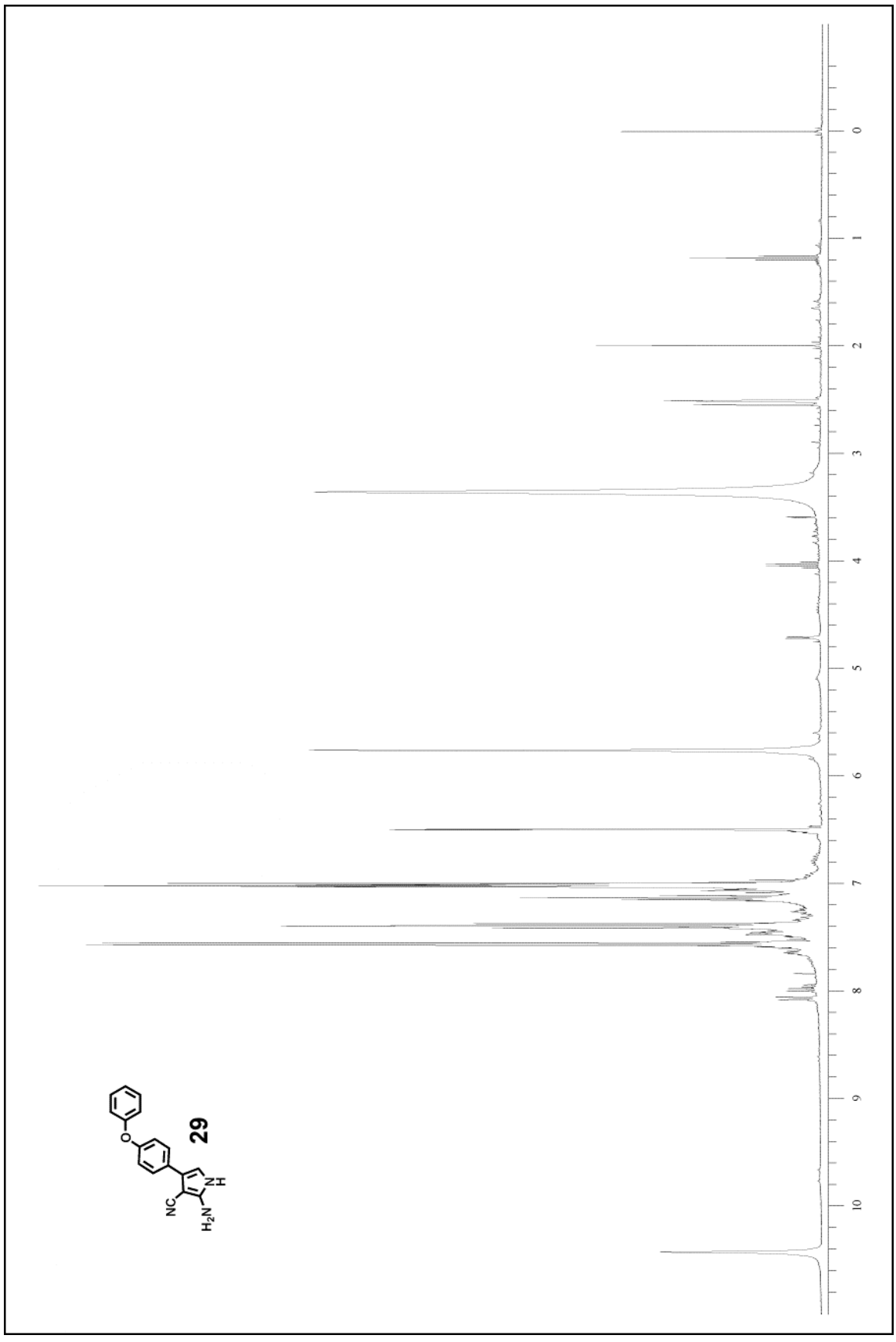

NBoc2
24

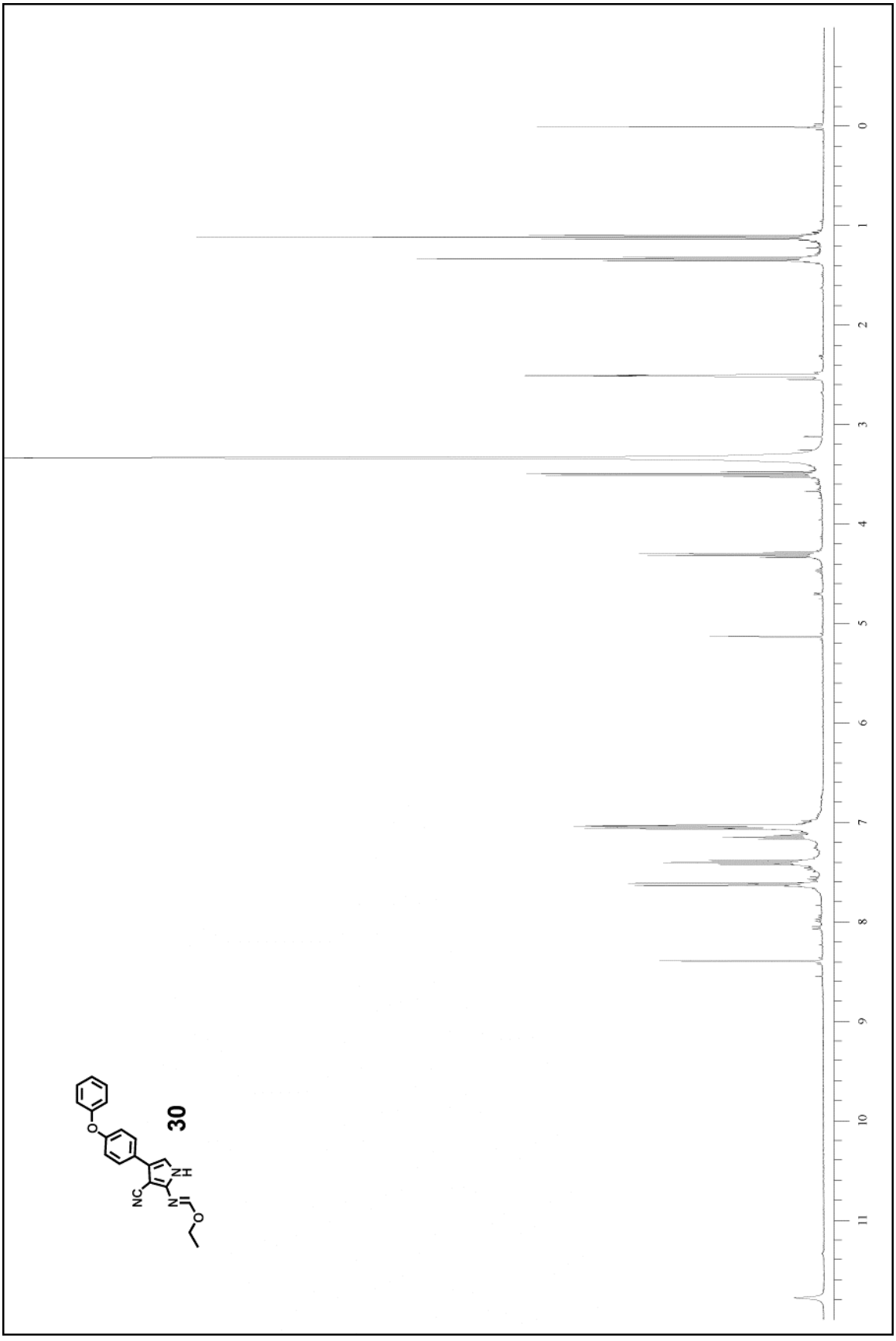

25

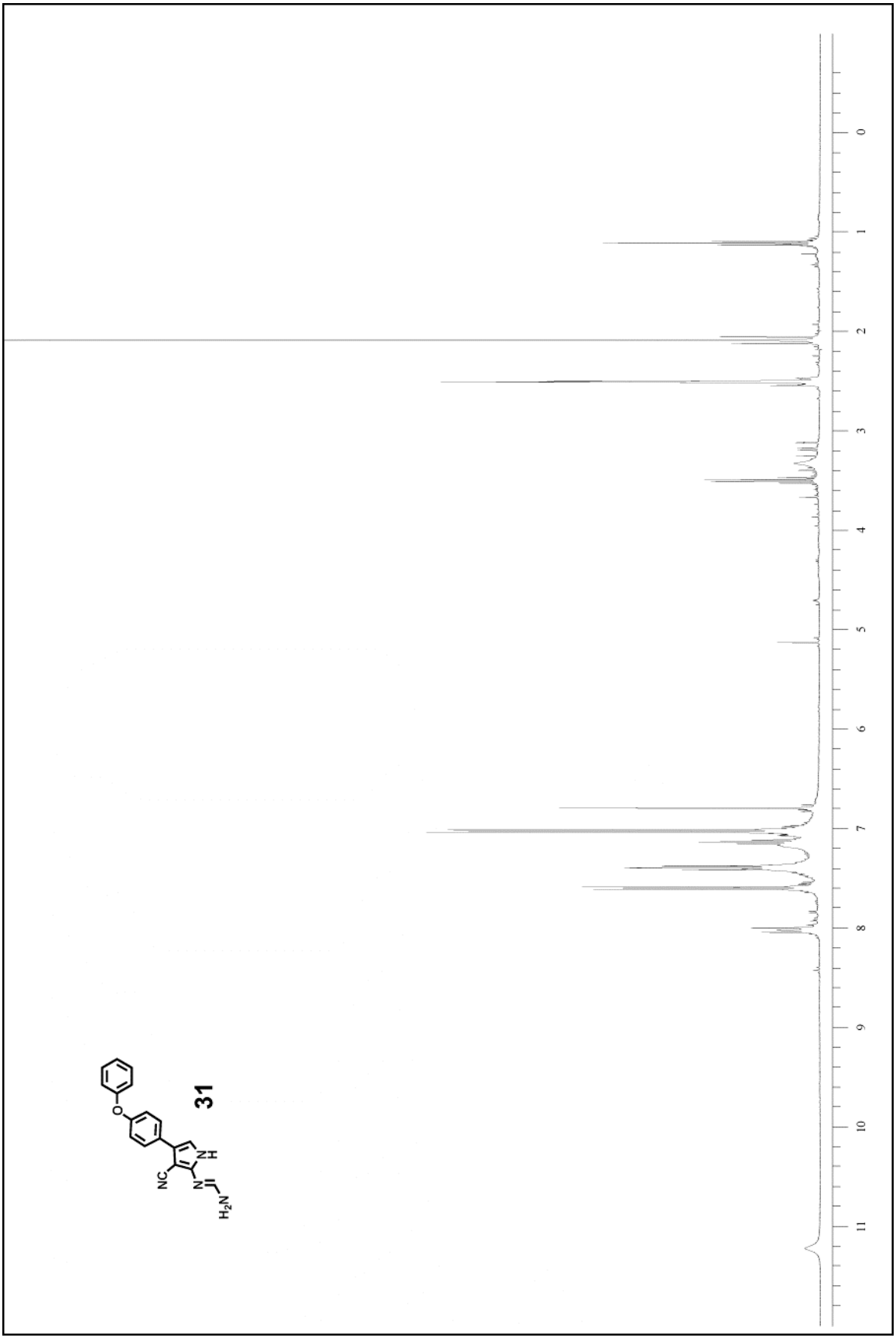

NBoc2
26

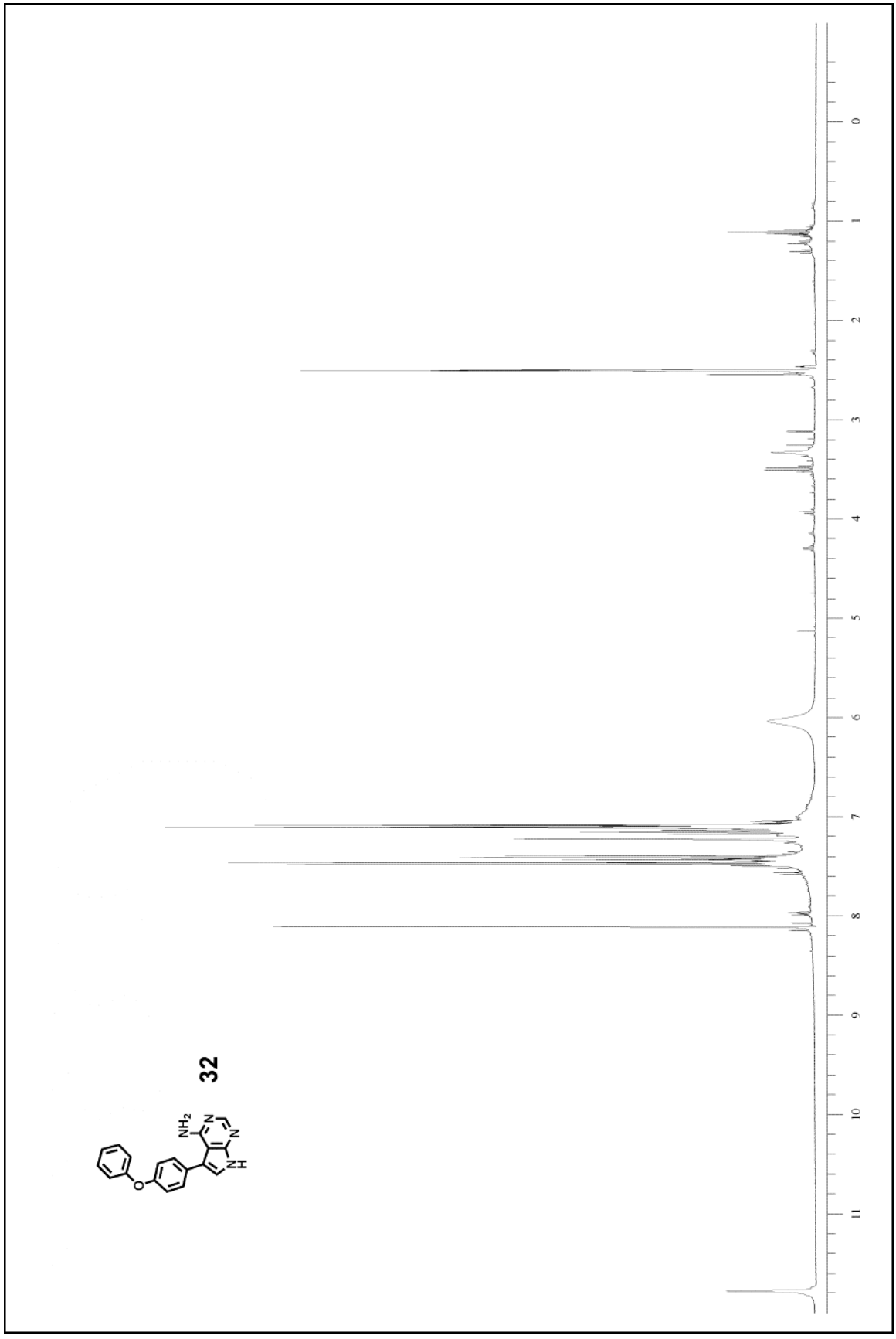

27
NH2
N
N
O
OH
HO
O
S
O
F

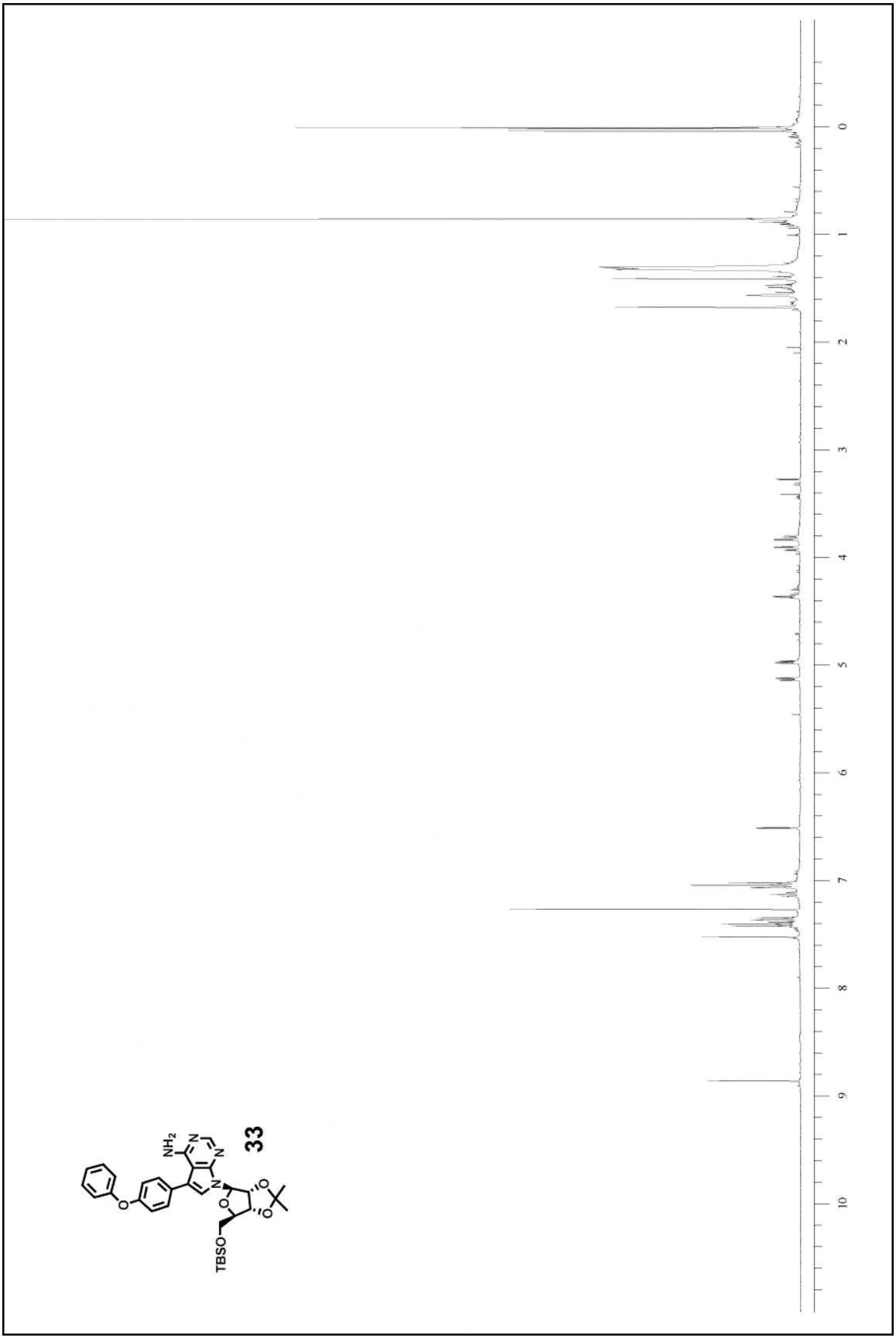

28

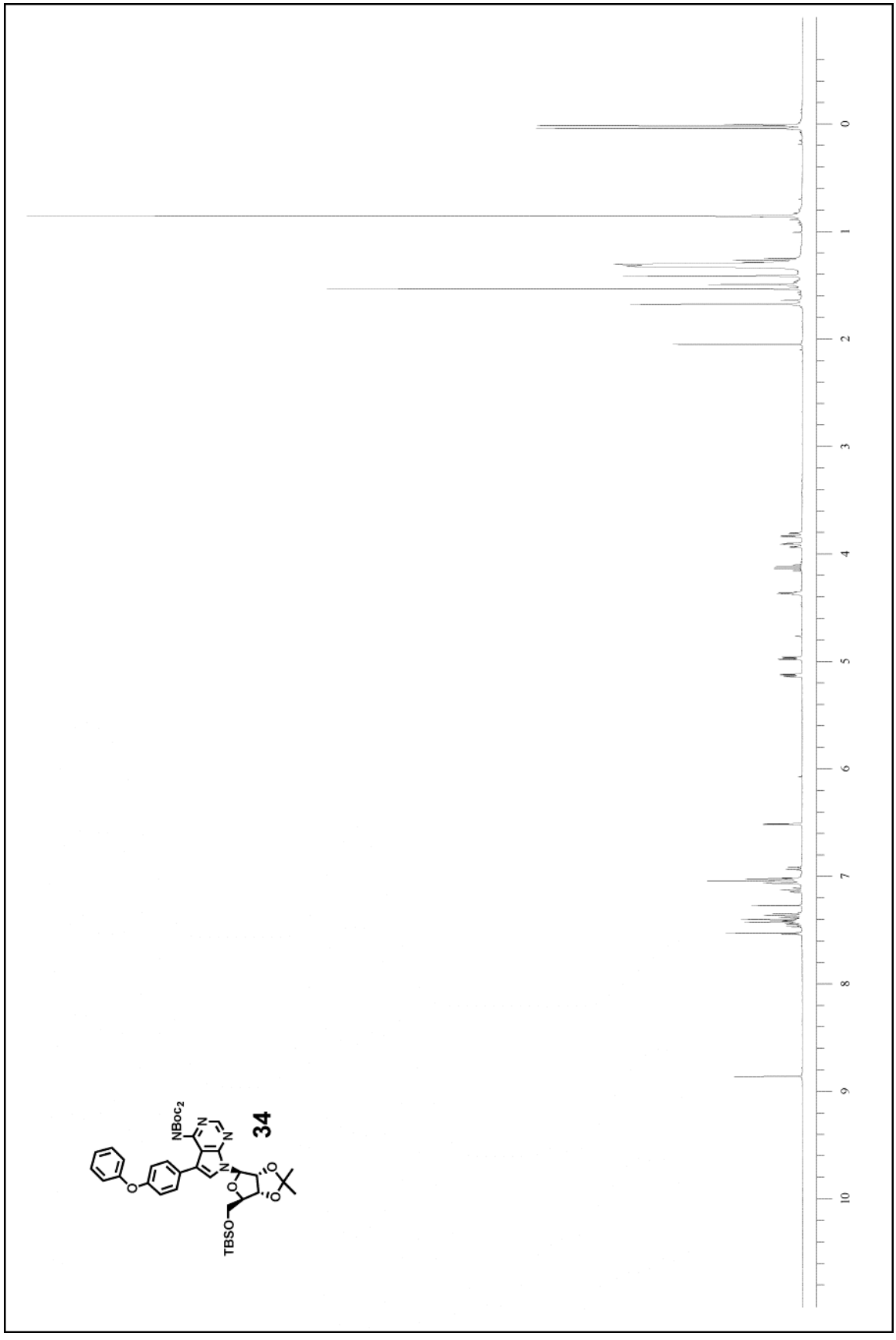

29
NC
H2N
N
H
O

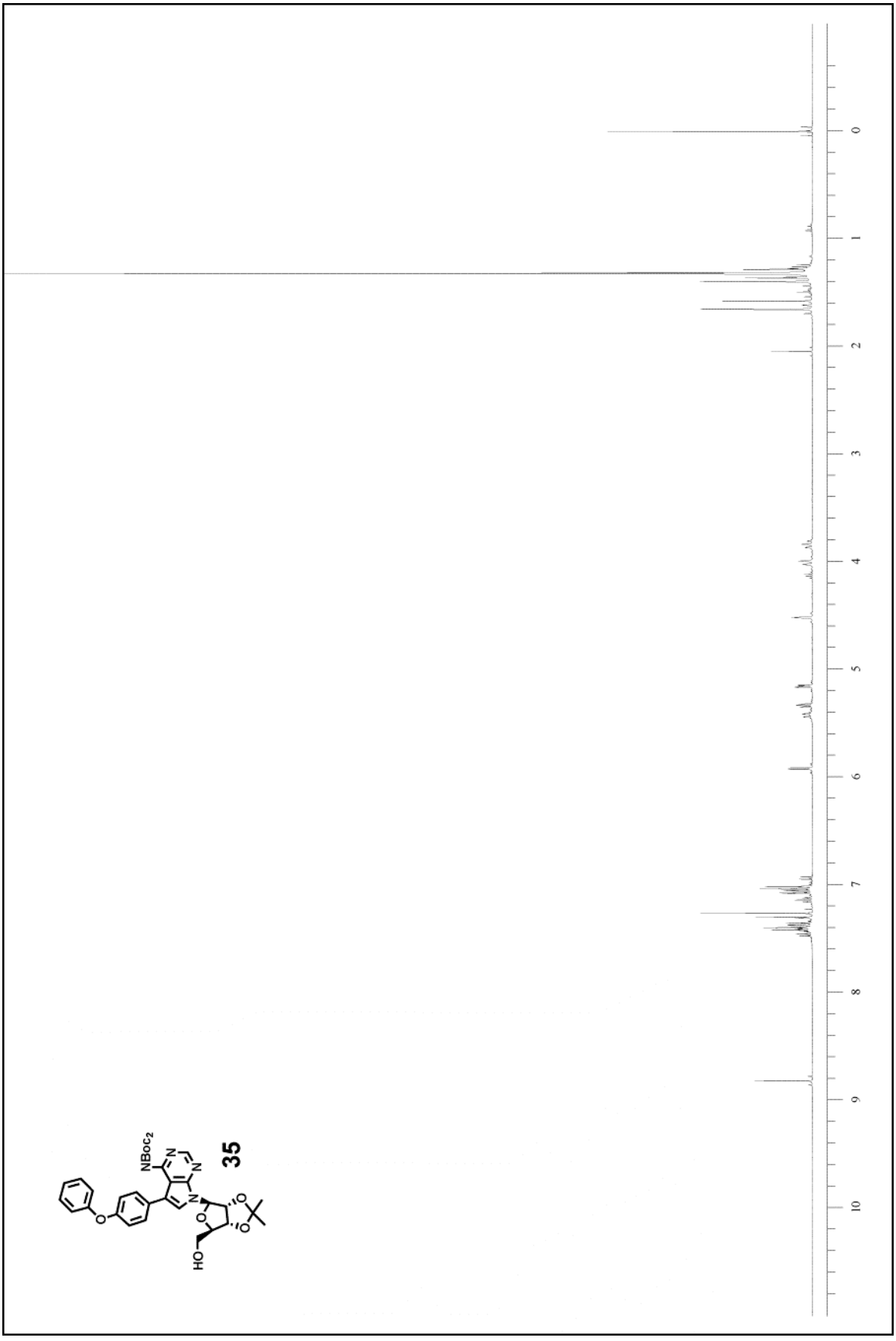

30

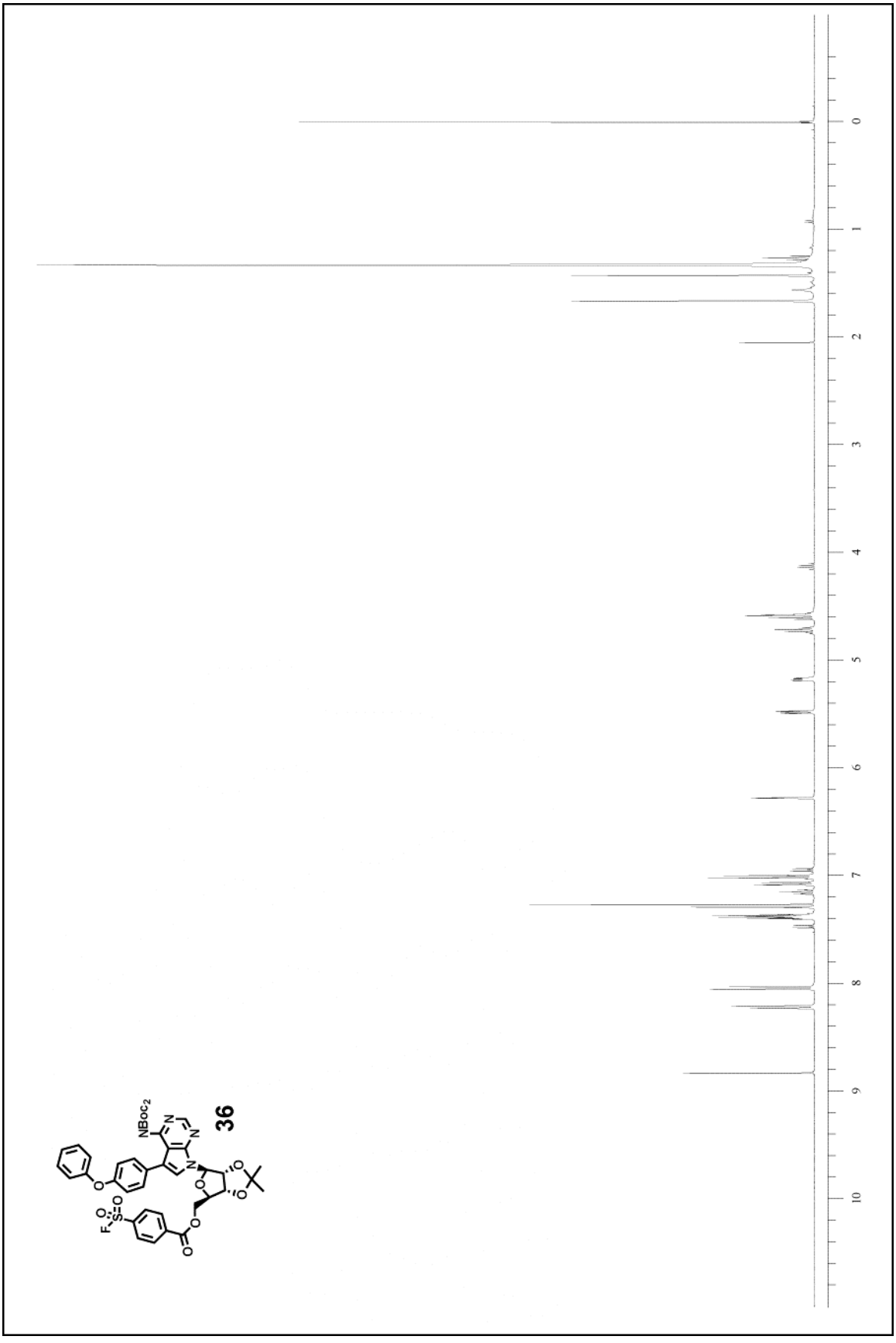

31
NC
N
H
N
H2N

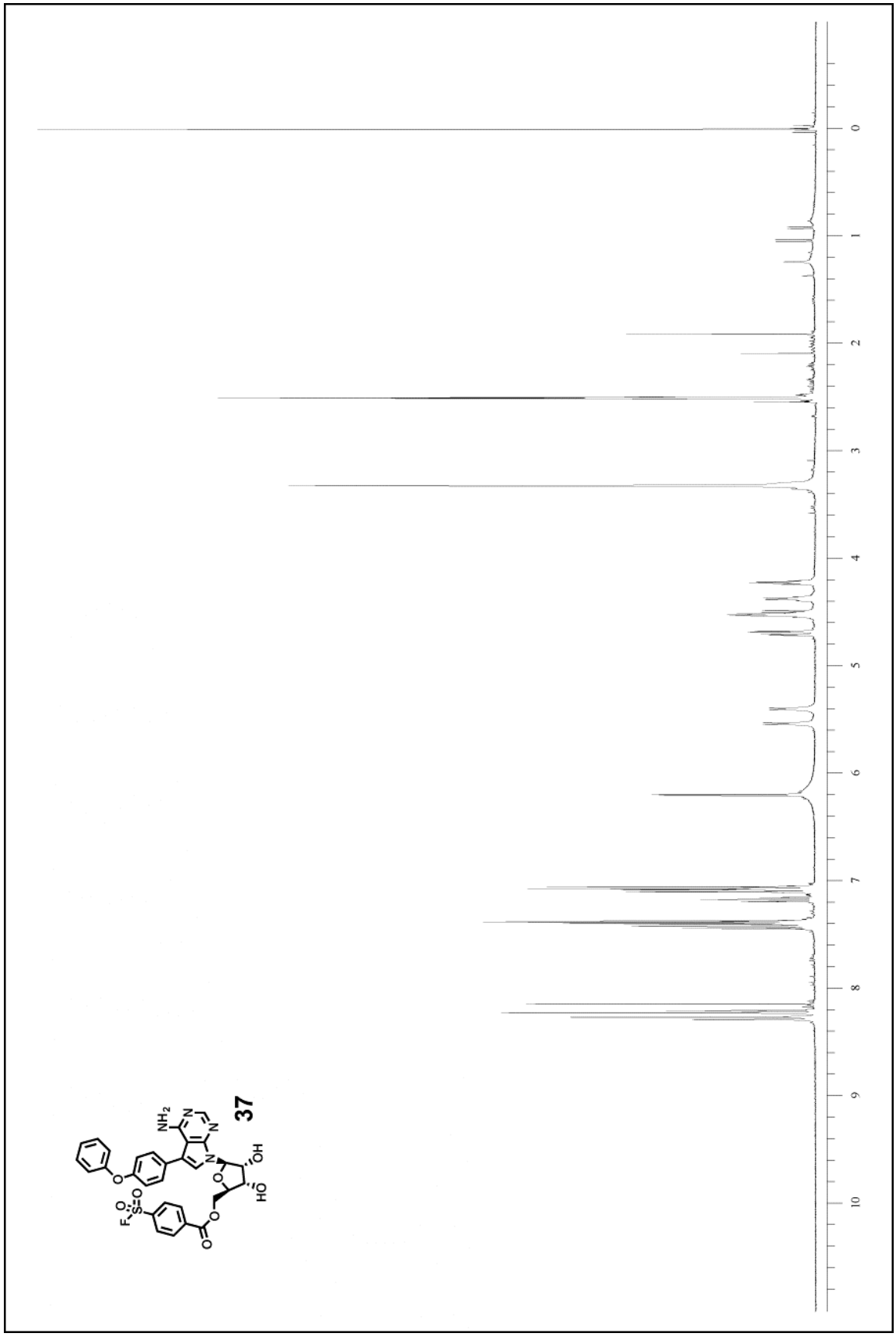

32
NH2
N
N
NH
O

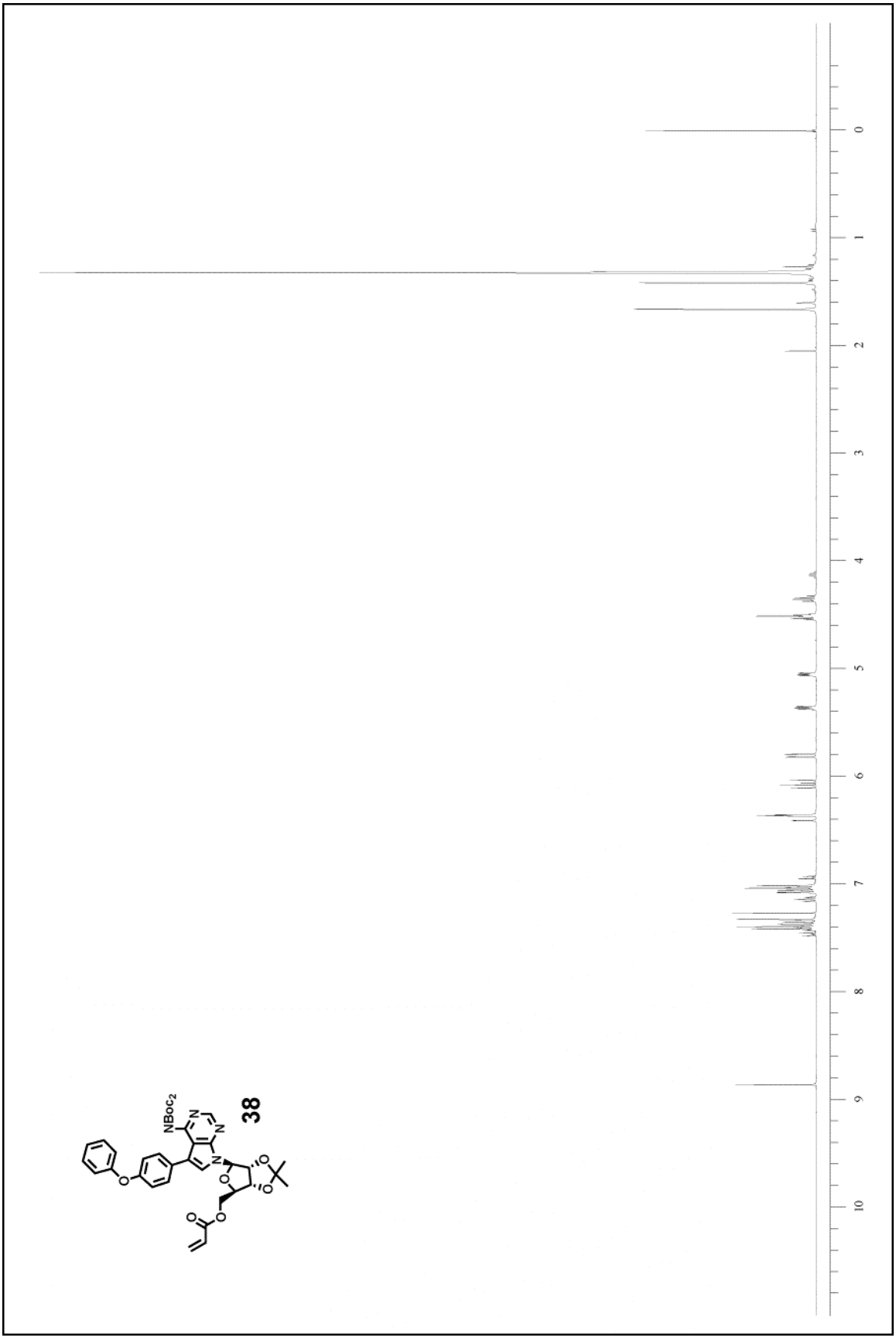

33
NH2
N
N
N
O
O
O
TBSO
0
1
2
3
4
5
6
7
8
9
10

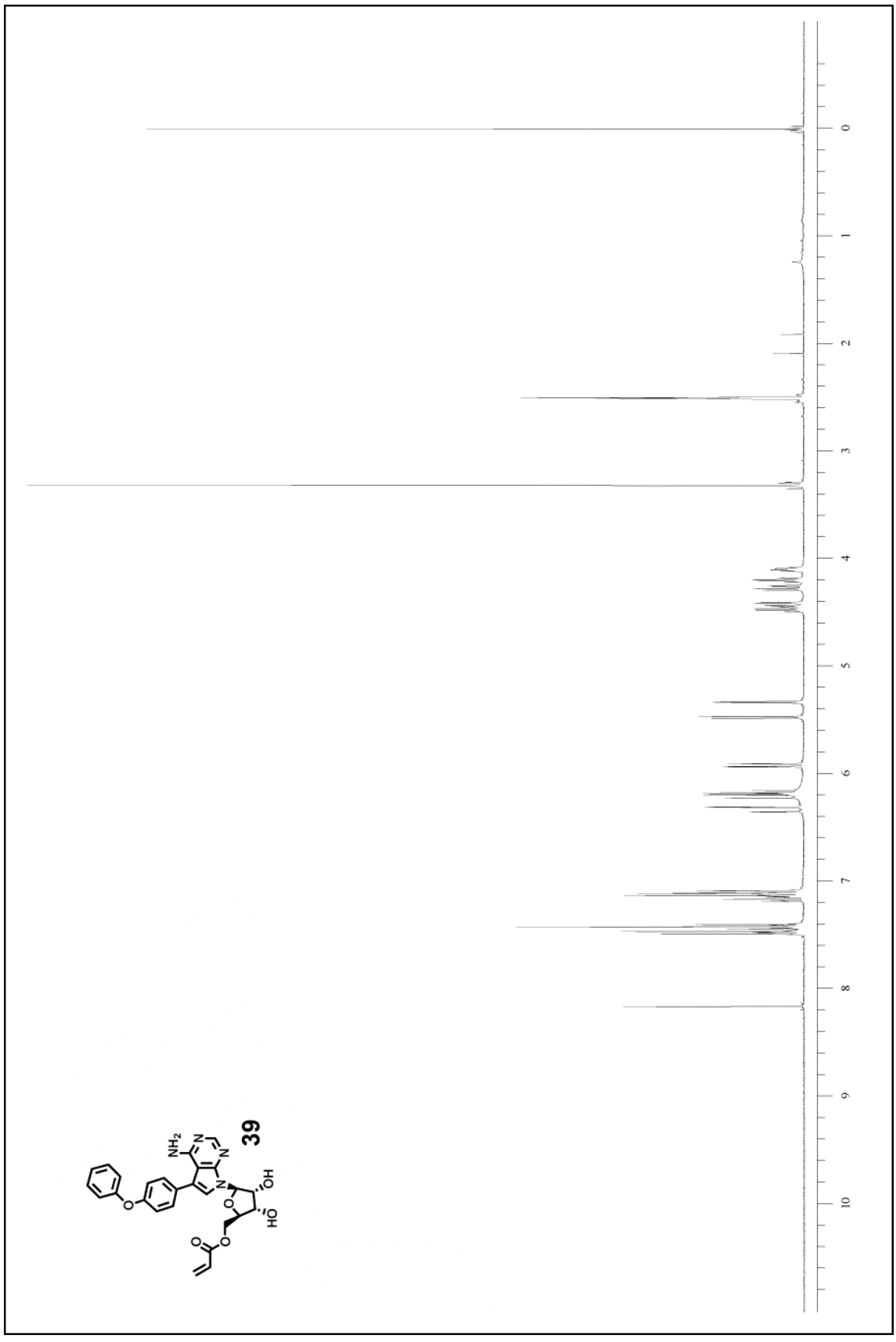

NBoc2
TBSO
34

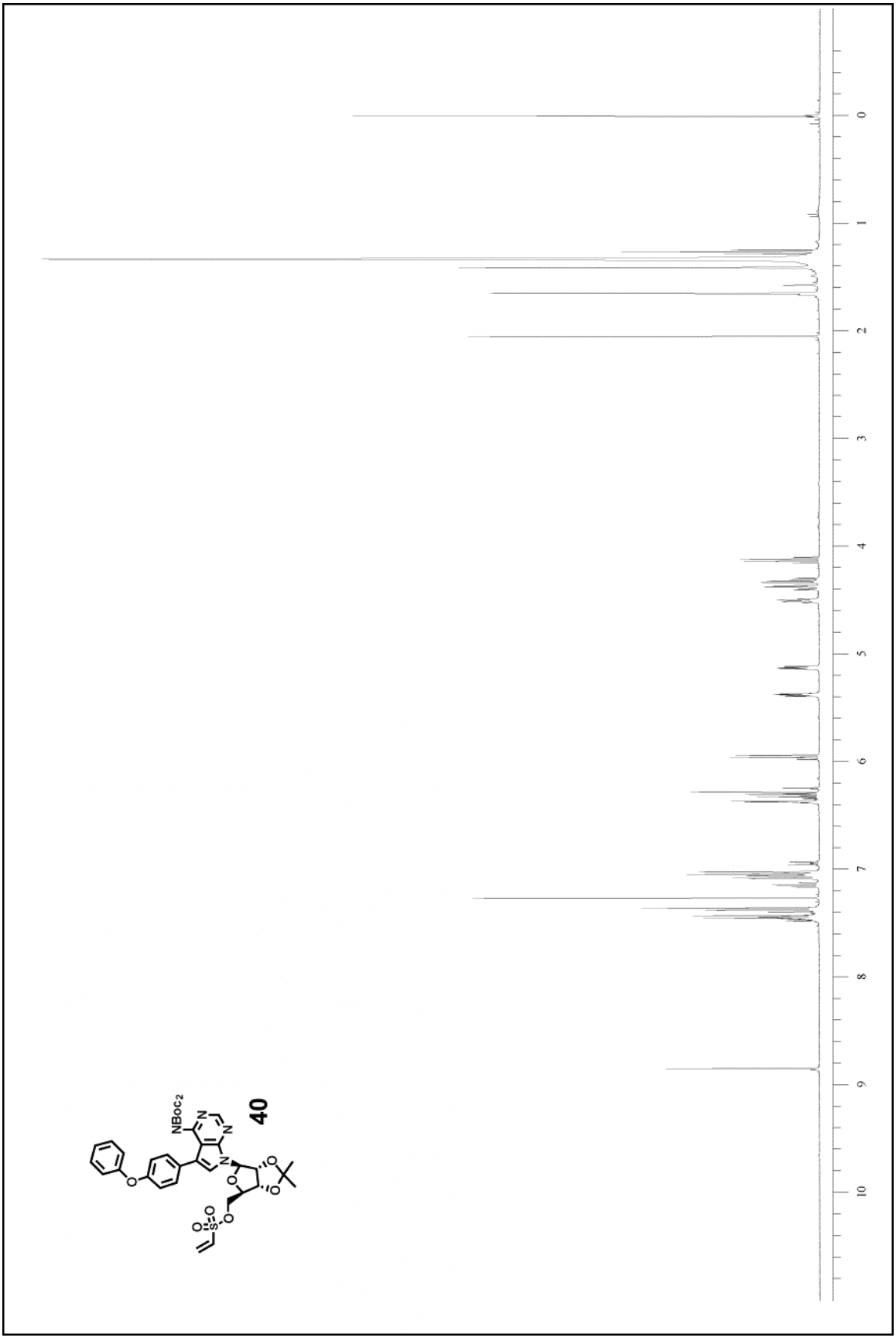

NBoc2
35
HO

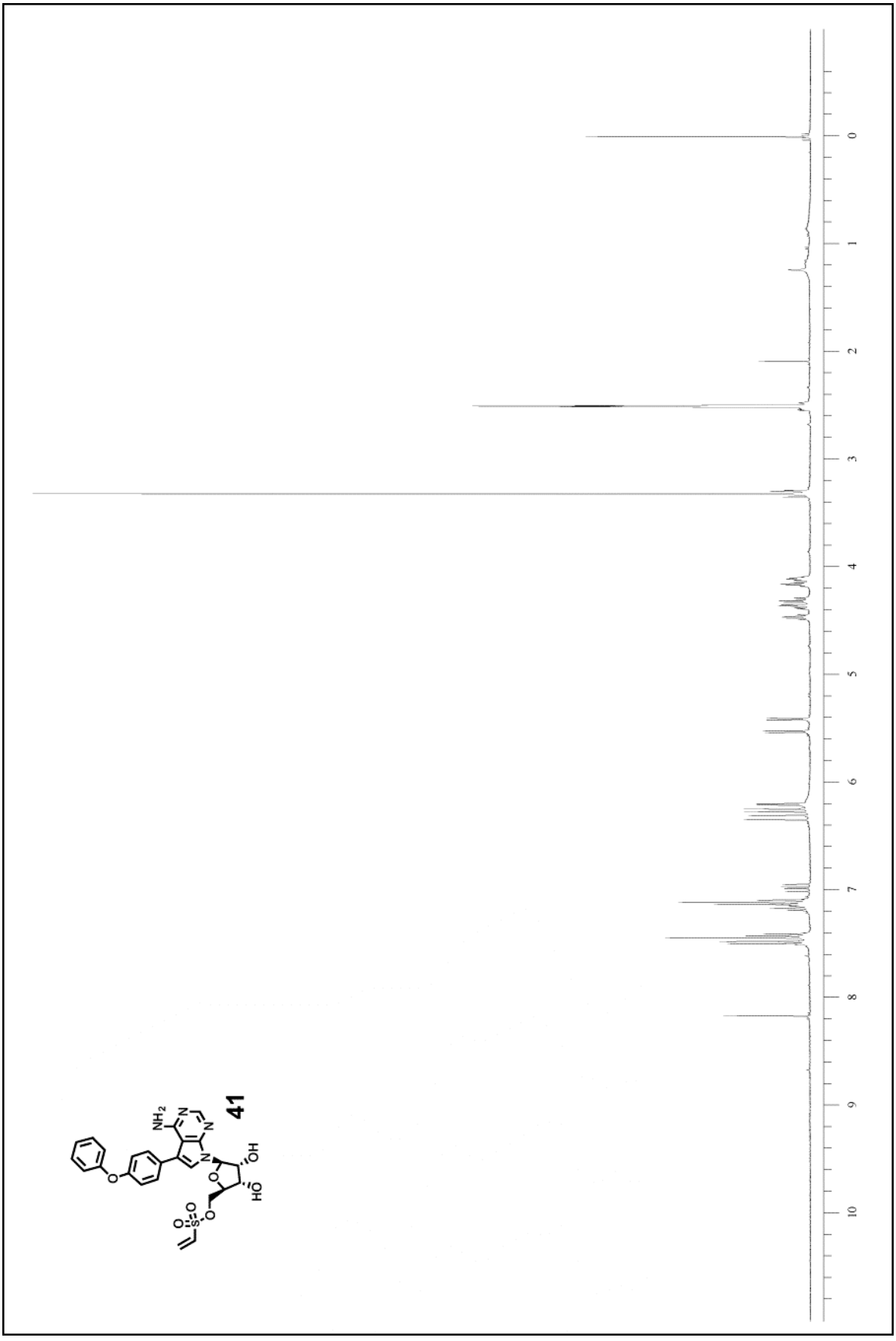

NBoc2
36

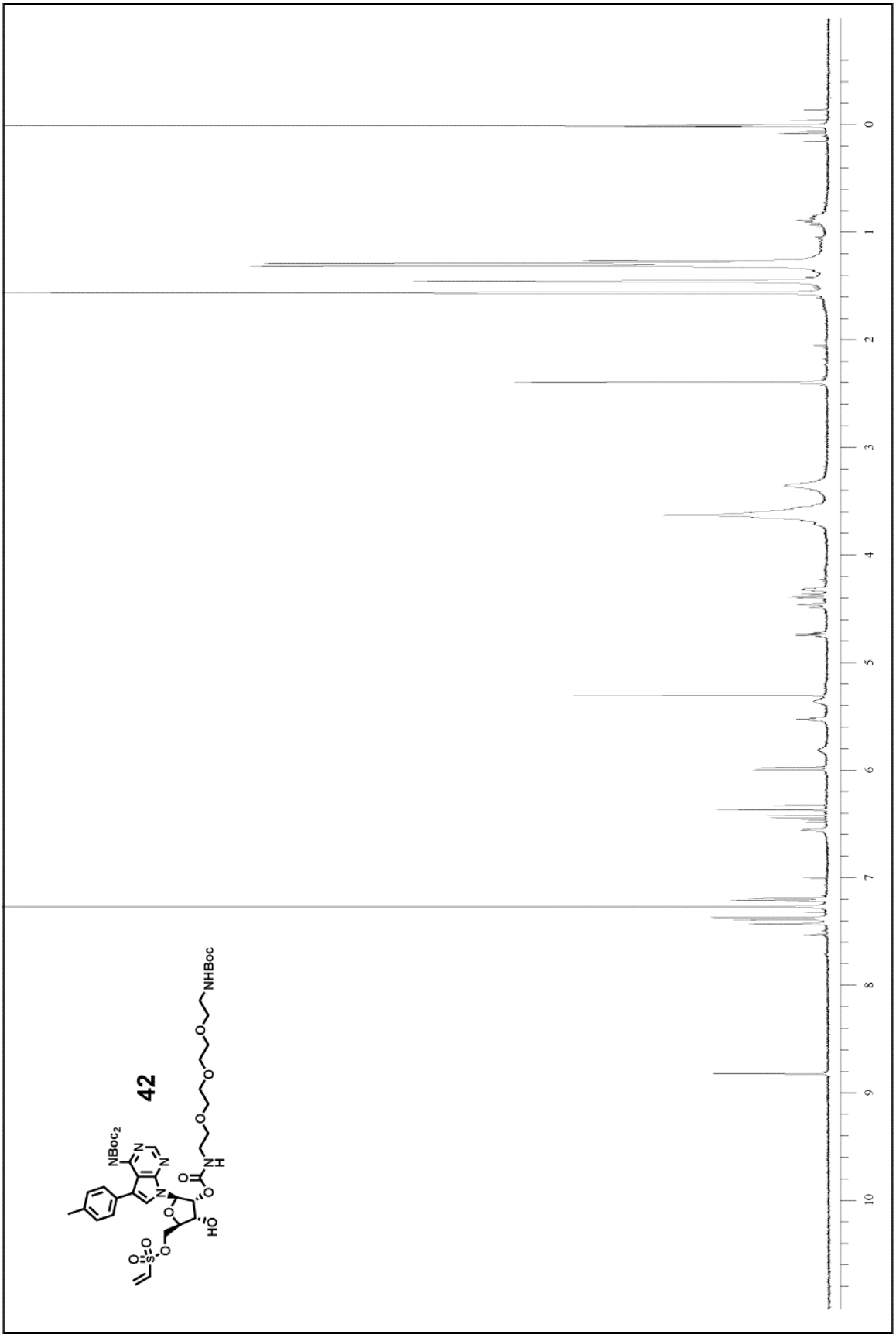

37

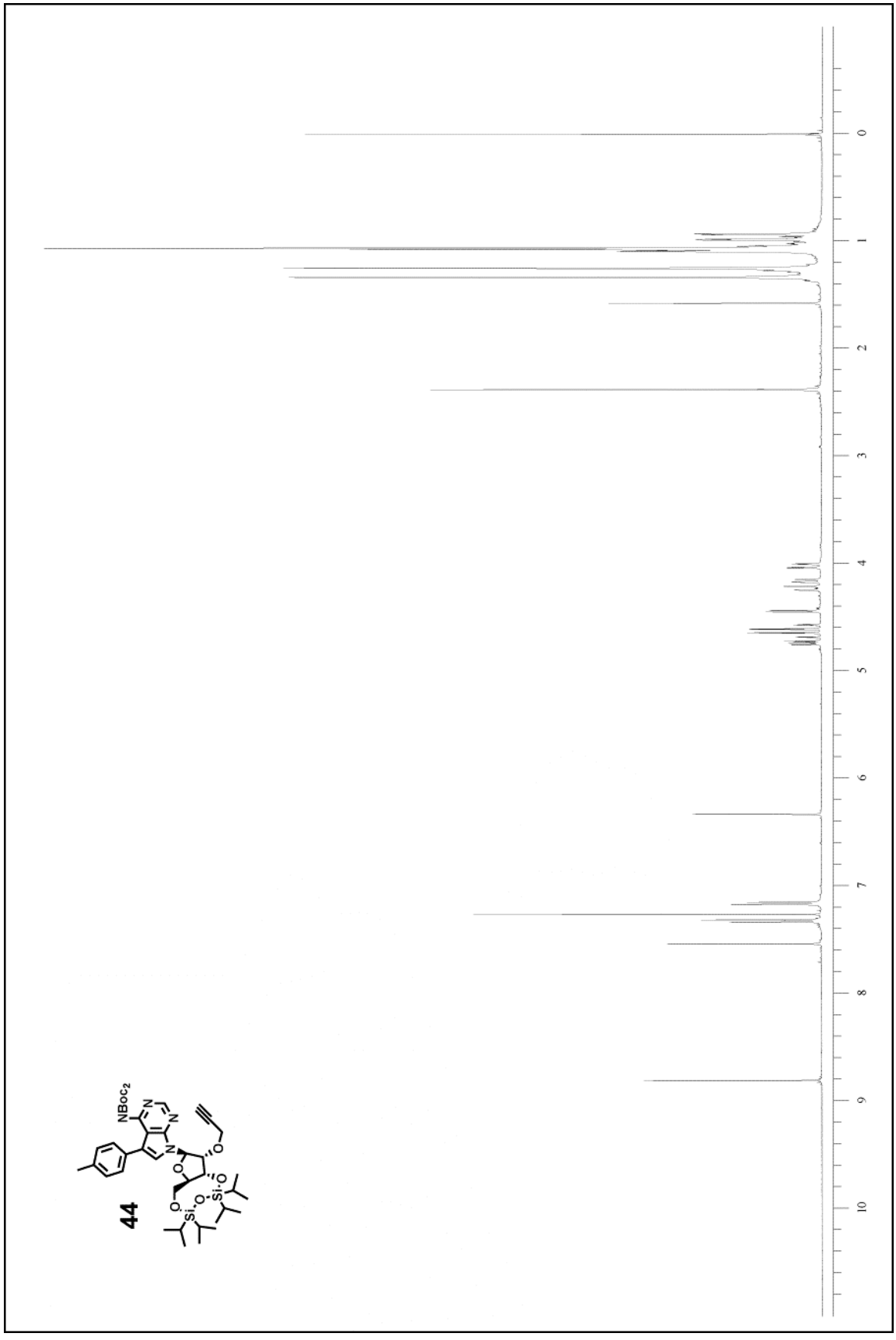

NBoc2
38

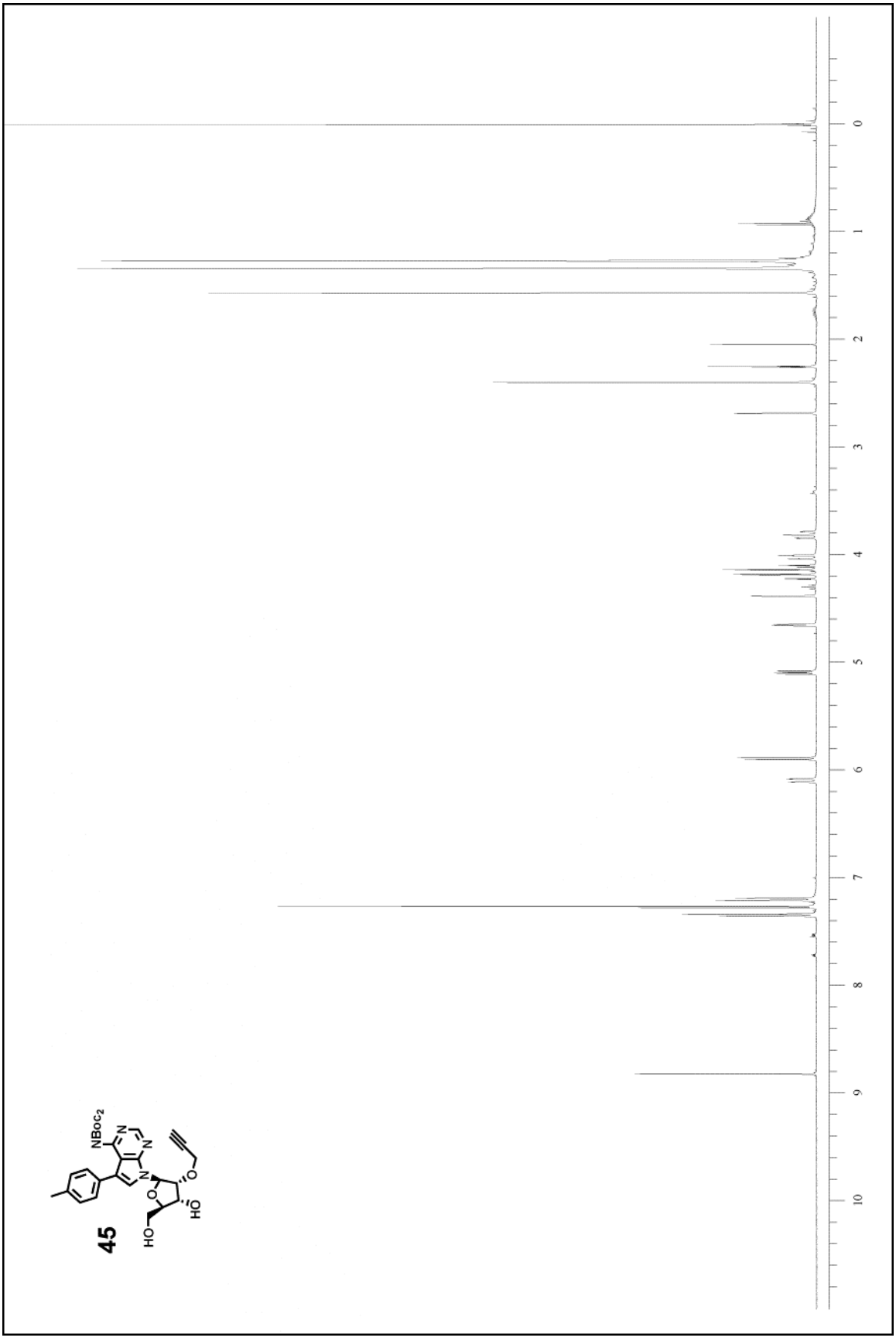

39
NH2
N
N
N
O
OH
HO
O
O
0
1
2
3
4
5
6
7
8
9
10

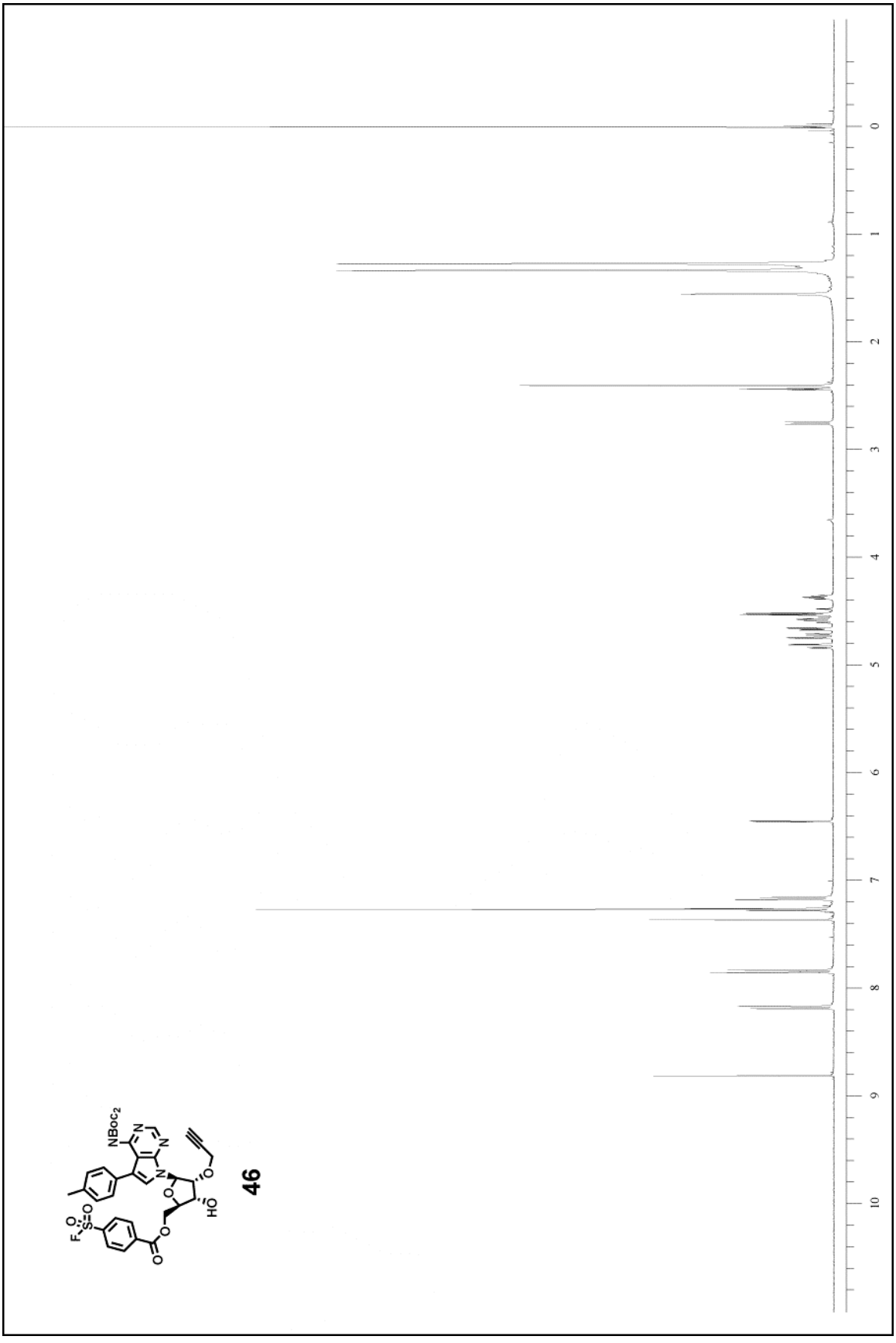

NBoc2
40

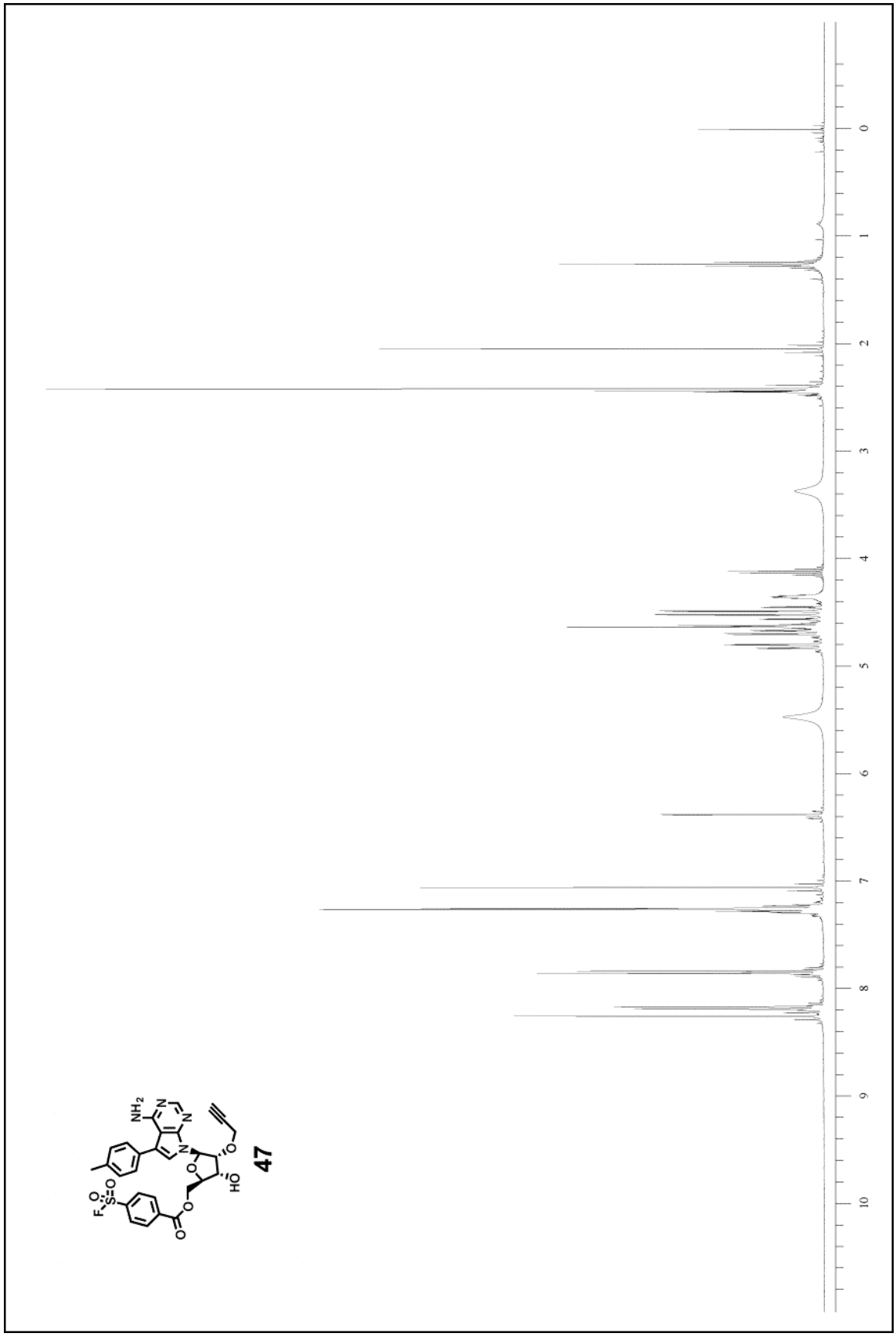

41
NH2
N
N
N
O
OH
HO
O
S
O
O
0
1
2
3
4
5
6
7
8
9
10

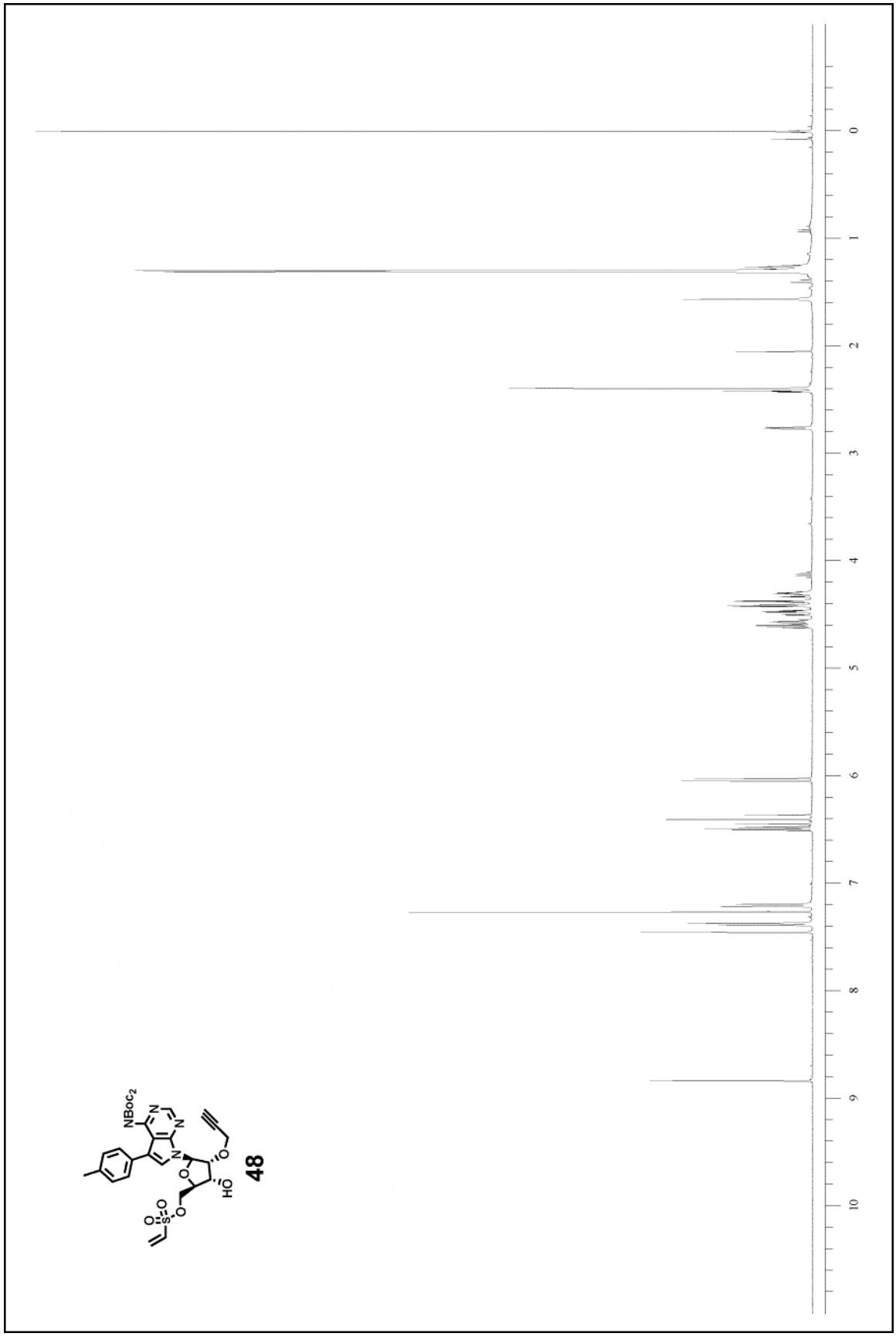

42
NBoc2
NHBoc
HO

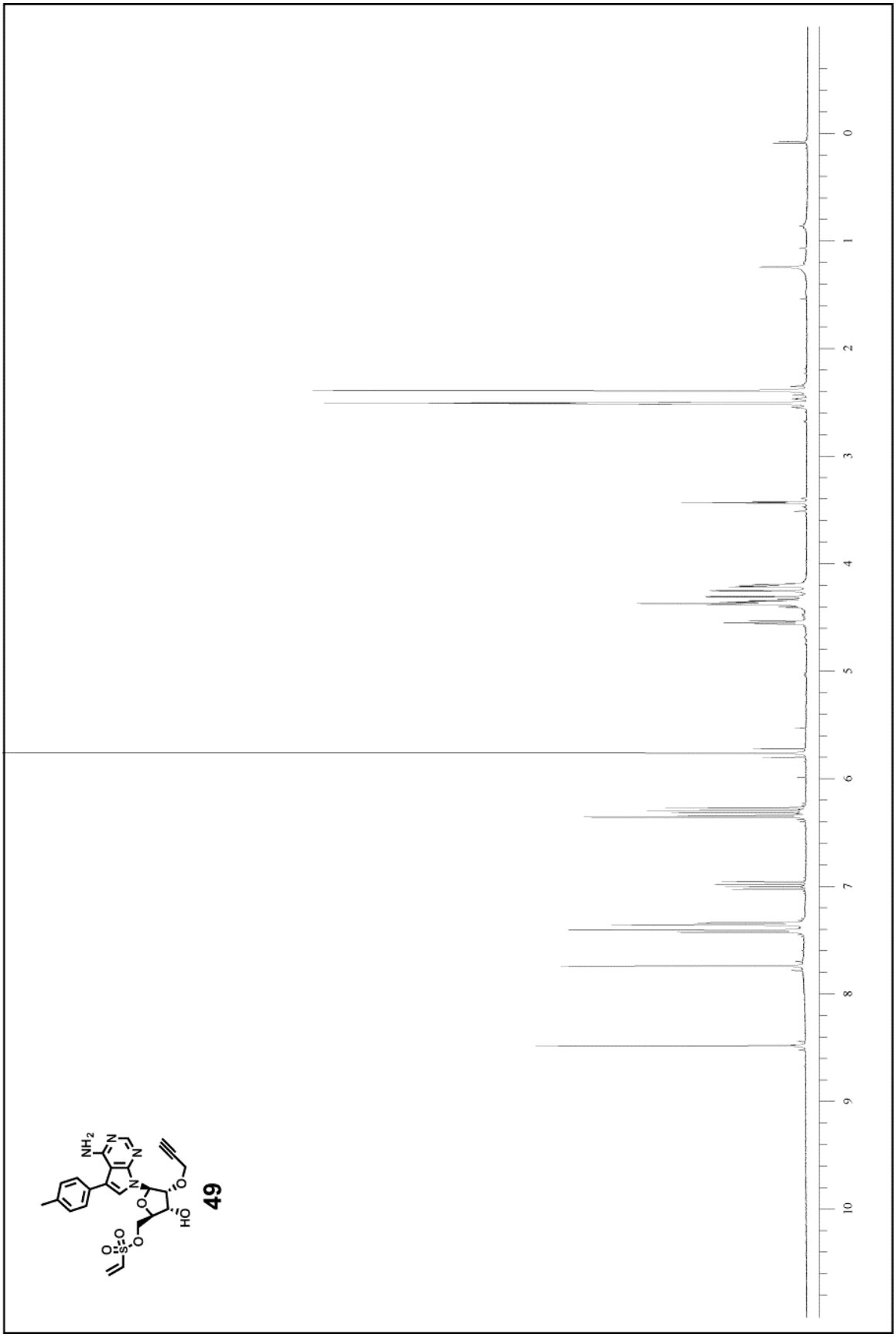

NBoc2
N
N
N
O
O
Si
O
Si
O
44
0
1
2
3
4
5
6
7
8
9
10

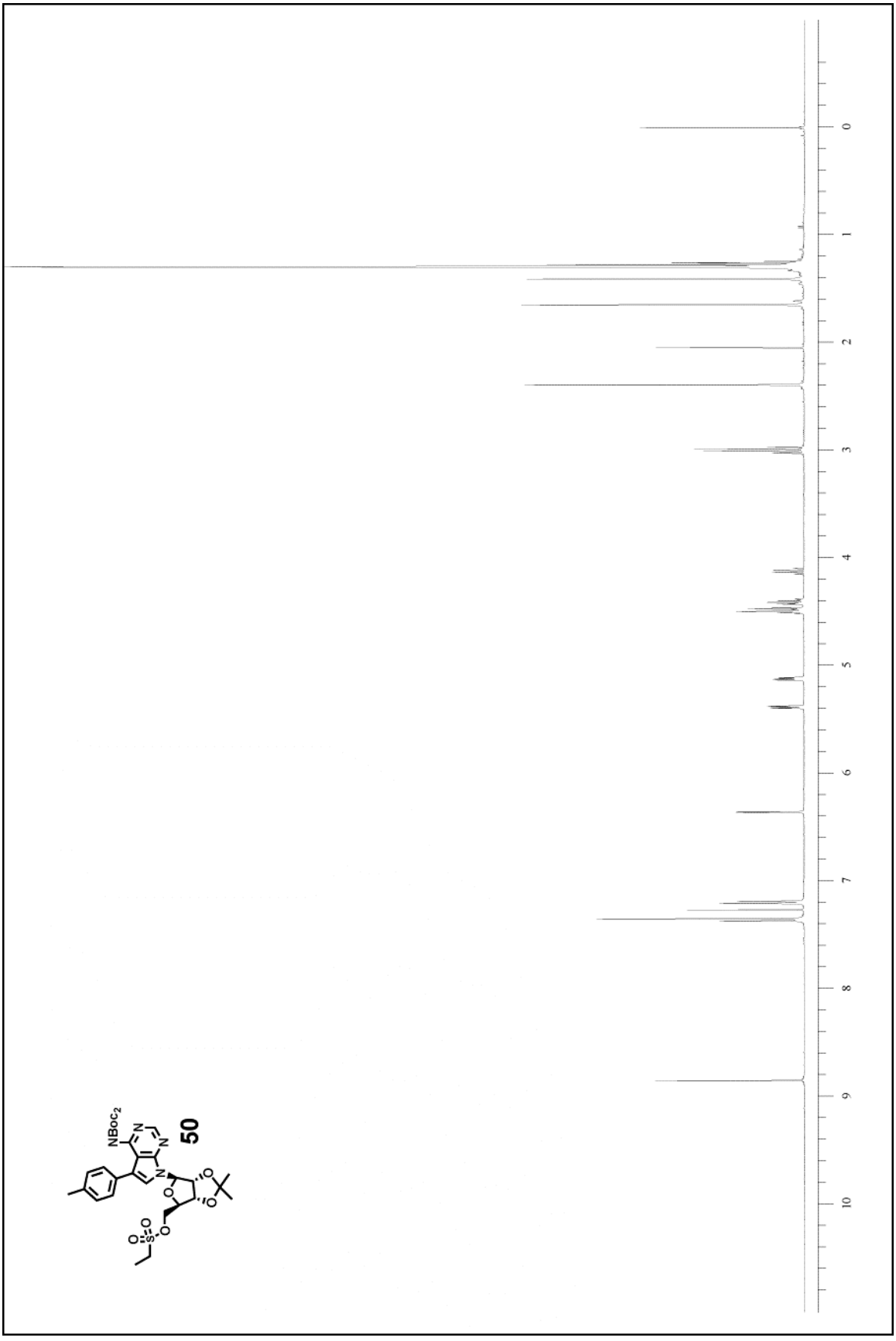

NBoc2
N
N
O
O
HO
HO
45

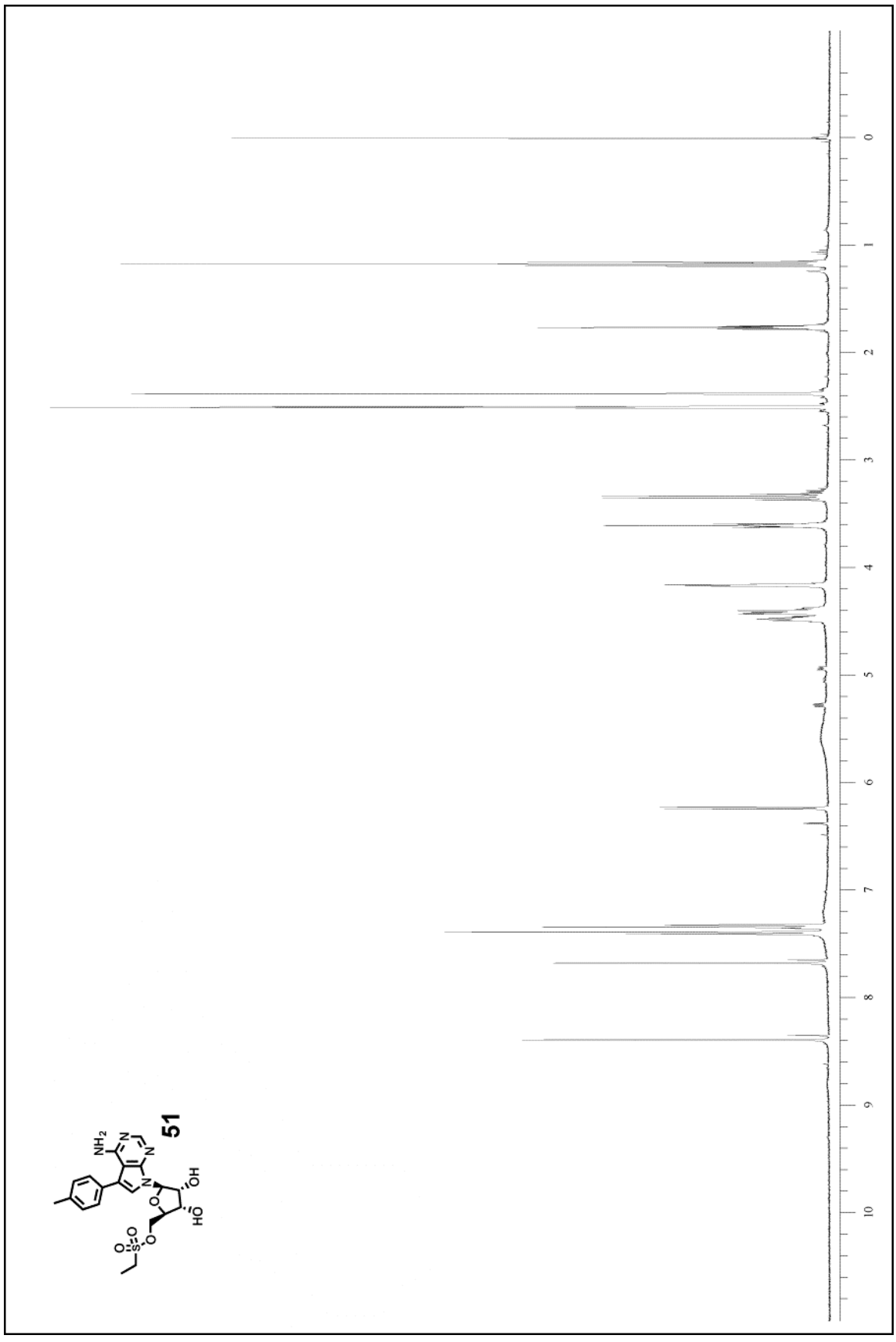

NBoc2
N
N
N
O
O
HO
F
S
O
O
O
46

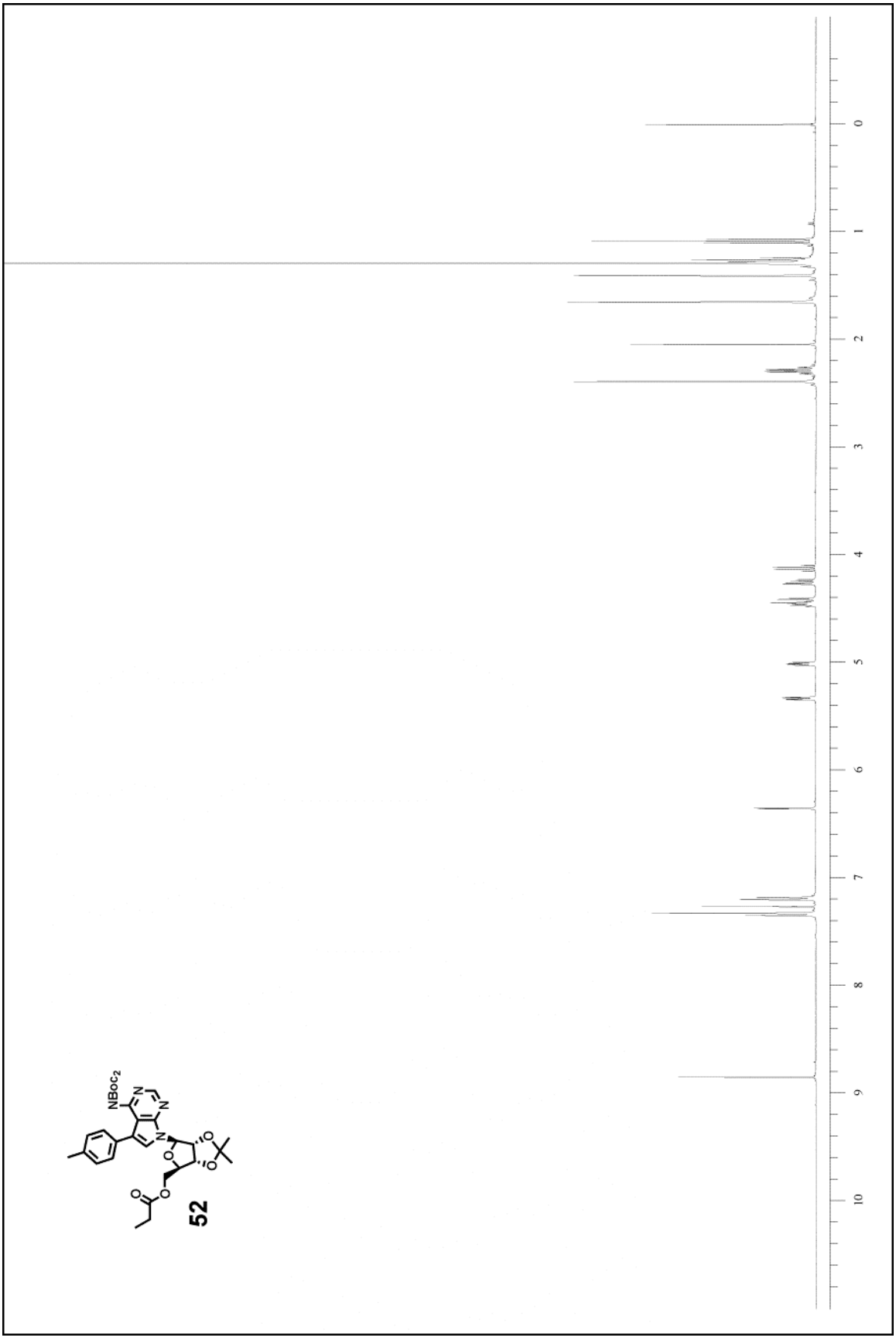

NH2
N
N
N
O
O
HO
F
S
O
O
O
O
47
0
1
2
3
4
5
6
7
8
9
10

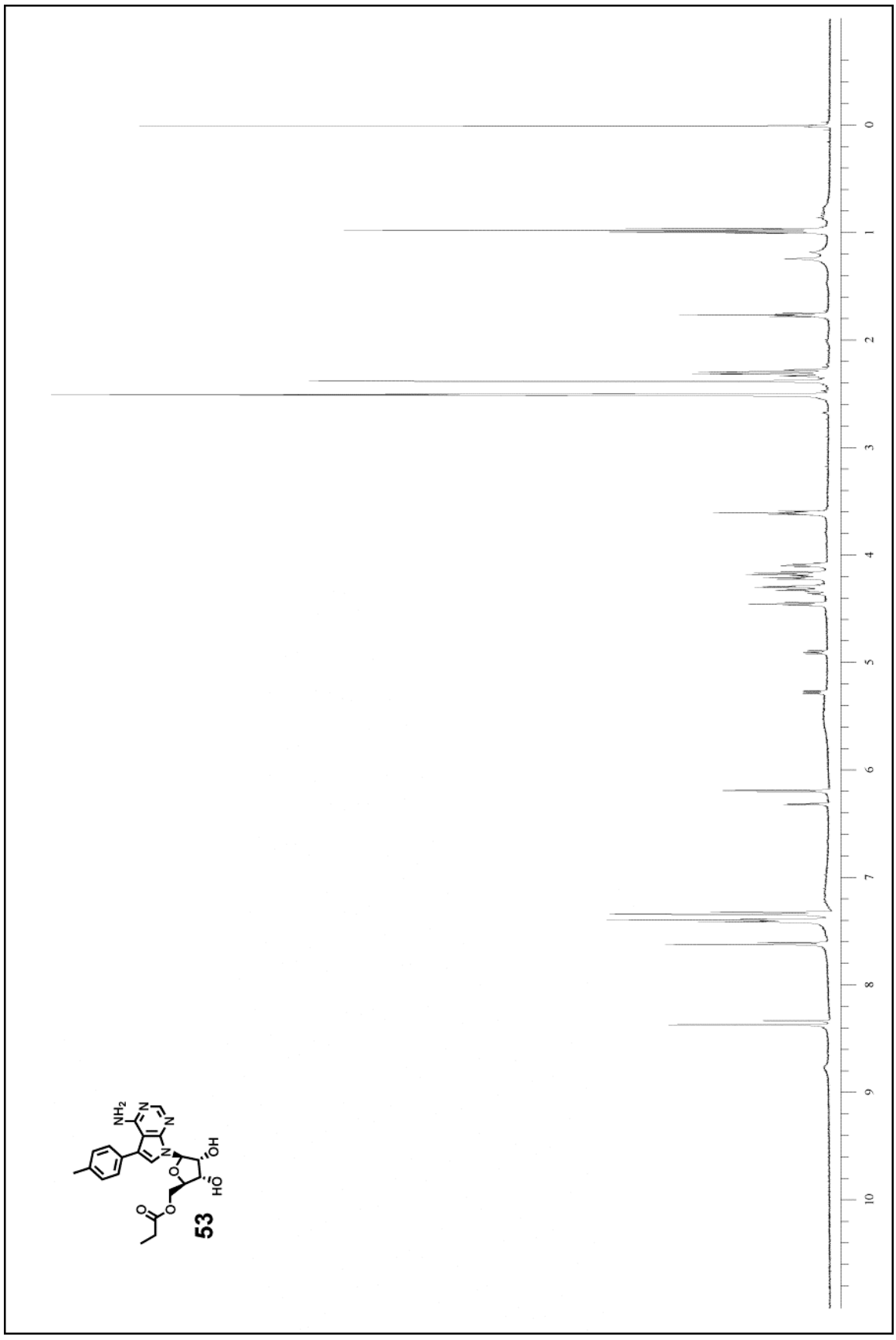

NBoc2
HO
48

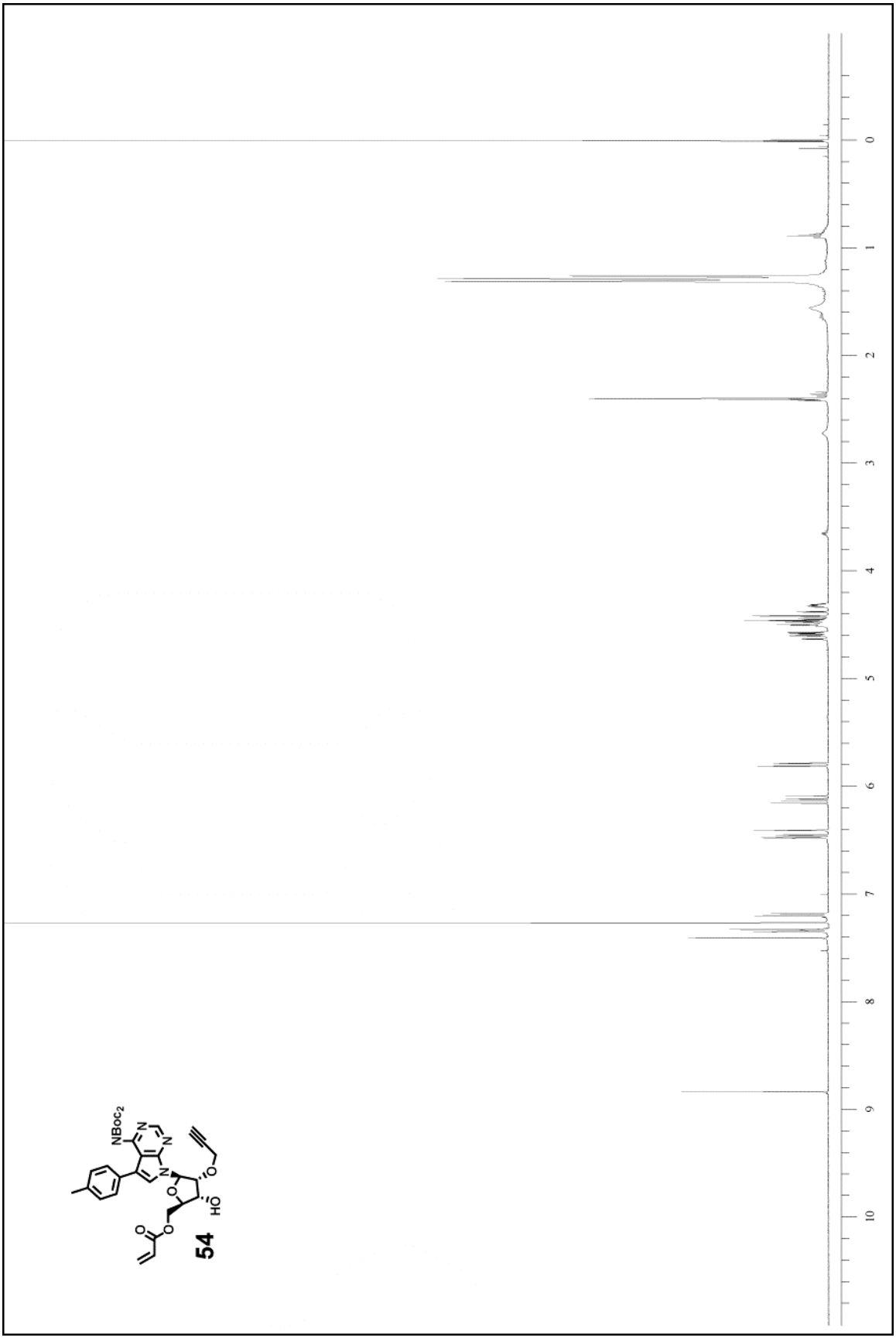

NH2
N
N
N
O
O
HO
O
S
O
O
49

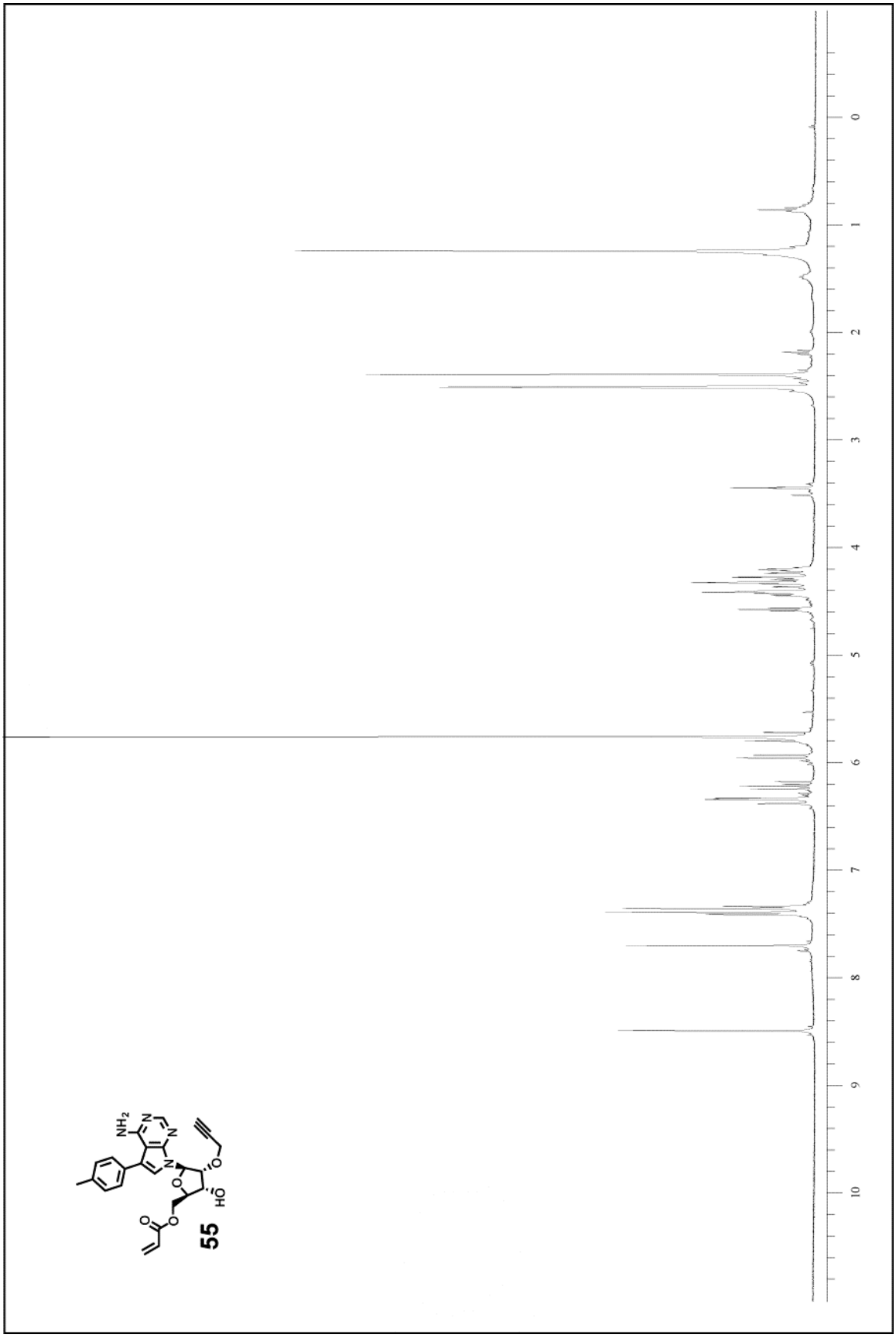

NBoc2
50


51
NH2
N
N
N
HO
OH
O
S
O
O


NBoc2
52

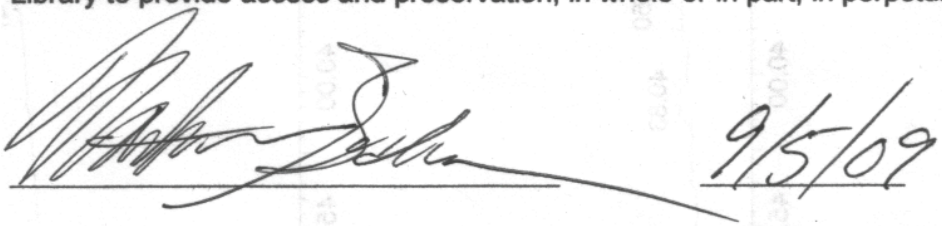

NH2
N
N
O
OH
HO
O
O
53


NBoc2
HO
54


NH2
N
N
N
O
HO
O
O
55
0
1
2
3
4
5
6
7
8
9
10

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

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



Author Signature

9/5/09

Date