Improvements to fluorescent affinity labels and the ribonuclease S...

Watkins, Rex Wayne

ProQuest Dissertations and Theses: 2010; ProQuest
pg. n/a

IMPROVEMENTS TO FLUORESCENT AFFINITY LABELS AND THE

RIBONUCLEASE S SYSTEM

AND

GENE EXPRESSION RESPONSE TO CLINICALLY RELEVANT

RIBONUCLEASES

by

Rex Wayne Watkins

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

Doctor of Philosophy

(Biochemistry)

at the

UNIVERSITY OF WISCONSIN–MADISON

2010
IMPROVEMENTS TO FLUORESCENT AFFINITY LABELS AND THE RIBONUCLEASE S SYSTEM AND GENE EXPRESSION RESPONSE TO CLINICALLY RELEVANT RIBONUCLEASES

submitted to the Graduate School of the University of Wisconsin-Madison in partial fulfillment of the requirements for the degree of Doctor of Philosophy

By

Rex Wayne Watkins

Date of final oral examination: 18 August 2010
Month and year degree to be awarded: August 2010

The dissertation is approved by the following members of the Final Oral Committee:
Ronald T. Raines, Professor, Biochemistry
Aseem Z. Ansari, Professor, Biochemistry
Laura L. Kiessling, Professor, Biochemistry
Douglas B. Weibel, Professor, Biochemistry
Jon S. Thorson, Professor, Pharmaceutical Sciences
IMPROVEMENTS TO FLUORESCENT AFFINITY LABELS AND THE RIBONUCLEASE S SYSTEM AND GENE EXPRESSION RESPONSE TO CLINICALLY RELEVANT RIBONUCLEASES

Rex Wayne Watkins

Under the supervision of Professor Ronald T. Raines

at the University of Wisconsin–Madison

An effective toolkit is essential in the quest to comprehend the remarkable complexities of nature and develop knowledge-based theories, models, and designs that lead to innovation and improved human conditions. Countless advances—including the development of novel chemical probes, the manipulation of model systems, and the rise of multiplex assays—have enabled important scientific discoveries. Many important scientific discoveries were made with the toolkit of the past. Increasingly, however, the frontiers of science are unexplored due to an inadequate armamentarium.

CHAPTER 1 serves as an introduction to the tools (affinity labels and protein-fragment complementation systems) and primary scientific question (cellular response to ribonuclease chemotherapeutics) addressed in this thesis.

CHAPTER 2 describes the synthesis and use of a novel fluorogenic affinity label to visualize a targeted protein in live cells. Unique features of the urea–rhodamine profluorophore facilitate rapid and specific labeling. Such fluorogenic affinity labels, which enlist catalysis by two cellular enzymes, may find utility in pulse-chase experiments, high-content screening, and other protocols.
In CHAPTER 3, we revisit the historic ribonuclease S protein-fragment complementation system. We avail ourselves of decades of experimental knowledge and technological improvements to overcome two limitations of the venerable system: the heterogeneous mixture of subtilisin-generated products and the loss of catalytic activity at low concentrations.

The gene-expression response of cancerous cells to a novel class of therapeutics is examined in CHAPTER 4. Certain pancreatic-type ribonucleases that evade the cytosolic ribonuclease inhibitor protein enter cancerous cells and destroy intracellular RNA, leading to cell death. The effects of two clinically investigated ribonucleases, onconase and QBI-139, on cancerous cells are observed by microarray analysis. The data show that the cellular response is likely due to RNA degradation. Interestingly, onconase and QBI-139 elicit remarkably different transcriptional responses, strongly suggesting that there are important differences in their mechanism of action. Furthermore, the most striking response to QBI-139 is an increase in polyadenylated histone mRNA perhaps due to an increase in trans-splicing.

In CHAPTER 5, I highlight several possible directions for future research. In particular I note several applications in which urea–rhodamine probes may be advantageous in labeling technologies. I also describe follow-up experiments from CHAPTER 4 that are designed help illuminate the connection between chemotherapeutic ribonucleases and chimeric RNA molecules.

Finally, in the appendix, I outline efforts to identify protein interaction partners with ribonuclease inhibitor protein.
Acknowledgements

Understandest thou what thou readest?

And he said, How can I, except some man should guide me? (Acts 8:30-31)

Mentoring is a brain to pick, an ear to listen, and a push in the right direction.

—John C. Crosby

I begin my acknowledgements with those who’ve been with me from the beginning—my parents—Wayne and Lynn Watkins. They have always been, and continue to be, sources of wisdom, love, encouragement, and support. They have fostered my love of learning and have demonstrated by example how to live life to the fullest.

I thank a myriad of neighbors, educators, friends, leaders, missionary companions, and family members who patiently endured me during my pre-college years and taught me invaluable skills such as optimism, hard work, diligence, and laughter.

My undergraduate advisor at Utah State University, Lisa Berreau, deserves special recognition. She is a model mentor and a wonderful friend. I count myself fortunate to begin laboratory research with someone of her enthusiasm, patience, and talent.

It has been a privilege to work in the laboratory of Ron Raines. I have benefitted greatly from his insight, experience, guidance, and professionalism. Additionally, he has assembled a fantastic scientific team. These fellow Raines lab members have become good friends and indispensable mentors. Greg Ellis deserves special recognition for tolerating me as a “baymate” on days when my experiments failed and my irritability increased.
My thesis committee—Aseem Ansari, Laura Kiessling, Doug Weibel, and Jon Thorson—has provided constructive criticism, new ideas, and encouragement.

The best science generally results from a team effort, and I certainly have had excellent collaborators and colleagues. Luke Lavis developed the synthetic methodology for the profluorescent ligand described in CHAPTER 2. He also gave excellent experimental guidance. Randy Dimond and Georgyi Los from Promega provided additional help on the same project. Uli Arnold pioneered the introduction of an enterokinase cleavage site for RNase S (CHAPTER 3). Darrell McCaslin and Gary Case assisted with instrumentation. Additionally, members from both Marv Wickens’ and Sam Butcher’s labs have provided technical assistance not available in our own.

Echelon Biosciences Inc. sponsored a ten-week internship for me in Salt Lake City, Utah. I greatly appreciate their kindness and the opportunity to experience research in a commercial setting.

I thank the taxpayers of the United States. I am well aware that significant federal funding has enabled both my research and stipend.

Last, and most importantly, I thank Natalie. Together we’ve laughed and cried. We’ve experienced profound joy and irksome uncertainty. She has been my best friend and confidant, and I’m lucky to be married to her. While in graduate school, our family has expanded to include two delightful little girls, Emelyn and Camilla. Natalie has sacrificed much for our family and my career. For this I will be forever grateful. I look forward to the road ahead with my best friend.
# Table of Contents

Abstract ................................................................................................................. i

Acknowledgements ............................................................................................... iii

List of Figures ....................................................................................................... x

List of Schemes and Tables .................................................................................. xii

List of Abbreviations ............................................................................................ xiii

CHAPTER 1

Introduction: Improvements to fluorescent affinity labels and the ribonuclease S system and gene expression response to clinically relevant ribonucleases ...................... 1

1.1 Overview ......................................................................................................... 2

1.2 Affinity labeling ............................................................................................... 3

1.3 Ribonuclease A as a model protein ................................................................... 5

1.4 Ribonucleases as chemotherapeutics ............................................................. 9

CHAPTER 2

Fluorogenic affinity label for the facile, rapid imaging of proteins in live cells .......... 25

2.1 Abstract ......................................................................................................... 26

2.2 Introduction ................................................................................................... 26

2.3 Results and discussion .................................................................................. 28

2.3.1 Design and synthesis of probe 1 ................................................................. 28

2.3.2 Cell imaging with probe 1 .......................................................................... 29

2.3.3 Probe lipophilicity ...................................................................................... 30

2.3.4 Recapitulation ............................................................................................ 32
CHAPTER 4

Gene expression response of cancerous cells to clinically relevant ribonucleases ....... 89

4.1 Abstract ............................................................................................................. 90
4.2 Introduction ........................................................................................................ 90
4.3 Results and discussion ..................................................................................... 92
   4.3.1 Primary cellular response is due to degradation of intracellular RNA .... 92
   4.3.2 Cellular response to onconase .................................................................... 93
   4.3.3 Histone mRNAs and other differences between onconase and QBI-139-
        treated cells .................................................................................................. 94
   4.3.4 Confirmation by quantitative PCR .............................................................. 95
   4.3.5 3’ rapid amplification of cDNA ends ......................................................... 95
   4.3.6 Interpretation .............................................................................................. 96
4.4 Conclusions ....................................................................................................... 97
4.5 Experimental ..................................................................................................... 98
   4.5.1 Materials ..................................................................................................... 98
   4.5.2 Instrumentation .......................................................................................... 99
   4.5.3 Ribonuclease production ............................................................................. 99
   4.5.4 Assay of cytotoxicity .................................................................................. 100
   4.5.5 Cell propagation ......................................................................................... 100
   4.5.6 Microarray analysis .................................................................................... 100
4.5.6.1 Isolation of RNA from cells exposed to ribonucleases.................101
4.5.6.2 First-strand cDNA synthesis.............................................101
4.5.6.3 Second-strand cDNA synthesis..........................................101
4.5.6.4 DNA labeling, hybridization and data extraction.......................102
4.5.6.5 Gene ontology and pathway analysis...................................102
4.5.7 Quantitative real-time PCR..................................................102
  4.5.7.1 RNA isolation ..................................................................103
  4.5.7.2 cDNA synthesis for quantitative real-time PCR.......................103
  4.5.7.3 Quantitative real-time PCR..............................................103
4.5.8 3' Rapid amplification of cDNA ends.......................................104

CHAPTER 5

Future directions............................................................................125
  5.1 Additional utility of membrane-permeant affinity labels................126
  5.2 Further investigations of trans-splicing..................................126

APPENDIX

Search for protein interaction partners of ribonuclease inhibitor protein......129
  A.1 Abstract...............................................................................130
  A.2 Introduction..........................................................................130
  A.3 Results and discussion..........................................................132
    A.3.1 Crosslinking of RI-interaction partners with a photoswitchable leucine
         analogue...............................................................................132
    A.3.2 Yeast two-hybrid screen with RI as bait...............................132
  A.4 Experimental..........................................................................133
A.4.1 Crosslinking with photo-leucine .........................................................133
A.4.2 Yeast two-hybrid screen .....................................................................134
A.4.3 Attempts to confirm GABARAPL2-Rl interaction ..............................135

REFERENCES ......................................................................................................139
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Mechanism of haloalkane dehalogenase</td>
<td>15</td>
</tr>
<tr>
<td>1.2</td>
<td>Proposed mechanism of RNase A</td>
<td>17</td>
</tr>
<tr>
<td>1.3</td>
<td>Transition state of RNase A</td>
<td>19</td>
</tr>
<tr>
<td>1.4</td>
<td>Irwin Geis’ painting of RNase S</td>
<td>21</td>
</tr>
<tr>
<td>1.5</td>
<td>Routing of cytotoxic ribonucleases</td>
<td>23</td>
</tr>
<tr>
<td>2.1</td>
<td>Fluorogenic labels of haloalkane dehalogenase</td>
<td>54</td>
</tr>
<tr>
<td>2.2</td>
<td>Live-cell labeling of HD variants</td>
<td>56</td>
</tr>
<tr>
<td>2.3</td>
<td>Lactone-quinoid equilibrium of 1, 2, and 3</td>
<td>58</td>
</tr>
<tr>
<td>2.4</td>
<td>Effect of the dielectric constant on the lactone-quinoid equilibrium</td>
<td>60</td>
</tr>
<tr>
<td>3.1</td>
<td>Notional structure of RNase–S</td>
<td>77</td>
</tr>
<tr>
<td>3.2</td>
<td>Protease digest of ribonucleases</td>
<td>79</td>
</tr>
<tr>
<td>3.3</td>
<td>HPLC separation of S-peptide and S-protein</td>
<td>81</td>
</tr>
<tr>
<td>3.4</td>
<td>SDS-PAGE of RNase–S</td>
<td>83</td>
</tr>
<tr>
<td>3.5</td>
<td>Zymogram of RNase–S</td>
<td>85</td>
</tr>
<tr>
<td>3.6</td>
<td>Initial velocities of RNase–S</td>
<td>87</td>
</tr>
<tr>
<td>4.1</td>
<td>Crystalline structures of onconase and RNase 1</td>
<td>107</td>
</tr>
<tr>
<td>4.2</td>
<td>Mechanism of ribonuclease cytotoxicity</td>
<td>109</td>
</tr>
<tr>
<td>4.3</td>
<td>Cytotoxicity of ribonucleases</td>
<td>111</td>
</tr>
<tr>
<td>4.4</td>
<td>Volcano plots for ribonuclease-treated cells</td>
<td>113</td>
</tr>
<tr>
<td>4.5</td>
<td>Protein interaction network of onconase-regulated genes</td>
<td>115</td>
</tr>
<tr>
<td>4.6</td>
<td>Protein interaction network of QBI-139-regulated genes</td>
<td>117</td>
</tr>
</tbody>
</table>
Figure 4.7: Volcano plots highlighting histone mRNA probesets ..........................119
Figure 4.8: Quantitative PCR from cDNA constructed using random and oligo(dT) primers .................................................................121
Figure 4.9: 3' Ends of polyadenylated histone sequences ..............................123
Figure A.1: Yeast two-hybrid screen with RI as bait ..................................137
List of Schemes and Tables

Table 2.1: Calculated logD values .................................................................52
Scheme 2.1: Synthesis of probe 1 .................................................................53
Table 4.1: Genes differentially regulated by onconase and QBI-139 ..............106
Table A.1: Putative protein interaction partners of ribonuclease inhibitor protein.....136
**List of Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3AT</td>
<td>3-Amino-1,2,4-triazole</td>
</tr>
<tr>
<td>Ala (A)</td>
<td>alanine</td>
</tr>
<tr>
<td>ANOVA</td>
<td>analysis of variance</td>
</tr>
<tr>
<td>Asp (D)</td>
<td>aspartic acid</td>
</tr>
<tr>
<td>Arg (R)</td>
<td>arginine</td>
</tr>
<tr>
<td>Boc (t-Boc)</td>
<td>\textit{tert}-butoxycarbonyl</td>
</tr>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>Cys (C)</td>
<td>cysteine</td>
</tr>
<tr>
<td>DEPC</td>
<td>diethylpyrocarbonate</td>
</tr>
<tr>
<td>DIPEA</td>
<td>diisopropylethylamine (Hünig’s base)</td>
</tr>
<tr>
<td>DMF</td>
<td>dimethylformamide</td>
</tr>
<tr>
<td>DMSO</td>
<td>dimethylsulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>dNTP</td>
<td>deoxyribonucleotide triphosphate</td>
</tr>
<tr>
<td>DTNB</td>
<td>5,5’-dithiobis(2-nitrobenzoic acid)</td>
</tr>
<tr>
<td>DTT</td>
<td>dithiothreitol</td>
</tr>
<tr>
<td>(\varepsilon)</td>
<td>extinction coefficient or dielectric constant</td>
</tr>
<tr>
<td>EDC</td>
<td>1-ethyl-3-(3-dimethylaminopropyl)carbodiimide hydrochloride</td>
</tr>
<tr>
<td>EDTA</td>
<td>ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>ESI</td>
<td>electrospray ionization</td>
</tr>
<tr>
<td>EtOAC</td>
<td>ethyl acetate</td>
</tr>
<tr>
<td>FBS</td>
<td>fetal bovine serum</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>FPLC</td>
<td>fast performance liquid chromatography</td>
</tr>
<tr>
<td>h</td>
<td>hour</td>
</tr>
<tr>
<td>HCl</td>
<td>hydrochloric acid</td>
</tr>
<tr>
<td>HD</td>
<td>haloalkane dehalogenase</td>
</tr>
<tr>
<td>His (H)</td>
<td>histidine</td>
</tr>
<tr>
<td>HPLC</td>
<td>high performance (pressure) liquid chromatography</td>
</tr>
<tr>
<td>HRMS</td>
<td>high-resolution mass spectrometry</td>
</tr>
<tr>
<td>IC₅₀</td>
<td>half maximal inhibitory concentration</td>
</tr>
<tr>
<td>IPTG</td>
<td>isopropyl-1-thio-β-D-galactopyranoside</td>
</tr>
<tr>
<td>kₜₐₜ</td>
<td>first-order enzymatic rate constant</td>
</tr>
<tr>
<td>Kₐ</td>
<td>equilibrium dissociation constant</td>
</tr>
<tr>
<td>kDa</td>
<td>kilodalton</td>
</tr>
<tr>
<td>Kₜ</td>
<td>Michaelis constant</td>
</tr>
<tr>
<td>λₘₐₓ</td>
<td>maximum wavelength</td>
</tr>
<tr>
<td>LB</td>
<td>Luria–Bertani medium</td>
</tr>
<tr>
<td>logD</td>
<td>log of the distribution coefficient</td>
</tr>
<tr>
<td>Lys (K)</td>
<td>lysine</td>
</tr>
<tr>
<td>MALDI–TOF</td>
<td>matrix-assisted laser desorption/ionization time-of-flight</td>
</tr>
<tr>
<td>min</td>
<td>minute</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>NaCl</td>
<td>sodium chloride</td>
</tr>
<tr>
<td>NHS</td>
<td>N-hydroxysuccinimide</td>
</tr>
<tr>
<td>NLS</td>
<td>nuclear localization signal</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>NMR</td>
<td>nuclear magnetic resonance</td>
</tr>
<tr>
<td>NTB</td>
<td>2-nitro-5-thiobenzoate</td>
</tr>
<tr>
<td>OD</td>
<td>optical density</td>
</tr>
<tr>
<td>PAGE</td>
<td>polyacrylamide gel electrophoresis</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate-buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>PDB</td>
<td>protein data bank</td>
</tr>
<tr>
<td>pI</td>
<td>isoelectric point</td>
</tr>
<tr>
<td>pKa</td>
<td>negative log of the acid dissociation constant</td>
</tr>
<tr>
<td>qPCR</td>
<td>quantitative real-time PCR</td>
</tr>
<tr>
<td>RI</td>
<td>ribonuclease inhibitor</td>
</tr>
<tr>
<td>RMA</td>
<td>robust multichip average</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RNase A</td>
<td>bovine pancreatic ribonuclease</td>
</tr>
<tr>
<td>RNase 1</td>
<td>human pancreatic ribonuclease</td>
</tr>
<tr>
<td>s</td>
<td>second</td>
</tr>
<tr>
<td>SDS–PAGE</td>
<td>sodium dodecyl sulfate poly(acrylamide) gel electrophoresis</td>
</tr>
<tr>
<td>TFA</td>
<td>trifluoroacetic acid</td>
</tr>
<tr>
<td>THF</td>
<td>tetrahydrofuran</td>
</tr>
<tr>
<td>T&lt;sub&gt;m&lt;/sub&gt;</td>
<td>temperature at the midpoint of the denaturation curve</td>
</tr>
<tr>
<td>Tris</td>
<td>2-amino-2-(hydroxymethyl)-1,3-propanediol</td>
</tr>
<tr>
<td>tRNA</td>
<td>transfer RNA</td>
</tr>
<tr>
<td>UV</td>
<td>ultraviolet</td>
</tr>
</tbody>
</table>
Y2H  

yeast two-hybrid

Z  

net molecular charge (Arg + Lys – Asp – Glu for proteins)
CHAPTER 1

Introduction:

Improvements to Fluorescent Affinity Labels

and the Ribonuclease S System

and

Gene Expression Response to Clinically Relevant Ribonucleases
1.1 Overview

Life on earth is astonishingly complex. For instance, a single human cell contains approximately 10,000 unique proteins (Morón and Devi, 2007) and 300,000 mRNA molecules (Velculescu et al., 1999). Making sense of this complexity requires suitable tools.

In Chapter 2, I present a urea–rhodamine affinity label as an important tool for the facile, rapid imaging of proteins in live cells. This fluorogenic label, together with an enzymic fusion tag, enables the specific labeling of a single protein from among the ~10,000 unique proteins in a cell. Detection instruments, such as confocal microscopy enable the investigation of protein localization, turnover, and function in a cellular context. We envision that this probe will be useful in pulse-chase experiments, and find utility in high throughput technologies such as high-content screening.

Chapter 3 describes improvements to a historic protein-fragment complementation system. A detailed understanding of protein structure and function requires an accessible, easily manipulable system that allows for perturbations that are not very accessible using the tools of molecular biology. The RNase S protein-fragment complementation system has for decades been the model system of choice in protein chemistry. Our improvements to this protein-fragment complementation system simplify the isolation of RNase S components and overcome difficulties inherent in studying a non-covalent complex.

Lastly, we use microarrays and amplification techniques as tools to investigate the response of cancerous cells to ribonucleases that show promise as cancer therapeutics. This multiplex experiment provides an important glimpse of the ribonuclease-induced changes that occur to the ~300,000 mRNA molecules within a cell.
In this chapter, I summarize the background that led me to each of these initiatives.

1.2 Affinity Labeling

In the 1990's, Douglas Prasher and Martin Chalfie cloned green fluorescent protein (Prasher et al., 1992) from the jellyfish *Aequorea victoria* and demonstrated its use in fusion protein systems (Chalfie et al., 1994). This work precipitated a rapid rise in the use of autofluorescent proteins that has revolutionized cell imaging. Common autofluorescent proteins have several advantageous features—genetic encodability for the facile generation of fusion proteins, relatively small size, high stability, and very low phototoxicity (Chalfie and Kain, 2006). These features have led to widespread adoption of autofluorescent proteins as gene reporters, and in determining the subcellular localization, dynamics, and chemical environment of many proteins (Chalfie and Kain, 2006).

Still, the use of autofluorescent proteins is limited in several respects. First, autofluorescent proteins are perpetually "on". The inability to initiate fluorescence with acute temporal control can obscure important information. Second, a single genetic construct generates only one fluorescent output. Cloning or mutagenesis is required to overcome this "one clone—one color" limitation. Furthermore, the spectral properties of autofluorescent proteins are limited, and many autofluorescent proteins suffer from other complications (Shaner et al., 2005).

More recently, a variety of site-specific labeling methods have been developed for fusion tags. For instance, Roger Tsien's group has developed a system in which biarsenical membrane-permeable non-fluorescent dyes bind to and create a fluorescent
complex with proteins that have been genetically fused to the CCPGCC tetracysteine motif (Griffin et al., 2000; Machleidt et al., 2006).

Still others, noting the high reactivity that has evolved in enzyme active sites, have modified two-step enzyme-catalyzed reactions such that turnover is prevented due to the formation of a non-labile covalent enzyme-substrate adduct (Johnsson and Johnsson, 2007; Los and Wood, 2007). These affinity labels react rapidly and specifically with the target proteins. By appending variously ligands to the core affinity label, a single genetic construct can be used to attach the fusion protein to a solid surface or conjugate it to affinity handles or fluorescent probes.

Haloalkane dehalogenase (HD; EC 3.8.1.5) is perhaps the most oft-used enzymic fusion tag. The enzyme catalyzes the hydrolysis of haloalkanes via a covalent enzyme-substrate intermediate that undergoes hydrolysis (Figure 1.1). An HD variant that cannot perform the second, hydrolysis step (HaloTag®, Promega) reacts rapidly and specifically with a wide variety of haloalkanes to form a covalent adduct (Los et al., 2008). Suitable haloalkanes include affinity handles (e.g., biotin), solid supports, and fluorophores with varied spectral and physiochemical properties.

One important advantage of fluorescent affinity labeling over autofluorescent proteins is the ability to control spatially and temporally the addition of the affinity label, and thus the onset of fluorescence. This attribute is crucial for pulse-chase experiments in which temporally disparate pools of protein must be differentially labeled.

Rapid labeling permits researchers to study processes that occur on short timescales. For many labeling technologies, the chemistry between the probe and the target are primarily responsible for slow labeling. Nonetheless, for HD (with a second-order rate
constant of approximately $2.7 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ (Los et al., 2008), the ability of the probe to
cross the membrane is often the rate-limiting step.

Many extant probes lack the membrane permeance needed for rapid labeling. In
Chapter 2, we describe a urea–rhodamine fluorogenic affinity label with desirable
membrane permeance that enables the rapid, facile labeling of fusion proteins in live
cells.

We envision that this probe will provide increased temporal resolution in pulse-chase
experiments and may find use in high throughput technologies, such as high-content
screening.

1.3 Ribonuclease A as a model protein

Bovine pancreatic ribonuclease (RNase A, EC 3.1.27.5) is perhaps the most studied
enzyme of the twentieth century. Discovered in 1920 (Jones, 1920), and first isolated in
1938 (Dubos and Thompson, 1938), RNase A attracted the attention of early protein
scientists. During a hiatus of ribonuclease research due to World War II, Edwin J. Cohn
of Harvard University convinced the research group at Armour, Inc. (a slaughterhouse
and meatpacking company with an abundant supply of bovine pancreata) that a large
quantity of purified protein was necessary for detailed study (Richards, 1997). In the
early 1950’s, using Moses Kunitz’s procedure (Kunitz, 1939), Armour purified over 1 kg
of RNase A and distributed it freely to interested scientists throughout the world.

The accessibility, purity, and remarkable stability of RNase A were a great boon to
protein chemists and spawned important research in enzymology, protein structure,
protein stability, and protein chemistry. In 1972, the Nobel Prize selection committee
recognized this historic work by awarding the Nobel Prize in chemistry to Chris Anfinsen (Anfinsen, 1973), Stanford Moore (Moore and Stein, 1973), and William Stein (Moore and Stein, 1973) for work on RNase A. Later, in 1984, another Nobel Prize would be awarded to Bruce Merrifield for work that also featured RNase A prominently (Merrifield, 1984).

RNase A catalyzes the cleavage of the P–O₅' bond of RNA, preferentially cleaving RNA on the 3'-side of pyrimidine residues (Raines, 1998). The rate of cleavage of the dinucleotide UpA is maximal at pH 6.0, where $k_{\text{cat}} = 1.4 \times 10^3$ s⁻¹ and $k_{\text{cat}}/K_{\text{M}} = 2.3 \times 10^6$ M⁻¹ s⁻¹ at 25 °C (delCardayré and Raines, 1994). This value of $k_{\text{cat}}$ corresponds to a $3 \times 10^{11}$-fold rate enhancement over the noncatalyzed reaction (Thompson et al., 1995; Raines, 1998). The active site contains two histidine residues (His12 and His119) that facilitate acid/base catalysis in the proposed mechanism (Figure 1.2) (Findlay et al., 1961) and stabilize the transition state (Figure 1.3). Alanine substitution at position 12 slows the enzyme considerably, decreasing the affinity of the enzyme for the transition state ~$10^4$-fold (Thompson and Raines, 1994).

Ribonuclease A is a remarkably stable enzyme ($T_m = 63$ °C) (Leland et al., 1998) as evidenced by Kunitz's early isolation procedure that demanded integrity under harsh conditions: 0.25 N sulfuric acid at 5 °C, and then, pH 3.0 at 95–100 °C (Kunitz and McDonald, 1953). Four native disulfide bonds are key to RNase A stability, as removal of any cysteine decreases the thermal stability (Klink et al., 2000). Additionally, previous work in our lab demonstrated that the introduction of a non-native disulfide bond (A4C/V118C) increased the thermal stability, and decreased the proteolytic susceptibility of an RNase A variant (Klink and Raines, 2000).
RNase A has served as a model protein system for groundbreaking research. In an effort to understand the principles that guide protein folding, Anfinsen and coworkers developed conditions (Anfinsen et al., 1961) to refold RNase A after complete reduction of its four disulfide bonds and its complete denaturation (Sela et al., 1957b). They discovered that although 105 possible pairings of the eight sulfhydryl groups of RNase A were possible, the protein refolded exclusively into the active conformation, suggesting that the amino acid sequence of a protein is sufficient to define its folded, active structure. This and other data led to the “thermodynamic hypothesis”, which postulates that “the three-dimensional structure of a native protein in its normal physiological milieu…is the one in which the Gibbs free energy of the whole system is the lowest; that is that the native conformation is determined by the totality of interatomic interactions, and hence the amino acid sequence, in a given environment” (Anfinsen, 1973).

Around the same time, Fred Richards received a sample of ribonuclease A from Armour via Chris Anfinsen (Richards, 1997). He used the enzyme as a substrate, finding that the protease subtilisin preferentially cleaves the peptide backbone of RNase A between residues 20 and 21 (Richards, 1958). The resulting complex, RNase S (wherein “S” refers to subtilisin), is composed of two fragments: S-peptide (residues 1–20) and S-protein (residues 21–124). Separation of the two components abolishes catalytic activity, while recombination in trans reconstitutes catalytic activity (Richards and Vithayathil, 1959; Richards, 1992; Richards, 1997).

This protein-fragment complementation system was useful in determining structure-function relationships prior to the advent of recombinant DNA technology. Chemists synthesized S-peptide analogues, including those with nonnatural amino acids, on solid
support and then studied their complexes with S-protein (Hirschmann et al., 1969; Richards et al., 1970; Gutte and Merrifield, 1971; Richards and Wyckoff, 1971). This early research yielded important insights into the principles of protein structure and function. Remarkably, the initial characterization of RNase S, and the genesis of the thermodynamic hypothesis occurred before the crystal structure of either RNase A or RNase S had been solved. Soon, however, both RNase A (Kerth et al., 1967) and RNase S (Wyckoff et al., 1967a; Wyckoff et al., 1967b) simultaneously succumbed to X-ray crystallography (Figure 1.4). Thus both have legitimate claim (Richards, 1997) as the third protein structure to be solved, after myoglobin (Kendrew and Parrish, 1956) and lysozyme (Blake et al., 1965).

RNase S has continued to be a choice system for protein engineering because of several important features of the complex. The interaction between the two fragments is generally stable, and can be tuned by pH (Schreier and Baldwin, 1976), ionic strength (Schreier and Baldwin, 1977), and temperature (Hearn et al., 1971). S-peptide is easily synthesized on solid support (enabling facile incorporation of nonnatural amino acids), and both fragments are soluble at physiological pH. Additionally, the topology of RNase S is such that appendages on the N- or C-termini of S-peptide do not compromise complex formation. This feature has led to the widespread adoption of the first 15 amino acid of S-peptide (which are sufficient for complex formation with S-protein) as a fusion protein tag (S-Tag®, Novagen Inc., Madison, WI) that facilitates purification (Raines et al., 2000) and sensitive detection (Kim and Raines, 1993; Kim and Raines, 1994b; Raines et al., 2000).
RNase S components have been commandeered further for signal amplification. For instance, attachment of S-peptide or S-protein to a solid support via a polycytidylic acid "leash" can result in >10⁴-fold activation upon addition of the fragment complementation partner (Cecchini et al., 1986; Ehrat et al., 1986), a useful feature for biosensors.

Additionally, Schultz and coworkers have created a sequence-specific ribonuclease by appending an oligonucleotide to the N-terminal cysteine of K1C S-peptide. Treatment of a 62-mer RNA molecule with the oligonucleotide–RNase S complex resulted in cleavage at a single location (Zuckermann and Schultz, 1988).

Although the RNase S protein-fragment complementation system has been the object or historic work in protein chemistry, two limitations compromise its utility. First, the traditional preparation of S-protein by proteolytic digestion with subtilisin yields a heterogeneous mixture of products (Doscher and Hirs, 1967). Secondly, the non-covalent nature of the S-peptide/S-protein interaction leads to the loss of catalytic activity at low concentrations. In Chapter 3, we report on the use of site-directed mutagenesis to overcome both limitations by the insertion of an enterokinase cleavage sequence and the installation of a nonperturbative disulfide bond between S-peptide and S-protein.

1.4 Ribonucleases as chemotherapeutics

RNA polymers play an indispensable role in biological systems. RNA not only serves as an essential conduit between DNA and protein, but also functions in several enzymatic systems. Ribonucleases catalyze the depolymerization of RNA, and thus play an important role in RNA turnover. In eukaryotic cells, intracellular ribonucleases such as XRNI and the exosome are primarily responsible for the degradation of RNA polymers.
(Parker and Song, 2004). Much less is known about biological influence of members of the RNase A superfamily which comprise a distinct group of secreted proteins.

Members of the RNase A superfamily are characterized by the ability to degrade RNA, an N-terminal signal sequence for secretion, the presence of conserved disulfide bonds, and a kidney-shaped three-dimensional structure (Beintema, 1987; Beintema et al., 1988a; Beintema et al., 1988b; Beintema et al., 1997; Dyer and Rosenberg, 2006). While RNase A, the prototypical family member, had been a favorite study object for decades, technological advances in the 1970’s led to a decline in ribonuclease research (Benner and Alleman, 1989). The discovery of other RNase A superfamily members with interesting biological properties has, however, led to a resurgence in ribonuclease research.

A comparative genomics analysis suggests that the RNase A superfamily first evolved as a host-defense mechanism in vertebrates (Cho et al., 2005). Since then, the functions of RNase A superfamily members have rapidly expanded (Cho et al., 2005; Dyer and Rosenberg, 2006) to include a wide range of biological functions including angiogenesis (Riordan, 1997), tumor cell toxicity (Youle and D’Alessio, 1997), immunosuppression, and embryotoxicity (D’Alessio et al., 1997).

Notably, in the early 1970’s, embryo extracts from the Northern Leopard frog (Rana pipiens) were found to possess antitumoral activity (Shogen and Yoan, 1973). Almost twenty years later, the antitumoral activity was attributed to a basic protein with a striking similarity to RNase A, onconase (Ardelt et al., 1991). Onconase is both cytostatic and cytotoxic towards many tumor cell lines (Darzynkiewicz et al., 1988; Rybak et al., 1996; Juan et al., 1998; Halicka et al., 2000; Leland et al., 2000; Lee and Raines, 2003;
Rodriguez et al., 2007) and, in mouse models, onconase treatment antagonizes xenograft tumor growth (Mikulski et al., 1990a; Rybak et al., 1996; Lee and Raines, 2003; Lee et al., 2007).

In Phase I clinical trials, onconase was well tolerated, although the dose was limited by reversible renal toxicity (Mikulski et al., 1993; Rodriguez et al., 2007). In Phase II, onconase treatment was effective against malignant mesothelioma, pancreatic, breast, and non-small cell lung cancers (Mikulski et al., 1995; Mikulski et al., 2002; Costanzi et al., 2005). Onconase is now in various stages of clinical trials and has been granted orphan-drug and fast-track status as a second-line therapy for patients with malignant mesothelioma. Encouragingly, onconase has shown synergy in combination with several other drugs (tamoxifen (Mikulski et al., 1990b; Lee et al., 2003), trifluoperazine (Mikulski et al., 1990b), vincristine (Rybak et al., 1996), lovastatin (Mikulski et al., 1992), rosiglitazone (Ramos-Nino and Littenberg, 2008)) and proteasome inhibitors (N-acetyl-leucinyl-leucinyl-norleucinal and the N-acetyl-leucinyl-valinyl-phenylalaninal) (Mikulski et al., 1998). Very recently, onconase has also shown favorable results against the dengue and SARS viruses (TAMIR, 2010).

The proposed mechanism of ribonuclease-mediated cytotoxicity is shown in Figure 1.5 (Lee and Raines, 2008). Ribonucleases bind to the cell surface and are internalized by endocytosis. Then, by a mechanism that is not well understood, ribonucleases translocate across the lipid bilayer to access the cytosol. Cytotoxic ribonucleases are then able to evade the cytosolic ribonuclease inhibitor protein and degrade intracellular RNA, leading to cell death.
The putative mechanism suggests several essential attributes of ribonucleases that elicit cytotoxic effects. First, ribonucleases must bind to the cell surface and be internalized efficiently. Second, ribonuclease must retain its three-dimensional structure (that is, have sufficient conformational stability). And most prominently, a cytotoxic ribonuclease must be able to degrade intracellular RNA in the presence of the ubiquitous ribonuclease inhibitor protein (RI) (Lee and Raines, 2008).

RI is a ubiquitously expressed ~50 kDa horseshoe-shape protein that binds to and inactivates pancreatic-type ribonucleases with 1:1 stoichiometry (Dickson et al., 2005). RI has been found in all cell types studied to date, and its intracellular concentration is estimated to be a nearly invariant 4 µM (Haigis et al., 2003). A broad spectrum of RNase A superfamily members, bind to RI with astounding affinity ($K_d \sim 10^{-15}$ M). Thus, RI is thought to act as an intracellular sentry to inactivate adventitiously internalized ribonucleases (Dickson et al., 2005).

The most distinguishing characteristic of onconase relative to non-cytotoxic ribonucleases is its ability to retain catalytic activity in the presence of RI (Wu et al., 1993; Dickson et al., 2005). Inspired by onconase and bovine seminal ribonuclease (which is a dimeric ribonuclease with a quaternary structure that also evades RI (D'Alessio et al., 1997)), mammalian pancreatic ribonucleases have been endowed with cytotoxic properties by disrupting their interaction with RI (Rutkoski and Raines, 2008).

Evasive mammalian ribonuclease variants, and human ribonuclease variants in particular, may possess advantageous chemotherapeutic properties. First, because enzymes with a low degree of sequence similarity tend to elicit an immune response (De Groot and Scott, 2007), mammalian ribonucleases (with greater sequence identity to the
human ribonucleases) may be less immunogenic. Still, onconase seems to be well-tolerated in a clinical setting (Pavlakis and Vogelzang, 2006). Second, onconase administration is limited by renal toxicity (Mikulski et al., 1993; Pavlakis and Vogelzang, 2006). The renal retention of mammalian ribonucleases, in contrast, is 50- to 100-fold lower than that of onconase (Vasandani et al., 1996). Furthermore, the catalytic efficiency of human RNase 1 is several orders of magnitude greater than that of onconase, suggesting a greater cytotoxic potential (Boix et al., 1996).

Recently an RI-evasive human RNase 1 variant, QBI-139, has entered into Phase I clinical trials. Onconase and human RNase 1, of course, share the general features of the RNase A superfamily (Beintema, 1987): the ability to degrade RNA, a kidney-shaped structure with multiple disulfide bonds, and conserved catalytic residues. Additionally, both are believed to elicit cytotoxic effects through the same general mechanism (Figure 1.5). Despite these similarities, however, important differences remain.

Onconase is smaller (11.8 kDa) (Lee and Raines, 2008) and has higher thermal stability (87 °C) (Notomista et al., 2000) than does human RNase 1 (~14.6 kDa, $T_m = 56$ °C) (Leland et al., 2001), making it highly resistant to proteolysis (Notomista, 2000). This increased stability may be advantageous, enabling onconase to maintain conformation integrity during endocytosis. Human RNase 1 and onconase share only 19.4% sequence identity (33.3% sequence similarity) (Rutkoski and Raines, 2008). Onconase also has two additional active-site residues, Lys9, and an N-terminal pyroglutamate that is formed by the co-translational cyclization of glutamine in the endoplasmic reticulum (Welker et al., 2007).
Furthermore, the ribonucleolytic activity of onconase with common substrates is ~5 orders of magnitude lower than that of mammalian homologues (Boix et al., 1996). Also, in vitro, onconase shows a unique preference for cleavage on the 5' side of a guanine nucleobase (Lee et al., 2008; Lee and Raines, 2008). tRNA is preferentially degraded in cellulo (Saxena et al., 2001) (perhaps cutting between a G–G bond in the variable loop or D-arm) (Suhasini and Sirdeshmukh, 2006). Recent data supports the notion that the tRNA cleavage by onconase may be responsible for its cytotoxic effects (Mei et al., 2010). Others, however, have suggested that the cytotoxic effects of onconase may be due to degradation of small RNA molecules (Zhao et al., 2008).

Finally, important details differentiate the cellular entry of onconase and mammalian ribonucleases. Onconase binds to broad array of cell-surface glycans, whereas RNase A, a bovine homologue, binds only with low affinity (Chao et al., 2010). Further, the internalization of mammalian ribonucleases, but not onconase, correlates with cell anionicity (Chao et al., 2010), suggesting that steps downstream from internalization (e.g., membrane translocation, substrate cleavage) must account for the high cytotoxicity of onconase. Moreover, it is unclear what role the increased cationicity of onconase (calculated pl > 9.5) (Ardelt et al., 1991) relative to human RNase 1 (Z = +6) (Johnson et al., 2007a) may play, although increased cationicity correlates with increased internalization efficiency for human RNase 1 variants (Johnson et al., 2007a).

The extent to which these dissimilarities elicit different cellular responses is not well understood. In chapter 4, we use microarray technology to investigate the transcriptional response of K-562 cells to clinically relevant ribonucleases.
Figure 1.1  A haloalkane dehalogenase variant that is incapable of the second hydrolysis step forms a covalent adduct with chloroalkanes (Los and Wood, 2007; Los et al., 2008).
Figure 1.1

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 1.2  Proposed mechanism of RNase A (Findlay et al., 1961).
Figure 1.2
Figure 1.3  Putative transition state of the transphosphorylation step of UpA cleavage by RNase A (Raines, 1998).
Figure 1.4 A detail of Irwin Geis’ painting of ribonuclease S. A dinucleotide substrate is found in the active site. The painting was commissioned as a gift in honor of Fredric M. Richards by his students at Yale University.
Figure 1.4
Figure 1.5  Putative routing of cytotoxic ribonucleases: cell surface binding, internalization, translocation, RI evasion, and degradation of cytosolic RNA.
CHAPTER 2

Fluorogenic affinity label for the facile,
rapid imaging of proteins in live cells

This chapter was published, in part, as:

affinity label for the facile, rapid, imaging of proteins in live cells. *Org. Biomol. Chem.* 7,
3969-3975.
2.1 Abstract

Haloalkane dehalogenase (HD) catalyzes the hydrolysis of haloalkanes via a covalent enzyme–substrate intermediate. Fusing a target protein to an HD variant that cannot hydrolyze the intermediate enables labeling of the target protein with a haloalkane in cellulo. The utility of extant probes is hampered, however, by background fluorescence as well as limited membrane permeability. Here, we report on the synthesis and use of a fluorogenic affinity label that, after unmasking by an intracellular esterase, labels an HD variant in cellulo. Labeling is rapid and specific, as expected from the reliance upon enzymic catalysts and the high membrane permeance of the probe both before and after unmasking. Most notably, even high concentrations of the fluorogenic affinity label cause minimal background fluorescence without a need to wash the cells. We envision that such fluorogenic affinity labels, which enlist catalysis by two cellular enzymes, will find utility in pulse–chase experiments, high-content screening, and numerous other protocols.

2.2 Introduction

The labeling of proteins with genetically encoded autofluorescent proteins has revolutionized cell imaging (Tsien, 2009). These “tags” can reveal subcellular localization, dynamics, and chemical environment (Chalfie and Kain, 2006). Nonetheless, the utility of autofluorescent proteins has notable limitations, including the restriction of “one clone–one color” and an inability to label temporally disparate pools of protein by using a single genetic construct, thereby precluding pulse–chase as well as other types of experiments.
A recent development has overcome some limitations of autofluorescent proteins. Specifically, genetically encoded enzymic tags now enable the use of small-molecule fluorophores to label proteins in cellulo (Lavis and Raines, 2008; Los et al., 2008; Johnsson, 2009). A key advantage of these systems is their modularity. With a single genetic construct, a wide variety of substrate analogs, including fluorophores with disparate spectral properties (Lavis and Raines, 2008), can be attached to a fusion protein of interest (i.e., one clone—many colors) at any time. In addition, this approach avails the high reactivity that has evolved within enzymic active sites. Second-order rate constants for enzyme-mediated labeling have been reported to be as high as $2.7 \times 10^6 \text{ M}^{-1}\text{s}^{-1}$ (Los et al., 2008), which exceeds by over a million-fold any rate constant reported for a chemoselective labeling reaction involving only small molecules (such as variations of the Huisgen 1,3-dipolar azide–alkyne cycloaddition or Staudinger ligation) (Soellner et al., 2006; Ning et al., 2008).

Haloalkane dehalogenase (HD; EC 3.8.1.5) is perhaps the most often-used enzymic fusion tag. This enzyme catalyzes the hydrolysis of haloalkanes via a covalent enzyme–substrate intermediate that undergoes hydrolysis. An HD variant that cannot perform the second, hydrolysis step (HaloTag®) reacts rapidly and specifically to form a covalent adduct with a wide variety of chloroalkanes, including assorted fluorophores of varied spectral and physiochemical properties (Los and Wood, 2007; Los et al., 2008).

Still, problems remain with HD and other enzymic fusion tags. Background fluorescence from excess probe and the inefficient passage of small-molecule probes through cellular membranes are especially problematic, necessitating long incubation times or vigorous washing steps. We reasoned that a latent fluorophore based on the
venerable trimethyl lock system (Borchardt and Cohen, 1972; Milstein and Cohen, 1972) could both diminish background fluorescence and enhance cellular delivery. Accordingly, we designed and synthesized probe 1 (Figure 2.1) as an affinity label for HD and tested its ability to label HD in living cells. The results are indicative of an advantageous means to label proteins in cellulo.

2.3 Results and discussion

2.3.1 Design and synthesis of probe 1

The design of probe 1 was based on previous work in which the trimethyl lock system was used to mask the intrinsic fluorescence or absorbance of a small molecule (Chandran et al., 2005; Huang and Lin, 2006; Lavis et al., 2006a; Lavis et al., 2006b; Johnson et al., 2007a; Levine et al., 2008; Mangold et al., 2008; Yatzeck et al., 2008; Turcotte et al., 2009). The trimethyl lock moiety is stable to spontaneous hydrolysis, but susceptible to intracellular esterase-catalyzed hydrolysis. The use of a urea moiety rather than a second trimethyl lock allows for single-hit kinetics and the facile addition of reactive groups for bioconjugation (Lavis et al., 2006a). The trimethyl lock–urea system is modular, accommodating a variety of dyes and appendages for bioconjugation.

We synthesized probe 1 by condensing two fragments: a fluorogenic substrate for a cellular esterase and a chloroalkane affinity label for HD (Scheme 2.1). Briefly, known \( t\)-Boc–rhodamine 4 was subjected to reaction with an \textit{in situ}-generated isocyanate from protected succinate 5 (Winkler et al., 2004) to generate \( t\)-Boc–rhodamine–urea 6. Deprotection with trifluoroacetic acid furnished the urea–rhodamine 7 that underwent
carbodiimide-mediated coupling with trimethyl lock acid 8 (Amsberry and Borchardt, 1990) to give benzyl-protected 9. Removal of the benzyl group by catalytic hydrogenation at −5 °C (Zaikova et al., 2001) afforded acid 10. Activation of the acid to the succimidyl ester, followed by reaction with alkyl chloride 11 gave the desired probe 1.

2.3.2  Cell imaging with probe 1

We compared the utility of probe 1 for live-cell imaging with that of known probes 2 (diAcFAM) and 3 (R110Direct™). In these experiments, we used a cell line, U2OS, that had been stably transfected to produce a nucleus-directed HD variant (HaloTag®–NLS3). The nuclear envelope is covered with pores that allow small molecules such as probes 1–3 to enter the nucleus by passive diffusion from the cytosol (Monnè, 1935).

We found that images of U2OS cells exposed to probes 1–3 differed significantly. Incubation of cells with probes 2 and 3 resulted in non-specific fluorescence after 15 min (Figure 2.1A). This unsolicited fluorescence was intracellular for 2 and extracellular for 3. In contrast, probe 1 showed remarkable specificity with virtually no background fluorescence. Additionally, imaging with 1 was strikingly rapid, as images had developed nearly fully after only 10 min (Figure 2.1B). Labeling with a ten-fold higher concentration of 1 retained specificity without a large increase in background fluorescence (Figure 2.1C).

We suspected that the variation in labeling between the fluorogenic affinity labels 1 and 2 is due to differing abilities of the masked and unmasked probes to cross the cell membrane. Although masked 1 and 2 have a net charge of \(Z = 0\), unmasked 1 has \(Z = 0\)
whereas unmasked 2 has \( Z = -2 \). Cell images suggest that unmasked 2 that has not reacted with the HD variant becomes trapped in the cell, leading to intracellular background fluorescence (Figure 2.1Aii and v). Conversely, unmasked 1 that has not reacted with the HD variant can exude from the cell, leading to low background fluorescence.

The low background fluorescence observed with probe 1 could be attributed solely to the latency of its fluorescence. To address this issue, we treated probes 1 and 2 with porcine liver esterase prior to the no-wash labeling procedure. As anticipated from its net charge of \( Z = -2 \), unmasked 2 was largely membrane impermeant (Figure 2.1Dii). Remarkably, unmasked 1 (Figure 2.1Dii) displayed much less intracellular and extracellular background fluorescence than did 3 (Figure 2.1Diii). These data demonstrate that the fluorescence-masking trimethyl lock moiety is not the singular determinant for preventing extracellular background fluorescence. Rather, the intrinsic cell-permeability of probe 1, in its masked and unmasked state, plays an important role.

2.3.3 *Probe lipophilicity*

To seek an explanation for the differences in probe internalization, we first calculated the value of \( \log D \) for the predominant form of relevant labels at pH 7.4 (Table 2.1). Because increased lipophilicity generally correlates with increased membrane permeability, probes with higher \( \log D \) values are expected to permeate the membrane more rapidly. The relatively low calculated \( \log D \) value of unmasked 2 corresponds with its slower rate of internalization. There is, however, only a small difference between the
calculated values of unmasked 1 and 3, even though cell-imaging experiments
demonstrate that unmasked 1 is much more membrane permeant (Figure 2.1D).

The inconsistency in the anticipated and demonstrated membrane permeability led us
to hypothesize that the increased rate of internalization of unmasked 1 relative to 3 is due
to perturbation of the lactone–quinoid equilibrium of the two rhodamine-based dyes
(Lavis et al., 2006a) (Figure 2.3). We reasoned that electron-withdrawal by the urea
moiety shifts the equilibrium somewhat from the relatively membrane-impermeant
quinoid (which is zwitterionic and hence polar) to the lactone.

To investigate the propensity of rhodamine–urea dyes to form the lactone, we
determined the effect of solvent dielectric constant (\(\varepsilon\)) on the lactone–quinoid equilibrium
of unmasked 1 and related compounds 12 (rhodamine 110) and 13 (a rhodamine–urea)
(Lavis et al., 2006a). The ultraviolet–visible spectrum of the quinoid is characterized by a
relative absorbance maximum at \(\sim 500\) nm, which is absent in the spectrum of the lactone
(Ioffe and Otten, 1965). The value of \(\varepsilon\) was varied by altering dioxane/water ratios (Kuila
and Lahiri, 2004). We found that formation of the less polar lactones of unmasked 1 and
13, which contain a urea moiety, is favorable in solutions of higher polarity than with
compound 12 (Figure 2.4). We conclude that the increased rate of internalization of
rhodamine–urea dyes is likely due to their increased propensity for lactone formation.
This feature ameliorates the difficulty of cell-membrane penetration and lowers
background fluorescence.
2.3.4 Recapitulation

Together, the data indicate that probe 3, which is always fluorescent, is not especially membrane-permeant and produces much extracellular, background fluorescence with our rapid-imaging protocol (Figure 2.1A, panels iii and vi). Although the acetyl groups of probe 2 mask its fluorescence, these groups are vulnerable to hydrolysis in cell-free medium and even PBS (Lavis et al., 2006a). At short time scales (i.e., minutes), the resultant unmasking is not a severe problem. The intracellular background is due to the anionic unmasked fluor that has not reacted with the HD protein but cannot exit the cell because of its membrane impermeance. Only probe 1 has the attributes necessary.

2.4 Conclusions

We have demonstrated that probe 1 is a useful cell-permeant affinity label for the facile, rapid labeling of target proteins in live cells. The omission of wash steps is particularly advantageous for the labeling of proteins in non-adherent cells, which are more tedious to separate from excess probe than are adherent cells. In addition, as probe 1 exhibits low background fluorescence, little effort is needed to determine a useful concentration. This attribute could be useful for the labeling of target proteins within a population of transiently transfected cells. The rapidity of cellular labeling with probe 1 enables interrogations with high temporal resolution. This attribute is particularly beneficial for pulse–chase experiments, which require differential labeling of temporally disparate pools of protein. Finally, we suspect that probe 1 will be useful in high-content screens, wherein wash steps, optimization of label concentration, and probe instability are especially problematic and costly. Accordingly, we envision that probe 1 and analogous
fluorogenic affinity labels will be a useful addition to the armamentarium for the labeling of proteins \textit{in cellulo}.

2.5 Experimental

2.5.1 \textit{Synthesis of 1}

\textit{General.} Amine 11 was a generous gift from the Promega Corporation (Madison, WI). Dimethylformamide (DMF), tetrahydrofuran (THF), and dichloromethane (\(\text{CH}_2\text{Cl}_2\)) were drawn from a Baker CYCLE-TAINER solvent delivery system. All other reagents were obtained from Sigma (Saint Louis, MO) or Fisher Scientific (Hanover Park, IL) and used without further purification.

Thin-layer chromatography was performed with aluminum-backed plates coated with silica gel containing F\(_{254}\) phosphor and visualized by UV illumination or staining with I\(_2\), ceric ammonium molybdate, or phosphomolybdic acid. Flash chromatography was performed by using open columns loaded with silica gel-60 (230–400 mesh), or on a FlashMaster Solo system (Argonaut, Redwood City, CA) with Isolute Flash Si II columns (International Sorbent Technology, Hengoed, Mid Glamorgan, UK). The term “concentrated under reduced pressure” refers to the removal of solvents and other volatile materials using a rotary evaporator at water-aspirator pressure (<20 mm Hg) while maintaining the water-bath temperature below 40 °C. The term “high vacuum” refers to vacuum achieved by a mechanical belt-drive oil pump.

NMR spectra were obtained with a Bruker DMX-400 Avance spectrometer at the National Magnetic Resonance Facility at Madison (NMRFAM). Mass spectrometry was
performed with a Micromass LCT (electrospray ionization, ESI) mass spectrometer in the Mass Spectrometry Facility in the Department of Chemistry.

**Synthesis of succinic acid monobenzyl ester (5).** Compound 5 was synthesized according to a published procedure (Winkler *et al*., 2004). The white crystalline material afforded in the published procedure was dissolved in a minimal amount of 1:1 hexanes/EtOAc and cooled to 4 °C. Crystallization was initiated by the dropwise addition of hexane (~20 drops). The mixture was stored overnight at −20 °C. The crystals were isolated by filtration, washed with cold hexane (4 °C), and dried in the air. Residual solvent was removed under high vacuum to yield compound 5 as a white crystalline solid (3.44 g, 83%). $^1$H NMR (400 MHz, DMSO-$d_6$) $\delta$ (ppm): 12.25 (s, 1H), 7.36 (m, 5H), 5.10 (s, 2H), 2.58 (m, 2H), 2.50 (m, 2H). $^{13}$C NMR (400 MHz, DMSO-$d_6$) $\delta$ (ppm): 173.47, 172.10, 136.25, 128.45, 128.00, 127.84, 65.53, 28.78, 28.72. HRMS (ESI): [M+Na]$^+$ calculated, 231.0628; found, 231.0635.

**Synthesis of t-Boc–rhodamine–urea 6.** Succinic acid monobenzyl ester (5; 193 mg, 0.929 mmol) was dissolved in anhydrous THF (2.0 mL) under Ar(g). Hünig's base (DIEA; 202 µL, 1.16 mmol) was then added, followed by the dropwise addition of diphenyl phosphoryl azide (DPPA; 256 mg, 0.929 mmol). The solution was stirred for 6 h and subsequently heated at reflux for an additional 2 h. Then, known t-Boc–rhodamine 4 (100 mg, 0.232 mmol) was added, and the reaction mixture was stirred at reflux for 18 h. The reaction mixture was then partitioned between 5% v/v HCl(aq) and CH$_2$Cl$_2$. The organic extract was washed consecutively with 5% v/v HCl, water (3×), 5% w/v NaHCO$_3$(aq) (2×), water (3×), and saturated brine, and dried over MgSO$_4$(s). The solution was concentrated under reduced pressure, and the residue was purified by
column chromatography (silica gel; first column: 5:3:2 hexanes/EtOAc/CH₂Cl₂→4:4:2 hexanes/EtOAc/CH₂Cl₂, second column: 0→2% v/v MeOH in CH₂Cl₂ containing AcOH (1% v/v), third column: 5:3:2 hexanes/EtOAc/CH₂Cl₂). Compound 6 was obtained as a pale yellow solid (137.2 mg, 93%). ¹H NMR (400 MHz, CDCl₃) δ (ppm): 7.99 (d, J = 7.0 Hz, 1H), 7.65–7.56 (m, 3H), 7.37–7.29 (m, 6H), 7.10 (s, 1H), 7.06 (m, 3H), 6.99 (d, J = 8.4 Hz, 1H), 6.66 (d, J = 8.6 Hz, 1H), 6.55 (d, J = 8.7 Hz, 1H), 6.00 (t, J = 5.6 Hz, 1H), 5.11 (s, 2H), 3.53 (dd, J = 11.5, 5.7 Hz, 2H), 2.62 (t, J = 6.1 Hz, 2H), 1.52 (s, 9H). ¹³C NMR (400 MHz, CDCl₃) δ (ppm): 172.42, 170.12, 155.20, 152.59, 151.98, 151.82, 142.15, 141.21, 135.67, 135.25, 129.82, 128.57, 128.28, 128.18, 126.32, 125.29, 124.34, 115.40, 114.41, 112.91, 112.17, 106.34, 106.15, 81.17, 66.47, 35.71, 34.76, 28.33. HRMS (ESI): [M+H]⁺ calculated, 636.2341; found, 636.2341.

Synthesis of rhodamine–urea 7. t-Boc–rhodamine–urea 6 (54.0 mg, 0.0848 mmol) was dissolved in CH₂Cl₂ (2.0 mL). Trifluoroacetic acid (TFA, 400 μL) was added, and the reaction mixture was stirred at 0 °C for 2 h. The reaction mixture was then allowed to warm to ambient temperature and stirred for an additional 3 h. The solution was concentrated under reduced pressure, and residual TFA was removed by azeotropick distillation with toluene. The residue was then dissolved in a minimal amount of acetone and purified by column chromatography (silica gel, 5→10% v/v MeOH:CH₂Cl₂) to afford compound 7 as an orange crystalline solid (40 mg, 88%). ¹H NMR (400 MHz, acetone-d₆) δ (ppm): 8.34 (s, 1H), 7.97 (d, J = 7.6 Hz, 1H), 7.81–7.76 (m, 2H), 7.70 (t, J = 7.1 Hz, 1H), 7.39–7.25 (m, 6H), 6.92 (dd, J = 8.6, 2.2 Hz, 1H), 6.61 (d, J = 8.6 Hz, 1H), 6.58 (d, J = 2.1 Hz, 1H), 6.47 (d, J = 8.5 Hz, 1H), 6.42 (dd, J = 8.5, 2.1 Hz, 1H),
6.09 (br, 1H), 5.14 (m, 2H), 3.50 (q, \( J = 6.3 \) Hz, 2H), 2.63 (t, \( J = 6.5 \) Hz, 2H), 2.09 (s, 2H). HRMS (ESI): [M+H]^+ calculated, 536.1822; found, 536.1816.

**Synthesis of trimethyl lock–rhodamine–urea 9.** Compound 7 (39.0 mg, 0.0728 mmol) was dissolved in 2.0 mL DMF and 2.0 mL pyridine. Trimethyl lock acid 8 (Amsberry and Borchardt, 1990) (38.5 mg, 0.146 mmol) and 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC; 28 mg, 0.146 mmol) were then added, and the reaction mixture was stirred overnight at ambient temperature. The solution was concentrated under reduced pressure, and the residue was dissolved in CH\(_2\)Cl\(_2\) (200 mL). The resulting solution was washed consecutively with 5% v/v HCl(aq), water, saturated NaHCO\(_3\)(aq), and saturated brine. The organic fraction was dried over Na\(_2\)SO\(_4\)(s). The residue was purified by column chromatography (silica gel, 4:6→5:5 hexanes/EtOAc) to give compound 9 as an orange solid (46 mg, 81%). \(^1\)H NMR (400 MHz, CDCl\(_3\)) \(\delta\) (ppm): 7.98 (d, \( J = 7.3 \) Hz, 1H), 7.66–7.56 (m, 3H), 7.38 (s, 1H), 7.32 (s, 5H), 7.20 (s, 1H), 7.06–7.02 (m, 3H), 6.80 (s, 1H), 6.67–6.61 (m, 2H), 6.57–6.53 (m, 2H), 5.74 (t, \( J = 5.7 \) Hz, 1H), 5.12 (s, 2H), 3.57–3.50 (m, 2H), 2.63–2.61 (m, 4H), 2.43 (s, 3H), 2.37 (s, 3H), 2.21 (s, 3H), 1.69–1.67 (m, 6H). \(^{13}\)C NMR (400 MHz, CDCl\(_3\)) \(\delta\) (ppm): 172.40, 171.96, 170.39, 170.24, 155.18, 153.15, 151.48, 149.97, 141.74, 139.91, 138.85, 137.18, 135.70, 135.31, 133.13, 132.96, 129.75, 128.55, 128.25, 128.13, 126.18, 124.90, 124.10, 123.42, 115.32, 114.93, 113.70, 111.65, 107.66, 106.22, 83.74, 66.38, 50.81, 40.25, 35.60, 34.77, 32.02, 25.53, 21.93, 20.16. HRMS (ESI): [M+H]^+ calculated, 782.3073; found, 782.3062.

**Synthesis of trimethyl lock–rhodamine–urea 10.** Ethanol (10 mL) was added to compound 9 (35 mg, 0.045 mmol) and Pd/C (10% w/w, 14 mg) at -5 °C. The resulting mixture was stirred under an H\(_2\)(g) atmosphere for 30 min. The reaction mixture was
filtered through celite and washed consecutively with CH$_2$Cl$_2$ and EtOH. The washings were combined and concentrated under reduced pressure. The residue was then adsorbed onto celite and purified by column chromatography (silica gel, 0→10% v/v MeOH in CH$_2$Cl$_2$ containing 1% v/v AcOH) to afford compound 10 as a pale-yellow solid (26 mg, 84%). $^1$H NMR (400 MHz, CDCl$_3$) $\delta$ (ppm): 8.00–7.90 (m, 2H), 7.84 (s, 1H), 7.59–7.49 (m, 2H), 7.34 (s, 1H), 7.22 (s, 1H), 6.95 (d, $J = 7.3$ Hz, 1H), 6.81 (d, $J = 7.2$ Hz, 1H), 6.74 (s, 1H), 6.64–6.55 (m, 2H), 6.53–6.42 (m, 2H), 6.10 (b, 1H), 3.39 (s, 2H), 2.62 (s, 2H), 2.46 (s, 2H), 2.39 (s, 3H), 2.32 (s, 3H), 2.16 (s, 3H), 1.63 (s, 6H). $^{13}$C NMR (400 MHz, CDCl$_3$) $\delta$ (ppm): 176.38, 171.97, 170.60, 170.04, 155.88, 152.83, 151.52, 151.48, 149.89, 141.58, 139.74, 138.76, 137.08, 135.31, 133.03, 132.93, 129.77, 128.00, 126.17, 124.87, 124.03, 123.35, 115.45, 114.93, 113.94, 111.88, 107.83, 106.40, 83.66, 50.65, 40.17, 35.53, 34.76, 32.00, 25.48, 21.88, 20.12. HRMS (ESI): [M+Na]$^+$ calculated, 714.2422; found, 714.2396.

Synthesis of probe 1. Compound 10 (26 mg, 0.038 mmol) was dissolved in DMF (1.0 mL) and pyridine (1.0 mL each). EDC (22 mg, 0.113 mmol) was added, and the resulting mixture was stirred for 1 h. N-Hydroxysuccinimide (NHS; 13 mg, 0.113 mmol) was then added, and the reaction mixture was stirred for 36 h under Ar(g). The solution was concentrated under reduced pressure, and the residue was partitioned between 10% w/v citric acid(aq) and CH$_2$Cl$_2$. The layers were separated, and the aqueous layer was extracted with CH$_2$Cl$_2$. The organic extracts were combined, washed consecutively with water and saturated brine, and dried over Na$_2$SO$_4$(s) to give the crude succinimide intermediate as a pale-yellow solid (35 mg). This intermediate (35 mg) was then dissolved in 2.0 mL of 9:1 DMF/DIEA. Amine 11 (10 mg, 0.044) was added, and the
reaction mixture was stirred under Ar(g) for 72 h. The solution was concentrated under reduced pressure, and the residue was partitioned between EtOAc and 6% v/v HCl(aq). The layers were separated, and the aqueous phase was extracted with EtOAc. The organic extracts were combined, washed consecutively with water and saturated brine, and dried over Na₂SO₄(s). Probe 1 was isolated by column chromatography (silica gel; first column: 5:3:2 hexanes/EtOAc/CH₂Cl₂, second column: 2% v/v MeOH in CH₂Cl₂) as an off-white solid (12.5 mg, 32%). ¹H NMR (400 MHz, CDCl₃) δ (ppm): 7.98 (d, J = 6.5 Hz, 1H), 7.87 (s, 1H), 7.60 (dt, J = 4.3, 1.4 Hz, 2H), 7.46 (s, 1H), 7.39 (dd, J = 10.3, 2.2 Hz, 2H), 7.09 (d, J = 7.3 Hz, 1H), 6.97 (dd, J = 8.6, 2.2 Hz, 1H), 6.79 (s, 1H), 6.62–6.52 (m, 4H), 6.17 (br, 1H), 5.80 (br, 1H), 3.62 (s, 4H), 3.53–3.45 (m, 10H), 2.56 (m, 2H), 2.42 (m, 5H), 2.38 (s, 3H), 2.23 (s, 3H), 1.75–1.64 (m, 8H), 1.63–1.54 (m, 2H), 1.46–1.26 (m, 4H). HRMS (ESI): [M+Na]⁺ calculated, 919.3656; found, 919.3617.

2.5.2 Cell imaging

General. Probe 2 and U2OS cells were generous gifts from Promega. Probe 3 was from Promega. Rhodamine 110 was from Sigma–Aldrich. Cell-culture medium was from Invitrogen or Hyclone (Fisher Scientific). Dulbecco’s phosphate-buffered saline (PBS) and fetal bovine serum were from Invitrogen. Porcine liver esterase, MW ~163 kDa, as a suspension in 3.2 M (NH₄)₂SO₄ was from Sigma Chemical (product number E2884). For labeling experiments, fluorogenic and fluorescent probes were stored as stock concentrations in DMSO and diluted such that DMSO concentration did not exceed 1%. Absorbance measurements were recorded at ambient temperature (23 ± 2 °C) in 1-cm path-length cuvettes on a Cary model 50 spectrometer from Varian.
Cell Preparation. U2OS cells (ATCC HTB-96TM) were plated on glass-bottom culture dishes (35 × 10 mm; Electron Microscopy Sciences) and grown to 60–90% confluence at 37 °C in McCoy’s 5A modified media containing FBS (10% v/v). Prior to the addition of probes, the medium was replaced with phenol red-free McCoy’s 5A medium (800 μL) that had been incubated at 37 °C. Probes were dissolved at a 5× working concentration in phenol red-free medium (200 μL) that had been incubated at 37 °C. For the experiments with unmasked probes (Figure 2.1D), porcine liver esterase (10 μL, 1.66 U/μL) was then added, and the resulting medium was incubated at room temperature for 3½ h. The probe-containing medium was added to the cell-containing medium, and the resulting medium was incubated at 37 °C. Cells were visualized by confocal microscopy after 10, 15, or 30 min.

Microscopy. Images of cells were obtained with a Nikon Eclipse TE2000-U confocal microscope equipped with a Zeiss AxioCam digital camera. Excitation at 488 nm was provided by an argon-ion laser, and emission light was passed through a filter centered at 515 nm and having a 40-nm band-pass. Brightfield images indicated that the cells were alive and appeared to have normal physiology, both before and after imaging.

2.5.3 Ultraviolet–visible spectroscopy

For each measurement, a probe was initially added to 20 μL PBS. To unmask the fluorescence of 1, porcine liver esterase (3.0 μL, 1.66 U/μL) was added, and the solution was left at ambient temperature for 4 h. Then, varying ratios of water and dioxane were added to a final volume of 200 μL, and absorbance was measured from 400–600 nm.
2.5.4 Calculation of logD

The parameter logD refers to the ratio of concentration of all probe microspecies (including both ionized and neutral forms) in octanol to that in water, according to eq 1 (van de Waterbeemd, 2009):

\[
\log D = \log \left( \frac{\sum [\text{microspecies}_{\text{octanol}}]}{\sum [\text{microspecies}_{\text{water}}]} \right)
\]  

(1)

Values of logD were calculated with ACD/PhysChem Suite, version 12 (ACD, 2009). Parameters of fluorescein (pKₐ 6.32 (Goldberg and Baldwin, 1998); logD ~0.99 (Cheruvu and Kompella, 2006)) and rhodamine 110 (logD 0.8) (Lampidis et al., 1989) were entered into the software training database.
2.5.5 \textit{NMR spectra}

400 MHz $^1$H NMR spectrum of compound 5 in DMSO-$d_6$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
400 MHz $^{13}$C NMR spectrum of compound 5 in DMSO-$d_6$. 

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
400 MHz $^1$H NMR spectrum of compound 6 in CDCl$_3$. 

[Diagram of the compound and its NMR spectrum with chemical shifts and peaks marked.]
400 MHz $^{13}$C NMR spectrum of compound 6 in CDCl$_3$. 

[Chemical structure image]
400 MHz $^1$H NMR spectrum of compound 7 in acetone-$d_6$. 
400 MHz \(^1\)H NMR spectrum of compound 9 in CDCl\(_3\).
400 MHz $^{13}$C NMR spectrum of compound 9 in CDCl$_3$. 
400 MHz $^1$H NMR spectrum of compound 10 in CDCl$_3$. 
400 MHz $^{13}$C NMR spectrum of compound 10 in CDCl$_3$. 
400 MHz $^1$H NMR spectrum of compound 1 in CDCl$_3$. 
Acknowledgements  We are grateful to the Promega Corporation for gifts of compounds 2 and 11, and U2OS cells. This work was supported by grant CA073808 (NIH). R.W.W. was supported by an NSF Graduate Research Fellowship and by Chemistry–Biology Interface Training Grant GM008505 (NIH). L.D.L was supported by an ACS Division of Organic Chemistry Fellowship sponsored by the Genentech Foundation and by Biotechnology Training Grant GM008349 (NIH). V.M.K. was supported by a Barry M. Goldwater Scholarship and a Hilldale Undergraduate/Faculty Research Fellowship. NMRFAM was supported by grant P41RR02301 (NIH).
Table 2.1 Calculated values of log$D$ for masked and unmasked probes at pH 7.4.

<table>
<thead>
<tr>
<th>Probe</th>
<th>log$D$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.82</td>
</tr>
<tr>
<td>2</td>
<td>3.93</td>
</tr>
<tr>
<td>3</td>
<td>2.37</td>
</tr>
<tr>
<td>![Diagram of probe 3]</td>
<td>2.63</td>
</tr>
<tr>
<td>![Diagram of probe 4]</td>
<td>0.12</td>
</tr>
</tbody>
</table>

$R = \text{H}$
Scheme 2.1  Synthesis of probe 1.
**Figure 2.1** Fluorogenic (1 and 2) and fluorescent (3) labels for haloalkane dehalogenase.
Figure 2.1
Figure 2.2  Labeling of an HD variant (HaloTag®-NLS2) in live, unwashed U2OS cells at 37 °C as visualized by confocal microscopy. Scale bars = 200 μm. (A) Effect of probe type. Probe (1.0 μM) was incubated with cells for 15 min; (i) 1, (ii) 2, (iii) 3, (iv, v, and vi) overlay with brightfield images. (B) Effect of incubation time. Probe 1 (1.0 μM) was incubated with cells for (i) 10 min, (ii) 15 min, and (iii) 30 min. (C) Effect of probe concentration. Probe 1 was incubated with cells for 15 min at (i) 1 mM, (ii) 10 mM, (iii and iv) overlay with brightfield images. (D) Effect of probe unmasking. Probe (1.0 μM) was incubated with cells for 15 min; (i) unmasked 1, (ii) unmasked 2, and (iii) 3.
Figure 2.2
Figure 2.3  Lactone–quinoid equilibrium of 1, 12, and 13.
Figure 2.3

quinoid $\xrightarrow{\text{ }}$ lactone

unmasked 1 $R^1 = H, R^2 = X$

12 $R^1 = R^2 = H$

13 $R^1 = H, R^2 = C(O)N(CH_3)_2$

$X = \begin{array}{c}
\text{N} \\
\text{O}
\end{array}$

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 2.4 Effect of dielectric constant on the lactone–quinoid equilibrium of unmasked 12, 13, and unmasked 1. Absorption spectra of (A) 12 (50 μM), (B) 13 (12.5 μM), and (C) unmasked 1 (12.5 μM) in mixtures of dioxane and water. (D) Absorption at $\lambda_{\text{max}}$ in the spectra in panels A–C. Values of $\varepsilon$ are from the literature (Kuila and Lahiri, 2004).
CHAPTER 3

Ribonuclease S redux

This chapter was prepared for submission to Chemical Communications as:

3.1 Abstract

The S-peptide and S-protein components of bovine pancreatic ribonuclease form a noncovalent complex with ribonucleolytic activity. Although this original of protein-fragment complementation systems has been the object of historic work in protein chemistry, two limitations compromise its utility. First, the traditional preparation of S-protein by proteolytic digestion with subtilisin yields a heterogeneous mixture of products. Secondly, the noncovalent nature of the S-peptide–S-protein interaction leads to the loss of catalytic activity at low concentrations. Here, we report on the use of site-directed mutagenesis to overcome both limitations. First, we install a site for enterokinase, which has much greater substrate specificity than does subtilisin and hence enables facile S-protein isolations. Secondly, we install a nonperturbative disulfide bond between the two components. The resulting complex is easy to prepare and retains its catalytic activity at low concentrations, enabling new uses for this venerable system.

3.2 Introduction

In the late 1950's, Fred Richards discovered protein-fragment complementation—the restoration of protein function by the noncovalent interaction of component polypeptides. Working in the renowned Carlsberg Laboratory in Copenhagen, he found that the protease subtilisin catalyzes the cleavage of bovine pancreatic ribonuclease (RNase A; EC 3.1.27.5) (Raines, 1998; Marshall et al., 2008; Marshall and Feng, D.J. Kuster) between residues 20 and 21 (Richards, 1958). The resulting complex, RNase S (wherein “S” refers to subtilisin), is composed of two fragments: S-peptide (residues 1–20) and S-protein (residues 21–124). Separation of these two components abolished ribonucleolytic
activity, which was restored by their mixing (Richards and Vithayathil, 1959; Richards, 1992; Richards, 1997). This work, which served to launch the field of molecular recognition, was done before the three-dimensional structure of any protein was known. Chris Anfinsen and coworkers later found that the first 15 residues of S-peptide (S15) yield a fully active complex (Potts et al., 1963). In the last fifty years, many other proteins have ceded to fragment complementation, including β-galactosidase, dihydrofolate reductase, β-lactamase, luciferase, ubiquitin, and the green fluorescent protein (Michnick et al., 2007).

The RNase S system has had a singular role in protein chemistry. Prior to the advent of recombinant DNA technology, Bruce Merrifield (Gutte and Merrifield, 1971), Ralph Hirschmann (Hirschmann et al., 1969), and others synthesized S-peptide analogues and studied their complexes with S-protein. A harbinger of current work on proteins containing nonnatural residues, these studies revealed important principles of protein folding (Kato and Anfinsen, 1969), protein–protein interactions (Schreier and Baldwin, 1976; Thomson et al., 1994), and enzymology (Dunn et al., 1974). More recently, the association of S-peptide with S-protein has found widespread use in fusion protein systems (Kim and Raines, 1993; Kim and Raines, 1994a). Vectors for making S-peptide (“S·tag”) fusions facilitate the purification and sensitive detection of fusion proteins (Raines et al., 2000). The RNase S complex has also been used as a molecular vehicle for targeted delivery (Gaidamakova et al., 2001; Backer et al., 2002; Backer et al., 2006).

Despite its venerable history, the traditional RNase S system is compromised in its utility. First, the isolation of the individual RNase S components is not trivial. A major problem is that subtilisin (Markland and Smith, 1971) is a non-specific protease that
cleaves RNase A not only between residues 20 and 21, but also at other peptide bonds (Doscher and Hirs, 1967; Richards and Wyckoff, 1971). Hence, digestion is commonly quenched prior to complete cleavage at the S-peptide/S-protein boundary. The result is an inefficient conversion to RNase S (Gaynutdinov et al., 2003), and a tedious isolation from intact RNase A (which contaminates commercial preparations). Moreover, subtilisin is not suitable for selective cleavage at the analogous S-protein/S-peptide boundary of RNase A homologues (Gaynutdinov et al., 2003). Likewise, amino-acid substitutions near residues 20 and 21 of RNase A can hinder digestion by subtilisin (Markert et al., 2001).

A second major problem relates to the stability of the S-protein-S-peptide complex, which has a $K_d$ value in the micromolar–nanomolar range (Schreier and Baldwin, 1976; Connelly et al., 1990). At lower concentrations, the complex is nearly fully dissociated, abolishing catalytic activity. Moreover, RNase S has a lower thermal stability than does RNase A (Ratnaparkhi and Varadarajan, 2001), and is more susceptible to chemical denaturation (Richards and Vithayathil, 1959) and proteolytic degradation (Richards, 1958; Allende and Richards, 1962).

Here, we revisit RNase S. We use modern methods of molecular biology and protein chemistry to facilitate the production and purification of S-protein and to install a covalent bond between the components, thereby generating “RNase–S” (Figure 3.1). These changes overcome the major limitations of the traditional system and provide new opportunities.
3.3 Results and Discussion

The non-specific proteolytic activity of subtilisin complicates the isolation of pure RNase S components. In our hands, digestion of RNase A with subtilisin resulted in a complex mixture of products and a daunting separation (Fig. 3.2). We reasoned that increased specificity for proteolytic cleavage between the S-peptide and S-protein regions of RNase A would simplify the isolation procedure.

Enterokinase is a digestive protease that catalyzes cleavage on the C-terminal side of the amino-acid sequence AspAspAspAspLys (Zheng et al., 2009). To avail the higher substrate specificity of enterokinase relative to subtilisin, we inserted an enterokinase cleavage site between residues 20 and 21 of RNase A variants.

Enterokinase digestion of wild-type RNase A with an inserted enterokinase cleavage site (DDDDK RNase A) resulted in essentially complete conversion to the desired RNase S product (Figure 3.2). S-protein and S-peptide (containing the C-terminal DDDDK sequence) were separated easily by high-performance liquid chromatography (HPLC) (Figure 3.3) to yield purified components.

Next, we took advantage of previous work in our laboratory, which demonstrated that the introduction of cysteine residues at positions 4 and 118 of RNase A results in the spontaneous formation of a stabilizing disulfide bond (Klink and Raines, 2000; Dickson et al., 2003; Plainkum et al., 2003). Analogous disulfides also stabilize homologous ribonucleases (Futami et al., 2000; Leland et al., 2001; Backer et al., 2006). We reasoned that the introduction of this disulfide bond in the RNase S complex would remove many of the drawbacks and complications of the noncovalent system. Accordingly, we replaced Val118 with a cysteine residue in DDDDK RNase A.
To enhance our system still further, we added another substitution, H12A, to DDDDK/V118C RNase A. His12 is in the active site of RNase A, and its replacement with an alanine residue decreases ribonucleolytic activity by $>10^4$-fold without perturbing the three-dimensional structure (Thompson and Raines, 1994; Park et al., 2001). As residue 12 will be discarded after digestion with enterokinase, the H12A substitution serves as a safeguard, diminishing catalytic activity from any trace contaminant of RNase A in an S-protein sample. Finally, to prevent the adventitious air oxidation of Cys118, we protected the purified V118C S-protein by reaction with 5,5'‐dithio‐bis(2‐nitrobenzoic acid) (DTNB), thereby forming a mixed disulfide.

To effect the semisynthesis of RNase–S, we added deprotected V118C S-protein to the A4C variant of S15, which had been activated as a mixed disulfide with 2‐nitro‐5‐thiobenzoic acid (NTB). After removal of the NTB byproduct, the presence of the disulfide linkage between A4C S15 and V118C S-protein was apparent by both SDS–PAGE (Fig. 4, lane 4) and MALDI–TOF mass spectrometry (m/z 13324; expected: 13317). The semisynthesis of RNase–S in the opposite manner (that is, by reacting unactivated A4C S15 with NTB-activated S-protein) was less effective.

We analyzed RNase A and the components of RNase S by non‐reducing zymogram electrophoresis, an extremely sensitive technique for detecting ribonucleolytic activity (Bravo et al., 1994). Clear bands on a dark background are indicative of ribonucleolytic activity.

Neither V118C S-protein (Fig 3.5, lane 2) nor V118C S-protein mixed with cysteine‐free S‐peptide (Fig 3.5, lane 3) displayed detectable activity. In marked contrast, RNase–
S had robust activity (Figure 3.5, lane 4). The activity due to RNase A contamination of commercial RNase S is apparent (Figure 3.5, lane 6).

Finally, we assessed the activities of RNase S and RNase−S as catalysts of RNA cleavage at high and low enzymic concentrations. RNase−S retains nearly all of the ribonucleolytic activity of RNase A across a concentration range of nearly $10^6$-fold (Figure 3.6). In contrast, RNase S loses detectable activity over that range.

### 3.4 Conclusions

Despite many advances since the 1950's, the isolation of RNase S components has remained an arduous challenge. In their initial isolation of S-protein and S-peptide, Richards and Vithayathil digested a whopping 730 mg of RNase A with subtilisin (Richards and Vithayathil, 1959). S-protein was separated from S-peptide by acid precipitation. In our hands, however, a subtilisin-digest of RNase A results in a complex mixture of products (Figure 3.2). S-protein can be isolated from such mixtures by using S-peptide-affinity chromatography (Gaynudtinov et al., 2003). This method requires an expensive custom-made column, and still suffers from the inefficient conversion of RNase A to RNase S. Furthermore, subtilisin is an ineffective means to generate RNase S from some RNase A variants (Markert et al., 2001) and homologues (Gaynudtinov et al., 2003). An alternative strategy could be the recombinant expression of S-protein. But because S-protein does not fold properly in the absence of S-peptide (Kato and Anfinsen, 1969), synthetic S-peptide must be added during the folding process which is wasteful and inefficient (Backer et al., 2002).
Our strategy takes advantage of the superior selectivity and general applicability of the protease enterokinase. Insertion of an enterokinase cleavage site into RNase variants enables complete cleavage of S-peptide from S-protein without unwanted side-products. This facilitates the facile separation of S-peptide and S-protein by HPLC. Historically, changes to the RNase S system have been restricted almost exclusively to S-peptide. This S-protein isolation methodology now makes S-protein variants more accessible, allowing manipulation and greater understanding of both parts of the RNase S system.

The noncovalent interaction of S-protein and S-peptide limits the use and complicates the characterization of RNase S complexes. RNase S is not active at low concentrations because RNase S has modest stability that is dependent on solution conditions (Schreier and Baldwin, 1976; Connelly et al., 1990). Further, the traditional RNase S system has less thermal stability ($\Delta T_m \sim 17^\circ$C) (Ratnaparkhi and Varadarajan, 2001), and less resistance to denaturants (Richards and Vithayathil, 1959) and proteases (Richards and Vithayathil, 1959; Allende and Richards, 1962; Nadig et al., 1996).

The covalent attachment of S-peptide to S-protein overcomes this intrinsic liability of the natural system. In particular, we demonstrate here that a nonnative disulfide bond between S-peptide and S-protein (linking residues 4 and 118) endows RNase–S with high enzymatic activity at low concentrations (Figure 3.6). We envision that this improvement will assist in the definition of structure–function relationships. Moreover, the disulfide bond in RNase–S could enable the discovery of higher affinity S-peptide ligands by the application of tethering strategies (Erlanson et al., 2000).
3.5 Experimental

3.5.1 Materials

DTNB, guanidine–HCl, ampicillin (sodium salt), isopropyl β-D-1-thiogalactopyranoside (IPTG), and dithiothreitol (DTT) were from Research Products International (Mount Prospect, IL). Enterokinase was from EMD Chemicals (Gibbstown, New Jersey). 6-carboxyfluorescein–dArU(dA)₂–6-tetramethylrhodamine was from Integrated DNA Technology (Coralville, IA). All other reagents, including subtilisin A (type VIII from bacillus licheniformis), RNase S, and RNase A were from Sigma–Aldrich (St. Louis, MO) and were used without further purification. Total yeast RNA was from Boehringer Mannheim (Damstadt, Germany).

3.5.2 Analytical methods

DNA was sequenced with a 3730xl DNA Analyzer (Applied Biosystems, Carlsbad, CA) at the Biotechnology Center of the University of Wisconsin–Madison. Peptides were synthesized with a Symphony (Protein Technologies, Tucson, AZ) automated synthesizer at the Biotechnology Center, and purified by semi-preparative HPLC using an UltiMate 3000 instrument (Dionex, Bannockburn, IL). Analytical HPLC was performed with a system from Waters (Milford, MA) equipped with two 515 pumps, a 717 plus autosampler, a 996 photodiode array detector, and a C-18 column from Varian (Palo Alto, CA). Non-kinetic ultraviolet/visible measurements were recorded with a Cary 50 spectrophotometer (Varian). An AKTA system (Amersham–Pharmacia, Piscataway, NJ) was used for fast protein liquid chromatography (FPLC), and the results were analyzed.
with the UNICORN Control System. A Voyager-DE-PRO Biospectrometry Workstation (Applied Biosystems) at the Biophysics Instrumentation Facility of the University of Wisconsin–Madison was used for matrix-assisted laser desorption ionization time of flight (MALDI–TOF) mass spectrometry. A Cary Bio400 spectrophotometer (Varian) and an Infinite M1000 microplate reader (Tecan, Männedorf, Switzerland) were used for kinetic assays.

3.5.3 Site-directed mutagenesis

Plasmids encoding variants of RNase A were generated from plasmid pBXR (delCardayré et al., 1995) by Quikchange site-directed mutagenesis (Stratagene, La Jolla, CA) using these oligonucleotides and their reverse complements from Integrated DNA Technology or the Biotechnology Center: DDDDK insertion between residues 20 and 21, GCACCTCCGCTGCCGATGATGATGATAAAAAAGCAGCTCCAACTAC; H12A, CCAAGTTTGAGCGGCAGGCTATGGACTCCAGCAGC; V118C, GCATCAAGTGACATGGGCACATACGGGTTCCC. All mutated genes were verified by DNA sequencing.

3.5.4 Production of ribonucleases

RNase A variants were produced and purified essentially as described elsewhere (Rutkoski et al., 2005). Briefly, transformed BL21(DE3) cells were grown at 37 °C with shaking in Terrific Broth containing ampicillin (400 μg/mL) until OD = 1.8–2.2 at 600 nm. Gene expression was induced by the addition of IPTG (0.5 mM). After 3–4 h, cells were harvested by centrifugation and lysed with a French pressure cell. Inclusion bodies
were recovered after centrifugation and solubilized for 2 h at room temperature with
denaturing solution (20 mM Tris–HCl buffer, pH 8.0, containing 7 M guanidine–HCl,
0.10 M DTT, and 10 mM EDTA). Solubilized inclusion bodies were then diluted 10-fold
with 20 mM HOAc. The precipitate was removed by centrifugation, and the solution was
dialeded overnight against 20 mM HOAc at 4 °C. After removal of further precipitate,
the ribonuclease solution was added dropwise to refolding buffer (0.10 M Tris–HCl
buffer, pH 7.8, containing 0.5 M l-arginine–HCl, 1.0 mM reduced glutathione, and 0.2
mM oxidized glutathione). After >3 days at 4 °C, the solution was adjusted to pH 5, and
concentrated using an Amicon YM10 membrane (Millipore, Billerica, MA). A 10-mL
sample was applied to a G75 gel filtration FPLC column (Amersham–Pharmacia). The
major peak after isocratic elution (50 mM sodium acetate buffer, pH 5.0, containing 0.10
M NaCl, 10 mM EDTA, and 0.02% w/v NaN₃) was collected and applied to a Mono S
cation-exchange FPLC column (Amersham–Pharmacia). Ribonucleases were eluted with
a linear gradient of NaCl (0–0.40 M) in 50 mM NaOAc buffer, pH 5.0, containing EDTA
(10 mM). Protein concentrations were determined by absorbance at 278 nm using \( e =
0.72 \text{ mg}^{-1}\cdot\text{mL}\cdot\text{cm}^{-1} \) (Sela et al., 1957a). The identity of each variant was verified by
MALDI–TOF mass spectrometry.

3.5.5 Protection of H12A/DDDDK/V118C RNase A with DTNB

After cation-exchange chromatography, the fractions that contained
H12A/DDDDK/V118C RNase A were combined, and the pH of the resulting solution
was increased by adding (to 8% v/v) 1.0 M Tris–HCl buffer, pH 8.3, containing EDTA
(10 mM). To this solution was added 50 mM Tris–HCl buffer, pH 8.0, containing DTNB
(5 mM) and EDTA (50 mM) such that the DTNB was in 4-fold molar excess to the protein. After incubation for 10 min, the solution was dialyzed overnight at 4 °C against 20 mM Tris–HCl buffer, pH 7.4, containing NaCl (50 mM).

3.5.6 Synthesis, activation with DTNB, and purification of peptides

S15 (KETAAAKFERQHMDS) and A4C S15 (KETCAAKFERQHMDS) were synthesized on a 50-μmol scale by standard fluorenylmethoxycarbonyl chemistry using HATU activation and an Fmoc-Ser(tBu) Wang resin (EMD Biosciences, La Jolla, CA). Peptides were deprotected and cleaved from the resin with 4 mL of 92.5:5:2.5 trifluoroacetic acid (TFA)/thioanisole/ethanediethiol for 4 h. Peptides were then precipitated with tert-buty methyl ether and dried under vacuum. S15 and A4C S15 were purified by reversed-phase HPLC on a C-18 semipreparative column. TFA (0.1%) was included as an ion-pairing agent in HPLC solvents. Peptides were eluted with a linear gradient of acetonitrile (10–30% v/v). One-quarter volume 0.10 M Tris–HCl buffer, pH 8.0, containing DTNB (5 mM) was added to fractions containing A4C S-peptide. The solvent was then removed under reduced pressure. NTB-activated A4C S-peptide was then purified again under the same HPLC conditions. The mass of purified peptides was confirmed by MALDI–TOF mass spectrometry.

3.5.7 Digestion of RNase A with subtilisin

Subtilisin (5 μg) was added to a solution of RNase A (0.5 mg, 88 μL) in phosphate-buffered saline (PBS) and incubated at 4 °C. Aliquots were removed over 24 h and subjected to SDS–PAGE (15% w/v acrylamide).
3.5.8 Digestion of ribonucleases with enterokinase

CaCl₂ (2 mM final concentration) was added to RNase A and its DDDDK and H12A/DDDDK/V118C variants (~5 mL of a 0.85 mg/mL solution) that had been dialyzed against 20 mM Tris–HCl buffer, pH 7.4, containing NaCl (50 mM). Enterokinase (5 μL, 5 U) was added, and the reaction mixture was incubated at room temperature. Aliquots were removed over 24 h and subjected to SDS–PAGE.

3.5.9 Separation of S-peptide and S-protein components

The S-peptide and S-protein fragment from the enterokinase-digestion of the DDDDK and H12A/DDDDK/V118C variants of RNase A were separated by analytical or semi-preparative reverse-phase HPLC with a C-18 column and two-step linear gradient (Step 1: 20–50% B over 3 min. Step 2: 50–100% B over 25 min. A: 50 mM sodium phosphate buffer, pH 2.7. B: 40% A + 60% acetonitrile) (Brems and Baldwin, 1984). Fractions containing S-protein were pooled and dialyzed overnight against PBS at 4 °C. The concentration of S-protein was determined by absorbance at 280 nm using ε = 9055 M⁻¹·cm⁻¹ (Gilmanshin et al., 1996). The identity of each peak on the HPLC trace was confirmed by MALDI–TOF mass spectrometry.

3.5.10 Fragment complementation

To V118C S-protein (100 μL of a 1.76 mg/mL solution) was added 1.0 M Tris–HCl buffer, pH 8.0, containing EDTA (10 mM) (8 μL), and DTT (25 mM) (2 μL). After 2 min, the mixture was desalted with a Zeba Spin column, 7K MWCO (Pierce, Rockford,
IL). To the collection tube was added S15 or NTB-activated A4C S15 (4 μL of 20 mg/mL). After ~10 min, the reactions were desalted again with the spin columns. SDS–PAGE and MALDI–TOF mass spectrometry confirmed the covalent linkage between A4C S15 and V118C S-protein (m/z 13324; expected: 13317).

3.5.11 Zymogram electrophoresis

The potassium salt of poly(cytidylic acid) (final concentration: 0.5 mg/mL) was included during the casting of a 15% w/v polyacrylamide gel. Laemmli buffer (no reducing agent) was added to each sample (2 ng). After SDS–PAGE, the gel was washed with isopropanol (20% v/v) in 10 mM Tris–HCl buffer, pH 7.0 (2×), 10 mM Tris–HCl buffer, pH 7.0 (2×), and 0.10 M Tris–HCl buffer, pH 7.5 (1×) for 10 min each. The gel was then stained for 10 min with 10 mM Tris–HCl buffer, pH 7.5, containing toluidine blue (0.02% w/v). The stained gel was rinsed several times in water, and soaked in water overnight.

3.5.12 Assays of ribonucleolytic activity

Initial velocities for catalysis of RNA cleavage were determined at high and low concentrations of RNase A, RNase S, and RNase–S in 50 mM NaOAc buffer, pH 6.0, containing NaCl (0.10 M). To avoid the contamination apparent in commercial preparations of RNase S (Figure 5, lane 6), the noncovalent complex assayed here was that of S15 and the deprotected S-protein derived from H12A/DDDDK/V118 RNase A.

A fluorogenic ribonuclease substrate, 6-carboxyfluorescein–dArU(dA)₂–6-tetramethylrhodamine (Kelemen et al., 1999), was used for assays at low ribonuclease
concentration. Briefly, a ribonuclease (50 μL of a 50 pM solution) was equilibrated at 37 °C in a 96-well plate. To each sample was added 6-carboxyfluorescein–dArU(dA)2–6-tetramethylrhodamine (50 μL of a 100 nM solution in sample buffer). Reaction progress was monitored at 37 °C by the increase in fluorescence emission at 515 nm upon excitation at 493 nm over 5 min.

A modified Kunitz assay was used at high ribonuclease concentration (Kunitz, 1946). Briefly, to total yeast RNA (100 μL of a 1 mg/mL solution) was added a ribonuclease (100 μL of a 0.29 μM solution). The decrease in absorbance at 300 nm was monitored at 37 °C for 10 min. Data from the first 2 min were used to determine initial velocities (v₀).

Acknowledgements and Dedication. This article is dedicated to the memory of Fred Richards (1925–2009). This work was supported by grants CA073808 and GM044783 (NIH). R.W.W. was supported by an NSF Graduate Research Fellowship and by Chemistry–Biology Interface Training Grant GM008505 (NIH).
Figure 3.1 Notional structure of "RNase-S", the covalent complex of A4C S15 and V118C S-protein. The image is based on the known structure of the noncovalent S15·S-protein complex (Taylor et al., 1981).
Figure 3.1

RNase-S

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
Figure 3.2  Analysis of the proteolytic digestion of ribonucleases with SDS–PAGE. Left, subtilisin-catalyzed cleavage of wild-type RNase A. Right, enterokinase-catalyzed cleavage of DDDDK RNase A and the wild-type enzyme.
Figure 3.2
Figure 3.3  Separation of S-peptide and S-protein components from DDDDK RNase A by reversed-phase HPLC.
Figure 3.3
Figure 3.4  Analysis of RNase-S semisynthesis with SDS-PAGE. Lane 1, RNase A; lane 2, RNase S; lane 3, S-protein derived from H12A/DDDDK/V118C RNase A; lane 4, A4C S-peptide + V118C S-protein of lane 3.
Figure 3.4
Figure 3.5  Analysis of RNase–S semisynthesis with zymogram electrophoresis. Lane 1, RNase A; lane 2, S-protein derived from H12A/DDDDK/V118C RNase A; lane 3, S15 + V118C S-protein of lane 2; lane 4, A4C S15 + V118C S-protein of lane 2; lane 5, RNase A; lane 6, commercial RNase S.
Figure 3.6  Initial velocities of RNA cleavage at high (0.15 μM) and low (25 pM) ribonuclease concentrations, relative to RNase A. “RNase S” refers to the noncovalent complex of S15 and the S-protein derived from H12A/DDDDK/V118 RNase A.
Figure 3.6
CHAPTER 4

Gene expression response of cancerous cells
to clinically relevant ribonucleases
4.1 Abstract

Pancreatic-type ribonucleases that evade the cytosolic ribonuclease inhibitor protein constitute a novel class of chemotherapeutic agents. These enzymes enter cancerous cells and destroy intracellular RNA, leading to cell death. Onconase, an amphibian ribonuclease, is in confirmatory Phase IIIb clinical trials and has been granted orphan-drug status for the treatment of malignant mesothelioma. A homologous human variant, QBI-139, is now in a Phase I clinical trial. Despite obvious similarities, mammalian ribonucleases and onconase differ in their substrate preferences, internalization efficiency and routing, conformational stability, and levels of catalytic activity. The extent to which these dissimilarities elicit different cellular responses is not well understood. Here, we use DNA microarrays and nucleic acid amplification techniques to investigate the transcriptional response of K-562 cells to clinically relevant ribonucleases. We find that onconase upregulates genes ATF3, IL-6, and TNFAIP3, and modifies genes associated with cytokine-cytokine interactions and the MAPK and JAK-STAT signaling pathways. We also observe increases in genes associated with apoptosis and the charging of tRNAs. Intriguingly, the most pronounced response in cells treated with QBI-139 is a relative increase of polyadenylated histone transcripts, many of which were chimeric histone mRNAs that were fused to unrelated RNA sequences encoded elsewhere in the genome.

4.2 Introduction

In the 1970’s, early frog embryo extracts were found to possess antitumoral activity (Shogen and Yoan, 1973). Later work attributed this activity to onconase, a pancreatic-
type ribonuclease (Ardelt et al., 1991). Onconase is cytostatic and cytotoxic towards tumor cells (Darzynkiewicz et al., 1988; Rybak et al., 1996; Juan et al., 1998; Halicka et al., 2000; Leland et al., 2000; Lee and Raines, 2003; Rodriguez et al., 2007) and, in mouse models, antagonizes xenograft tumor growth (Mikulski et al., 1990a; Rybak et al., 1996; Lee and Raines, 2003; Lee et al., 2007). The enzyme is now in confirmatory Phase IIIb clinical trials and has been granted orphan-drug status as a second-line therapy for patients with malignant mesothelioma.

Onconase and other cytotoxic ribonucleases cause cell death by degrading cellular RNA. A necessary feature of these ribonucleases is the ability to evade the cytosolic ribonuclease inhibitor protein (RI) that serves as a sentry to protect intracellular RNA. RI-evasive bovine and human variants have been created and shown to display potent cytotoxicity (Leland et al., 1998; Leland et al., 2001; Rutkoski et al., 2005; Johnson et al., 2007a). Notably, QBI-139 is an RI-evasive variant of human pancreatic ribonuclease (RNase 1) that is in a Phase I clinical trial with patients having advanced, refractory, solid tumors.

Onconase and cytotoxic mammalian ribonucleases share many features typical of members of the RNase A superfamily (Beintema, 1987). These cationic enzymes catalyze RNA degradation, share a similar tertiary structure and key active site residues (Figure 4.1), and are remarkably stable. Additionally, the same general mechanism has been proposed to explain their cytotoxic effects (Figure 4.2). Nevertheless, important differences remain. The ribonucleolytic activity of onconase with common substrates is ~5 orders of magnitude lower than that for mammalian homologues (Boix et al., 1996). Also, in vitro, onconase shows a unique preference for cleavage on the 5′ side of a
guanine nucleobase (Lee et al., 2008; Lee and Raines, 2008), and tRNA is preferentially degraded in cellulo (Saxena et al., 2001). Other evidence suggests that the cytotoxic effects of onconase may be due to degradation of small RNA substrates (Zhao et al., 2008). Finally, important details differentiate the cellular entry of onconase and mammalian ribonucleases. Onconase binds to a broad array of cell-surface glycans, whereas RNase A, a bovine homologue, binds with only low affinity. Further, the internalization of mammalian ribonucleases, but not that of onconase, correlates with cell anionicity (Chao et al., 2010).

To determine the extent to which these dissimilarities elicit a different biological response, we use DNA microarrays and nucleic acid-amplification techniques to investigate the gene expression response of cancerous cells to both onconase and QBI-139. We find distinct differences in the cellular response to an assault by these homologous cytotoxins.

4.3 Results and Discussion

4.3.1 Primary cellular response is due to degradation of intracellular RNA

To evaluate the ribonuclease-dependent gene expression response of K-562 cells, we added RNase 1, onconase, or QBI-139 to a final concentration of 0.15 μM for 24 h. Under these conditions, cell proliferation of onconase and QBI-139-treated cells (as measured by S-phase incorporation of [methyl-3H]thymidine into cellular DNA) is ~80–90% of the PBS control (Figure 4.3), and ribosomal RNA is largely intact (data not shown). This ribonuclease concentration was arbitrarily chosen because a cellular
response is apparent in the cell proliferation assay after 24 h (Figure 4.3). After isolation of the RNA and cDNA synthesis, gene expression changes were investigated by analysis with DNA microarrays. The data show that cells treated with RI-evasive ribonucleases elicit a much more pronounced gene-expression response than do those treated with RNase 1. ANOVA-derived $p$-value versus fold-change scatter plots ("volcano plots") show that treatment with onconase or QBI-139 results in both larger fold changes and increased statistical significance relative to RNase 1 (Figure 4.4). The more dramatic transcriptional response of RI-evasive ribonucleases strongly suggests that the primary cellular response to clinical ribonucleases results from the degradation of intracellular RNA.

4.3.2 Cellular response to onconase

There is notable agreement between previous microarray results of onconase-treated malignant mesothelioma cells (Altomare et al., 2010) and those reported here. In particular, three of the eight (37.5%) most highly onconase-regulated genes (>4-fold increase) in the earlier work were also differentially regulated by onconase in this study (>2-fold increase, $p$-value <0.05, Table 4.1). These genes have been implicated in tumor growth inhibition (ATF3) (Fan et al., 2002; Lu et al., 2006) and inflammation (IL-6, TNFAIP3) (Vereecke et al., 2009; Nishimoto, 2010). Further, we note that protein interaction networks derived from the onconase-regulated genes (Figure 4.5) highlight important nodes that correspond to the MAPK signaling (MAP2K1/2, ERK, Ras), cytokine-cytokine receptor interactions (IL1, interferon, interferon $\alpha/\beta$, interferon $\beta$, TNFAIP3, IL20RB), and JAK-STAT signaling pathways (STAT5a/b) that were
implicated previously in an onconase-dependent response (Altomare et al., 2010). Network and gene ontology analyses also reveal that onconase-regulated genes are associated with apoptosis \((p\text{-value } = 1.99 \times 10^{-5})\), and tRNA aminoacylation \((p\text{-value } = 5.3 \times 10^{-4})\). It is unclear if there is any association between the upregulation of genes associated with tRNA aminoacylation \((e.g.,\) tRNA synthetases) and the preferential cleavage of tRNA by onconase \textit{in vivo}. The selective cleavage of charged tRNAs by onconase would, however, explain why onconase preferentially degrades tRNA in both cells and cells lysates but not \textit{in vitro} with purified total RNA (Saxena et al., 2001). Additionally, onconase treatment leads to a modest apparent increase in histone mRNA (see below).

4.3.3 \textit{Histone mRNAs and other differences between onconase- and QBI-139-treated cells}  

Onconase and QBI-139 elicited remarkably different transcriptional responses, as only one gene, JUN (jun oncogene), was differentially upregulated (>2-fold difference in gene expression, \(p\text{-value } < 0.05\)) in both treatments (Table 4.1). Furthermore, network analyses of QBI-139 and onconase-regulated genes show that each ribonuclease produces a distinct cellular response (compare Figures 4.5 and 4.6). Most dramatically, 20 of the 26 genes that are differentially regulated by QBI-139 encode cell cycle-regulated histone proteins (Table 4.1 and Figure 4.7). This result was quite unexpected. While several replication-independent histone mRNAs are polyadenylated, cell cycle-regulated histone mRNAs generally lack poly(A) tails (Marzluff et al., 2008), so their inclusion in our oligo(dT) primer-derived cDNA library was not anticipated. To a much lesser extent,
onconase and RNase 1 treatment also resulted in an apparent increase in histone mRNA expression (Figure 4.7).

4.3.4 Confirmation by quantitative PCR

Quantitative PCR was used to confirm the microarray data and probe further the puzzling histone mRNA results. For amplification, we selected genes that microarray analysis showed were upregulated in either QBI-139 (HIST1H1E and HIST1H2AE) or onconase-treated cells (DDIT4 and FNDC6). Quantitative PCR using cDNA samples generated with random primers confirmed the microarray results for the onconase-regulated genes, but showed no increased expression of HIST1H1E or HIST1H2AE mRNAs. Yet, with cDNA generated using an oligo(dT) primer, quantitative PCR not only confirmed the results of the onconase-regulated genes, but also showed a pronounced increase in histone mRNA in QBI-139-treated cells (see Figure 4.8). The ratio of non-polyadenylated to polyadenylated histone mRNA was estimated to be 16:1–50:1 by comparison of apparent histone mRNA expression in QBI-139 and onconase-treated cells relative to the control GAPDH mRNA. These data, combined with the microarray results, strongly suggest that treatment with QBI-139 leads to a significant increase in polyadenylated histone mRNA, which is, nonetheless, a small fraction of the overall histone mRNA population.

4.3.5 3' rapid amplification of cDNA ends

To better understand the 3' processing of polyadenylated histone mRNAs, we amplified the 3' end of HIST1H1E, HIST1H2AB, and HIST1H2AE mRNAs using 3'
rapid amplification of cDNA ends (RACE). Surprisingly, 3' RACE revealed that a significant number of the polyadenylated histone mRNAs were chimeric—the RNA sequence of the histone mRNA was fused to a different, unrelated RNA molecule encoded elsewhere in the genome (Figure 4.6). More specifically, the following RNAs from QBI-139-treated cells were fused to histone transcripts as identified by sequencing: RPLP1 (NM_001003.2, Chromosome (Chr.) 15), MRPL52 (NM_181306.2, Chr. 14), INTS4 (NM_033547.3, Chr. 11), NDUFB9 (NM_005005.2, Chr. 8), NASP (NM_172164.1, Chr. 1), LOC100128191 (NR_027157.1, Chr. 12), RPL41 (NM_021104.1, Chr. 12), and TIGD5 (NM_032862.3, Chr. 1). From both untreated and QBI-139-treated cells we found histone mRNAs fragments that were polyadenylated without any evidence of trans-splicing. Additionally, chimeric histone mRNAs were also identified from the PBS-treated control: HINT1 (NR_024611.1, Chr. 5), EIF4A1 (NM_001416.2, Chr. 17), MED25 (NM_030973.2, Chr. 19), TXNL1 (NM_004786.2, Chr. 18), HNRNPM (NM_005968.3, Chr. 19), OAZ1 (NM_004152.2, Chr. 19), LOC100130107 (XM_001718888.2, Chr. 15), and KHDRBS1 (NM_006559.1, Chr. 1).

4.3.6 Interpretation

Taken together, the data hint that exposure to QBI-139 leads to an increase in trans-spliced polyadenylated histone mRNAs. It is not known if this increase is histone-specific or if QBI-139 treatment leads to a global increase in trans-splicing. Since trans-splicing occurs at a relatively low level, the poly(A) tails of conventional mRNAs contribute far too much background noise to detect a global increase in trans-splicing using the facile detection methods that we use here. Furthermore, the data from the
untreated cells suggest that there is a low abundance of polyadenylated histone mRNAs under normal conditions.

The mechanism by which ribonuclease-dependent trans-splicing occurs is not clear. We are skeptical of spliceosome involvement, as these chimeric transcripts lack the traditional cis elements involved in spliceosome-mediated splicing (Walsh and Yang, 2005; Wang and Burge, 2008). It is tempting, however, to speculate that endonucleolytic cleavage of mRNA transcripts by QBI-139 results in RNA fragments that are subsequently joined by endogenous ligases. This mechanism is analogous to the nonspliceosomal IRE1p- and Methanococcus jannaschii endoribonuclease-dependent splicing of the HAC1 mRNA and bulge-helix-bulge-containing tRNAs (Gonzalez et al., 1999; Deidda et al., 2003). Several candidate RNA ligases have been identified in mammalian cells (Deidda et al., 2003).

4.4 Conclusions

Here, we show that the gene expression response of K-562 cells to ribonucleases is due predominantly to the degradation of intracellular RNA. Also, our data generally confirm previous findings of the cellular response to onconase. The gene expression response to QBI-139 is, however, markedly different. Treatment with the RI-evasive mammalian ribonuclease may lead to an increase in chimeric histone mRNAs. Although the mechanism for this process is not well understood, the RNA-degrading ability of QBI-139 may lead to increased concentrations of mRNA fragments that are subsequently ligated within cells.
Recent work in C. elegans revealed the presence of polyadenylated histone mRNAs prompting speculation that histone polyadenylation may occur in higher eukaryotes as well (Mangone et al., 2010). Our data here confirm that hypothesis, demonstrating the expression of polyadenylated cell-cycle-dependent histone mRNAs in human cells.

Finally, our work suggests that RI-evasive mammalian ribonucleases may be useful tools to increase the abundance of chimeric RNA transcripts and to help understand the mechanisms of trans-splicing. The isolation of chimeric histone mRNAs using oligo(dT) primers is facile, making histone mRNAs prime candidates for trans-splicing studies. We anticipate that this work will prompt a more in-depth investigation of whether histone mRNAs or the RNA fragments fused to the 3' end of histone mRNA fragments are privileged trans-splicing substrates, or whether the proposed cleavage/ligation activity operates stochastically. Further research should also reveal whether this response is confined to particular cell types and establish the kinetics and concentration dependence of ribonuclease-induced histone trans-splicing.

4.5 Experimental

4.5.1 Materials

E. coli BL21(DE3) cells were from Novagen (Madison, WI). K-562 and DU145 cells were obtained from the American Type Culture Collection (Manassas, VA). Cell culture media and supplements were from Invitrogen (Carlsbad, CA) or ATCC. [methy1-

$^3$H]Thymidine (6.7 Ci/mmol) was from Perkin–Elmer (Boston, MA). QBI-139 was a gift from Quintessence Biosciences (Madison, WI). TaqMan gene expression assays and PCR
Master Mix (No AmpErase) were from Applied Biosystems (HIST1H1E: Hs00271195_s1; HIST1H2AE: Hs00368307_s1; DDIT4: Hs00430304_g1; FNDC6: Hs01023484_m1; GAPDH: 4352934E). Oligo(dT)12-18 primer was from Invitrogen (Carlsbad, CA). All other oligonucleotides were obtained from the Biotechnology Center of the University of Wisconsin–Madison. Phase Lock tubes (light, 1.5 mL) were from 5 PRIME (Gaithersburg, MD). All other chemicals used were of commercial reagent grade or better, and were used without further purification.

4.5.2 Instrumentation

Cellular DNA for the cytotoxicity assay was isolated using a PHD cell harvester (Cambridge Technology). A Microbeta TriLux liquid scintillation counter (Perkin–Elmer, Wellesley, MA) was used to quantify [methyl-3H]thymidine incorporation. A Voyager-DE-PRO Biospectrometry Workstation (Applied Biosystems, Foster City, CA) at the University of Wisconsin-Madison Biophysics Instrumentation Facility was used for matrix-assisted laser desorption ionization time of flight (MALDI–TOF) mass spectrometry. RNA and cDNA concentrations were determined using a Nanodrop ND–1000 instrument. DNA sequences were determined with a 3730xl DNA Analyzer (Applied Biosystems, Carlsbad, CA) at the Biotechnology Center of the University of Wisconsin–Madison.

4.5.3 Ribonuclease production
Human RNase 1 and onconase were produced, purified, and quantified as described elsewhere (Leland et al., 2001). The identity of each was verified by MALDI–TOF mass spectrometry.

4.5.4 Assay of cytotoxicity

The effect of ribonucleases on cell proliferation was determined by measuring [methyl-\(^3\)H]thymidine incorporation into the DNA of asynchronous log-phase K-562 cells. The assay was carried out as described previously (Leland et al., 1998; Rutkoski et al., 2005) with the following modifications. Cells were delivered to wells at an initial concentration of \(1 \times 10^5\) cells/mL. After a 24-h incubation with the ribonucleases, cells were pulsed with radiolabeled thymidine for 4 h. Cellular DNA was recovered and [methyl-\(^3\)H]thymidine incorporation was measured by scintillation counting. Results show the percentage of [methyl-\(^3\)H]thymidine incorporation compared to PBS controls and represent the average of at least three measurements at each concentration. Values for IC\(_{50}\) were calculated as described previously (Rutkoski et al., 2005).

4.5.5 Cell propagation

Cell culture medium was supplemented with fetal bovine serum and antibiotics (streptomycin, 100 \(\mu\)g/mL; penicillin, 100 U/mL). K-562 cells were propagated in RPMI 1640.

4.5.6 Microarray analysis

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
4.5.6.1 Isolation of RNA from cells exposed to ribonucleases

Ribonucleases were added to K-562 cells (25 mL, 1.4 × 10^6 cells/mL) to a final concentration of 0.15 μM and incubated at 37 °C in a humidified atmosphere containing 5% CO₂ (v/v). After 24 h, cells were harvested by centrifugation, and total RNA from PBS-, RNase1-, onconase-, and QBI-139-treated cells (one batch) was isolated using the RNeasy Mini Kit (QIAGEN, Germantown, MD). RNA from three batches was isolated on separate days. For all samples, both \( A_{260}/A_{280} \) and \( A_{260}/A_{230} \) were greater than 1.8. RNA integrity was verified by agarose electrophoresis (1% w/v).

4.5.6.2 First-strand cDNA synthesis

cDNA was produced using the Invitrogen SuperScript Double-Stranded cDNA synthesis kit. Briefly, oligo(dT)\(_{12-18}\) primer (1 μL, 0.5 μg) was added to 10 μg RNA. After addition of water to a final volume of 11 μL, the sample was heated at 70 °C for 10 min. After cooling on ice, 4 μL 5X first-strand buffer, 2 μL 0.1 M DTT, and 1 μL 10 mM dNTP mix were added, and the resulting solution was heated to 42 °C for 2 min. Then, SuperScript II (400 U) was added, and the reaction mixture was incubated again at 42 °C for 1 h.

4.5.6.3 Second-strand cDNA synthesis

The following were added to each first strand cDNA synthesis reaction: 91 μL DEPC-treated water, 30 μL 5X second-strand buffer, 3 μL 10 mM dNTP mix, 1 μL (10 U) DNA ligase, 4 μL (40 U) DNA polymerase I, and 1 μL (2 U) RNase H. After incubation at 16 °C for 2 h, T4 DNA polymerase was added (10 U), and the reaction
mixture was incubated at 16 °C for an additional 5 min. RNA was degraded by incubating the reaction mixture at 37 °C with 1 μg of RNase A. cDNA was isolated by phenol:chloroform:isoamyl alcohol extraction using Phase Lock tubes. After ethanol precipitation, cDNA was resuspended in 20 μL water. cDNA was quantified using a NanoDrop ND–1000 instrument. For all samples, A_{260}/A_{280} and A_{260}/A_{230} were both greater than 1.8.

4.5.6.4 DNA labeling, hybridization and data extraction

DNA labeling, hybridization to Nimblegen HG18_60mer_expr arrays, data extraction, and pre-processing of hybridization signals were performed by Nimblegen (Madison, WI). All data manipulation was done with Partek Genomics Suite Software. Data were normalized by quantile normalization (Bolstad et al., 2003), and gene calls were made using the RMA (robust multichip average) algorithm (Irizarry et al., 2003). Ribonuclease-regulated genes were identified by applying a filter (>2-fold gene expression change and unadjusted p-value <0.05) to an ANOVA analysis.

4.5.6.5 Gene ontology and pathway analysis

Gene ontologies that were responsive to QBI-139 and onconase were identified using the gene ontology ANOVA feature of the Partek software suite. Pathway analyses (Figures 4.5 and 4.6) were generated by importing QBI-139- and onconase-responsive gene lists into Ingenuity Pathways Analysis web-based software analysis program.

4.5.7 Quantitative real-time PCR
4.5.7.1 RNA isolation

Ribonucleases were added to K-562 cells (25 mL, $1.4 \times 10^6$ cells/mL) to a final concentration of 0.15 µM and incubated at 37 °C in a humidified atmosphere containing 5% CO$_2$ (v/v). After 24 h, cells were harvested by centrifugation, and total RNA was isolated using the RNeasy Mini kit (QIAGEN, Germantown, MD). For all samples, both A$_{260}$/A$_{280}$ and A$_{260}$/A$_{230}$ were greater than 1.8.

4.5.7.2 cDNA synthesis for quantitative real-time PCR

cDNA was synthesized from 1.5 µg total RNA using the high capacity cDNA reverse transcription kit (Applied Biosystems) according to the manufacturer’s protocol. For oligo(dT)-primed reverse transcription, 1 µg oligo(dT)$_{12-18}$ primer (Invitrogen) was used in place of the kit-supplied random primer.

4.5.7.3 Quantitative real-time PCR

All quantitative PCR reactions used TaqMan® technology and were processed according to the manufacturer’s protocol on a 7500 Fast Real-Time PCR System instrument (Applied Biosystems). All samples were denatured (95 °C for 10 min) and then cycled 40 times between 95 °C (15 s) and 60 °C (1 min). Data acquisition was performed with Sequence Detection Software v.1.3 (Applied Biosystems, Foster City, CA). For each assay, four technical replicates for each of three biological replicates were run. Relative gene expression was calculated using the comparative C$_T$ method.
(Schmittgen and Livak, 2008), and error bars represent the standard error of the biological replicates.

4.5.8 3' Rapid amplification of cDNA ends

The 3' ends of histone genes were investigated using the classic 3' rapid amplification of cDNA ends (RACE) protocol (Scotto-Lavino et al., 2007). cDNA was generated from 5 µg total RNA (isolated as for quantitative PCR) using the Q† primer (Scotto-Lavino et al., 2007) to amplify only poly(A)-containing RNAs. The following gene-specific primers were used in consecutive amplification steps: HIST1H1E, ATGTCGGAGACTGCGCTGCGG, GCCGGCCGAGAAGACTCCCGTGAAG, and TCCAGAGCTATTACTAAAGCT; HIST1H2AB, TGCAGTTCTCTCTGCGGCGGA, TATCTCGCCGGTGCTTTG; HIST1H2AE, TGCTGTTAGGAAGCCTATGTCTG, AAGCCACTATGTCTGGACGTGAAA. 3' RACE was carried out as specified except (Scotto-Lavino et al., 2007) that Pfu Turbo (Stratagene) or GoTaq (Promega) polymerases were used for amplification steps. The annealing temperature for all amplification steps was 60 °C. Following the last amplification step, DNA was isolated by using the Wizard SV gel and PCR clean-up system (Promega). DNA was then cloned into a pGEM-T Easy vector (Promega) and sequenced by using the following oligonucleotide primers: HIST1H1E, GCCGCCTCGGTTTCTTCAA; HIST1H2AB, CATCGGCAGGGTGGCGTTT, HIST1H2AE, CCTGGCAGCGGTGCTGGAATAT. Pfu Turbo-amplified DNA was "A-tailed" with GoTaq polymerase before cloning according to manufacturer's instructions (Promega).
Acknowledgements. We thank Brian Yandell, Allan Attie, Mark Keller, Aimee Broman, and Jean-Yves Sgro for experimental and statistical guidance. This work was supported by grants CA073808 and GM044783 (NIH). R.W.W. was supported by an NSF Graduate Research Fellowship and by Chemistry–Biology Interface Training Grant GM008505 (NIH).
Table 4.1  Genes with >2-fold change and p-values <0.05 relative to PBS controls.

<table>
<thead>
<tr>
<th><strong>OBI-139-responsive genes</strong></th>
<th><strong>Onconase-responsive genes</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gene Symbol</strong></td>
<td><strong>Gene description</strong></td>
</tr>
<tr>
<td>C10orf10</td>
<td>chromosome 10 open reading frame 10</td>
</tr>
<tr>
<td>HIST1H1B</td>
<td>histone cluster 1, H1b</td>
</tr>
<tr>
<td>HIST1H1D</td>
<td>histone cluster 1, H1d</td>
</tr>
<tr>
<td>HIST1H1E</td>
<td>histone cluster 1, H1e</td>
</tr>
<tr>
<td>HIST1H2AB</td>
<td>histone cluster 1, H2a</td>
</tr>
<tr>
<td>HIST1H2AE</td>
<td>histone cluster 1, H2ae</td>
</tr>
<tr>
<td>HIST1H2AH</td>
<td>histone cluster 1, H2ah</td>
</tr>
<tr>
<td>HIST1H2AM</td>
<td>histone cluster 1, H2am</td>
</tr>
<tr>
<td>HIST1H3B</td>
<td>histone cluster 1, H3b</td>
</tr>
<tr>
<td>HIST1H3F</td>
<td>histone cluster 1, H3f</td>
</tr>
<tr>
<td>HIST1H3J</td>
<td>histone cluster 1, H3j</td>
</tr>
<tr>
<td>HIST1H4A</td>
<td>histone cluster 1, H4a</td>
</tr>
<tr>
<td>HIST1H4B</td>
<td>histone cluster 1, H4b</td>
</tr>
<tr>
<td>HIST1H4K</td>
<td>histone cluster 1, H4k</td>
</tr>
<tr>
<td>HIST2H2AB</td>
<td>histone cluster 2, H2ab</td>
</tr>
<tr>
<td>HIST2H2AC</td>
<td>histone cluster 2, H2ac</td>
</tr>
<tr>
<td>HIST2H3C</td>
<td>histone cluster 2, H3c</td>
</tr>
<tr>
<td>HIST2H3A</td>
<td>histone cluster 2, H3a</td>
</tr>
<tr>
<td>HIST2H4</td>
<td>histone cluster 2, H4a</td>
</tr>
<tr>
<td>ID2B</td>
<td>inhibitor of DNA binding 2B</td>
</tr>
<tr>
<td>ID3</td>
<td>inhibitor of DNA binding 3</td>
</tr>
<tr>
<td>JUN</td>
<td>jun oncogene</td>
</tr>
<tr>
<td>LOC200261</td>
<td>hypothetical protein LOC200261</td>
</tr>
<tr>
<td>BC017361</td>
<td>histone cluster 1, H4j</td>
</tr>
<tr>
<td>AB088847</td>
<td>BPA-1 mRNA for brain peptide A1</td>
</tr>
<tr>
<td>RP5-998N21.6</td>
<td>histone cluster 2, H3, pseudogene 2</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.1  Crystalline structures of (A) onconase (PDB entry 1onc) (Mosimann et al., 1994) and (B) human RNase 1 (PDB entry 1z7x) (Johnson et al., 2007b). Key catalytic residues are shown in red. Images were generated with MacPyMOL software (DeLano Scientific, South San Francisco, CA).
Figure 4.2  Mechanism of ribonuclease cytotoxicity. Ribonucleases bind to the cell surface, are internalized by endocytosis, and translocate across the lipid bilayer to enter the cytosol. RI-evasive RNases degrade intracellular RNA leading to cell death.
Figure 4.2

- cell-surface binding
- endocytosis and translocation of the lipid bilayer
- RI protects cells from RNA degradation
- RI evasion and cell death

= ribonuclease

= ribonuclease inhibitor
Figure 4.3  Effect of ribonucleases on the proliferation of K-562 cells. Cells were incubated in the presence of ribonucleases for 24 h, followed by a 4 h pulse with [methyl-\textsuperscript{3}H]thymidine (0.20 μCi/well). Results show the percentage of [methyl-\textsuperscript{3}H]thymidine incorporation compared to PBS controls and represent the average of at least three measurements at each concentration.
Figure 4.3

- RNase 1
- onconase
- QBI-139

Cell proliferation (% of control)

μM ribonuclease

Onconase IC$_{50}$ = 0.37 μM
QBI-139 IC$_{50}$ = 0.34 μM
Figure 4.4 ANOVA-derived $p$-value versus fold-change scatter plots ("volcano plots") for ribonuclease-treated cells. Each data point corresponds to a specific gene probeset.
Figure 4.4

A  QBI-139 v. PBS

B  onconase v. PBS

C  RNase 1 v. PBS

D  QBI-139 v. onconase
Figure 4.5  Protein interaction network of genes differentially regulated by onconase. Red shapes denote genes that were upregulated >2-fold relative to the PBS control and were statistically significant (ANOVA-derived p-value <0.05). The diagram was generated by entering the onconase-responsive gene list (>2-fold upregulation and ANOVA-derived p-value <0.05) into Ingenuity Pathway Analysis software.
Figure 4.5
Figure 4.6  Protein interaction network of genes differentially regulated by QBI-139. Red shapes denote genes or gene families that were upregulated >2-fold relative to the PBS control and were statistically significant (ANOVA-derived p-value <0.05). The diagram was generated by entering the QBI-139-responsive gene list (>2-fold upregulation and ANOVA-derived p-value <0.05) into Ingenuity Pathway Analysis software.
Figure 4.6
Figure 4.7  ANOVA-derived $p$-value versus fold-change scatter plots ("volcano plots") for ribonuclease-treated cells. Blue triangles represent histone mRNA probe sets.
Figure 4.8  Quantitative real-time PCR from cDNA constructed using random and oligo(dT) primers. Microarray data from cDNA derived from oligo(dT) primers suggests that DDIT4 and IL-20 are upregulated in onconase-treated cells, while HIST1H1E and HIST1H2AE were upregulated in QBI-139-treated cells.
Figure 4.8

The figure shows bar graphs comparing the fold-change of various gene expressions under different ribonuclease treatments using random primers and oligo(dT) primers. The genes compared include hRNase 1, onconase, and QBI-139, with the following treatments:

- **hRNase 1**
- **onconase**
- **QBI-139**

The graphs are color-coded as follows:

- DDIT4
- IL-20
- HIST1H1E
- HIST1H2AE

The y-axis represents the fold-change, ranging from 0 to 4.
Figure 4.9  Sequences of cDNA derived from chimeric histone mRNAs as identified by 3' RACE. RNA was isolated from PBS- and QBI-139-treated cells.
CHAPTER 5

Future directions
5.1 Additional utility of membrane-permeant affinity labels

In Chapter 2, we demonstrated the use of a urea-rhodamine affinity label that enables the facile, rapid labeling of a HaloTag\textsuperscript{®} fusion protein in live cells.

The rapid labeling now possible with probe 1 should prove useful in pulse-chase experiments, which require differential labeling of temporally disparate pools of protein. Probe 1 may allow increased temporal resolution when used in conjunction with a cell-permeable red fluorophore that is already available commercially (HaloTag\textsuperscript{®} TMRDirect\textsuperscript{TM} Ligand, Promega).

Additionally, the utility of this probe can be expanded to include other affinity labeling systems. Because the synthesis of the urea-rhodamine label is modular, a separate bioreactive handle may be used in place of the chloroalkane. In particular, a diverse assortment of amines can be conjugated to probe 10 with ease.

This attribute may prove especially advantageous in the SNAP- and CLIP-tag labeling systems (Keppler et al., 2004) that allow for two separate fusion proteins to be labeled in a single cell.

5.2 Further investigations of trans-splicing

In Chapter 4, we demonstrated that treatment of K-562 cells with QBI-139 results in an apparent increase of polyadenylated histone mRNAs. This observation may have been possible only because eukaryotic histone mRNAs are not usually polyadenylated (Marzluff et al., 2008) (although there is evidence of polyadenylated histone mRNAs in Caenorhabditis elegans (Mangone et al., 2010)) and thus a small increase in trans-splicing resulted in a relatively large increase of polyadenylated histones. Of note,
however, our experiments do not distinguish between a global increase in *trans*-splicing and a histone-specific increase.

RNA-seq (Wang *et al.*, 2009) is likely the best tool to distinguish between a histone-specific and a global increase in *trans*-splicing. RNA-seq could provide us with nucleotide resolution, whereas our microarray analysis is limited by the decreased resolution provided by hybridization techniques. These data may distinguish definitively between global and histone-specific changes.

If QBI-139 treatment results in a large increase in global *trans*-splicing, mammalian ribonucleases may become very useful reagents in *trans*-splicing experiments.

Interestingly, in K-562 cells, equal concentrations of QBI-139 and onconase did not elicit similar transcriptional responses, although they behave nearly identically in a cell-proliferation assay. The reasons for this disparity are unclear, but several experiments may shed additional light.

The different cellular responses to QBI-139 and onconase may be due to differing kinetics. That is, the cellular response to onconase at one timepoint may resemble more closely the cellular response to QBI-139 at a later point of time (or vice versa). Microarray, qPCR, and RNA-seq analysis of RNA harvested at various timepoints after treatment may help explain the existing disparity. Analogous experiments varying ribonuclease concentration may also provide increased clarity.

Also, it is not yet clear that all tumor cells will respond in the same fashion as K-562 cells. A repetition of the work in Chapter 4 with different cancer cell lines would provide important information about the generality of this response.
Likely, further research into the routing, translocation, and intracellular substrate specificity will likely be needed to explain fully the different cellular responses to onconase and QBI-139 treatment.
APPENDIX

Search for protein interaction partners
of ribonuclease inhibitor protein
A.1 Abstract

Ribonuclease inhibitor protein binds adventitiously internalized ribonucleases with 1:1 stoichiometry and renders them inactive. Although the role of ribonuclease inhibitor protein as an intracellular sentry is well established, several observations suggest additional biological roles. Here, I describe attempts to shed light on the biological role of ribonuclease inhibitor protein by determining additional protein interaction partners.

A.2 Introduction

Mammalian ribonuclease inhibitor protein (RI) is a 50-kDa cytosolic protein that binds secretory ribonucleases in a 1:1 ratio with very high affinity ($K_d \sim 10^{-15}$ M) and renders them inactive (Hofsteenge, 1997; Dickson et al., 2005). RI is composed almost entirely of leucine-rich repeats (Kobe and Deisenhofer, 1994), which give it a remarkable horseshoe shape (Papageorgiou et al., 1997). These repeats are known to foster protein–protein interactions by displaying large surface areas for protein binding (Kobe and Deisenhofer, 1994).

The biological role(s) of RI are not understood fully. Clearly, ribonuclease inhibitor protein acts as an intracellular sentry (Haigis et al., 2003; Dickson et al., 2005), protecting cytosolic RNA from degradation by adventitiously internalized secretory ribonucleases. This role, however, does not fully account for several observations. First, RI is found exclusively within cells, yet all known high affinity ligands for RI are secreted proteins. This paradox, in conjunction with the relatively high concentration ($\sim 4 \mu$M, $\sim 0.1\%$ of cytosolic protein) of RI (Haigis et al., 2003) within the cytosol (much higher than would be anticipated to merely inactivate adventitiously internalized
ribonucleases) suggests additional role(s) (Rutkoski and Raines, 2008). Moreover, RI is found in all cell types studied to date, even red blood cells, which are essentially devoid of RNA to protect (Nadano et al., 1995; Moenner et al., 1998). Furthermore, the interaction of RI with the neovascularization-inducing ribonuclease angiogenin is one of the tightest known interactions in all of biology (Chen and Shapiro, 1997), yet an anti-angiogenic role for ribonuclease inhibitor protein, although documented (Dickson et al., 2009), is not well understood. Finally, RI has an uncommonly high percentage of cysteine residues and is quite susceptible to oxidative inactivation through cooperative disulfide bond formation (Fominaya and Hofsteenge, 1992; Kim et al., 1999).

Several alternative/additional roles have been proposed to account for these observations. Some have advanced that RI monitors the oxidative state of a cell and/or protects it from oxidation (Blázquez et al., 1996; Cui et al., 2003; Monti et al., 2007). Others have suggested that RI antagonizes angiogenesis (Dickson et al., 2005; Dickson et al., 2009). The presence of RI in erythrocytes has also led to the hypothesis that RI plays a role in the differentiation of red blood cells (Moenner et al., 1998).

Clearly, further investigation is needed to decipher the true function(s) of RI within a cell. Determining the protein interaction partners of ribonuclease inhibitor protein may help answer important unanswered questions: Why has the $K_d$ value of the RI·ribonuclease complex evolved to be so low? What is the fate of the bound complex? Is RI localized to a specific location within the cytosol? Why is RI so sensitive to oxidation? I anticipate that answers to these and related questions will both further basic science and help in the development of ribonuclease-based chemotherapeutic agents.
interest is fused to the DNA-binding domain of a transcription factor (known as the "bait" protein), and a library of potential interaction partners is fused to the activation domain of the transcription factor ("prey" protein).

A prey vector is then introduced into a yeast strain containing the bait vector. Binding of the prey protein to the bait protein brings together the activation and DNA-binding domains of the transcription factor and activates expression of a downstream reporter gene (Figure A.1).

I submitted the cDNA sequence of human RI to the Molecular Interaction Facility (MIF) at UW–Madison for a yeast two-hybrid screen (Table A.1). Human RI bait fused to the GAL4 DNA binding domain was screened against a combined brain/testes cDNA library (human library B). Four proteins were identified as putative interaction partners (Table A.1).

A.4 Experimental

A.4.1 Crosslinking with photo-leucine

L-Photo-leucine was synthesized according to a published procedure (Suchanek et al., 2005). The amino acid was incorporated into HeLa cells by incubating 4 μM L-photo-leucine in DMEM devoid of leucine for 24 h. Cells were then placed in a photochamber and exposed to UV light for 20 min. RI was subsequently immunoprecipitated from HeLa cell lysates and then electrophoresed. After silver staining of the gel, isolated bands were submitted for trypsin digestion and mass-spectrometry. However no human
proteins were identified by mass spectrometry, perhaps due the large background of antibody proteins.

A.4.2 Yeast two-hybrid screen

A yeast two-hybrid (Y2H) screen (Fields and Song, 1989) was performed in the Molecular Interaction Facility (MIF) at the University of Wisconsin–Madison according to their protocols. MIF uses yeast strains developed by Phillip James (James et al., 1996). A combined brain/testes human library is in pACT and pACT2 (Durfee et al., 1993) or pGAD-T7Rec (Clonetech, Mountain View, CA) prey vectors.

DNA encoding the full-length sequence of the human ribonuclease inhibitor protein was cloned in-frame with the GAL4 DNA-binding domain of bait vector pBUTE (a kanamycin-resistant version of GAL4 bait vector pGBDUC1 (James et al., 1996)). The resulting bait vector was sequenced to confirm an in-frame fusion, then transformed into mating type A of strain PJ694 and tested for autoactivation of the β-galactosidase reporter gene.

The Y2H screen was conducted using a human brain/testes cDNA library. Approximately 18 million clones were screened via mating. Controls included medium only, empty prey (activation domain) vector (pGADC1) (James et al., 1996), prey construct expressing mouse epsin, prey construct expressing human Fbox3, and two premated interaction pairs: SH3:SOS and EH:Epsin. Following selection, nine yeast wells tested positive for interaction (via selection on histidine drop-out medium containing 1 mM 3-amino-1,2,4-triazole (3AT) and a β-galactosidase assay). Prey plasmids were re-transformed into the alpha mating type of PJ694 and validated in a
A.3 Results and Discussion

A.3.1 Crosslinking of RI-interaction partners with a photoactivatable leucine analogue

Since RI contains many leucine residues, I synthesized a photoactivatable leucine analogue (Suchanek et al., 2005) that can be incorporated naturally into the proteins of live cells to enable identification of protein interaction partners by photocrosslinking. The photoactivatable amino acid was added to mammalian cells in culture medium devoid of leucine. Under such conditions, the amino acid escapes the stringent identity control of the mammalian translation machinery and is incorporated naturally into cellular proteins (Suchanek et al., 2005). I then exposed the cells to UV light, which generates a reactive carbene that crosslinks protein interaction partners. Control experiments showed that the photo-leucine reagent crosslinks the dimerized EEA1 (Simonsen et al., 1998; Suchanek et al., 2005) protein by immunoblot.

In an attempt to identify proteins that interact with RI, I irradiated photo-leucine treated cells with UV light and then immunoprecipitated the RI. The immunoprecipitate was electrophoresed and silver stained. Although several bands were visible, attempts to identify the crosslinked proteins by mass spectrometry (in-gel digest) were unsuccessful. This reagent may, however, prove useful to confirm the interaction of RI with candidate interaction partners identified by other means.

A.3.2 Yeast two-hybrid screen with RI as bait

Yeast two-hybrid screening uses the modular nature of transcription factors to discover protein-protein interactions (Fields and Song, 1989). In general, the protein of
**Table A.1**  Putative protein-interaction partners of RI identified from a yeast two-hybrid screen.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Blast ID</th>
<th>Putative function</th>
</tr>
</thead>
<tbody>
<tr>
<td>GABARAPL2</td>
<td>NM_007285.6, Homo sapiens GABA (A) receptor-associated protein-like 2</td>
<td>Marker of autophagy</td>
</tr>
<tr>
<td>Clusterin</td>
<td>NM_0018321.2, Homo sapiens clusterin (CLU), transcript variant 1, mRNA</td>
<td>Chaperone</td>
</tr>
<tr>
<td>RCC2</td>
<td>NM_018715.1, Homo sapiens regulator of chromosome condensation 2 (RCC2) mRNA</td>
<td>Guanine nucleotide exchange factor that helps kinetochore associate with microtubules</td>
</tr>
<tr>
<td>SCARF1</td>
<td>NM_003693.2, Homo sapiens scavenger receptor class F, member 1 (SCARF1)</td>
<td>Membrane receptor that internalizes low density lipoproteins</td>
</tr>
</tbody>
</table>
REFERENCES
Figure A.1  Yeast two-hybrid screen with RI as bait.
Advanced Chemical Development. www.acdlabs.com/physcemsuite. 29 April 2009


parallel mating and selection assay with the human ribonuclease inhibitor protein bait and
the empty bait vector. Four clones were positive, that is, they grew in interaction
selection medium and exhibited β-galactosidase activity, and subsequently were
identified by sequencing.

β-galactosidase activity was assayed in yeast culture lysed by the addition of yeast
protein extraction reagent (Pierce Biotechnology, Rockford, IL) combined with
chlorophenylred-β-D-galactopyranoside (Roche, Mannheim, Germany) as a substrate.

A.4.3 Attempts to confirm GABARAPL2-RI interaction

Preliminary immunoprecipitation experiments with HeLa and DU145 cell lysates
were carried out with appropriate antibodies (rabbit anti-human ribonuclease inhibitor
protein polyclonal antibody, Harlan Laboratories, Madison, WI; mouse anti-
GABARAPL2 polyclonal antibody, Abnova, Taipei, Taiwan). These blots did not
confirm the interaction, even when authentic GABARAPL2 protein was spiked into cell
lysates.


