Exploring and Exploiting Ribonuclease 1: from Protein Biochemistry to Protein Engineering

by

Evans Christian Wralstad

B.Sc. Chemistry
University of Virginia, 2012

Submitted to the Department of Chemistry
in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy in Chemistry

at the

Massachusetts Institute of Technology

February 2024

Authored by: _____

Evans C. Wralstad
Department of Chemistry
December 7, 2023

Certified by: _____

Ronald T. Raines
Roger and Georges Firmenich Professor of Natural Products Chemistry
Thesis Supervisor

Accepted by: _____

Adam P. Willard
Professor of Chemistry
Graduate Officer

This doctoral thesis has been examined by a committee of professors
from the Department of Chemistry as follows:

Matthew D. Shoulders
Professor
Thesis Committee Chair

Ronald T. Raines
Roger and Georges Firmenich Professor of Natural Products Chemistry
Thesis Supervisor

Sangeeta N. Bhatia
John and Dorothy Wilson Professor
Thesis Committee Member

Exploring and Exploiting Ribonuclease 1: from Protein Biochemistry to Protein Engineering

by

Evans Christian Wralstad

Submitted to the Department of Chemistry on
December 7, 2023 in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in
Chemistry

ABSTRACT

Ribonuclease (RNase) 1 is a human protein with a remarkable ability to indiscriminately hydrolyze RNA. RNase 1 and its bovine homologue RNase A exhibit ubiquitous expression across tissues, a catalytic efficiency within the diffusion-limited regime, and minimal substrate sequence requirements. RNase A has been a favorite model protein of biochemists for over half a century; due to the high level of sequence conservation between RNase A and RNase 1, many observations made for RNase A have corollaries for RNase 1.

RNase 1 and RNase A are members of the pancreatic-type ribonuclease (ptRNase) superfamily, a class of enzymes which share many biophysical features, including a small molecular weight, high cationicity, and a secretory nature. Historical elucidation of ribonuclease biochemistry describes their susceptibility to oxidation-induced inactivation. This raises the question: how are these secretory enzymes able to preserve catalytic competency in oxidatively challenging extracellular environments such as blood serum and even epidermal skin?

In Chapter 2 of this thesis, the intrinsic antioxidative capacity of RNase 1 is described. Chemical biology and biomimetic techniques corroboratively implicate two methionine residues as sacrificial antioxidants to protect the enzymic active site, allowing catalysis to persist in the presence of reactive oxygen species. *In silico* studies suggest evolutionary patterns to install these antioxidative features across the ptRNase superfamily. Sulfur–arene interactions appear to tune the reactivity of methionine residues in a manner consistent with rates of oxidation. These findings highlight an underappreciated role for methionine—to protect catalytic histidine residues—and indicate a means by which ptRNases remain functional in oxidatively challenging physiological environments.

The desirable biophysical features of RNase 1 and the wealth of biochemical knowledge regarding it have also made it a favored model system of protein engineers, as exemplified by RNase S and cyclic RNase-based zymogens, two systems which reversibly attenuate ribonucleolytic activity. In particular, RNase-based zymogens can be activated by exogenous proteases; this schema has biotherapeutic potential, as demonstrated by zymogens which activate in response to viral infection and exert cytotoxic ribonucleolytic activity.

Efforts to establish a zymogen directed toward the coronavirus SARS-CoV-2 are described in two parts of this thesis. In Chapter 3, the main protease $3CL^{pro}$ of SARS-CoV-2 is enzymologically characterized. This work clarifies reported inconsistencies in enzymological features of this key viral protease and relies on a non-Michaelis–Menten, Bayesian inference-based analytical technique to circumvent some of the causes of the inconsistent prior reports. Then, in Chapter 4, the newfound knowledge of $3CL^{pro}$ enzymology is applied toward the design

of an RNase 1-based, $3CL^{pro}$-directed zymogen. The zymogen is inactivated by steric occlusion and conformational distortion of the active site, and site-specific activation by $3CL^{pro}$ results in a multi-order of magnitude increase in ribonucleolytic activity. $3CL^{pro}$ action upon the zymogen leads to ribonucleolytic turnover of a fluorescent RNA substrate by the activated species, affording signal amplification that enables detection of nanomolar $3CL^{pro}$ concentrations in a timeframe comparable to rapid antigen detection testing.

Thesis supervisor: Ronald T. Raines
Title: Roger and Georges Firmenich Professor of Natural Products Chemistry

ACKNOWLEDGMENTS

A great number of people stand behind my success in graduate school. These people contributed to innumerable moments that invigorated me over half a decade, through many long nights, and during a concerning number of coffee runs. It'd be impossible to fully articulate the extent of my gratitude, but I can certainly try.

First and foremost, I need to extend my gratitude to Professor Ronald T. Raines for his years of mentorship and expert guidance as I conducted my doctoral research. Many advisors can provide research funding, a desk in the office, and a set of micropipettes. Fewer have the equanimity to allow a student the intellectual freedom to stumble, to hit dead-ends, to consume reagents on experiments doomed to failure, to excitedly declare discoveries only to find that other researchers had already reported the same findings—summarily, to experience all the parts of research that don't make their way into a scientific publication. Here's one example: as a still-inexperienced second-year graduate student, I noticed a strange data pattern while analyzing what should have been a routine protein sample. I naively asked Ron if I could purchase a not-inexpensive cell line to test a hypothesis I had to explain my observation, which he agreed to. When my experiment crashed and burned, rather than tell me to abandon ship and focus on my other projects, Ron allowed me to continue pursuing my hypothesis. His willingness to stoke my scientific curiosity led to the most expansive of my graduate research projects. I also have Ron to thank for my newfound awareness of the finer points of the em dash, the en dash, the hyphen, and the minus sign—my typography has never been better.

I'm also incredibly grateful to Dr. Breanne Gjurich, my mentor and supervisor during my pre-doctoral research career in industry R&D. Breanne had the esteem for me to encourage me to submit applications to graduate school despite knowing that she stood to lose a member of her research team. She's periodically checked on me over the years, and I'm glad to be able to call her my former boss and current friend. This dissertation quite literally would not exist without her support.

The MIT research community has also been a source of tremendous knowledge and advice. My thesis committee chair, Professor Matt Shoulders, has provided considerable guidance throughout graduate school. Notably, he emphasized a simple concept that many graduate students are guilty of neglecting: that it's okay to take breaks. My other committee member, Professor Sangeeta Bhatia, has offered insightful perspective as a faculty member outside the Chemistry department, in the process helping me see the forest in spite of the trees. One of my projects owes its success to her advice that I look at a stalled experiment from a different perspective. And to the Raines lab as a whole: thank you for providing expertise in so many skills, for making research feel less frustrating and more exciting, and for constantly keeping me on my toes with quips about my Starbucks habit.

There are certain people I've had the privilege of becoming good friends with by virtue of being at MIT. Some of these people joined graduate school at the same time I did, others decided to join the Raines lab at the same time I did, and some very special individuals did both! In no certain order, I am incredibly grateful to Nile Abularrage, JoLynn Giancola, Jinyi Yang, Jess Sayers, Victoria Marando, and Mikaila Hoffman for their friendship. These people are united in their ability to make me laugh and their eagerness to keep my ego grounded. There's nobody I'd rather spend evenings with, drinking on the Esplanade.

A number of friends outside of MIT have reminded me that there's more to life than graduate school. Special recognition is owed to Jeff Mailhot for being the source of many martinis and memes, both of which kept me sane throughout this half-decade process—thank you for sending me Instagram posts out of the blue and sparking a friendship in the process. Chris Wyeth, one of my oldest friends, has kept me laughing over twelve years, two degrees, and one pandemic. Nat Mahabir has been a phenomenal confidant, co-captain, and co-conspirator whose ability to compassionately listen is something I don't know how I ever lived without. I also need to extend a tremendous thank you to Dennis Adey, to my Stonewall Sports SOMB family, to Matt Smedal and Lyndon James, and to K.M. and J.H. Your surprise pastries, elaborate garden dinners, cheers of "you got this!", meanderings inside of bookstores, evenings spent at Walden Pond, and long conversations about nothing in particular have provided a lot of joy that kept me going.

Finally, I owe great appreciation to my family: to my Mom and Dad, to my brother Garrett, and to my sister Isabella. I'm most certainly the product of my upbringing, both in terms of my sense of curiosity and my sense of humor.

I often joke with friends and family that "it's all about me." Just this once, it's entirely about other people. To all those mentioned above, and to the many others that have touched my life: thank you. I wouldn't be the scientist and person I am without your positive influence.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

DIGITAL SUPPLEMENTARY MATERIAL

wralstad-evansw-phd-chem-2024-supplemental_1.mp4

MPG movie file relevant to Chapter 4, depicting protein–protein docking of the $3CL^{pro}$-directed ribonuclease zymogen with SARS-CoV $3CL^{pro}$, as predicted by ClusPro (rank 3 model).

ABBREVIATIONS

$[D]_x$: molar concentration of $3CL^{pro}$ dimer
$3CL^{pro}$: 3-chymotrypsin-like protease
6-FAM: 6-carboxyfluorescein
6-TAMRA: 6-carboxytetramethylrhodamine
AA: amino acid
ACE2: angiotensin converting enzyme 2
ACF: autocorrelation function
ACN: acetonitrile
AEX: anion-exchange chromatography
AFPS: automated fast-flow peptide synthesis
Ala: L-alanine
Arg: L-arginine
Asn: L-asparagine
Asp: L-aspartic acid
ATP: adenosine triphosphate
BCA: bicinchoninic acid
C: carboxy-terminal, with respect to a polypeptide/protein
COVID-19: coronavirus disease 2019
CPK: Corey–Pauling–Koltun (*coloring scheme of the atomic elements*)
CuAAC: Cu(I)-catalyzed alkyne–azide cycloaddition
Cys: L-cysteine
dA: 2′-deoxyadenosine
DABCYL: 4-((4-(dimethylamino)phenyl)azo)benzoic acid
DCM: dichloromethane
DEPC: diethyl pyrocarbonate
dG: 2′-deoxyguanosine
Dha: dehydroalanine
DIC: diisopropylmethanediimine
DIPEA: *N*,*N*-diisopropylethylamine
DMF: dimethylformamide
DMSO: dimethyl sulfoxide
DNA: deoxyribonucleic acid
DnaE: DNA polymerase III
DPBS: Dulbecco's phosphate-buffered saline
DSF: differential scanning fluorimetry
dT: 2′-deoxythymidine (*d included to emphasize deoxyribonucleotide*)
DTNB: 5,5′-dithio-(2-nitrobenzoic acid)
DTT: dithiothreitol
E. coli: *Escherichia coli*
EDANS: 5-((2-aminoethyl)amino)naphthalene-1-sulfonic acid
EDTA: ethylenediaminetetraacetic acid
EKMCMC: enzyme kinetics Markov chain Monte Carlo
ESI: electrospray ionization
$E_{T,x}$: molar concentration of total $3CL^{pro}$ protomer

Ext: extein
FPLC: fast protein liquid chromatography
FRET: Förster resonance energy transfer
GFC: gel-filtration chromatography
Gln: L-glutamine
Glu: L-glutamic acid
Gly: L-glycine
GSH: reduced glutathione
GSSG: oxidized glutathione
GST: glutathione *S*-transferase
H. sapiens: *Homo sapiens*
HATU: hexafluorophosphate azabenzotriazole tetramethyl uronium
HEPES: 2-(4-(2-hydroxyethyl)piperazin-1-yl)ethane-1-sulfonic acid
His: L-histidine
HIV: human immunodeficiency virus
HPLC: high-performance liquid chromatography
HRP: horseradish peroxidase
IAM: iodoacetamide
IC_{50}: inhibitory concentration that results in 50% attenuation of enzyme activity
ICEKAT: Interactive Continuous Enzyme Kinetics Analysis Tool
I_0: background fluorescence intensity measured during an RNase activity assay
I_{max}: maximal fluorescence intensity measured during an RNase activity assay
Insol: insoluble protein fraction
Int: intein
IPTG: isopropyl β-D-1-thiogalactopyranoside
ITC: isothermal titration calorimetry
k_{cat}: enzymatic turnover rate
K_M: enzymatic Michaelis constant
k_{cat}/K_M: catalytic efficiency
K_d: equilibrium dissociation constant
kDa: kilodaltons
LB: Luria–Bertani (*bacterial growth medium*)
LC/MS: liquid chromatography/mass spectrometry
Leu: L-leucine
Lys: L-lysine
m/z: mass-to-charge ratio
MALDI–TOF MS: matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
MCMC: Markov chain Monte Carlo
mCPBA: *meta*-chloroperoxybenzoic acid
MES: 2-(*N*-morpholino)ethanesulfonic acid
Met: L-methionine
MM: Michaelis–Menten
MSR: methionine sulfoxide reductase
MUSCLE: Multiple Sequence Comparison by Log-Expectation
MW: molecular weight
MWCO: molecular weight cut-off

N: amino-terminal, with respect to a polypeptide/protein
ND: not determined
NIH: National Institutes of Health
Nle: L-norleucine
NMR: nuclear magnetic resonance
Npu: *Nostoc punctiforme* PCC73102
***Npu*N**: N-terminal fragment of the split intein from *Nostoc punctiforme* PCC73102
***Npu*C**: C-terminal fragment of the split intein from *Nostoc punctiforme* PCC73102
NR: not reported
NS3: nonstructural protein 3
NSF: National Science Foundation
NSP: nonstructural protein
NTB: 2-nitro-5-thiobenzoic acid
OD_{600}: optical density at 600 nm
OVS: oligo(vinylsulfonic acid)
PAF: platelet-activating factor
PBS: phosphate-buffered saline
PCR: polymerase chain reaction
PDB: Protein Data Bank
PES: polyethersulfone
Phe: L-phenylalanine
pK_a: negative base-10 logarithm of the acid dissociation constant
plDDT: predicted local Distance Difference Test
Poly(A:U): double-stranded polyadenylic–polyuridylic acid
pp1a/pp1ab: SARS-CoV-2 polyprotein 1a/1ab open reading frame gene products
ppm: parts-per-million (*unit of mass difference in mass spectrometry*)
ptRNase: pancreatic-type ribonuclease
PyBOP: (benzotriazol-1-yloxy)tripyrrolidinophosphonium hexafluorophosphate
Pyr: pyrimidine
Q–TOF MS: quadrupole time-of-flight mass spectrometry
RADT: rapid antigen detection testing
ReACT: redox-activated chemical tagging
RFU: relative fluorescence units
RMSD: root-mean-square deviation
RNA: ribonucleic acid
RNase: ribonuclease
RPM: rotations per minute
RT: room temperature
rU: uridine (*r included to emphasize ribonucleotide*)
S protein: spike glycoprotein (*of coronaviruses*)
SARS: severe acquired respiratory syndrome
SARS-CoV: SARS-associated coronavirus 1 (*responsible for the 2003 SARS epidemic*)
SARS-CoV-2: SARS-associated coronavirus 2 (*responsible for the COVID-19 pandemic*)
SASA: solvent-accessible surface area
SAXS: small-angle X-ray scattering
SD: standard deviation

SDS–PAGE: sodium dodecylsulfate polyacrylamide gel electrophoresis
SE-AUC: sedimentation equilibrium analytical ultracentrifugation
SE: standard error
SEM: standard error of the mean
Ser: L-serine
Sol: soluble protein fraction
SPPS: solid-phase peptide synthesis
ssRNA: single-stranded ribonucleic acid
SV-AUC: sedimentation velocity analytical ultracentrifugation
TBST: Tris-buffered saline with Tween 20
TCEP: tris(2-carboxyethyl)phosphine
TFA: trifluoroacetic acid
Thr: L-threonine
$\boldsymbol{T}_{\mathbf{m}}$: thermal denaturation midpoint
tQSSA: total quasi-steady state approximation (*of enzyme kinetics*)
Tris: 2-amino-2-(hydroxymethyl)propane-1,3-diol
Tyr: L-tyrosine
U: unknown orientation, with respect to the placement of a tag on a polypeptide/protein
UpA: uridylyl(3′→5′)adenosine
UV-Vis: ultraviolet-visible (*spectroscopy*)
VADAR: Volume, Area, Dihedral Angle Reporter
Val: L-valine
$\boldsymbol{V}_{\mathbf{max}}$: maximal enzymatic velocity
WT: wild-type
$\boldsymbol{\delta}$: partial charge (*sign indicated by superscript*)
$\boldsymbol{\Delta I/\Delta t}$: slope of the fluorescence intensity versus time measured during an RNase activity assay
$\boldsymbol{\Delta m}$: mass difference (*between theoretical and empirical molecular weights*)
$\boldsymbol{\varepsilon}$: molar extinction coefficient
$\boldsymbol{\lambda}$: wavelength
σ*: sigma-star antibonding molecular orbital

Chapter 1:
Introduction

1.1 The endogenous and engineered potential of proteins

Proteins present a vast array of biological possibilities. Among the classes of biomacromolecules, proteins may afford the most versatile array of functions: from providing structural support to cells and tissues;[1] to facilitating communication between proximal and distal cells, whether they be distinct entities or components of a multicellular organism;[2] to enabling countless chemical reactions crucial to organismal fitness that would otherwise be energetically hindered.[3,4] Built from 20 canonical amino acid building blocks, proteins stitch together various functional groups in a bewildering number of structural conformations;[5] a single polypeptide sequence of a modest 100 amino acids can take approximately 1.7×10^{47} conformations.[6] Considering the sequence space and conformational space of proteins, it does not seem an exaggeration to suggest that the biological possibilities of proteins appear limitless.

To further expand on the complexity of proteins, their biophysical nature is made dynamic by endogenous chemical modifications. Through posttranslational modification of the various reactive functional groups of amino acid sidechains, the number of amino acid isoforms comprising proteins is expanded to approximately 140.[7] Some modifications, such as phosphorylation and glycosylation, are deliberately installed to alter the chemistry and/or sterics of an extant protein.[8] Other modifications, such as oxidation, are adventitious and result from a protein's unanticipated encounter with reactive species.[9] Even in the case of stochastic posttranslational modifications, biological mechanisms exist to direct the outcome of the modifying reaction and/or reverse the modification.

In recent decades, researchers have sought to unlock the full potential of proteins by increasing their scope of functions beyond those already established by Nature. The field of protein engineering, in which a protein is exogenously modified to impart non-native

functionality or outright designed *de novo*, has taken advantages of research technologies in directed evolution, chimeragenesis, rational design site-directed mutagenesis, and bioconjugation to establish adaptive and novel functions for proteins.[10] Protein engineering has broad utility in therapeutics,[11] biofuels,[12] and synthetic biology for basic research,[13] to name only a few areas of active research. Within the Raines lab, protein engineering research has resulted in the development of cytotoxic variants of the enzyme ribonuclease A and its human homologue,[14-17] strategies to bioreversibly conjugate proteins for facilitated cell delivery,[18,19] and semisynthetic ribonuclease constructs bearing reporter tags.[20]

As stated by the biologist François Jacob, proteins are subject to "constraints that, at every level, specify the rules of the game and define what is possible with those systems; and the historical circumstances that determine the actual course of events and control the actual interactions between the systems."[21] Herein, I describe two aspects of the protein science of ribonucleases that capture the essence of Jacob's dichotomy: the fundamental protein biology of ribonuclease oxidation with specific emphasis on methionine, and applied bioengineering to construct a novel, SARS-CoV-2-directed, ribonuclease-based zymogen. The former topic begins with a historical accounting of ribonuclease research and culminates in a basic scientific inquiry I sought to address in my doctoral work: how do ribonucleases preserve their function in the face of oxidative insult? The latter topic recounts the means by which organisms spatiotemporally regulate their enzymes and asks the question: can ribonuclease regulation be engineered such that the construct acts as a useful tool, in this case as a diagnostic for viral disease? This bioengineering effort necessitated additional work to characterize the viral enzyme to be detected, bringing even this applied research back to the basic research of protease enzymology.

1.2 Fundamental protein biology of ribonucleases

1.2.1 Ribonuclease A

Bovine ribonuclease A (RNase A) is one of the most well-studied enzymes in all of biochemistry, chemical biology, and related disciplines.[22] The existence of RNase A was first alluded to in 1920, where it was noted that an extract of porcine pancreas was capable of degrading yeast nucleic acid: "An active agent is present in animal pancreas which is specifically adapted to plant nucleic acid. It suggests evolutionary matters."[23] The crystalline protein was first isolated in 1939,[24] and its putative catalytic mechanism was reported in 1961.[25] The protein would go on to be the subject of multiple Nobel Prizes centered on achievements in the structural elucidation and chemical synthesis of proteins.[26-28] RNase A owes its prestige as a model system for protein scientists to a fortuitous decision by the Armour & Company meatpacking company, who chose to isolate one kilogram of crystalline RNase A from industrial byproduct and distribute the purified protein to labs around the world. This decision globally enabled researchers to investigate then-disputed topics within enzymology, structural biology, and other biochemical areas of research using a consistent model system, thereby removing one confounding influence from interlaboratory results.[29]

As depicted in **Figure 1.1**, the ribonucleolytic activity of RNase A relies on a catalytic triad of His12, Lys41, and His119. His12 and His119 engage in general acid/base catalysis to cleave a phosphodiester bond of the RNA backbone, and Lys41 interacts with a nonbridging oxygen atom vicinal to the scissile bond to stabilize the buildup of negative charge on the phosphoryl group.[30] Notably, while features of the ribonucleolytic mechanism have been elucidated in the intervening span of time—including recognition of nucleotide-binding and phosphoryl-binding subsites, and

the identification that RNase A acts in a non-processive manner—the mechanism of RNase catalysis as first proposed has remained accepted for over fifty years.[31-33]

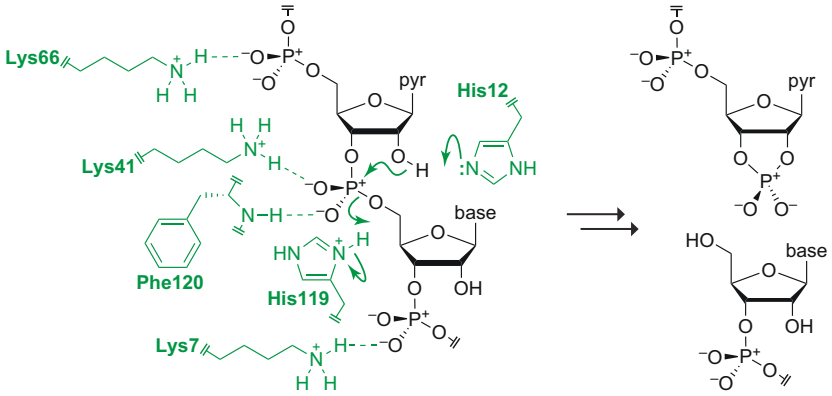


Figure 1.1. Putative mechanism of catalysis of RNA cleavage by RNase A. Relevant protein residues are depicted in green, the RNA substrate is depicted in black. *Pyr*, pyrimidine.

RNase A is a remarkable biological catalyst that exists in the privileged class of diffusion-limited enzymes.[34] In near-physiological ionic strength solutions, RNase A exhibits a catalytic efficiency exceeding 10^7 M^{-1} s^{-1} as measured using the FRET tagged, chimeric tetranucleotide 6-FAM–dArU$(dA)_2$–6-TAMRA;[35,36] the enzyme's catalytic efficiency increases to over 10^9 M^{-1} s^{-1} in assay buffers of lower ionic strength, hinting at structural features of the enzymic active site.[37] As shown in **Figure 1.2**, RNase A possesses multiple cationic surface patches near the active site cleft. These patches, and more specifically the residues that comprise them, interact with the polyanionic phosphate backbone of RNA substrates and greatly facilitate substrate turnover.[38]

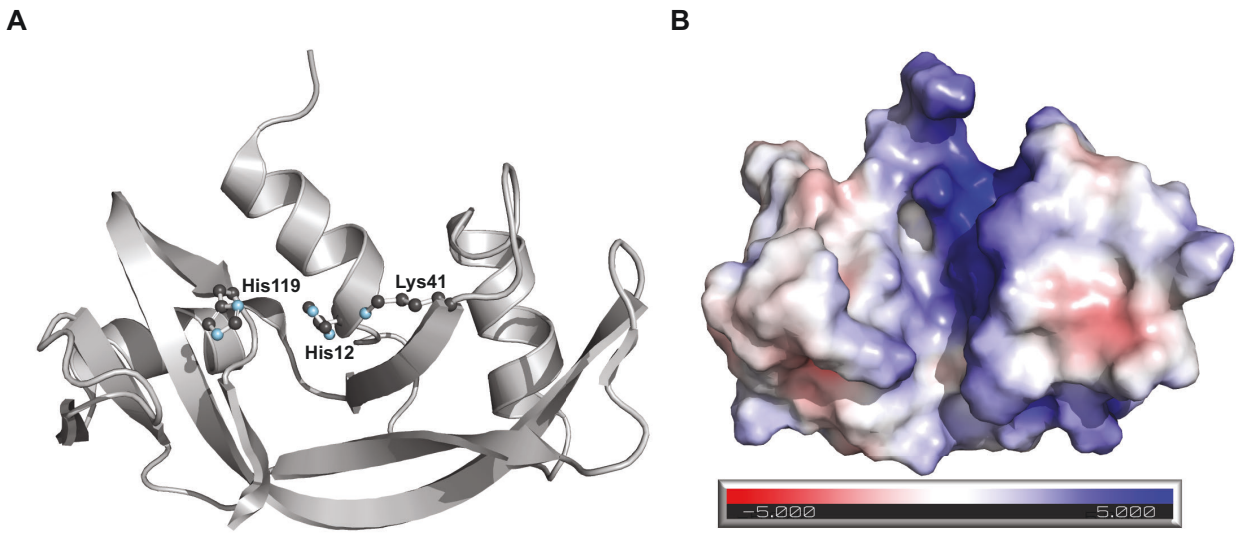


Figure 1.2. Bovine RNase A structure, from PDB ID: 7RSA.[39] (A) Cartoon structure with active site residue sidechains shown in ball-and-stick form (CPK color scheme). (B) Electrostatic surface mapping. Blue: cationic surface patches, red: anionic surface patches.

1.2.2 The RNase superfamily

Ribonuclease 1 (RNase 1) is the human orthologue of bovine RNase A and possesses many of the same features: a small and stable structure possessing four disulfide bonds, a high level of cationicity, and expedient, promiscuous cleavage of RNA.[40] Other vertebral species possess their own orthologue of RNase A, and within a single species many paralogues of RNase A are found which constitute the so-called pancreatic-type RNase (ptRNase) superfamily.[41] In humans, eight canonical ptRNases are known to exist and have a distinct array of biochemical features. All eight human ptRNases possess four disulfide bonds (except RNase 5, which possesses three), a catalytic triad of His–Lys–His, and an N-terminal signal sequence that directs posttranslational secretion of the nascent polypeptide; conversely, the human RNases each have very divergent tissue specificities and catalytic competencies, both in terms of substrate preferences and in their maximal turnover rates.[42,43] Originally believed only to aid in digestion (owing to the original discovery of RNase A from pancreatic extract and a particularly high level of tissue secretion from ruminant species), RNases are now appreciated to serve regulatory, angiogenic, immunoprotective, and antimicrobial roles.[22,44] Phylogenetic analyses implicate RNase 5, also known as angiogenin, as the most ancestral of the RNase A superfamily members; the other canonical RNases are believed to have arisen from gene duplication events.[41,42] This phenomenon, along with the unusually high rate of amino acid substitution in RNases, is consistent with proteins involved in host defense.[43,44] The frequency of sequence mutations among the RNase family members may in part explain the divergent catalytic roles exhibited by particular paralogues, and it underscores the importance of appreciating conserved sequence and structural features among RNase homologues within and across species.

1.2.3 Early work characterizing the chemical modification of pancreatic-type RNases

In 1951, Weil and coworkers embarked on an exploration of the impact of photo-oxidation on various proteins, beginning with a thorough investigating of the amino acid-specific nature of photo-oxidation.[45-49] Their characterization of isolated amino acids demonstrated that the aromatic residues tyrosine, histidine, and tryptophan, as well as the sulfur-containing residues methionine and cystine, were the most susceptible to oxidation upon irradiation with visible light in the presence of the photosensitizer methylene blue.[46] In 1955, Weil and Seibles chose bovine RNase A as a model system for their photo-oxidation studies, and they characterized the impact that oxidation had on the enzyme's structure and catalytic competency. They observed that 50% of RNase A's enzymatic activity was abolished upon uptake of one mole equivalent of oxygen, and that nearly all catalytic competency was lost by two mole equivalents of oxygen uptake. Concomitant with this, they reported a direct proportionality between oxygen uptake and a decrease in native histidine content; no changes were observed in the protein's optical rotation (which would suggest protein denaturation), α-amino nitrogen content (which would suggest photolysis of peptide bonds), or sulfhydryl content (which would suggest disruption of native disulfide bonds). Given that the enzyme's overall architecture was not observably perturbed, Weil and Seibles attributed the photo-oxidation induced inactivation of RNase A to the modification of histidine residues.[48] While unsurprising with our modern understanding of the catalytic mechanism of RNase A, Weil and Seibles' findings predate elucidation of the enzymatic mechanism (reported in 1961) by six years.[25]

Attempting to identify residues relevant to RNase A catalysis, Gundlach, Stein, and Moore performed seminal work in 1959 characterizing the modification of RNase A with iodoacetate. They observed expeditious loss of ribonucleolytic activity upon iodoacetate-induced alkylation,

and the most rapid rate of inactivation occurred when RNase A and iodoacetate were co-incubated at pH 5.5. Amino acid analysis of the isolated, inactivated RNase A species revealed a reduction in native histidine content and the appearance of (1/3)-carboxymethylhistidine, again implicating histidine in the ribonucleolytic mechanism before RNase catalysis was mechanistically understood.[50] The expedient alkylation of histidine with corresponding loss of catalytic competency at pH 5.5 was counterintuitive, given that the reaction of histidine with iodoacetate most efficiently occurs at alkaline pH, when the imidazole ring is deprotonated and can effectively engage in nucleophilic attack; at pH 5.5, the sidechain of histidine should be protonated, and the cationic imidazolium species is a poor nucleophile. Additional work identified that the iodoacetate-induced alkylation of methionine is distinctly pH-independent and forms a reactive carboxymethylsulfonium species, leading Gundlach and coworkers to speculate that if methionine and histidine were neighboring one another, the iodoacetate-induced inactivation of RNase A could occur by a two-step mechanism in which (1) carboxymethylation of one or more methionine residues proximal to catalytic histidine is followed by (2) subsequent transfer of the alkyl group from the activated carboxymethylsulfonium species to the nearby histidine (see **Figure 1.3**).[50,51]

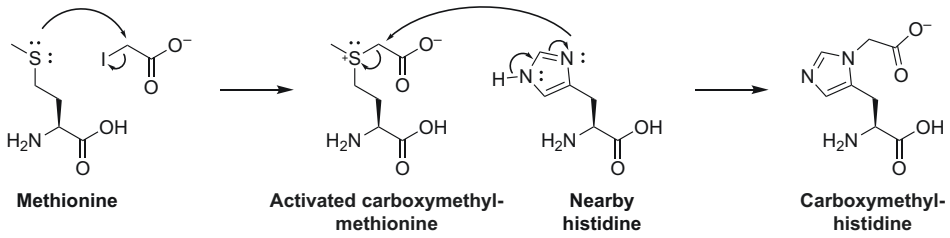


Figure 1.3. Proposed mechanism of carboxymethylation of histidine in bovine RNase A at pH 5.5.[50]

In 1962, Neumann, Moore, and Stein attempted to investigate the impact of methionine oxidation on RNase A's structure and its catalytic competency. Believing methionine to only exist within the hydrophobic core of a protein, they hypothesized that methionine modification must perturb the native fold of RNase A and prevent its active site from maintaining the conformation needed for catalysis. In contrast to this expectation, they noted that "Oxidation with hydrogen peroxide has yielded evidence for the presence of at least one *active* derivative which contains one methionine sulfoxide residue per molecule. […] [T]he amino acid analyses suggest […] that oxidation of methionine sulfur is the sole chemical modification that has been brought about."[52] Thus, it was established that RNase A catalysis is catalytically tolerant of methionine oxidation.

Proceeding to 1967, Fruchter and Crestfield definitively characterized the amino acids implicated in the alkylation-induced inactivation of RNase A and were able to contextualize their results with the then-proposed enzymatic mechanism. Using iodoacetamide to slow the rate of alkylation of RNase A relative to its reaction with iodoacetate, they observed carboxamidomethylation of His12 with resultant loss of ribonucleolytic activity. Intriguingly, they noted a pH-independent, ionic strength-independent alkylation of RNase A methionine; the methionine-modified RNase exhibited equivalent catalytic efficiency to the unreacted enzyme, similar to the results of Neumann and coworkers.[52,53] Fruchter and Crestfield concluded that broad conformational changes must not be occurring during the reaction of RNase A with alkylating reagents. Collectively, these early studies suggested that histidine and methionine were somehow associated in the context of RNase A catalysis, and that direct modification of methionine residues was on its own insufficient to hinder catalysis.

1.2.4 On the nature of methionine utilization in proteins

Methionine is principally recognized as a hydrophobic amino acid which plays a crucial role in translation initiation in eukaryotes.[54,55] Methionine is the second most infrequently utilized amino acids in human proteins, second only to tryptophan;[56,57] it comprises 2.2% of the global amino acid content of human proteins, a frequency which includes its retention as the translation initiator on proteins for which methionine aminopeptidase is not able to cleave the N-terminal methionine.[58] Methionine is also the most biosynthetically expensive amino acid in organisms, costing a net 18 equivalents of ATP for every equivalent of the amino acid synthesized from glucose; the majority of this expense is attributed to the installation of the sulfur atom in the residue sidechain.[59] Early biochemists believed that methionine serves limited roles among the canonical amino acids, serving only to pack the hydrophobic interior of proteins and signal the start of protein translation.

This simplified perspective of methionine has been challenged in recent years, and methionine has even been branded the "Cinderella" of the proteinogenic amino acids owing to its indifferent energetic preference for sidechain torsion angles and the relative polarizability of the sulfur atom.[60] Levine and coworkers championed the premise that methionine might serve an additional role: that of an antioxidant defense system.[61] Using glutamine synthetase as a model system, they demonstrated that surface-resident methionine residues were readily oxidized under physiological conditions; these oxidizable methionines were specifically clustered in a "phalanx" around the active site, and their oxidation did not impact enzyme catalysis. Methionine sulfoxide—the preliminary product of methionine oxidation—is readily corrected in eukaryotes by methionine sulfoxide reductases (MSRs), meaning the antioxidative behavior of methionine can function in a catalytic rather than a stoichiometric manner.[62,63] Levine and coworkers later

documented a number of proteins in which methionine oxidation had been observed with varying levels of impact on function, including the 1962 observations in RNase A made by Neumann, Moore, and Stein.[64] Since then, many researchers have investigated features of methionine oxidation, including its reaction kinetics;[65] relevancy to metal-binding proteins;[66] steady-state levels with respect to MSR-mediated correction;[67] sequence-dependent, one-electron redox potential;[68] and relationship to local conformational flexibility in a protein.[69] Methionine oxidation has been implicated in a number of medical conditions, including Alzheimer's disease,[70] Parkinson's disease,[71] emphysema,[72] hepatocellular carcinoma,[73] breast cancer,[74] and congenital deafness.[75] Conversely, given the bioreversible nature of methionine oxidation, interconversion of methionine and its sulfoxide has been proposed to act as a regulatory switch to trigger or inactivate protein function.[60,76] No longer relegated to an amino acid of limited utility, methionine is now understood to contribute significant redox sensitivity and antioxidative protection to proteins.

1.2.5 The *S*–aromatic motif

Proteins employ a number of secondary forces beyond the well-known primary forces of hydrogen bonding, Coulombic interactions, and van der Waals interactions.[77] One such secondary force is the sulfur–arene interaction, in which the sulfur-bearing amino acids methionine, cysteine, and/or cystine interact with the aromatic rings of phenylalanine, tyrosine, tryptophan, and/or histidine.[78] The energy of this interaction is on the order of 1–3 kcal mol^{-1}, on par with that of a salt bridge.[79] This force can significantly contribute to the stability of a protein's native fold,[76] and as I will explore in this work, sulfur–arene interactions may also tune the reactivity of sulfur atoms.

This two-residue interaction, also referred to as the *S*–aromatic motif, has two sub-types. In the first, dubbed the *S-toward interaction*, sulfur lone pair electrons are directed toward the aromatic ring; this orientation is most energetically favorable when the sulfur atom approaches the aromatic ring along the partly electron-deficient (δ^+) ring edge (*i.e.*, approach along the plane of the aromatic residue). Alternatively, the sulfur atom may approach along either face of the aromatic ring; in this form, it is most energetically favorable for the sulfur atom to point away from the aromatic ring (dubbed the *S-away interaction*), both to prevent electrostatic repulsion between sulfur lone pairs and the electron-rich (δ^-) face of the aromatic ring as well as to direct empty C–S/H–S/S–S σ* orbitals toward regions of electron density.[80] In a survey of 393 PDB-deposited structures, 9% of methionine residues were found to interact with an aromatic ring edge and 8% were found to interact with a ring face.[81] Perhaps unsurprisingly, an interplay exists between the inter-residue phenomenon of *S*–aromatic motifs and the intra-residue phenomenon of methionine oxidation.

The work of Aledo, Cantón, and Veredas suggests that methionine residues engaged in *S*–aromatic motifs are relatively more recalcitrant to oxidation, implying that *S*–aromatic motifs have a deactivating nature with respect to the reactivity of the methionyl sulfur atom.[82] A significant limitation of this work is that the motif was identified based on a suitable sulfur–arene interatomic distance, but no attention was paid to the directionality of the interaction, *i.e.*, whether the interaction was *S*-toward or *S*-away. The complementary orbital interactions of sulfur lone pairs with the δ^+ ring edge or C–S σ* orbitals with the δ^- ring face suggest that *S*–aromatic motifs can be activating *or* deactivating with respect to the sulfur atom.

1.2.6 Motivation for investigating the oxidation of *H. sapiens* RNase 1

In Chapter 2 of this thesis, I describe my work characterizing how human RNase 1 utilizes methionine as an antioxidant defense system, as broadly proposed for proteins by Levine and coworkers.[61] This work was inspired by a simple observation: RNase 1 utilizes methionine at a frequency (3.9%, 5 of 128 residues) nearly twice that expected for a protein of its size. As often happens in science, this observation came about purely by chance—while conducting an unrelated, mass spectrometry-based experiment using RNase 1, a surprising number of oxidation peaks were observed (**Figure 1.4**). This single dataset would inspire an entire project.

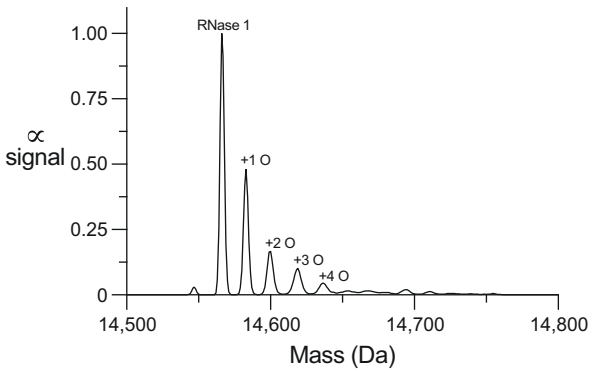


Figure 1.4. Graph of a deconvoluted Q–TOF mass spectrum for *H. sapiens* RNase 1 depicting its pattern of oxidation. Following the parent protein peak, a series of peaks separated by 16 Daltons are observed (+O), consistent with the addition of oxygen to the protein.

I apply orthogonal bioinformatics, chemical biology, and biomimetic strategies to interrogate RNase 1 and bolster the rigor of my findings implicating methionine as antioxidative residues. My bioinformatic analyses interrogate the evolutionary patterns of methionine usage in the ptRNase superfamily, and I note distinct patterns of methionine accumulation across homologues as well as the occurrence of *S*–aromatic motifs that suggest a residue-specific tuning of methionine reactivity such that Met29 and Met35 are readily oxidizable residues. In my chemical biology approach, I employ a previously reported, methionine selective, oxaziridine-based probe in a novel manner—as a chemical capping group—to effectively pre-oxidize solvent-accessible RNase 1 methionine residues; the labeled enzyme is more easily inactivated via oxidation than the unlabeled one, implying that RNase 1 methionine residues protect the active site. In the complementary biomimetic strategy, I obviate the antioxidative capability of the suspected reactive methionine residues (Met29, Met35) by site-specific installation of the noncanonical amino acid norleucine; like the results from labeled RNase 1, the norleucine variant enzyme is more easily oxidatively inactivated than the wild-type protein. Notably, in the course of pursuing this biomimetic strategy, I report the first one-pass chemical synthesis of RNase 1.

1.3 Opportunities for ribonuclease-based engineered proteins

1.3.1 Zymogens in human physiology and as a bioengineering tool

A core facet of multicellular biology is that of specialization, in which not every cell behaves in an equivalent fashion; cell types and functions may be distinguished spatially, temporally, or both.[83,84] Analogous to cellular phenotype changing in response to stimuli or environmental demands, the functional behavior of proteins can be altered in response to external factors. One way in which proteins can attenuate their function is by evolving toward a zymogen, a pro-enzyme that acts as the inactive precursor to a cognate functional protein.[85] Proteins may be rendered inactive by appending peptidic tags to the functional protein unit in order to cause a steric clash or trigger conformational changes; by modifying amino acid sidechains through the addition of interfering pendant groups or the removal of catalytically necessary ones; or by sequestering protein cofactors.[86] Many classes of protein exist in a zymogen form, including digestive enzymes such as trypsin, chymotrypsin, and pepsin, which are initially produced as zymogens to prevent premature, off-target hydrolytic activity at the initial site of secretion;[87] signaling proteins such as the caspases, which are maintained as zymogens until needed to initiate apoptotic pathways;[88] and coagulation proteins such as thrombin, fibrin, and Factor X, whose zymogen forms help funnel the cascade of enzymatic reactions that lead to blood clotting.[89]

Inspired by natural zymogens, the field of protein engineering has designed artificial zymogens from native proteins in order to exogenously render the protein inactive until an appropriate activating signal is introduced.[11] Methods used to reversibly inactivate a protein of interest include site-directed mutagenesis to hinder conformational changes,[90] reversible complexation and encasement within macromolecular protecting groups,[91] and fusion to

inhibitory elements via protease-cleavable linkers.[92] These artificial zymogens hold great promise as basic research tools as well as biotherapeutics, as exemplified by the work of the Raines lab.

1.3.2 Engineering cyclic ribonuclease zymogens

The Raines lab has explored the design of artificial, ribonuclease-based zymogens using the inactivation strategies of active site occlusion and conformation distortion. To allow for an activating signal to restore ribonuclease function, the Raines lab chose to install a peptidic linker recognized by a protease of interest to occlude the enzymic active site; the protease therefore behaves as the activating stimulus by freeing the RNase active site from its occluding linker. Early efforts to engineer a zymogen based on RNase A relied on circular permutation, in which the native termini of a protein are bridged by a covalent peptidic linkage and new termini are established elsewhere in the protein sequence to effectively permute the protein's sequence with respect to its wild-type one.[93] The inspiration to use circular permutation to occlude RNase A's active site was based on the fortuitous fact that the termini of RNase A point in opposite directions and flank its active site (see **Figure 1.2A**); it was reasoned that a peptidic linker bridging the native termini might therefore span the active site and interfere with the binding of substrate RNA. Using circular permutation and an tetradecapeptide linker recognized by plasmepsin II from *Plasmodium falciparum*, the Raines successfully created an RNase A-based zymogen that was specifically activated upon protease administration with a >900-fold increase in catalytic efficiency following linker cleavage.[94] Other RNase A-based zymogen constructs followed suit with peptidic linkers recognized by the NS3 protease (of hepatitis C virus) and the HIV-1 protease. In each case, a >50-fold change in catalytic efficiency was achieved between

zymogen and activated states.[95,96] More recently, the Raines lab has designed an orthologous zymogen based on human RNase 1 which is also recognized by HIV-1 protease. For the RNase 1-based zymogen, truncation of the native termini and an octapeptide linker were optimal for balancing a large fold-change in catalytic efficiency (between zymogen and activated states) with maintained thermostability.[97] The RNase 1-based zymogen is specifically activated by the cognate HIV-1 protease, and the zymogen can induce cytotoxicity in HIV-infected cells using MT-4 cells as a model system.[98] Notably, this construct did not utilize circular permutation in its synthetic design, but rather, was created using intein chemistry to engineer a truly cyclic protein (*vide infra*).[97]

1.3.3 RNase S: a fragmented ribonuclease zymogen

Distinct from converting a protein into a zymogen by appending additional peptides or domains, one might envision creating a zymogen by instead fragmenting the protein between structural units. Often, protein fragmentation results in irreversible loss of structure and function, for example following proteolytic digestion with trypsin or peptide bond hydrolysis at extreme pH.[99,100] Under certain conditions, however, it is possible to site-specifically fragment a protein to perturb its function in a manner that allows for re-complementation of the fragments and restoration of activity, as in protein complementation assays.[101,102]

In 1959, it was first reported that bovine RNase A can be specifically fragmented into two constituents by treatment with the serine protease subtilisin. Proteolysis results in a 20-residue, amino-terminal fragment, referred to as *S-peptide*, and a remaining 104-residue fragment, referred to as *S-protein*. Neither fragment alone exhibits ribonucleolytic activity, likely due to

physical separation of the catalytic residue His12 (within S-peptide) from Lys41 and His119 (within S-protein) (see **Figure 1.1** and **1.5**).

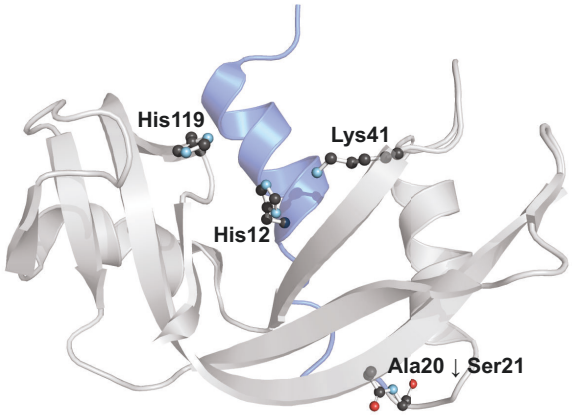


Figure 1.5. Structure of RNase S, which derives from bovine RNase A. Subtilisin cleaves RNase A at the Ala20–Ser21 peptide bond (shown in ball-and-stick form, CPK color scheme), resulting in liberation of S-peptide (blue cartoon) from S-protein (gray cartoon) and physically separating the three active site residues (sidechains shown in ball-and-stick form, CPK color scheme). This structure is prepared from PDB ID: 7RSA.[39]

Remarkably, coincubation of S-peptide and S-protein results in fragment complementation to produce so-called *RNase S* with restoration of ribonucleolysis.[103] Subsequent work has demonstrated that the RNase S complex does, in fact, recapitulate the structure and conformational dynamics of its progenitor, RNase A.[104,105] Given the synthetic accessibility of S-peptide and the ease of producing S-protein, RNase S and its components have been a prime model system within protein engineering to investigate the impact of unnatural amino acid incorporation,[106,107] to produce novel purification strategies,[108-110] to demonstrate photo-control of enzymatic activity,[111] to design a targeted drug delivery platform,[112,113] and to develop an anticancer therapeutic.[114]

1.3.4 Intein chemistry

Inteins (*in*tervening pro*teins*) are auto-processing domains in polypeptide sequences which excise themselves from the polypeptide and seamlessly ligate the flanking exteins (*ex*ternal pro*teins*) present on either side, acting in a somewhat analogous manner to introns in pre-mRNA. While intein sequences are varied, core features are retained to preserve the conserved reaction mechanism shown in **Figure 1.6** and described below:

1. An N-terminal intein nucleophile (often cysteine or serine) attacks the C-terminal main-chain carbonyl of the N-extein fragment to form a linear (thio)ester.
2. An N-terminal, C-extein fragment nucleophile (cysteine/serine/threonine) attacks the newly formed (thio)ester carbonyl in a *trans*-(thio)esterification reaction to form a branched (thio)ester intermediate.
3. A highly conserved, C-terminal intein asparagine cyclizes at its carbonyl to form a stable, five-membered succinimide, thereby excising the intein.

4. The non-native adduct of the extein fragments spontaneously rearranges to form a stable, native peptide bond in an intein-independent fashion, thereby ligating the extein fragments to form the intended protein.[115]

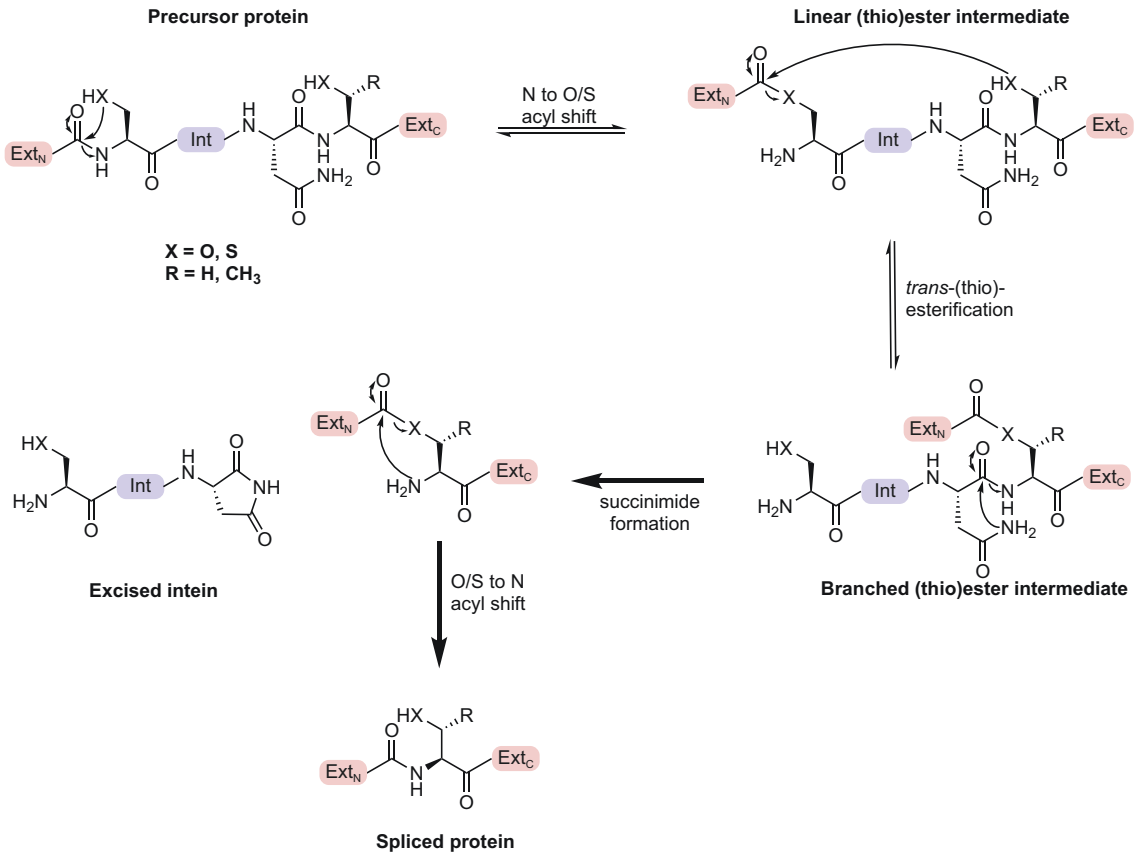


Figure 1.6. Canonical mechanism of intein splicing. N-terminal extein (Ext_N) and C-terminal extein (Ext_C) fragments are spliced during the excision of the intein (*Int*). Mechanism is as described by Shah and Muir.[115]

The biological role for inteins remains an active area of investigation, but they are often described as "selfish" or parasitic gene elements.[116] Conversely, inteins have been a boon to protein engineering and synthetic biology.[115,117] Of particular note are the split inteins, in which the intein exists as two fragments which cannot separately enact their autocatalytic splicing function; the fragments must noncovalently associate, at which point the splicing reaction occurs as previously described.[118] It is possible to cyclize a desired protein with strategic fragmentation of the protein such that a suitable nucleophile (*i.e.*, cysteine/serine/threonine) is properly positioned for step (2) of the splicing reaction above, along with proper concatenation of split intein fragments. A general scheme for protein cyclization using split inteins is shown in **Figure 1.7**.

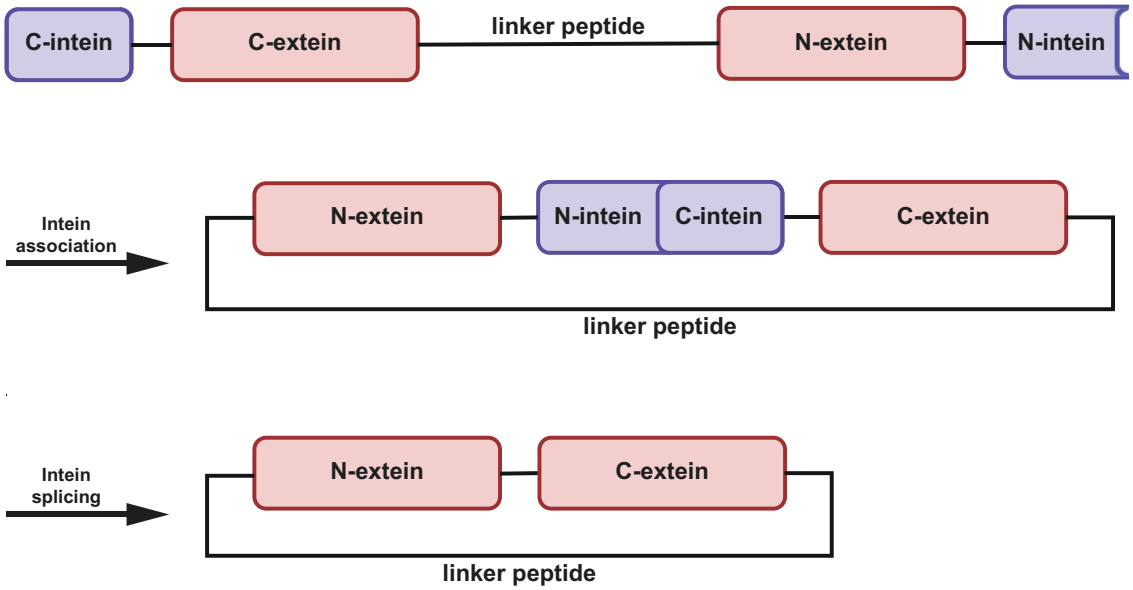


Figure 1.7. Schematic of protein cyclization using a split intein system. The initial, linear construct is produced as a polypeptide in the order (N- to C-terminal): (1) C-terminal intein fragment; (2) C-terminal extein fragment; (3) linker peptide (if included); (4) N-terminal extein fragment; (5) N-terminal intein fragment. Following noncovalent association of the split intein fragments, the intein splicing reaction occurs as shown in **Figure 1.6**, excising the intein and ligating the extein fragments into a contiguous, cyclic construct.

The Raines lab used the naturally split DNA polymerase III (DnaE) intein from *Nostoc punctiforme* PCC73102 (*Npu*) to produce the cyclic, RNase 1-based, HIV-1 protease-recognized zymogen described above.[97] *Npu* split intein fragment association is rapid (k_{on} = (8.9 ± 0.3) × 10^5 M^{-1} s^{-1}) and thermodynamically favorable (K_d = 1.2 ± 0.8 nM).[119] Reported first-order rate constants for the splicing reaction of this split intein at 37 °C range from 1.1 × 10^{-2} to 3.7 × 10^{-2} s^{-1}, meaning that the half-life of the splicing reaction is less than one minute.[120,121] The yield of properly spliced protein is >90%, making the *Npu* split intein system very efficient for circularization of proteins.[115]

1.3.5 SARS-CoV-2 main protease $3CL^{pro}$

To further the Raines lab's work engineering ribonuclease zymogens, I chose to design a cyclic, RNase 1-based zymogen which would be recognized by the main protease of Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2), the etiological agent of the COVID-19 pandemic.[122,123] Following SARS-CoV-2 virion attachment to the infected host cell ACE2 receptor and membrane fusion, injected viral RNA is rapidly translated to produce the polyproteins pp1a and pp1ab.[124] These polyproteins code for 16 nonstructural proteins (NSPs) necessary for virion maturation, and the chymotrypsin-like SARS-CoV-2 3C-like main protease $3CL^{pro}$ (contained within pp1a) is responsible for 11 of the cleavage events to liberate the NSPs.[125-128] SARS-CoV-2 $3CL^{pro}$ is 96% sequence identical to the cognate main protease of SARS-CoV, the etiological agent of the 2003 SARS epidemic, and the majority of the twelve nonidentical residues either consist of conservative mutations or occur in nonstructured regions of the two proteins (see **Figure 1.8A**).

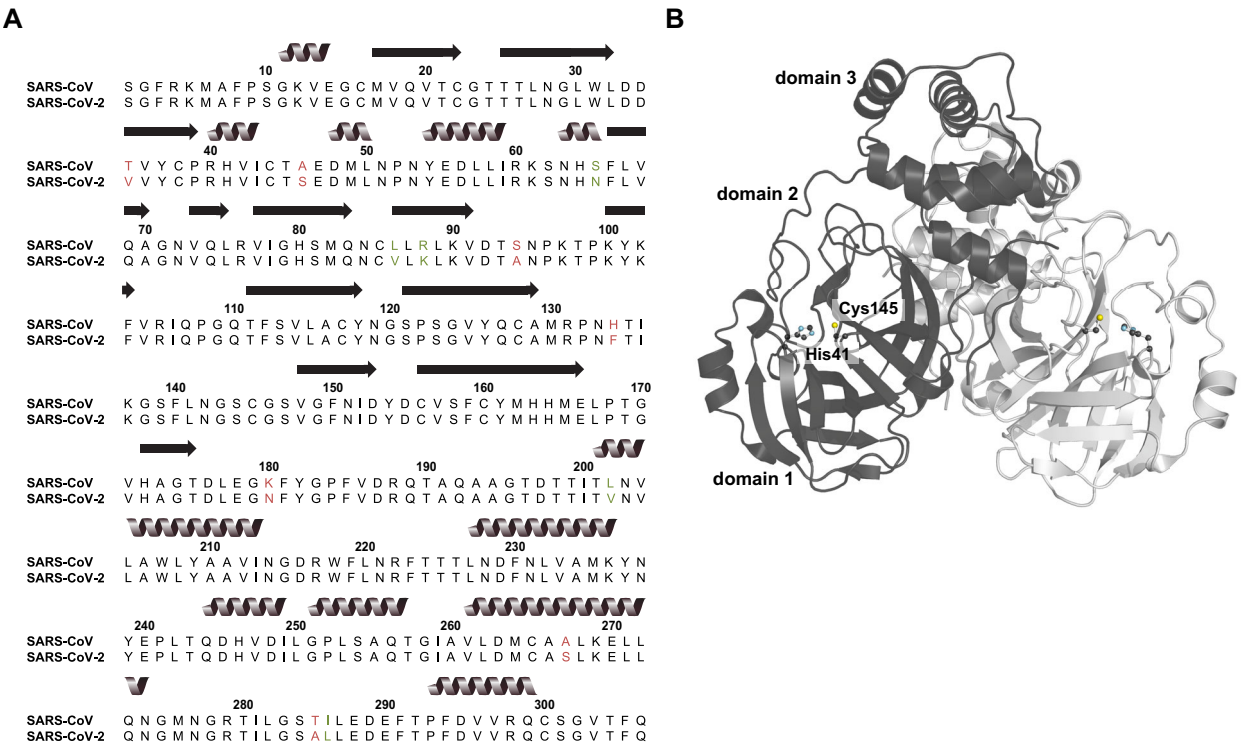


Figure 1.8. SARS-CoV-2 3CL^{pro} sequence and structure. (A) ClustalW sequence alignment of SARS-CoV-2 3CL^{pro} (UniProt accession ID: P0DTD1, positions 3264−3569) with the SARS-CoV homologue (UniProt accession ID: P0C6X7, positions 3241−3546). Conservative mutations are identified by green font; nonconservative mutations are identified by red font. (B) SARS-CoV-2 3CL^{pro} structure (PDB ID: 6Y2E[129]). Homodimer protomers are distinguished by light and dark gray. Catalytic dyad residues at the interface of domains 1 and 2 are depicted in ball-and-stick form, and the dyad and domains of one protomer are labeled.

SARS-CoV and SARS-CoV-2 $3CL^{pro}$ are each a homodimer of 306-residue subunits arranged perpendicular to one another (**Figure 1.8B**). Each subunit consists of three domains; domains I and II each possess a β-barrel structure and collectively form the active site cleft, while domain III is a globular, α-helix rich region implicated in dimerization.[130] Unlike the canonical chymotrypsin catalytic triad Ser/Cys–His–Asp/Glu possessed by many coronavirus proteases, $3CL^{pro}$ instead possesses a catalytic dyad of His41 and Cys145.[126,131] It is hypothesized that a bound water molecule additionally hydrogen bonds to His41 and could be the third member of a catalytic pseudo-triad.[132] Dimerization is strictly necessary to establish catalytic competency; empirical and computational evidence shows that the N-terminal residues of one protomer (the "N-finger") weave between domains I and II of the other protomer, completing the active site of the latter and properly orienting cleft residues to prevent steric clash with polypeptide substrates.[132,133] In particular, the N terminus of one protomer hydrogen bonds with Glu166 of the other protomer to prevent the latter from blocking access to the active site, and the main chain of Ser1′ (′ denoting the sibling protomer) hydrogen bonds with the main chain of Phe140A to properly orient the oxyanion loop of which Phe140 is a part.[132] This loop is responsible for stabilizing the tetrahedral transition state of the nascent acyl-enzyme intermediate, and a collapsed active site has been observed in monomeric $3CL^{pro}$ as a result of missing interprotomer interactions.[132,134] Data suggest that the two active sites of dimeric $3CL^{pro}$ are asymmetric and that only one is active at any given moment (*i.e.*, half-site reactivity).[131,133,135]

1.3.6 Motivation for investigating $3CL^{pro}$ and designing a $3CL^{pro}$-targeted zymogen

At-home, rapid antigen detection testing (RADT) is a crucial tool for mitigating the spread of COVID-19, as it allows individuals to confirm a suspected infection and self-isolate without the

need for comparatively costly PCR-based testing or contact with the public.[136] Unfortunately, RADT often detects the same antigens targeted by anti-COVID-19 biotherapeutics, and the resultant selective pressure exerted by the latter can induce viral mutations; a potential consequence of viral mutations is reduced diagnostic sensitivity in cases where the viral epitope is altered.[137,138] This is perhaps most exemplified by the rise of the SARS-CoV-2 Omicron variant in late 2021, during which a multi-day lag period emerged between positive diagnosis with RADT versus PCR-based testing.[139] This begged the question: could a diagnostic that relied on the *function* of a viral enzyme be more robust than one based on the *detection* of a viral protein, considering that mutations are unlikely to alter the fundamental chemical reaction catalyzed by an enzyme?

Given the indispensable nature of the main protease $3CL^{pro}$ to SAR-CoV-2 virion replication, I believed it would be a prime target for detection of a viral infection. Unfortunately, the body of scientific literature for SARS-CoV-2 $3CL^{pro}$ and its highly conserved SARS-CoV homologue are rife with inconsistencies. The reported values of the catalytic efficiency k_{cat}/K_M vary by nearly three orders of magnitude, and the reported values of the $3CL^{pro}$ dimer dissociation constant K_d vary by nearly six orders of magnitude. Closer investigation of reported studies revealed a host of discrepancies in enzyme assays between researchers, including the use of enzyme constructs bearing exogenous tags likely to interfere with enzyme function, detection methods lacking adequate sensitivity, and solution conditions which impede enzymatic catalysis (see Chapter 3 for a comprehensive review of the literature). It became evident that developing a SARS-CoV-2 $3CL^{pro}$-based diagnostic needed to be preceded by prudent characterization of the core enzymology of $3CL^{pro}$ itself.

In Chapter 3 of this thesis, I describe my enzymological characterization of SARS-CoV-2 $3CL^{pro}$. This work began with production of authentic $3CL^{pro}$, *i.e.*, enzyme free of any exogenous purification tags or fusion proteins. I then rationally designed a FRET-tagged protease substrate and characterized the limits of its utility in relation to the inner filter effect, a documented impediment to fluorescence-based assays.[125,140,141] With these tools in hand, I used an enzymology-based method to determine the value of the dimerization K_d for SARS-CoV-2 $3CL^{pro}$, relying on the comparative sensitivity of enzyme activity assays in relation to other bioanalytical techniques in order to accurately interrogate enzyme dimerization. My measured K_d value is among the lowest reported for $3CL^{pro}$ from SARS-CoV and SARS-CoV-2. I then characterized the kinetic parameters k_{cat} and K_M for the turnover of my substrate by authentic $3CL^{pro}$. Notably, I relied on a Bayesian algorithm to circumvent the assay conditions demanded by traditional Michaelis–Menten methodology, most pertinently the need for saturating substrate concentrations. Finally, I showed that $3CL^{pro}$ is highly sensitive to solution ionic strength. This observation may partly explain the confounding results from prior research of $3CL^{pro}$. My work sets a firm foundation of knowledge for the enzymology of SAR-CoV-2 $3CL^{pro}$ and highlights the need to diligently consider the biochemical features of an enzyme when investigating it.

In Chapter 4 of this thesis, I describe my work using split intein chemistry to construct a cyclic RNase 1 zymogen with an installed peptide sequence matching that previously shown in Chapter 3 to behave as a suitable $3CL^{pro}$ substrate. Using both empirical and *in silico* methodologies, I demonstrate that the $3CL^{pro}$-directed zymogen is catalytically inactivated by a combination of steric occlusion and conformational distortion of the active site. The zymogen is specifically acted upon by $3CL^{pro}$ at the installed peptidic linker, resulting in expeditious restoration of ribonucleolytic activity. I demonstrate that by functioning as an enzyme itself, the

activated RNase zymogen is able to amplify the signal of $3CL^{pro}$ activity, resulting in a system that can detect nanomolar concentrations of $3CL^{pro}$ in a timeframe comparable to RADT. This proof-of-principle diagnostic enables function-based detection of a viral infection, which holds the potential to obviate the diagnostic retooling that has been necessary because of the biological arms race between viral evolution and human intervention.

Chapter 2:
Methionine as a Vanguard Against Oxidation in the Ribonuclease Superfamily: Interrogation of Antioxidative Protection in Ribonuclease 1

This chapter will be submitted for publication as:

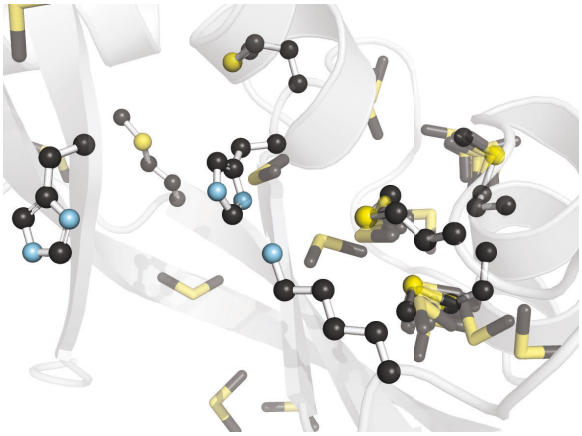
Wralstad, E. C.; Sayers, J.; Fittolani, G.; Sarabia-Castillo, D.; Callahan, A. J.; Pentelute, B. L.; Raines, R. T. Methionine as a vanguard against oxidation in the ribonuclease superfamily: interrogation of antioxidative protection in ribonuclease 1.

Manuscript in preparation.

Contributions: Conceptualization, E.C.W.; experimental methodology, E.C.W., J.S., G.F., D.S.-C., and A.J.C.; experimental investigation, E.C.W., J.S., G.F., D.S.-C., and A.J.C.; writing—original draft preparation, E.C.W., G.F., and D.S.-C.; writing—review and editing, B.L.P. and R.T.R.; supervision, B.L.P. and R.T.R.; funding acquisition, E.C.W., B.L.P., and R.T.R.

2.1 Abstract

Ribonucleases (RNases) constitute an important class of secretory proteins. RNases statistically overutilize methionine, poise the residue near the active site, and possess *S*–aromatic motifs consistent with empirical patterns of methionine oxidation, suggesting that methionine flanks the active site to preserve catalysis in oxidatively challenging environments. Using a chemical biology approach, we pre-labeled *H. sapiens* RNase 1 methionine and implicated Met29 and Met35 in providing antioxidative protection to the active site. Our findings are corroborated by a biomimetic approach in which we, for the first time, chemically synthesized RNase 1 in one pass with site-specific installation of norleucine and demonstrate this variant is more oxidatively sensitive. Collectively, our findings demonstrate the antioxidative utility of methionine in RNase 1, implicate methionine-based protection as a conserved feature of the RNase superfamily, and suggest the residue broadly acts to sacrificially protect key residues.



2.2 Introduction

Multicellular organisms broadly rely on secretory proteins to facilitate communication between specialist cell types and coordinate complex physiological functions; indeed, approximately 11% of the human proteome is calculated to consist of soluble secretory proteins.[142] Ribonucleases (RNases) constitute a diverse class of secretory enzymes which share the common ability to cleave RNA, a function with digestive, regulatory, and immunoprotective purposes.[22,42,44] A subset of RNases comprise the RNase A superfamily, which are among the most well-studied proteins in chemical biology; these secreted, vertebral proteins are small, highly stable, cationic proteins with a common ability to nonspecifically degrade RNA at an efficiency of 10^6–10^7 M^{-1} s^{-1} for ssRNA.[40]

RNase-mediated transphosphorylation relies upon a catalytic triad of two histidines and one lysine. In bovine RNase A and its human homologue, RNase 1, His12 deprotonates the 2′ hydroxyl of RNA concomitant with nucleophilic attack of the 3′ phosphoryl group, and His119 protonates a bridging oxygen of the same phosphoryl group, thereby cleaving the RNA backbone (**Figure 2.1**). Lys41, meanwhile, stabilizes the negatively charged phosphoryl group in a Coulombic interaction. Additional residues, including Lys7, Lys66, and Phe120 also interact with the phosphate backbone of RNA to facilitate substrate binding.[22]

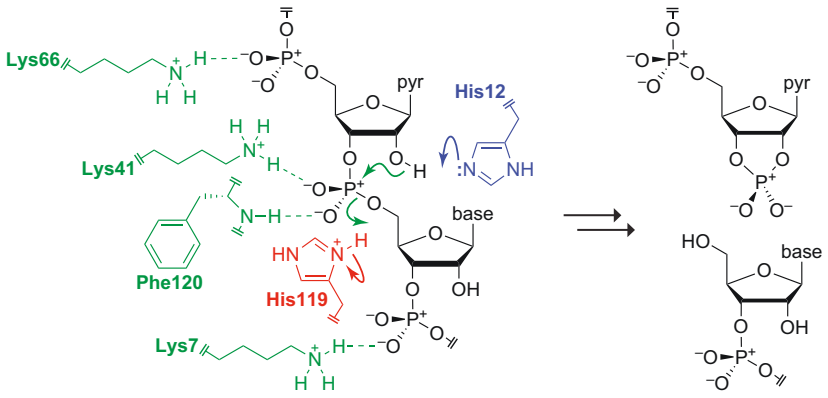


Figure 2.1. Putative mechanism of catalysis of RNA cleavage by RNase 1. The catalytic residues His12 and His119 are shown in blue and red, respectively; other relevant RNase 1 residues are shown in green. *Pyr*, pyrimidine.

Photo-oxidation of active site histidines has been shown to result in loss of catalytic competency and is likely the result of forming 2-oxo-histidine, as has been observed in the metal-catalyzed, oxidation-induced inactivation of other enzymes.[48,143] Delocalization of imidazole nitrogen electrons into the nascent carbonyl of 2-oxo-histidine likely perturbs the pK_a of the sidechain and thereby disrupts catalysis (**Figure 2.2**), analogous to the effect seen when installing synthetic 4-fluorohistidine in RNase A.[144] Lysine is also susceptible to oxidation, resulting in deaminated species unable to contribute to RNase catalysis.[145]

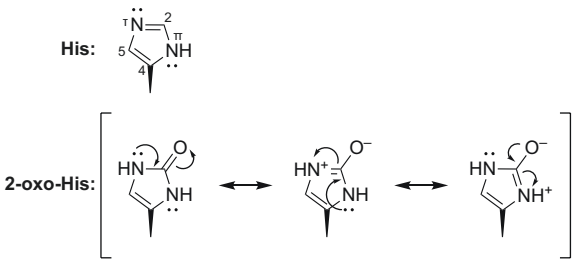


Figure 2.2. Comparison of histidine and 2-oxo-histidine. The numbering of the imidazole atoms of histidine is explicitly shown according to IUPAC nomenclature. Amidic electron delocalization into the nascent carbonyl of 2-oxo-histidine obviates the acid/base character of the imidazole nitrogen atoms.

Unlike intracellular proteins, secretory proteins must be robust to the relatively oxidizing environment of the extracellular space if they are to properly enact their function;[146] this requirement becomes all the more crucial in the case of a protein with oxidation-prone residues such as RNase 1, the prototypical RNase A superfamily member in *Homo sapiens*. We observed that RNase 1, which circulates at approximately 0.5 μg mL^{-1} (≈40 nM) in serum,[147] contains methionine at nearly twice the frequency of that for all human proteins, and that these residues are well-conserved across homologues.[56,57] In RNase 1, the methionine residues cluster in proximity to the active site. In orthogonal proteins, methionine "phalanxes" can line an active site, where the residue reversibly acts as a sacrificial antioxidant to protect other active site residues.[64] We hypothesized that the methionines of RNase 1 serve a similar protective role. Glycoxidation and lipoxidation of methionine have been reported for bovine RNase A (68% sequence identity to RNase 1), though no functional characterization of the oxidized protein was performed.[148] To our knowledge, no work has investigated if conserved methionines serve as antioxidant residues or contribute other functionality to RNase A homologues, including RNase 1.

In the present work, we report that methionine indeed confers oxidative resistance to RNase 1. We employed complementary bioinformatics, chemical biology, and biomimetic approaches to perturb methionine's ability to act as an antioxidant, and we observed the resultant effect of oxidative insult on RNase 1 activity. Our bioinformatic analyses interrogated the evolutionary patterns of methionine usage in the RNase superfamily, and we identified distinct patterns of methionine accumulation across homologues as well as structural motifs that suggest a residue-specific tuning of methionine reactivity. Our chemical biology approach applied a methionine-selective oxaziridine probe to cap the methionyl thioether and hinder its ability to act

as an antioxidant; our work represents the first reported usage of this probe in such a manner. Our biomimetic approach reports the first one-pass chemical synthesis of RNase 1, and we specifically prepared an RNase 1 variant in which select methionines were replaced with the bioisostere norleucine, which lacks the redox-active sulfur atom. We hypothesize that the results of our case study with RNase 1 are relevant to other important secretory proteins, including calprotectin, platelet-activating factor acetylhydrolase, and trypsin 1. Broadly, our findings highlight an underappreciated role that methionine plays in the proper maintenance of multicellular processes in the face of oxidative insult.

2.3 Results

2.3.1 Characterizing methionine conservation patterns among homologues

To begin our bioinformatic analyses, we used the Dali protein structure comparison server to identify a collection of non-redundant RNase 1 homologues.[149] Our homologue set consists of 19 proteins across vertebral species with RNase 1 sequence identity ranging from 24–72% (see **Table 2.1** for the list of homologues). Despite this broad range of sequence identity, all 19 proteins are structurally similar (consensus RMSD: 3.5 Å; see **Figure 2.3**).

Table 2.1. Homologues of *H. sapiens* RNase 1 identified by the Dali protein structure conservation server.[149]

Species	Protein	PDB ID	RMSD[a]	% identity[b]
M. musculus	RNase 1	3TSRA	0.7	71
R. norvegicus	RNase A	1RRAA	0.7	66
S. scrofa	RNase 4	5AR6A	1.2	46
H. sapiens	RNase 4	1RNFA	1.2	44
C. mydas	Egg white RNase	2ZPOA	1.4	42
B. taurus	RNase 5	1AGIA	1.7	34
R. norvegicus	RNase 5	4QFIB	1.7	41
B. taurus	Seminal RNase	11BAA	0.6	72
M. musculus	RNase 5	2BWLA	1.6	38
H. sapiens	RNase 6	4X09A	1.6	36
H. sapiens	RNase 5	1A4YB	1.7	38
H. sapiens	RNase 3	1H1HA	1.8	34
D. rerio	RNase 5	3LJEA	1.7	35
G. gallus	RNase 1	4PERB	2.1	34
H. sapiens	RNase 7	2HKYA	2.1	39
L. pipiens	Amphinase	2P6ZA	1.9	29
L. catesbeianus	RC-RNase 4	1KVZA	2.6	24
L. catesbeianus	RC-RNase 2	1M58A	2.2	27

Proteins all have ≤90% sequence identity with one another. [a]*RMSD*, backbone root-mean-square deviation to *H. sapiens* RNase 1, reported in Ångstroms. [b]*% identity*, sequence identity to *H. sapiens* RNase 1.

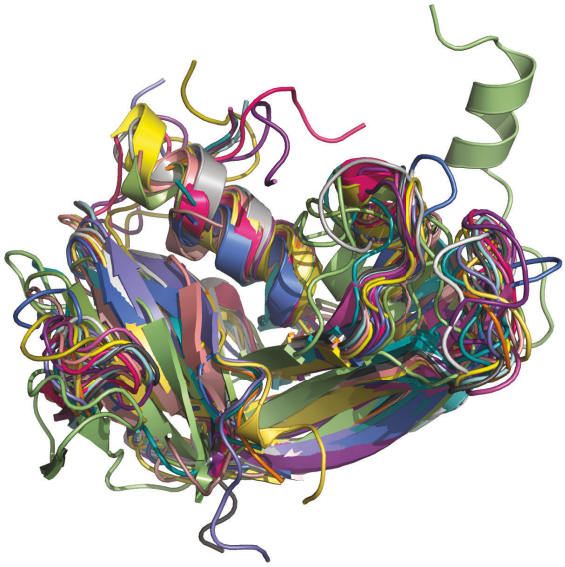


Figure 2.3. Structural alignment of *H. sapiens* RNase 1 and the 18 homologues identified by the Dali protein structure conservation server.[149]

We surveyed our homologue set for patterns of methionine conservation relative to RNase 1 and determined the frequency of conserved residue solvent accessibility as well as the average distance of each conserved methionine to the strictly conserved active site residues homologous to His12, Lys41, and His119. The data in **Table 2.2** reveal several trends: (1) methionines aligned to RNase 1 Met29 and Met30 are frequently present; (2) methionines aligned to Met30 and Met35 are consistently proximal to active site residues; and (3) methionines aligned to Met29 and Met35 are frequently solvent-exposed.

Table 2.2. Methionine conservation patterns among the RNase 1 homologue set.

Residue	Conservation	Distance to… (Å)[a]			Frequency exposed[b]
		His12	Lys41	His119	
Met13	6/19	10 ± 1	14 ± 1	15 ± 2	5/6
Met29	11/19	13 ± 1	14 ± 1	20 ± 1	10/11
Met30	19/19	9 ± 2	7 ± 1	16 ± 2	3/19
Met35	7/19	6 ± 1	4 ± 1	12 ± 1	7/7
Met79	7/19	15 ± 5	19 ± 4	15 ± 1	6/7

[a]For histidine, distance is measured from the methionyl sulfur to the imidazole carbon C2 (see **Figure 2.2**); for lysine, distance is measured from the methionyl sulfur to the terminal sidechain carbon. Distances are mean ± SD (*n* = number of conserved methionine residues). [b]Exposure is determined by solvent accessibility based on the consensus criteria of Marino and Gladyshev with Aledo, Cantón, and Veredas.[82,150]

Structural alignment of the homologue set reveals that despite the conformational flexibility of the methionine sidechain, conserved methionines adopt consistent conformations with retention of the orientation of the methionyl sulfur atom (**Figure 2.4**).

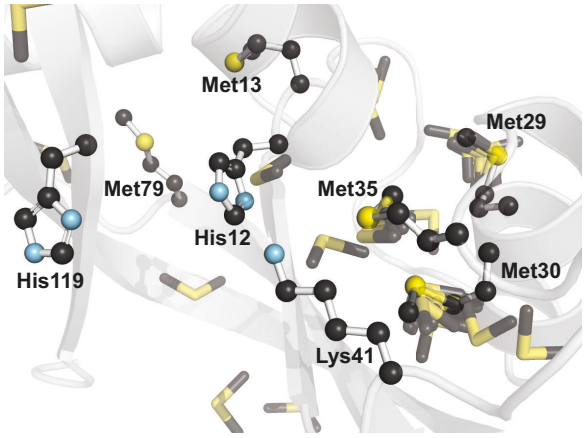


Figure 2.4. Conserved methionines of the homologue set relative to *H. sapiens* RNase 1. The cartoon structure of RNase 1 is depicted with active site residue and methionine sidechains in ball-and-stick form (CPK color scheme); the methionine residues of the homologues are shown in stick form.

Phylogenetic analysis of the homologue set elucidates patterns of methionine conservation.[151] As shown in **Figure 2.5**, methionine appears to accumulate during protein evolution from the more ancestral RNase 5 to the more descendant RNases, which are believed to have originated from gene duplication events.[41-43]

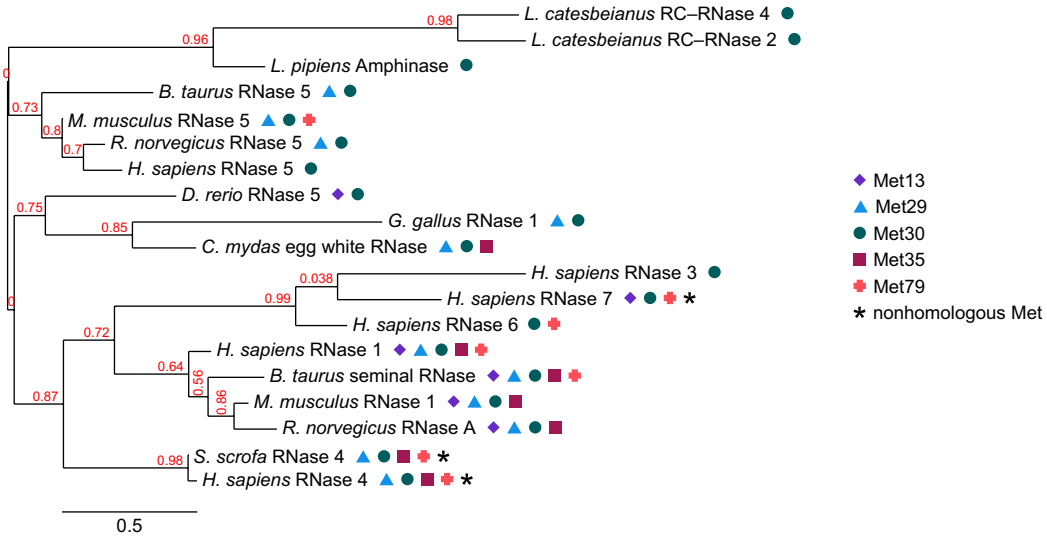


Figure 2.5. Phylogenetic tree of the *H. sapiens* RNase 1 homologue set with symbols representing conserved methionine residues present in a homologue. Residues are numbered according to their equivalent position in *H. sapiens* RNase 1. Node support is in red font.

Intriguingly, greater methionine utilization is seen in the enzymes of the terrestrial species than in those of the aquatic or amphibious species (*cf.*, *L. catesbeianus*, *L. pipiens*, *D. rerio*, and *H. sapiens*). Greater incidence of methionine is also seen among homologues with higher ribonucleolytic activities, both when comparing RNase family members across species—*e.g.,* *G. gallus*, k_{cat}/K_M <1 mM^{-1} min^{-1} versus *H. sapiens* RNase 1, k_{cat}/K_M =12,000 mM^{-1} min^{-1} for the cleavage of poly(A:U)—and when comparing RNase family members within a species—*e.g.,* *H. sapiens* RNase 1, v = 0.108 mmol min^{-1} versus RNase 5, v = 0.000239 mmol min^{-1} for the cleavage of UpA.[40,43] Collectively, evidence suggests that methionine incorporation has been a part of RNase evolution within and across species, and these methionines are frequently installed in the vicinity of the active site.

2.3.2 Identifying *S*–aromatic motifs among homologues

We sought to determine whether structural features relevant to methionine reactivity were present within our homologue set. A logo plot of the aligned homologue set sequences reveals that several aromatic residues are highly conserved (**Figure 2.6**), some nearly as strictly as the active site residues and disulfide bonds responsible for much of the structural stability of RNases.[152,153]

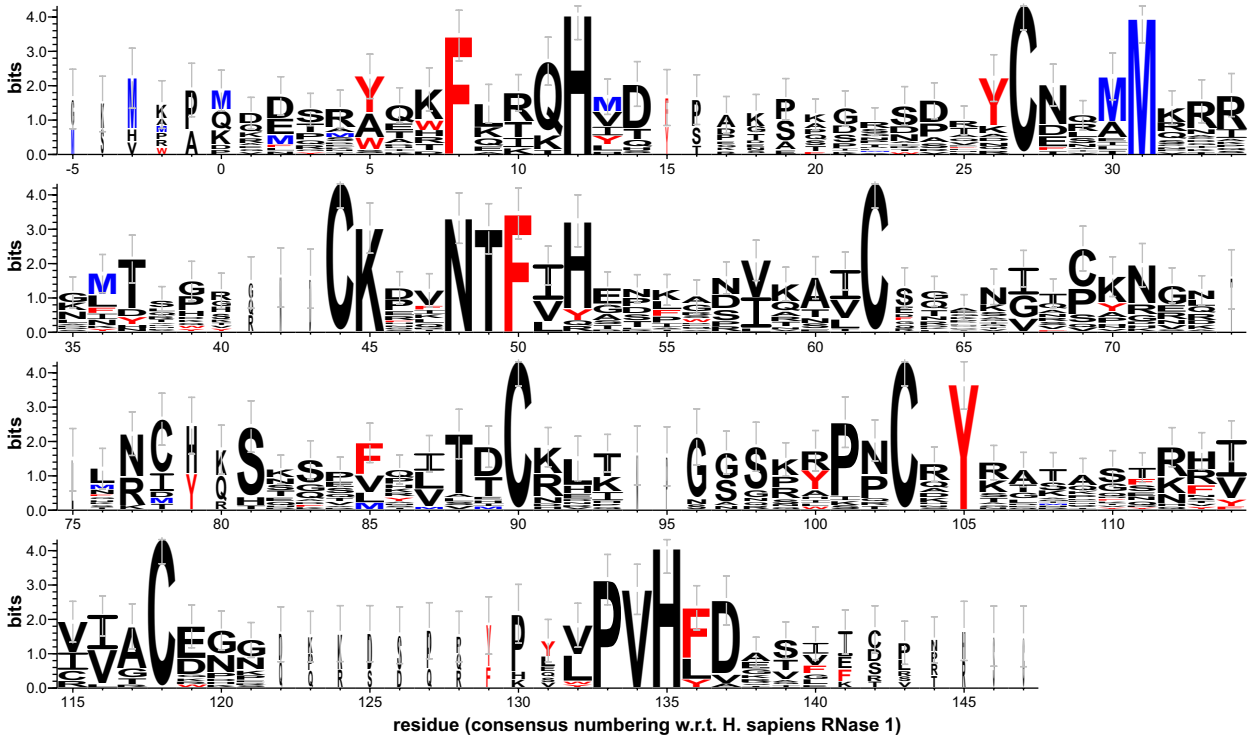


Figure 2.6. Sequence logo of the *H. sapiens* RNase 1 homologue set. Residues are numbered with respect to the sequence of RNase 1; numbering is approximate due to the presence of sequence gaps in the alignment of the homologues. Conserved methionine residues are indicated by blue font; conserved phenylalanine and tyrosine residues are indicated by red font. This plot was produced using WebLogo v3.7.4.[154]

Sulfur–arene interactions, referred to as the *S*–aromatic motif in the context of proteins, have recently been recognized as a significant secondary force in proteins with interaction energies of 1–3 kcal mol^{-1}, on par with salt bridges.[76,78-80] We hypothesized that the conservation of aromatic residues in our homologue set may be related to the patterns of methionine conservation.

We applied a 7-Å cut-off to screen for *S*–aromatic motifs in the structures of our homologue set, consistent with prior work.[79] Several aromatic residues of the homologue set are within the consensus cut-off distance for an *S*–aromatic motif; as shown in **Figure 2.7**, most of these aromatic residues are vicinal to the RNase active site and are highly conserved among homologues.

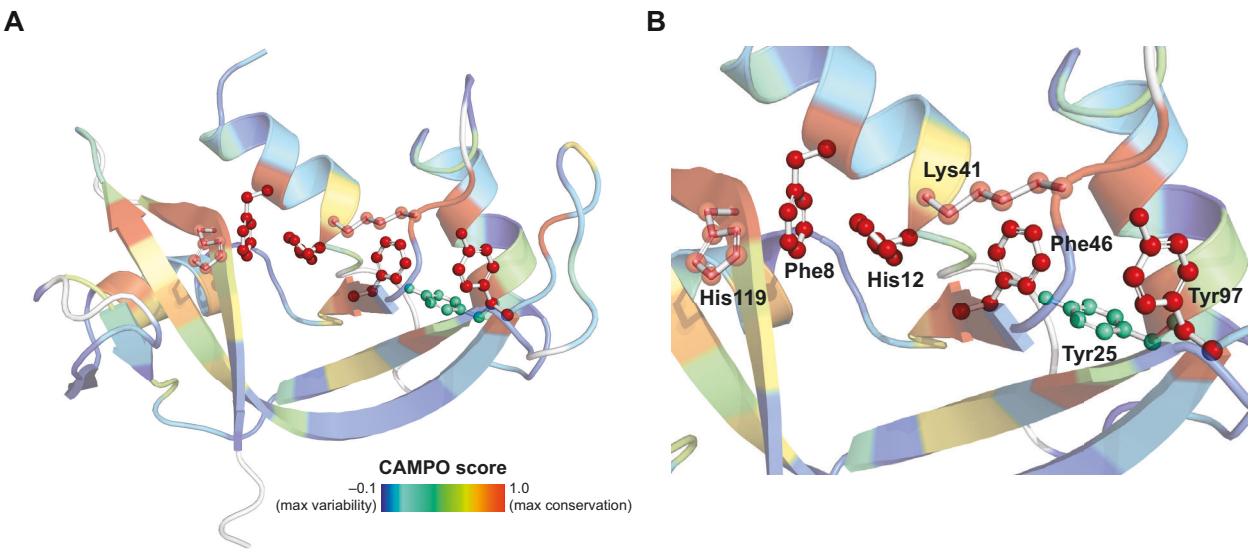


Figure 2.7. *H. sapiens* RNase 1 colored by the level of sequence conservation among the homologue set.[155] (A) Cartoon structure and (B) active site region. Aromatic residues within 7 Å of a methionyl sulfur atom are shown in solid ball-and-stick form; active site residues His12, Lys41, and His119 are shown semi-transparent ball-and-stick form.

We characterized these *S*–aromatic motifs based on (1) the distance between the methionyl sulfur atom and the nearest aromatic ring atom; (2) the angle of approach between the methionyl sulfur and the plane of the aromatic ring; and (3) whether the methionyl thioether group was oriented such that the sulfur atom or the thioether carbon atoms were pointed toward the aromatic ring ("*S*-toward" and "*S*-away", respectively). The *S*-toward interaction directs sulfur lone pair electrons toward the aromatic ring and is best complemented by the partially electron-deficient (δ^+) ring edge, while the *S*-away interaction directs methionyl C–S σ* orbitals toward the aromatic ring and is best complemented by the electron-rich (δ^-) ring faces.[80] As shown in **Figure 2.8**, the sulfur orientation and angle of approach largely complement one another for the *S*–aromatic motifs present in our homologue set: approach angles <30° (*i.e.*, approach along the ring edge) tend to correlate with an *S*-toward orientation, and approach angles >70° (*i.e.*, approach along the ring face) tend to correlate with an *S*-away orientation.

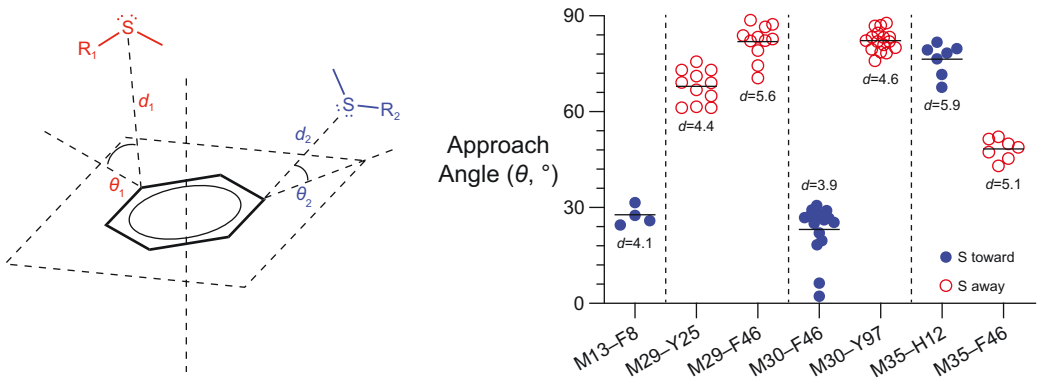


Figure 2.8. Graph of metrics for the *S*–aromatic motifs found in the homologue set. Methionyl sulfur distance (d), angle of approach (θ), and orientation toward aromatic residues are shown for each *S*–aromatic motif. The schematic on the left represents two model *S*–aromatic motifs with one "toward" (in blue) and one "away" (in red). Solid, horizontal lines on the graph indicate the mean approach angle of each motif. Labels on the graph indicate the mean motif distance for the members of the motif. All residues are numbered according to their position in *H. sapiens* RNase 1.

Of the four methionines involved in *S*–aromatic motifs, Met29 is engaged in two interactions (with Tyr25 and Phe46) that direct electron density into the thioether; Met35 is engaged in one donating interaction (with Phe46) and in one abstracting interaction (albeit with the active site residue His12); Met30 is engaged in one donating interaction (with Tyr97) that is offset by an even closer abstracting interaction (with Phe46); and Met13 is engaged in one abstracting interaction (with Phe8). We hypothesize that interactions which donate electron density to methionyl sulfur should render that methionine more predisposed to oxidation, and interactions which abstract electron density from the sulfur atom should render that residue more recalcitrant to oxidation. While a prior report indicates that methionine residues engaged in *S*–aromatic interactions are in general less prone to oxidation, we note that neither the angle of approach or the sulfur orientation were considered in that work;[82] we believe that structural nuances enable the *S*–aromatic interaction to be either activating or deactivating with respect to oxidation. On this basis, we hypothesized that Met35 and Met29 should be more easily oxidized than Met30 or Met13. This is supported by Baynes and coworkers, who determined that the rates of methionine oxidation in bovine RNase A are Met29 > Met30 > Met13 (Met79 was not oxidized, and RNase A lacks Met35).[148] Bovine RNase A was not a part of our homologue set, but investigation of the structure of that protein reveals that it possesses homologous *S*–aromatic motifs (**Figure 2.9**) which conserve the interatomic distances, approach angles, and sulfur orientations observed in our homologue set (**Table 2.3**). Taken collectively, our bioinformatics analyses enabled us to hypothesize that in *H. sapiens* RNase 1, Met29 and Met35 serve as sacrificial antioxidative residues meant to protect active site residues from oxidation.

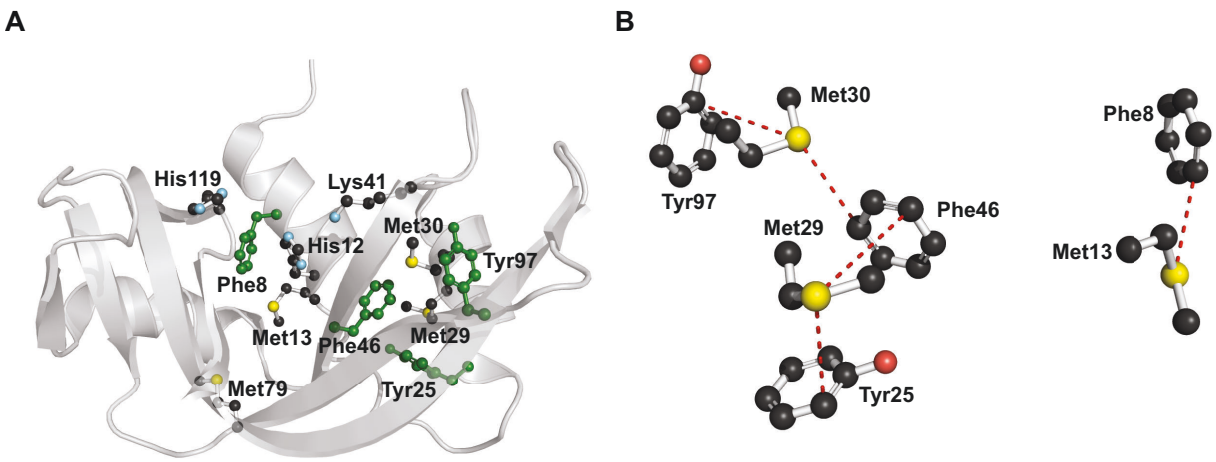


Figure 2.9. Bovine RNase A *S*–aromatic motifs. (A) Structure of RNase A (PDB ID: 7RSA[39]) with active site, methionine, and aromatic residues engaged in *S*–aromatic motifs represented in ball-and-stick form (active site and methionine residues in CPK color scheme; aromatic residues in green). (B) Detailed representation of RNase A *S*–aromatic motifs, indicated by red dashed lines (methionyl sulfur to closest aromatic residue sidechain carbon). Only methionine sidechain and aromatic ring atoms are shown (CPK color scheme).

Table 2.3. *S*–aromatic motifs present in bovine RNase A.

Interaction	Distance (Å)	Approach angle (θ)	Sulfur orientation
Met13–Phe8	4.2	28.2	Toward
Met29–Tyr25	4.4	70.4	Away
Met29–Phe46	5.4	83.5	Away
Met30–Phe46	4.0	27.9	Toward
Met30–Tyr97	4.7	82.0	Away

Metrics are for the motifs depicted in **Figure 2.9**.

2.3.3 Labeling RNase 1 with ReACT Ox2

In our chemical biology approach to interrogating methionine function, we sought to chemically cap RNase 1 methionine residues in a manner analogous to the utilization of synthetic protecting groups, hypothesizing that doing so would render RNase 1 more susceptible to oxidation-induced inactivation by hindering the ability of methionine to act as a sacrificial antioxidant in the protection of catalytic residues. To prevent exogenous methionine (*i.e.*, the N-terminal translation initiator methionine) from confounding our enzyme studies, we prepared authentic RNase 1 by recombinantly expressing the fusion construct M–DDDDK–RNase 1, treating the construct with enterokinase, and purifying the reaction mixture by HPLC to remove the initiator Met-containing tag. Mass spectrometry confirmed that the resultant protein had a molecular weight consistent with authentic RNase 1 (**Figure 2.10**).

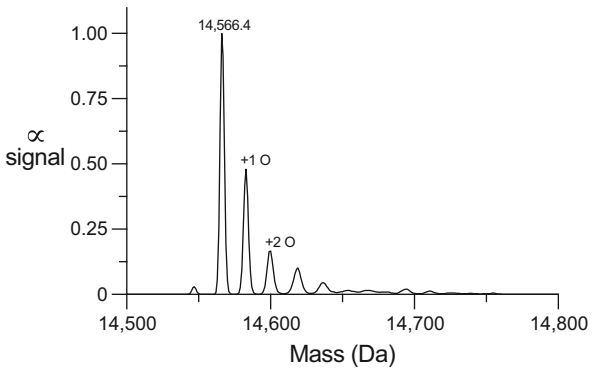


Figure 2.10. Deconvoluted Q–TOF mass spectrum of authentic, wild-type RNase 1 (expected mass: 14,566.2 Da; Δm = 14 ppm). Closely spaced peaks correspond to oxidation events (+O, 16 Da).

We subsequently reacted authentic RNase 1 with the methionine-selective chemical probe ReACT Ox2 reported by Toste, Chang, and coworkers, which effectively pre-oxidizes methionine according to the mechanism shown in **Figure 2.11** and hinders its ability to later react with an introduced oxidant.[156]

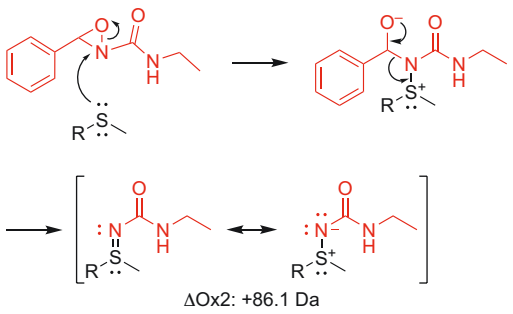


Figure 2.11. Mechanism of methionine labeling by ReACT Ox2 (in red), as described by Toste, Chang, and coworkers.[156] R group: remainder of protein.

We predominantly observe two labeling events (**Figure 2.12**).

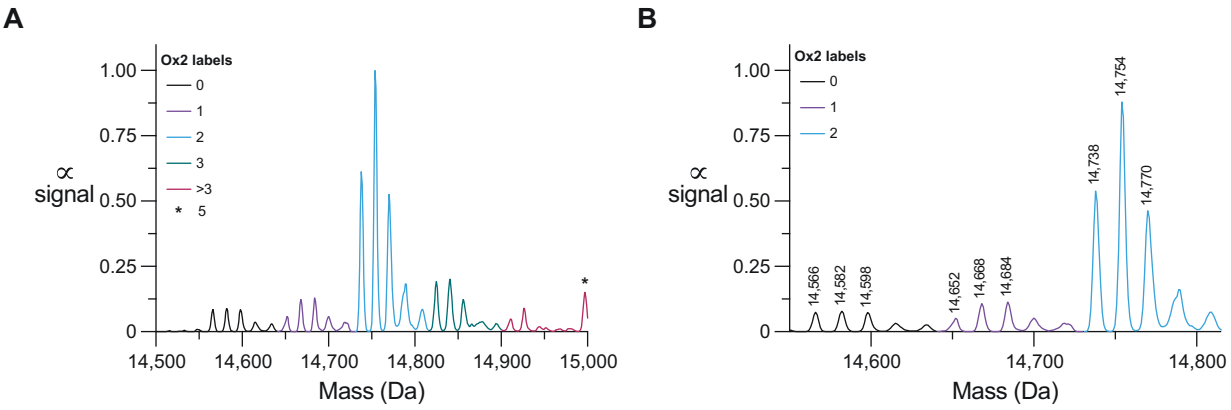


Figure 2.12. Deconvoluted Q–TOF mass spectrum of Ox2-labeled RNase 1. (A) Full spectrum and (B) inset to explicitly show detected masses. Oxidation events correspond to a +16 Da mass shift; Ox2 labeling corresponds to a +86 Da mass shift, enabling its distinction from oxidation events. The predominant labeled species contains two Ox2 labels.

Of the five RNase 1 methionine residues, tryptic digestion of Ox2-labeled protein clearly identifies Met35 and eliminates Met79 as solvent-accessible targets, while a single labeling event is detected on the peptide harboring Met13, Met29, and Met30 (**Figure 2.13**).

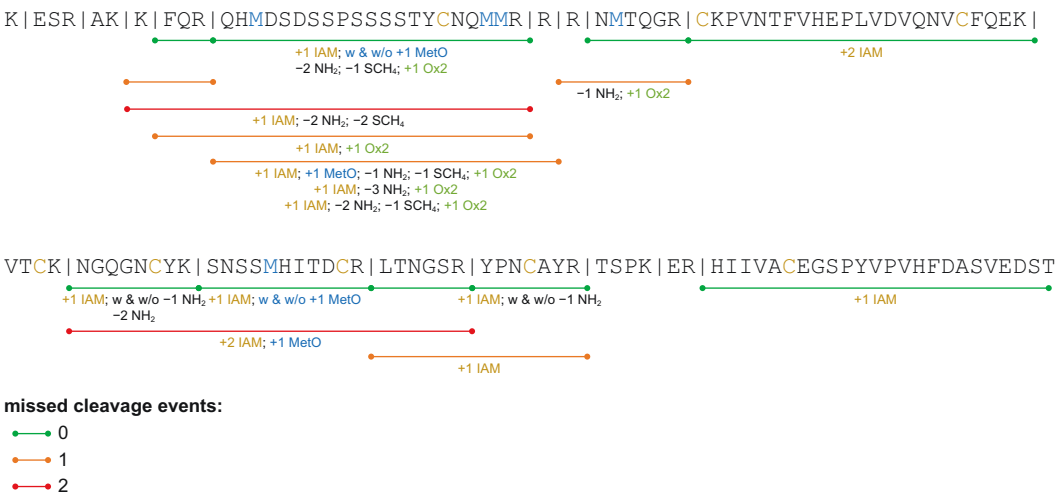


Figure 2.13. Tryptic peptide mapping for Ox2-labeled RNase 1, analyzed by Q–TOF mass spectrometry. Colored lines beneath the RNase 1 sequence indicate detected peptides (green: no missed tryptic cleavage; orange: one missed cleavage; red: two missed cleavages), with each line of text indicating a unique peptide and its predicted modifications. +IAM, cysteine alkylation with iodoacetamide; +MetO, methionine oxidation; +Ox2, methionine labeling with ReACT Ox2; $-NH_2$, asparagine/glutamine deamidation; $-SCH_4$, methionine dethiomethylation.[157] Peptides were detected at a Δ10 ppm cut-off. Sequence coverage: 90.6%.

Heat denaturation experiments show that RNase 1 stability is minimally destabilized by Ox2 labeling ($\Delta T_m = -4.0 \pm 0.1$ °C, mean ± SE, $n = 4$; **Figure 2.14**).

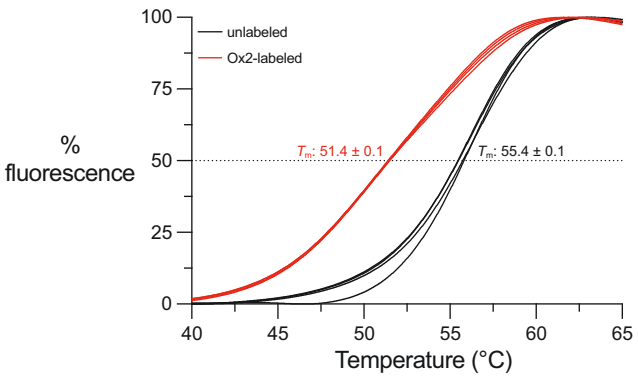


Figure 2.14. Heat denaturation curves for wild-type, unlabeled RNase 1 and Ox2-labeled RNase 1. $\Delta T_m = -4.0 \pm 0.1$ °C, mean ± SE, $n = 4$.

As measured with the chimeric, FRET-tagged RNase oligonucleotide substrate 6-FAM–dArU$(dA)_2$–6-TAMRA, the catalytic efficiency of Ox2-labeled RNase 1 is (20 ± 1)% that of wild-type, unlabeled RNase 1 (mean ± SE, n = 3). Thus, labeling RNase 1 with Ox2 perturbs enzymic stability and activity only to a minor extent, enabling our use of Ox2 as a tool to interrogate methionine's ability to act as a sacrificial antioxidant.

2.3.4 Preparing a biomimetic variant of RNase 1

Based on our bioinformatics analyses and Ox2 labeling results, we chose to rationally design an RNase 1 variant in which the putative antioxidative capability of Met29 and Met35 was obviated by amino acid substitution. Considering that Ox2 labeling of RNase 1 methionines introduces steric bulk that could possibly confound the results of our subsequent oxidation-induced activation assay (*e.g.*, by altering the conformation of the RNase 1 active site), we sought to install an amino acid that would remove the methionyl sulfur atom but not otherwise alter the residue sidechain. While no canonical amino acid meets these criteria, the noncanonical amino acid norleucine (Nle) is a bioisostere of methionine in which the sulfur atom is replaced with a methylene group (**Figure 2.15**); substitution of norleucine for methionine has successfully been used by others to interrogate the function of methionine in proteins.[158-160]

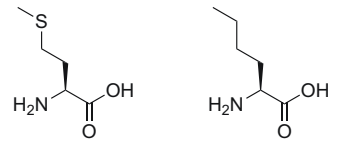


Figure 2.15. Comparison of L-methionine (L-Met) and L-norleucine (L-Nle).

While prior work has installed norleucine into recombinant proteins using chemically-defined expression systems, our desire to site-selectively install norleucine necessitated chemical synthesis of our variant Met29Nle Met35Nle RNase 1, which we refer to as *Nle–RNase 1*. Using the methods recently reported by Pentelute and coworkers, we successfully prepared the polypeptide for Nle–RNase 1 in one pass using automated fast-flow peptide synthesis (AFPS).[161] We dissolved and folded the polypeptide, purified the folded protein using HPLC, and confirmed that the protein mass was consistent with the molecular weight of folded Nle–RNase 1 (**Figure 2.16**).

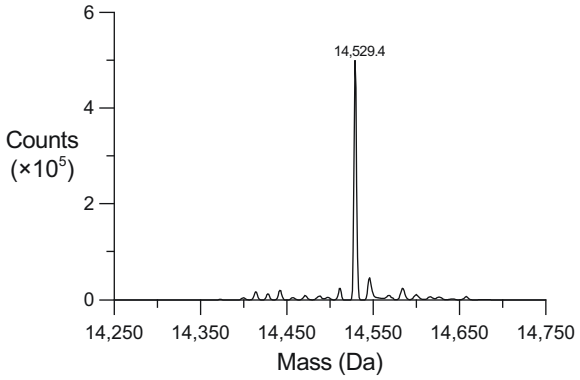


Figure 2.16. Deconvoluted Q–TOF mass spectrum of purified Nle–RNase 1 (expected mass: 14,530.1 Da; Δm = −48 ppm). The observed molecular weight corroborates that Nle–RNase 1 is properly folded; the expected and observed pre-folding molecular weight is 14,538.2 Da due to the presence of eight reduced cysteine residues rather than the four disulfide bonds present in folded RNase 1.

Our yield of refolded, purified Nle–RNase 1 was 24% relative to the starting mass of synthesized polypeptide, which is consistent with results for the synthetic preparation of other proteins via AFPS.[161] The catalytic efficiency of Nle–RNase 1, as measured with 6-FAM–dArU(dA)$_2$–6-TAMRA, is (26 ± 1)% that of wild-type RNase 1 (mean ± SE, n = 3); our results are consistent with what we observed for Ox2-labeled RNase 1 and suggest that Nle–RNase 1 is suitable for examining the function of Met29 and Met35 in RNase 1.

2.3.5 Challenging RNase 1 in an oxidation-induced inactivation assay

To investigate whether methionine acts as a sacrificial antioxidant to protect the oxidation-prone catalytic residues His12, Lys41, and His 119, we challenged RNase 1, Ox2-labeled RNase 1, and Nle–RNase 1 in an oxidation assay. Each protein was incubated with varying concentrations of hydrogen peroxide and iron(II) sulfate (100:1 H_2O_2:$FeSO_4$) at 37 °C for 3 h. Iron(II) sulfate acts as a catalyst for the decomposition of hydrogen peroxide in a Fenton-type reaction to produce the hydroxyl radical, which can go on to react with histidine and lysine to produce 2-oxo-histidine and aminadipic semialdehyde, respectively;[145,162] the use of iron(II) sulfate is particularly important, as the oxidation of biomolecules by hydrogen peroxide is thermodynamically favored but kinetically hindered, while the reaction of the hydroxyl radical with amino acids has a second-order rate constant approximately 10^6–10^9-fold greater.[163] Following oxidation, the proteins were assayed for ribonucleolytic activity using 6-FAM–dArU(dA)$_2$–6-TAMRA. We hypothesized that because Ox2-labeled RNase 1 contains chemically capped Met29 and Met35, the enzyme should be rendered less protected against inactivation by oxidation of catalytic residues and should therefore lose catalytic competency at a lower number of equivalents of hydrogen peroxide than is required to inactivate wild-type, unlabeled RNase 1. We likewise

hypothesized that Nle–RNase 1, in which the sulfur atom of Met29 and Met35 has been obviated by residue substitution with norleucine, should also be more sensitive to inactivation via oxidation. Our hypothesis is graphically presented in **Figure 2.17**.

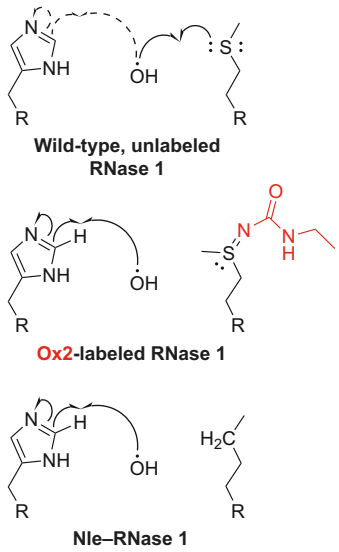


Figure 2.17. Proposed mechanism of attack of each RNase 1 variant by an oxidant, here represented by the hydroxyl radical. In all three enzymes, an oxidant may react with the catalytic residues His12 or His119. In wild-type, unlabeled RNase 1, the oxidant may instead react with a proximal methionine, thereby protecting the enzyme's catalytic residues. In both Ox2-labeled RNase 1 and the norleucine variant Nle–RNase 1, this putative, sacrificial antioxidative function of methionine is hindered, rendering the enzyme more susceptible to inactivation.

As shown in **Figure 2.18**, Ox2-labeled RNase 1 is indeed more susceptible to oxidation-induced inactivation: IC_{50} = 111 ± 8 equivalents, versus IC_{50} = 222 ± 18 equivalents for wild-type, unlabeled RNase 1 (mean ± SE, n = 3). Similarly, Nle–RNase 1 is more easily inactivated by oxidation: IC_{50} = 89 ± 17 equivalents (mean ± SE, n = 3). The equivalents required for attenuation of ribonucleolytic activity are consistent (within error) for both Ox2-labeled and norleucine-substituted RNase 1; wild-type RNase 1 is approximately two- and threefold more stable to the loss of ribonucleolytic activity due to oxidation than are Ox2-labeled RNase 1 and Nle–RNase 1, respectively. Our chemical biology and biomimetic approaches support the hypothesis that Met29 and Met35 serve an antioxidative, protective role for RNase 1 catalysis.

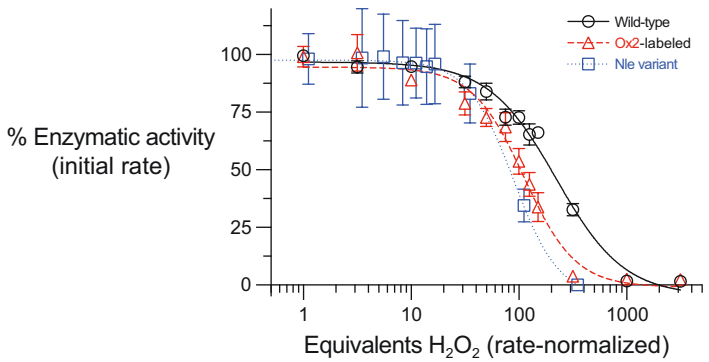


Figure 2.18. Graph showing the oxidation-induced inactivation assay inhibition curves for wild-type, unlabeled RNase 1 (black solid line), Ox2-labeled RNase 1 (red dashed line), and Nle–RNase 1 (blue dotted line). Equivalents of hydrogen peroxide are rate-normalized to account for differing protein concentrations.

2.4 Discussion

The correlation between oxidation of catalytic residues and loss of enzymatic activity has been appreciated for many decades; the work of Weil and coworkers in the 1950s carefully characterized the relationship between the oxidation of histidine/tryptophan and the loss of enzymatic activity in lysozyme, chymotrypsin, and bovine RNase A.[45,47,48] Their work demonstrating that oxidation of two molar equivalent of histidine correlated with near-total loss of ribonucleolytic activity is particularly noteworthy, given that the mechanism of RNase A catalysis would not be reported for another six years.[25] In addition, early work regarding the enzymology of RNase A by Gundlach, Stein, and Moore demonstrated loss of catalytic activity upon modification of histidine residues, this time by alkylation of the histidine sidechain with iodoacetate. Intriguingly, they observed that alkylation of RNase A most expeditiously led to loss of catalytic activity at pH 5.5 and specifically correlated with a decrease in native histidine content (determined by amino acid analysis), but alkylation of the imidazole nitrogen is slow at a non-alkaline pH.[50] Further work identified that the iodoacetate-induced alkylation of methionine is distinctly pH-independent and forms a reactive carboxymethylsulfonium species, leading Gundlach and coworkers to speculate that the inactivation of RNase A with iodoacetate may occur via a two-step mechanism in which (1) carboxymethylation of one or more methionine residues proximal to catalytic histidine is followed by (2) subsequent transfer of the alkyl group from the activated carboxymethylsulfonium group to the nearby histidine (**Figure 2.19**).[50,51,164] Thus, even early research into ribonucleases alluded to a relationship between histidine and methionine within the enzymic active site.

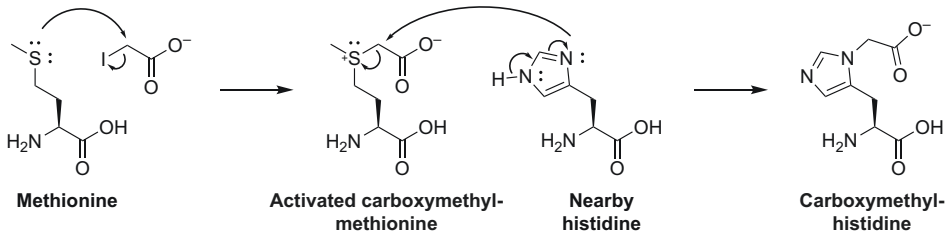


Figure 2.19. Proposed mechanism of carboxymethylation of histidine in bovine RNase A at pH 5.5. Gundlach, Stein, and Moore propose that methionine is first carboxymethylated by iodoacetate in a pH-independent manner, generating an activated carboxylmethylsulfonium species that subsequently alkylates nearby histidine residues.[50]

Our findings for *H. sapiens* RNase 1 and its homologues are more striking when framed in the context of amino acid utilization frequencies. Among human proteins, methionine is one of the most infrequently utilized amino acids, second only to tryptophan.[56,57] Methionine comprises 2.2% of the global amino acid content of human proteins, a frequency which includes its retention as the translation initiator on proteins for which methionine aminopeptidase is not able to cleave the N-terminal methionine.[58] Histidine, meanwhile, is greatly overrepresented in enzymic active sites: 21% of enzymes utilize histidine in their catalytic function, while the amino acid globally comprises 2.6% of protein residues.[57,165] RNase 1 has five methionine residues (the translational initiator is removed alongside a signal peptide during posttranslational processing) and thus utilizes the amino acid at a frequency of 3.9%, nearly double that expected for a protein of its size (128 residues). Statistical overrepresentation of methionine within an enzyme's sequence should be cause for further investigation to determine whether that enzyme uses methionine for antioxidative protection, particularly if the enzyme uses oxidation-prone residues such as histidine in its catalytic mechanism or other function.

As a brief demonstration of this premise, we queried human secretory proteins known to utilize histidine and surveyed their incidence of methionine, reasoning that the reduction potential of the extracellular environment is much lower than the cytosol, and secretory proteins should therefore endure greater oxidative insult.[166] The first of our query secretory proteins, calprotectin, is a heterotetrameric calcium- and zinc-binding protein implicated in several inflammatory diseases.[167] Methionine comprises 3.9% of its amino acid content, and one methionine, $Met94_D$ (subscript denoting protomer), is found within ~7 Å of the zinc-coordinating residues $His91_D$ and $His95_D$; this same methionine is also proximal to the zinc-coordinating residues $His17_B$ and $His27_B$ (**Figure 2.20**).

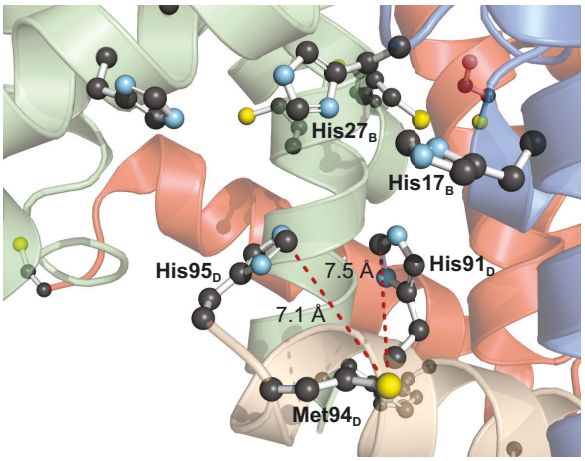


Figure 2.20. Methionine utilization near a calprotectin zinc-binding site (PDB ID: 1XK4).[167] Calprotectin is depicted in cartoon form with zinc-binding residue and proximal methionine sidechains in ball-and-stick form (CPK color scheme). Each cartoon color corresponds to one protomer of the heterotetrameric protein (also indicated by subscript letters). Distances between proximal methionyl sulfur atoms and histidine carbon C2, which is susceptible to oxidative attack, are shown in Ångstroms.

In addition to being secreted, calprotectin is found in high abundance in the cytoplasm of neutrophils, which use a "respiratory burst" to produce reactive oxygen species for their antimicrobial effect;[168] it may be that $Met94_D$ sacrificially protects zinc-coordinating histidine residues in their native, reduced state in the face of such oxidizing conditions, allowing the latter to preserve their metal ion coordinating function. We note that methionine oxidation can be corrected by methionine sulfoxide reductases, enabling a methionine residue to act as a sacrificial antioxidant more than once.[76]

As a second example, we identified platelet-activating factor (PAF) acetylhydrolase as an enzyme in which methionine is proximal to the active site. PAF acetylhydrolase, as the name implies, is responsible for hydrolyzing platelet-activating factor, a phospholipid messenger produced as a part of host defense that is implicated in many inflammatory diseases, including those associated with oxidative stress.[169] PAF acetylhydrolase possesses 11 methionine residues, which is slightly overrepresented at 2.6% utilization. As shown in **Figure 2.21**, two of these methionines are found within approximately 10 Å of the catalytic residue His351, and a third methionine is found only slightly further away.

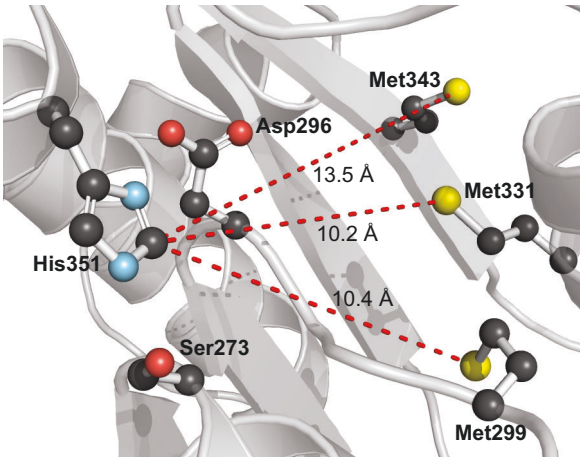


Figure 2.21. Methionine utilization near the PAF acetylhydrolase catalytic triad (PDB ID: 3D59A).[169] PAF acetylhydrolase is depicted in cartoon form with catalytic residue and proximal methionine sidechains in ball-and-stick form (CPK color scheme). Distances between proximal methionyl sulfur atoms and histidine carbon C2, which is susceptible to oxidative attack, are shown in Ångstroms.

PAF acetylhydrolase relies on a classical His–Asp–Ser hydrolase catalytic triad, and oxidation of His351 would disrupt the enzyme's catalytic mechanism due to the amidic resonance of 2-oxo-histidine hindering its ability to abstract protons (**Figure 2.2**) and thereby increase the nucleophilicity of the catalytic residue Ser273. Because PAF acetylhydrolase is associated with both high-density and low-density lipoprotein in plasma and is implicated in the oxidative stress response, the enzyme natively encounters oxidizing environments and must be able to preserve its catalytic function in the face of such conditions;[170,171] thus, we hypothesize that methionine serves as a sacrificial antioxidant in this protein.

As a counterexample of methionine overrepresentation being a strictly necessary indicator of antioxidative utility in a protein, we surveyed the quintessential protease trypsin 1, which is secreted in the pancreas and likewise uses a His–Asp–Ser catalytic triad to proteolyze its many substrates.[172] This enzyme possesses only two methionine residues and in fact underutilizes the amino acid (0.89% amino acid content). In spite of this underrepresentation, the two methionines are found on a single face of the protein, proximal to the enzymic active site and approximately 10 Å from the catalytic residue His57 (**Figure 2.22**).

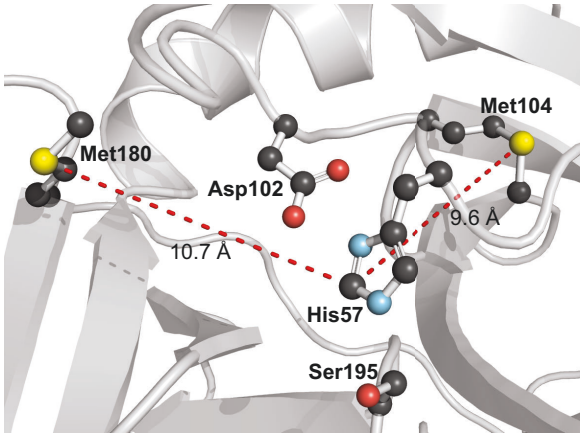


Figure 2.22. Methionine utilization near the trypsin 1 catalytic triad (PDB ID: 1TRNA).[172] Trypsin 1 is depicted in cartoon form with catalytic residue and proximal methionine sidechains in ball-and-stick form (CPK color scheme). Distances between proximal methionyl sulfur atoms and histidine carbon C2, which is susceptible to oxidative attack, are shown in Ångstroms.

Trypsin, like RNase 1 and PAF acetylhydrolase, relies on histidine as a general acid/base, and we believe oxidation of His57 would perturb trypsin's catalytic mechanism in much the same way we speculate it would disrupt PAF acetylhydrolase catalysis.[172] Thus, although methionine overrepresentation within a protein can be suggestive of antioxidative functionality, methionine installation within secretory proteins which rely upon histidine may broadly be a fruitful area of scientific exploration regardless of the amino acid's sequence frequency.

We note that not all methionine residues are created equal, either in terms of their function within a protein or in terms of their oxidative reactivity. Baynes and coworkers corroborated this premise by observing differential rates of methionine oxidation in bovine RNase A with one residue, Met79, being entirely recalcitrant to oxidation;[148] we similarly observed that Met79 did not react with the redox-activated probe Ox2 (**Figure 2.13**), despite the fact that Met79 tends to be solvent-exposed (**Table 2.2**). As we show, the *S*–aromatic motif appears to be implicated in tuning the reactivity of methionine to render certain residues more susceptible to oxidation while inducing robustness against oxidation in others. For example, we observed in our oxidation-induced inactivation assay that perturbation of Met29 and Met35, whether by Ox2 labeling or norleucine installation, renders RNase 1 more susceptible to inactivation (**Figure 2.18**); these same methionine residues are engaged in *S*–aromatic motifs which push electron density toward the methionyl sulfur and should therefore be activating (**Figure 2.8**). While certain studies have suggested that the *S*–aromatic motif only decreases the likelihood of methionine oxidation,[82] more nuanced investigations suggest that the *S*–aromatic motif can be activating *or* deactivating. Sachs and coworkers recognized this fact in their computational studies of dimethylsulfide interacting with benzene, phenol, and indole; they demonstrated differential interaction energies in sulfur–arene interactions depending on whether the sulfur atom or the adjacent carbon atoms

were pointed toward the aromatic ring face.[79] Conversely, the oxidation of methionine may alter the nature of an existing *S*–aromatic motif. Work by Orabi and English showed that an *S*–aromatic motif is destabilized by up to 80% upon methionine oxidation to the corresponding sulfoxide or sulfone, which they attribute to the greater preference of the sulfoxide/sulfone (relative to the sulfide) to hydrogen bond with bulk water instead of interacting with the aromatic residue.[173] Methionine oxidation events could, therefore, trigger conformational changes that expose new parts of the protein. This possibility is particularly interesting for RNase 1, which possesses the adjacent residues Met29 and Met30. Among the homologue set, Met29 tends to be solvent-exposed (**Table 2.2**), and we have implicated it in antioxidative utility for RNase 1; conversely, Met30 is very infrequently exposed among the homologues. Oxidation of Met29 may cause it to disengage from *S*–aromatic interactions with Tyr25 and Phe46 to instead hydrogen bond with bulk water or other protein residues, triggering conformational changes that expose Met30 and thereby facilitate continued antioxidative function to protect the RNase active site. Such possibilities merit further study.

2.5 Conclusions

In this work, we have used bioinformatics, chemical biology, and biomimetic strategies to empirically determine that Met29 and Met35 of *H. sapiens* RNase 1 function as antioxidative residues that flank the active site, a phenomenon seen in other proteins.[64] Our work demonstrates a novel use of the oxaziridine probe ReACT Ox2—as a way to install a chemical capping group on methionine to interrogate a residue's function in a protein. In the course of interrogating methionine function, we also report the first one-pass synthesis of RNase 1 and demonstrate that it can be used to site-selectively install non-canonical amino acids without the need for piecemeal procedures such as native chemical ligation. Because of the observed patterns of methionine utilization in RNase superfamily homologues, we believe that our findings for RNase 1 may be applicable to homologues. Broadly, we also believe that methionine serves an underappreciated function within the proteinogenic toolkit: to protect catalytic histidine residues in the face of oxidatively challenging environments.

2.6 Acknowledgments

E.C.W. was supported by a Graduate Research Fellowship from the NSF. This work was supported by Grants R35 GM148220 and R01 CA07308 (NIH).

The formulation of M9 medium was kindly provided by the Drennan lab at MIT. The authors are grateful to Dr. Nile S. Abularrage for advice regarding *in silico* analyses, JoLynn B. Giancola for intellectual contributions related to the use of Ox2 as a capping group, Jinyi Yang for assistance with preparative HPLC, and Dr. Joomyung (Vicky) Jun for technical expertise with ^{1}H NMR.

2.7 Experimental Details

2.7.1 Characterizing methionine conservation patterns among homologues

H. sapiens RNase 1 was queried by the Dali protein structure comparison server using PDB ID 1Z7XZ.[149,174] Preliminary query results were manually curated to remove entries that did not correspond to an endogenous protein (*e.g.*, PDB-deposited structures of RNases harboring heterologous mutations). Using the PyMod v3.0.2 plugin in a PyMOL environment, the curated homologue set sequences (n = 19 proteins) were aligned using MUSCLE and structurally aligned using PyMOL.[175,176] For the distance between conserved methionine and catalytic residues, PyMOL was used to measure the distance from the methionyl sulfur atom to the histidine imidazole carbon atom between the two nitrogen atoms (C2), or from the methionyl sulfur atom to the terminal sidechain carbon atom for lysine, as these atoms have been identified as the oxidative targets for those two amino acids.[145]

To classify homologue set methionine residues as solvent-exposed or buried, two methods were applied. In the first, we used PyMOL to measure the solvent-accessible surface area (SASA) of the methionyl sulfur atom; applying the criterion set forth by Marino and Gladyshev, we characterized a methionine residue as exposed if the SASA of its sulfur atom is ≥0.1 $Å^2$.[150] In the second method, we used the VADAR model protein structure analysis server (v1.8) to measure the fractional accessible surface area (ASA) of all methionine residues;[177] applying the criterion established by Aledo, Cantón, and Veredas, we characterized a methionine residue as exposed if its fractional ASA is ≥5%.[82] To harmonize these methods, a methionine residue was classified as exposed if at least one of these two criteria were met; otherwise, the residue was classified as buried.

Phylogenetic analysis of the homologue set was performed using the Phylogeny.fr web server with the previously-generated MUSCLE sequence alignment.[151] Branch support was determined using the approximate likelihood-ratio test.

2.7.2 Identifying *S*–aromatic motifs among homologues

A logo plot of the aligned homologue set sequences was generated using WebLogo v3.7.4.[154] The structures of the homologue set were colored by the sequence alignment using the CAMPO algorithm (Blosum62 scoring matrix; gap score: −1; gap-to-gap score: 0; gaps tossed).[155] *S*–aromatic motifs were identified using PyMOL to screen for distances ≤7 Å between conserved methionyl sulfur atoms and the nearest sidechain ring atom of aromatic residues. Approach angles were calculated using PyMOL and are defined as the altitude angle between the plane of the aromatic ring and the line from the nearest ring atom to the methionyl sulfur atom (see schematic in **Figure 2.8**). The sulfur orientation was defined based on which methionyl sidechain atoms were closest to the aromatic ring: *toward* if the methionyl sulfur atom was closest, *away* if the methionyl thioether carbon atoms were closest.

2.7.3 Preparing authentic RNase 1

Our preparation of M–DDDDK–RNase 1 follows methods described previously.[178] Briefly, site-directed mutagenesis of a pET-22b(+)/*H. sapiens* RNase 1 plasmid was conducted using the Q5 site-directed mutagenesis kit (New England BioLabs) to install codons for Asp–Asp–Asp–Asp–Lys between the translation initiator Met codon and the gene sequence of mature RNase 1 using the following primers (Integrated DNA Technologies):

Forward: 5′-TGA-TAA-GAA-AGA-ATC-TCG-TGC-TAA-AAA-ATT-C-3′

Reverse: 5′-TCA-TCA-TCC-ATA-TGT-ATA-TCT-CCT-TC-TTA-AAG-3′

PCR amplification of the RNase 1 plasmid with the above primers was followed by 5′ phosphorylation of the linear amplicon, circularization by DNA ligase, and digestion of the template plasmid by DpnI (New England BioLabs). NEB 10-beta *E. coli* (New England BioLabs) were transformed by heat shock, grown overnight on 2% w/v LB-Miller agar plates containing ampicillin (200 μg mL^{-1}), and subsequently used to inoculate small-scale cultures (also LB-Miller containing ampicillin (200 μg mL^{-1})). The recombinant plasmid was purified by Miniprep (Thermo Scientific); plasmids were confirmed by Sanger sequencing (Quintara Biosciences) to contain the N-terminal enterokinase recognition tag.

Electrocompetent BL21(DE3) *E. coli* were transformed with the assembled plasmid by electroporation and positive selection using 200 μg mL^{-1} ampicillin. LB-Miller containing ampicillin (200 μg mL^{-1}) and glucose (1% w/v) was inoculated with a single transformant and grown overnight at 37 °C with orbital shaking at 250 RPM. Bacteria were pelleted and resuspended in M9 medium (42.3 mM Na_2HPO_4, 22 mM KH_2PO_4, 18.7 mM NH_4Cl, 8.6 mM NaCl, 2 mM $MgSO_4$, 100 μM $CaCl_2$, 31 μM $FeCl_3$, 3.67 μM $ZnCl_2$, 0.74 μM $CuCl_2$, 0.42 μM $CoCl_2$, 81 nM $MnCl_2$, 1.61 μM H_3BO_3) containing riboflavin (1 μg mL^{-1}), nicotinamide (1 μg mL^{-1}), pyridoxine hydrochloride (1 μg mL^{-1}), thiamine (1 μg mL^{-1}), glucose (0.4% w/v), and ampicillin (200 μg mL^{-1}). Cultures were inoculated to a starting OD_{600} = 0.05 and grown at 37 °C as before until OD_{600} = 1.00. Cultures were then supplemented with L-amino acids at the following concentrations: 100 mg L^{-1} each lysine, threonine, and phenylalanine; 50 mg L^{-1} each leucine, isoleucine, valine, and methionine. Protein overexpression was induced with isopropyl

β-D-1-thiogalactopyranoside (1 mM) for 4 hours with incubation as before. Cultures were pelleted by centrifugation, and pellets were stored at –70 °C.

Pelleted *E. coli* cells were thawed and then resuspended in 20 mM Tris-HCl buffer, pH 7.6, containing EDTA (10 mM). Cells were passed through a cell disruptor (Constant Systems) at 22 kPSI, the lysate was centrifuged, and inclusion bodies were denatured in 20 mM Tris-HCl buffer, pH 8.0, containing guanidinium-HCl (7 M), EDTA (10 mM), and DTT (100 mM). Solubilized protein was diluted tenfold with acetic acid (20 mM) and clarified by centrifugation; the supernatant was dialyzed approximately 100-fold against acetic acid (20 mM) overnight at 4 °C to further clarify. Following centrifugation of the retentate, the supernatant was dropwise added to a refolding buffer of 100 mM Tris-HCl, pH 7.8, containing EDTA (10 mM), L-arginine (0.5 M), reduced glutathione (1.0 mM), and oxidized glutathione (0.2 mM). The protein was allowed to refold at 4 °C without agitation for 5 days. The refolded protein was purified by gel filtration (HiLoad 26/600 Superdex 75 prep grade, Cytiva) and cation exchange (HiTrap Q, Cytiva) fast protein liquid chromatography; all chromatography buffers were treated with DEPC prior to use, except for Tris, which was added from ribonuclease-free stocks. The purified protein was buffer-exchanged into 20 mM Tris-HCl buffer, pH 7.2, containing NaCl (50 mM). The protein was concentrated, quantified by BCA assay, flash-frozen in liquid nitrogen, and stored at –70 °C.

To prepare authentic RNase 1, M–DDDDK–RNase 1 was diluted to 1 mg mL^{-1} using 20 mM Tris-HCl buffer, pH 7.2, containing NaCl (50 mM) and $CaCl_2$ (2 mM). Bovine enterokinase light chain (New England BioLabs) was added to the reaction mixture at 50 U enterokinase per 1 mg M–DDDDK–RNase 1. The reaction mixture was gently agitated at room temperature for 24 h. Completion of digestion with specific cleavage only at the N-terminal enterokinase recognition sequence was confirmed by mass spectrometry. Enterokinase-digested RNase 1 was buffer-

exchanged into 50 mM sodium phosphate buffer, pH 2.7, and purified by semi-preparative reversed-phase HPLC on a C4 column (Vydac, 10 × 250 mm, 5.0–μm particle, 300-Å pore) that was equilibrated with the same buffer at room temperature. The column was washed with 12% v/v acetonitrile in 50 mM sodium phosphate buffer, pH 2.7, for 5 min followed by a 12–60% v/v gradient (again with 50 mM sodium phosphate buffer, pH 2.7) for 50 min. Fractions were manually collected based on detection at A_{218} and A_{280}. HPLC fractions confirmed by mass spectrometry to contain authentic RNase 1 were combined in regenerated cellulose dialysis tubing (Spectrum Laboratories) and dialyzed 500-fold against 20 mM Tris-HCl buffer, pH 7.2, containing NaCl (50 mM). The protein was concentrated, quantified by BCA assay, flash-frozen in liquid nitrogen, and stored at –70 °C.

2.7.4 Synthesis of ReACT Ox2

Our synthesis of the oxaziridine probe Ox2 follows the methods reported by Toste, Chang, and coworkers.[156] Benzaldehyde (12 mmol) and urea (10 mmol) were added to tetrahydrofuran (20 mL) along with titanium(IV) isopropoxide (14 mL). The reaction mixture was gently stirred at room temperature for approximately 16 h; the crude mixture was then concentrated under vacuum and dissolved into dichloromethane (30 mL). A solution of *m*-chloroperoxybenzoic acid (mCPBA, 30 mmol) was prepared in equal parts saturated potassium carbonate and dichloromethane (60 mL total); after stirring the mCPBA solution at room temperature for 10 min, the crude product mixture was dropwise added to the mCPBA solution. The reaction mixture was stirred for a further 6 h, at which point water (100 mL) was added. The reaction mixture was extensively extracted with dichloromethane, and the organic layers containing Ox2 were combined. The product mixture was washed with brine, dried with sodium sulfate, and

concentrated under vacuum. The product residue was purified by column chromatography with 100:1 dichloromethane:diethyl ether. The ^{1}H NMR spectrum for Ox2 was as reported: (400 MHz, $CDCl_3$) δ 7.50–7.39 (m, 5H), 6.18 (s, 1H), 5.01 (s, 1H), 3.33 (m, 2H), 1.20 (t, J = 7.3 Hz, 3H) (**Figure 2.23**).

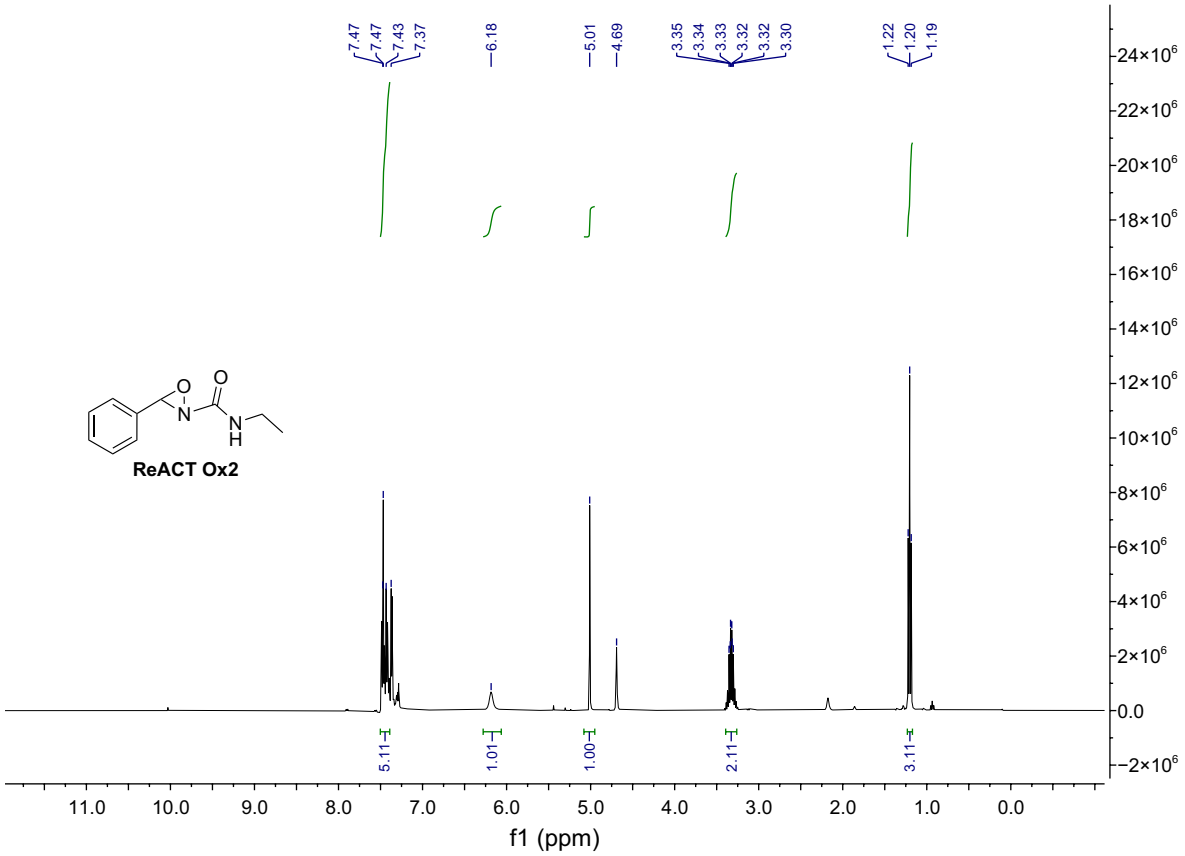


Figure 2.23. [1]H NMR spectrum for ReACT probe Ox2. Spectrum is as reported by Toste, Chang, and coworkers.[156]

2.7.5 Labeling RNase 1 with ReACT probe Ox2

Authentic RNase 1 was diluted to 1 mg mL^{-1} protein (≈69 mM) in 20 mM Tris-HCl buffer, pH 7.2, containing NaCl (50 mM). A stock solution of ReACT probe Ox2 was prepared in dimethylformamide at 350 mM; 50 equivalents Ox2 were then added to the RNase 1 sample ([Ox2] ≈3.4 mM). The reaction mixture was incubated at room temperature for 30 min with gentle mixing on an orbital shaker. To remove excess Ox2 probe and terminate labeling, the reaction mixture was passed through a Zeba spin desalting column (Thermo Scientific) that had been pre-equilibrated with Dulbecco's phosphate-buffered saline (Gibco); the eluate was passed through a second pre-equilibrated desalting column to ensure probe removal. Ox2-labeled RNase 1 was concentrated using a Vivaspin 500 centrifugal concentrator (Sartorius); protein concentration was quantified by UV absorption at A_{215} (ε_{215} = 227,448 M^{-1} cm^{-1}), and labeling was confirmed by mass spectrometry. Ox2-labeled RNase 1 was held at 4 °C and used within 24 h with verification of label retention.

2.7.6 Tryptic digestion of Ox2-labeled RNase 1

Tryptic peptides of Ox2-labeled RNase 1 were prepared using the Rapid-Digestion Trypsin kit from Promega. Digestion was conducted following the manufacturer's protocol with the following specifications: (1) 2 mM TCEP (Thermo Scientific) was used for sample reduction; (2) 5 mM iodoacetamide (Sigma) was used for sample alkylation; (3) RNase 1 was digested with Rapid Trypsin at a 10:1 protein:protease ratio with incubation at 70 °C for 45 min; (4) digestion was terminated by addition of trifluoroacetic acid (Sigma) to 0.5% v/v. Digests were desalted using C18 pipette tips (Thermo Scientific), flash-frozen in liquid nitrogen, and lyophilized overnight.

2.7.7 Determining protein thermostability by differential scanning fluorimetry

Differential scanning fluorimetry (DSF) was used to determine the thermostability of Ox2-labeled RNase 1 in relation to unlabeled RNase 1. SYPRO Orange protein gel stain (Supelco) was added to Ox2-labeled RNase 1 and unlabeled RNase 1 to achieve 1 mg mL^{-1} protein and 50× SYPRO Orange (vendor stock: 5000× in DMSO) in Dulbecco's phosphate-buffered saline (Gibco). The resulting samples were heated from 15–95 °C at 1 °C/min using a QuantStudio 7 Flex Real-Time PCR system (Applied Biosystems); fluorescence was monitored in real-time with λ_{ex} = 470 ± 15 nm and λ_{em} = 586 ± 10 nm. Data were processed with Protein Thermal Shift software (Applied Biosystems) using the Boltzmann fitting method. Values of T_m represent the temperature at which fluorescence reached 50% of its maximum.

2.7.8 Preparing Nle–RNase 1

The synthesis of Met29Nle Met35Nle RNase 1 (*Nle–RNase 1*) was performed using the AFPS system with H-rink amide ChemMatrix LL resin (142 mg, 0.17 mmol g^{-1}).[161] After synthesis, the resin was washed with dichloromethane and dried under a stream of nitrogen gas. The polypeptide was cleaved using a freshly prepared reducing cleavage mixture (trifluoroacetic acid:thioanisole:triisopropylsilane:trimethylsilyl chloride:dimethyl sulfide:ethanedithiol 77.5:5:5:5:5:2.5 % v/v, with 0.1% w/v triphenylphosphine, 0.8 mL μmol^{-1}) with incubation at room temperature for 3.5 h.[179] The cleavage mixture was drawn up with the reaction syringe containing the peptidyl resin, and the syringe was sealed with a Luer-tip syringe tip cap. The assembly was placed on a nutating mixer and incubated for 3.5 h; meanwhile, a 50-mL aliquot of diethyl ether was chilled on dry ice for >1 h. The cleavage supernatant in the reaction syringe was expelled into a 50-mL conical tube, and the resin was retained by the internal frit of the

reaction syringe. An additional 2.5 mL trifluoroacetic acid was drawn into the reaction syringe, and the syringe was gently mixed; the aliquot was expelled into the conical tube containing the cleavage supernatant. The pre-chilled diethyl ether was added to the cleavage supernatant, resulting in the formation of a white precipitate. The conical tube was centrifuged at 4000*g* for 10 min at 4 °C. The supernatant was removed, and the solid pellet was briefly dried under a stream of nitrogen gas. The dried peptide was dissolved in 5 mL of a 50:50 water:acetonitrile mixture with 0.1% v/v trifluoroacetic acid with vortexing and sonication as needed. The mixture was flash-frozen in liquid nitrogen, and the frozen material was lyophilized.

The resulting crude material was dissolved in 100 mM Tris-HCl buffer, pH 7.0, containing guanidinium-HCl (6 M). The crude material was purified by preparative reversed-phase HPLC (Agilent Zorbax 300SB-C3, 9.4 × 250 mm, 5.0–μm particle, 300-Å pore) thermostatted to 50 °C. The column was washed with 5% v/v acetonitrile in water (containing 0.1% v/v trifluoroacetic acid) for 10 min, followed by a 5–20% v/v gradient for 15 min, and finally a 20–35% v/v gradient for 60 min. Fractions were collected in 20-s increments. HPLC fractions confirmed by mass spectrometry to be of high purity were combined and lyophilized to afford pure Nle–RNase 1 polypeptide.

Pure Nle–RNase 1 polypeptide was dissolved at 2 mg mL^{-1} in 20 mM acetic acid containing EDTA (10 mM); the mixture was gently agitated at 37 °C for 30 min to ensure full dissolution of polypeptide. The polypeptide solution was analyzed by mass spectrometry to confirm a molecular weight consistent with eight reduced cysteine residues (folded RNase 1 has four disulfide bonds). The polypeptide solution was dropwise added to degassed folding buffer (100 mM Tris-HCl, pH 7.8, containing EDTA (10 mM), L-arginine (0.5 M), reduced glutathione (1.0 mM), and oxidized glutathione (0.2 mM)) while gently stirring at 4 °C, affording a final

polypeptide concentration of approximately 50 mg mL^{-1}. After complete addition of the polypeptide solution to the folding buffer, the resultant solution was allowed to incubate without agitation at 4 °C for 5 days.

Following protein folding, the protein solution was buffer-exchanged into 50 mM sodium phosphate buffer, pH 2.7, and was concentrated to approximately 3 mg mL^{-1} using an Amicon Ultra-15 centrifugal concentrator (EMD Millipore). The buffer-exchanged protein solution was passed through a Spin-X centrifugal tube filter (0.22-µm, cellulose acetate membrane; R&D Systems). The folded protein was then purified by preparative reversed-phase HPLC using a C4 column (Vydac, 10 × 250 mm, 5.0–µm particle, 300-Å pore) that was equilibrated with the same sodium phosphate buffer at room temperature. The column was washed with 12% v/v acetonitrile in 50 mM sodium phosphate buffer, pH 2.7, for 5 min followed by a 12–36% v/v gradient (again with 50 mM sodium phosphate buffer, pH 2.7) for 60 min. Fractions were manually collected based on detection at A_{218} and A_{280}. HPLC fractions confirmed by mass spectrometry to contain folded Nle–RNase 1 were placed under a gentle stream of nitrogen overnight to reduce sample volume and acetonitrile content. The protein solution was adjusted to pH ~5 using dibasic sodium phosphate, loaded into regenerated cellulose dialysis tubing (Spectrum Laboratories), and dialyzed 500-fold against 20 mM Tris-HCl buffer, pH 7.2, containing NaCl (50 mM). The protein was concentrated, quantified by UV absorption at A_{215} (ε_{215} = 227,448 M^{-1} cm^{-1}), flash-frozen in liquid nitrogen, and stored at −70 °C.

2.7.9 Mass spectrometry of intact proteins & tryptic peptides

High-resolution mass spectrometry of authentic RNase 1, Ox2-labeled RNase 1, and Nle–RNase 1 in solution was performed using ESI mass spectrometry on a 6530C Accurate-Mass Q–

TOF mass spectrometer equipped with a PLRP-S column (Agilent Technologies, 2.1 × 50 mm, 5-μm particle, 1000-Å pore). A gradient of 5–95% v/v acetonitrile (0.1% v/v formic acid) in water (0.1% v/v formic acid) over 7 min was used. Before Q–TOF LC-MS analysis, all proteins were diluted to 2 mM in 0.1% v/v formic acid and passed through a Spin-X centrifugal tube filter (0.22-μm, cellulose acetate membrane; R&D Systems).

For high-resolution mass spectrometry of tryptic peptides, lyophilized peptides were redissolved in ultrapure water with 0.1% v/v formic acid to approximately 5 μM (based on the starting protein concentration). The peptide solution was passed through a Spin-X centrifugal tube filter as for intact protein mass spectrometry. Samples were analyzed using ESI mass spectrometry on a 6530C Accurate-Mass Q–TOF mass spectrometer equipped with a C18 column (Agilent Technologies, 3 × 150 mm, 2.7-μm particle, 120-Å pore). A gradient of 5–40% v/v acetonitrile (0.1% v/v formic acid) in water (0.1% v/v formic acid) over 15 min was used.

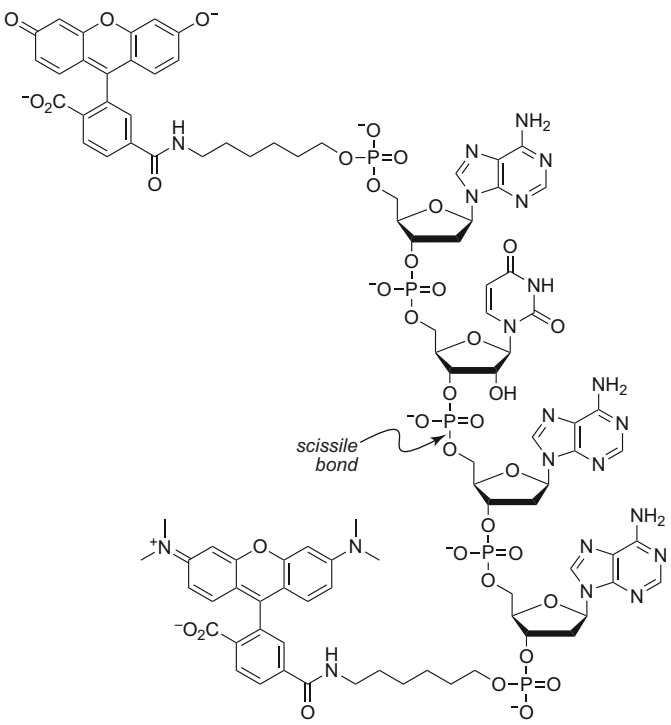
2.7.10 Challenging RNase 1 in an oxidation-induced inactivation assay

For authentic RNase 1 and Ox2-labeled RNase 1, 1 mg mL^{-1} protein was mixed with 1–3160 equivalents hydrogen peroxide (with respect to protein) in Dulbecco's phosphate-buffered saline; iron(II) sulfate was included in all oxidation samples at a ratio of 100:1 hydrogen peroxide to iron(II) sulfate. Samples were incubated at 37 °C for 3 h to oxidize the protein. Bovine liver catalase (Sigma) was then added to 50 μg mL^{-1}, and samples were incubated at 37 °C for a further 10 min to degrade any remaining hydrogen peroxide and halt oxidation.

For Nle–RNase 1, there was not an adequate amount of folded, purified protein to oxidize samples at 1 mg mL^{-1} Nle–RNase 1. Due to the limited yield, protein samples were instead oxidized at 0.333 mg mL^{-1} with 1–3160 equivalents hydrogen peroxide (with respect to protein)

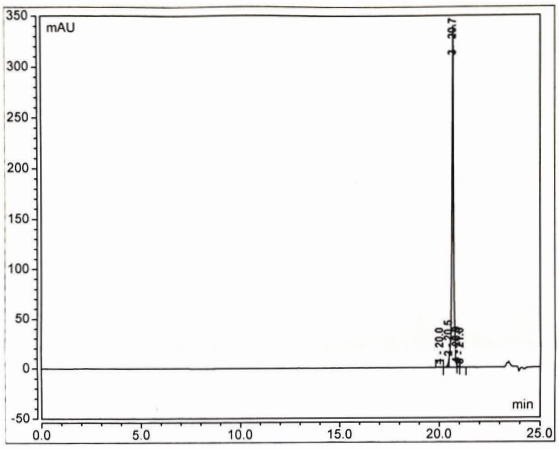
in Dulbecco's phosphate-buffered saline; the molar concentrations of hydrogen peroxide were therefore also threefold lower than their corresponding concentrations for authentic RNase 1 and Ox2-labeled RNase 1. Iron(II) sulfate was included in all oxidation samples at a ratio of 100:1 hydrogen peroxide to iron(II) sulfate. Nle–RNase 1 samples were otherwise oxidized under the same conditions as for authentic RNase 1 and Ox2-labeled RNase 1.

To assay authentic RNase 1 for ribonucleolytic activity, oxidized protein samples were diluted to 100 pM with RNase-free assay buffer (100 mM Tris-HCl buffer, pH 7.5, containing NaCl (100 mM)). To render the assay buffer RNase-free, 100 mM NaCl was DEPC-treated before the addition of Tris (from RNase-free stocks), due to the reactivity of its primary amino group with DEPC. Equal volumes of oxidized protein and FRET tetranucleotide substrate 6-FAM–dArU$(dA)_2$–6-TAMRA (Microsynth AG, Switzerland; **Figure 2.24**) were mixed to afford 50 pM oxidized protein, 200 nM substrate in assay buffer.



Oligo-ID 04778720 **Microsynth**

Order-ID:	821713	*Injection Volume:*	**5 µl**
Chromatography:	**Reverse Phase HPLC**	*Wavelength:*	**260 nm**
Buffer system:	**TEAA; 50°C**	*Recording Time:*	**03.04.2023 15:37**

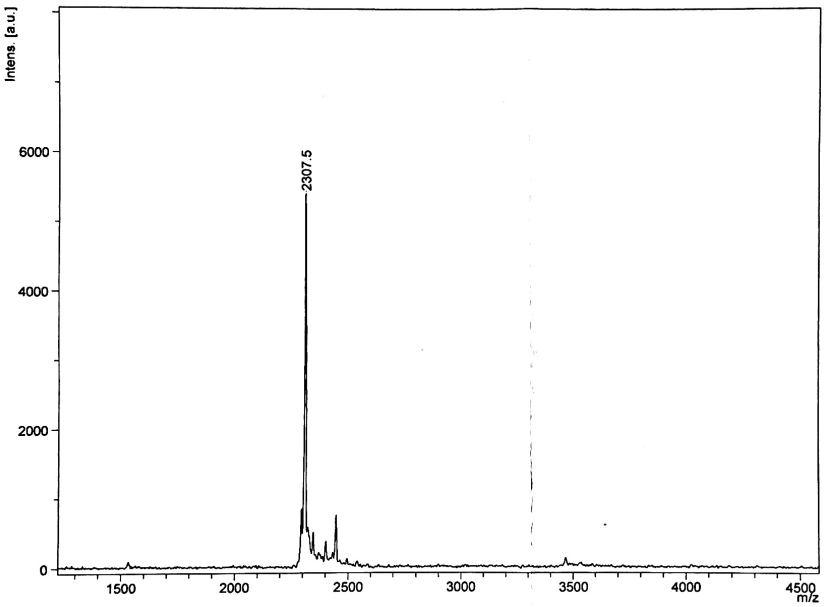


No.	Ret.Time min	Peak Number	Peak Height mAU	Peak Area mAU*min	Relative Area %
1	20.03	*1*	0.9	0.1	0.28
2	20.47	*2*	8.3	0.4	1.17
3	20.68	*3*	332.7	31.9	97.90
4	20.87	*4*	1.8	0.1	0.41
5	21.04	*5*	0.6	0.1	0.24
Total:			344.3	32.6	100.0

Oligo ID # 4778720

Theoretical Mass [g/mol]: 2312.66

Acceptance Criteria: Mw ± 0.5%



Microsynth

Figure 2.24. Structure of 6-FAM–dArU(dA)$_2$–6-TAMRA, which is the fluorogenic substrate for assays of ribonucleolytic activity and has only one scissile phosphodiester bond. The cleavage of that bond increases the fluorescence of the FAM moiety with λ_{ex} = 493 ± 5 nm and λ_{em} = 515 ± 5 nm. Insets: HPLC chromatogram (top) and MALDI–TOF mass spectrum (bottom) of synthetic 6-FAM–dArU(dA)$_2$–6-TAMRA.

Using a Tecan Spark multimode plate reader thermostatted to 25 °C, substrate fluorescence was measured at λ_{ex} = 493 ± 5 nm and λ_{em} = 515 ± 5 nm before addition of oxidized protein to determine the background fluorescence I_0. After adding protein, the reaction was incubated at 25 °C and continuously monitored for 4 minutes. Bovine RNase A was then added to ~10 μM to cleave all of the substrate and determine I_{max}. The value of k_{cat}/K_M was calculated with the equation:

$$\frac{k_{\mathrm{cat}}}{K_{\mathrm{M}}} = \frac{\Delta I/\Delta t}{(I_{\mathrm{max}} - I_0)[\mathrm{E}]} \tag{2.1}$$

where $\Delta I/\Delta t$ is the slope of the fluorescence vs. time data collected during the 4-minute observation period and [E] is the concentration of enzyme.

To assay Ox2-labeled RNase 1 for ribonucleolytic activity, oxidized protein samples were assayed under the same conditions as those used for authentic RNase 1, except oxidized samples were diluted to 400 pM and mixed with substrate to afford 200 pM oxidized protein, 200 nM substrate in assay buffer. This higher protein concentration was used to ensure an adequate rate of product generation over the 4-min monitoring period, as Ox2 labeling lowers the catalytic efficiency of RNase 1.

To assay Nle–RNase 1 for ribonucleolytic activity, oxidized protein samples were assayed under the same conditions as those used for authentic RNase 1 and Ox2-labeled RNase 1, except oxidized samples were diluted to 250 pM and mixed with substrate to afford 125 pM protein, 200 nM substrate in assay buffer. This higher protein concentration was used to ensure an adequate rate of product generation over the 4-min monitoring period, as norleucine incorporation lowers the catalytic efficiency of RNase 1.

For each of the three proteins, the catalytic efficiency of each oxidized sample was normalized to the catalytic efficiency of an equivalent sample of the same protein which had not

undergone oxidation. Plots of normalized reaction rates vs. equivalents hydrogen peroxide were fitted to a four-parameter logistic function using GraphPad Prism v10.0.3. Because the concentrations of protein and hydrogen peroxide were each threefold lower for Nle–RNase 1 than for the other two proteins, the rate of the bimolecular reaction of Nle–RNase 1 with hydrogen peroxide was ninefold slower than the corresponding reaction for the other two proteins; to account for this, equivalents of hydrogen peroxide were normalized by the rate of the oxidation reaction. IC_{50} values represent the normalized number of equivalents of hydrogen peroxide that are necessary to obviate ribonucleolytic activity by 50%.

Chapter 3:
Bayesian Inference Elucidates the Catalytic Competency of the SARS-CoV-2 Main Protease 3CL^{pro}

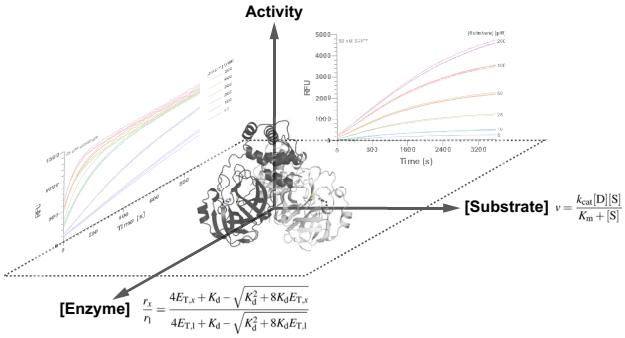
Reprinted with permission from ref. 180.

Wralstad, E. C.; Sayers, J.; Raines, R. T. Bayesian inference elucidates the catalytic competency of the SARS-CoV-2 main protease 3CL^{pro}. *Anal. Chem.* **2023**, *95*, 14981-14989.

Contributions: Conceptualization, E.C.W.; experimental methodology, E.C.W. and J.S.; experimental investigation, E.C.W.; writing—original draft preparation, E.C.W. and J.S.; writing—review and editing, R.T.R.; supervision, R.T.R.; funding acquisition, E.C.W. and R.T.R.

3.1 Abstract

The main protease of SARS-CoV-2, 3CL^{pro}, is a dimeric enzyme that is indispensable to viral replication and presents an attractive opportunity for therapeutic intervention. Previous reports regarding the key properties of 3CL^{pro} and its highly similar SARS-CoV homologue conflict dramatically. Values of the dimeric K_d and enzymic k_{cat}/K_M differ by 10^6- and 10^3-fold, respectively. Establishing a confident benchmark of the intrinsic capabilities of this enzyme is essential for combating the current pandemic as well as potential future outbreaks. Here, we use enzymatic methods to characterize the dimerization and catalytic efficiency of the authentic protease from SARS-CoV-2. Specifically, we use the rigor of Bayesian inference in a Markov Chain Monte Carlo analysis of progress curves to circumvent the limitations of traditional Michaelis−Menten initial rate analysis. We report that SARS-CoV-2 3CL^{pro} forms a dimer at pH 7.5 that has K_d = 16 ± 4 nM and is capable of catalysis with k_{cat} = 9.9 ± 1.5 s^{-1}, K_M = 0.23 ± 0.01 mM, and k_{cat}/K_M = (4.3 ± 0.7) × 10^4 M^{-1} s^{-1}. We also find that enzymatic activity decreases substantially in solutions of high ionic strength, largely as a consequence of impaired dimerization. We conclude that 3CL^{pro} is a more capable catalyst than appreciated previously, which has important implications for the design of antiviral therapeutic agents that target 3CL^{pro}.



3.2 Introduction

The COVID-19 pandemic, caused by the novel severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2), first emerged in Wuhan, China, in December 2019.[122,123] In addition to causing 770 million infections and 7.0 million deaths (as of August 30, 2023; reported by the World Health Organization), this public health crisis has disrupted the global economy and significantly impaired the physical and mental health of billions of people worldwide. Significant progress in mitigating the perpetuation of the pandemic has been achieved through the development of vaccines, neutralizing antibody treatments, and other antiviral therapeutics to stunt viral transmission and improve the prognosis of infected individuals.[127,130,181-183]

The trimeric spike glycoprotein (S protein) present on the surface of the coronavirus virion, which is responsible for docking to the ACE2 receptor of host cells and facilitating membrane fusion with the host cell as the causative mechanism of infection, has been an attractive target in the development of these therapeutics.[124] Although vaccines and therapeutics targeting the S protein have been developed, selective pressure exerted on the virus can result in S protein mutations that alter epitopes that are targeted by antibodies, dampening therapeutic efficacy.[184,185] This problem is exemplified by the emergence of SARS-CoV-2 variants with reduced inhibition by vaccine-stimulated neutralizing antibodies and/or therapeutic antibody treatment.[186,187] It is evident that complementary strategies are necessary to combat SARS-CoV-2.

An alternative target for therapeutic intervention is the chymotrypsin-like SARS-CoV-2 3C-like main protease $3CL^{pro}$.[127,128] As a result of the strict conservation between SARS-CoV and SARS-CoV-2 $3CL^{pro}$ (96% sequence identity), as well as the greater body of research available for the former protease, it is tempting to rely on enzymological values reported for SARS-CoV

3CL^{pro} when studying the SARS-CoV-2 homologue. Unfortunately, studies of each enzyme report key parameters that differ by many orders of magnitude. For example, reported values of the catalytic efficiency (k_{cat}/K_M) range from ~10^2 to nearly 10^5 M^{-1} s^{-1} (see **Table 3.1**), even within a series of similar substrates; likewise, reported values of the dimer dissociation constant, K_d, vary from low nanomolar to high micromolar (see **Table 3.2**). These studies are confounded by the use of enzymic constructs bearing artificial affinity tags (which could inhibit the association of 3CL^{pro} subunits), suboptimal substrates unable to occupy the enzymic binding pocket C-terminal to the scissile bond, assay conditions that impair 3CL^{pro} activity, and enzymatic activity assays that lack sensitivity. The design and assessment of 3CL^{pro} inhibitors rely on the robust characterization of SARS-CoV-2 3CL^{pro} catalysis.

Here, we circumvent these many experimental limitations and employ a novel Bayesian analytical technique to facilitate the interpretation of 3CL^{pro} enzymology from data obtained during the full-time course of the reaction, that is, a "progress curve". We report that SARS-CoV-2 3CL^{pro} is an undervalued catalyst. The dimer has a K_d value in the low-nanomolar regime. The authentic enzyme is highly active, with a measured catalytic efficiency that is larger than reported previously. We also find that SARS-CoV-2 3CL^{pro} activity is attenuated by increased ionic strength, likely due to the disruption of dimerization. We hypothesize that the use of artificial enzymic constructs and the inclusion of high NaCl concentrations in assay buffers led to the preponderance of low k_{cat}/K_M and high K_d values in the literature. Our findings provide a foundation for applied research, including the design of anti-SARS-CoV-2 therapeutics that disrupt the activity of a key viral enzyme.

Table 3.1. Survey of Literature SARS-CoV and SARS-CoV-2 3CL^{pro} Steady-State Kinetic Parameters.

Virus	Enzyme[a]	[Enzyme] (nM)	Substrate[b]	pH	[NaCl] (mM)	T (°C)[c]	k_{cat} (s^{-1})	K_M (μM)	k_{cat}/K_M ($M^{-1}s^{-1}$)	Reference
CoV-2	Native	25–500	R–E(EDANS)–ATLQ↓SGNA–K(DABCYL)–R	7.5	0	25	9.8	229	4.3×10^4	This Work
CoV-2	Native	200	MCA–AVLQ↓SGFR–K(DNP)–K	7.3	0	30	NR	NR	2.9×10^4	130
CoV-2	Native	100	(MCA)–AVLQ↓SGFR–K(DNP)–K–NH_2	7.3	0	27	0.54	19	2.8×10^4	188
CoV-2	Native	250	K(DABCYL)–TSAVLQ↓SGFRKM–E(EDANS)	8.0	100	25	0.52	41	1.3×10^4	189
CoV-2	+His (U)	200	K(DABCYL)–TSAVLQ↓SGFRKM–E(EDANS)-NH_2	8.0	150	NR	0.04	11	3.6×10^3	190
CoV-2	Native	2000	K(DABCYL)–TSAVLQ↓SGFRKM–E(EDANS)–NH_2	7.3	100	NR	NR	NR	3.4×10^3	129
CoV	Native	0–3000	$(ARLQ↓NH)_2$-rhodamine	8.0	0	RT/37	NR	NR	2.1×10^8	191
CoV	Native	50	K(DABCYL)–TSAVLQ↓SGFRKM–E(EDANS)	7.0	0	25	1.9	17	1.1×10^5	192
CoV	Native	25	K(DABCYL)–TSAVLQ↓SGFRKM–E(EDANS)	7.0	0	RT	1.5	45	3.4×10^4	193
CoV	Native	200	MCA–AVLQ↓SGFR–K(DNP)-K–NH_2	7.3	0	30	1.1	40	2.7×10^4	194
CoV	Native	200	SITSAVLQ↓SGFRKMA	7.5	0	25	8.5	600	1.4×10^4	193
CoV	Native	200	S(TAMRA)–ITSAVLQ↓SGFRKMA–K(DABCYL)	7.0	0	RT	NR	NR	1.4×10^4	193
CoV	Native	1000	K(DABCYL)–TSAVLQ↓SGFRKM–E(EDANS)	7.4	10	25	0.64	56	1.1×10^4	195
CoV	+GPH_6 (C)	200	MCA–AVLQ↓SGFR–K(DNP)-K–NH_2	7.3	0	30	0.41	61	6.8×10^3	194
CoV	+GS (N)	1000	MCA–AVLQ↓SGFR–K(DNP)-K–NH_2	7.3	0	30	0.14	129	1.1×10^3	196
CoV	Δ305–306	750	SWTSAVLQ↓SGFRKWA	7.0	0	25	NR	NR	1.0×10^3	132
CoV	+His (U)	1000	NMA– TSAVLQ↓SGFR–K(DNP)–M	8.0	150	37	0.0046	7	7.1×10^2	197
CoV	+His (C)	1070–17,100	TSAVLQ↓SGFRK–NH_2	7.3	0	RT	0.20	1150	1.8×10^2	198
CoV	+GPLGS (N)	2000	MCA–AVLQ↓SGFR–K(DNP)-K–NH_2	7.3	0	30	0.021	126	1.7×10^2	194
CoV	+His tag (C)	2800	TSAVLQ↓pNA	7.4	0	25	0.05	690	6.5×10^1	199

[a]Homologues are ordered by reported values of k_{cat}/K_M. His, polyhistidine tag; N and C, N- and C-terminal tags, respectively (U, tag orientation unknown). [b]Cleavage site indicated by ↓. [c]RT, room temperature (value not indicated by authors). NR, value not reported by authors.

Table 3.2. Survey of Literature SARS-CoV and SARS-CoV-2 3CL^{pro} Dimer Dissociation Constants.

Virus	Enzyme[a]	[Enzyme] (nM)	pH	[NaCl] (mM)	T (°C)[b]	Method[c]	K_d (nM)	Reference
CoV-2	Native	25–500	7.5	0	25	Activity	16	This Work
CoV-2	Native	230–18,100	7.3	150	20	SV-AUC	2500	129
CoV	Native	28,600	7.5	120	25	SV-AUC	0.35	200
CoV	+GKFKKIVKGT (C)	28,600	7.5	120	25	SV-AUC	5.6	200
CoV	Native	5–150	7.0	0	25	Activity	15	192
CoV	+QTSITSAVLQ (N)	28,600	7.5	120	25	SV-AUC	17.2	200
CoV	+His (C)	1440–21,600	8.0	NR	20	SV-AUC	190	201
CoV	Unknown	100–3000	7.3	NR	NR	Activity	810	133
CoV	+M(−1)	0–1000	7.5	0	25	Activity	1000	125
CoV	Native	7400–29,500	7.4	100	20	SE-AUC	1300	195
CoV	Native	≤3000	8.0	0	RT/37	Activity	5200	191
CoV	Native	2700–217,000	8.0	150	NR	SAXS	5800–6800[d]	191
CoV	Native	990–25,000	7.5	75	RT	SDS–PAGE	12,700	191
CoV	+His (C)	5700; 11,400	8.0	100	RT	GFC	100,000	198
CoV	+MRGSH$_6$GSTM (U)	4000–100,000	7.5	100	25	ITC	227,000	202

[a]Homologues are ordered by reported values of K_d. His, polyhistidine tag; N and C, N- and C-terminal tags, respectively (U, tag orientation unknown). [b]RT, room temperature (value not indicated by authors). NR, value not reported by authors. [c]SV-AUC, sedimentation velocity analytical ultracentrifugation; ITC, isothermal titration calorimetry; GFC, analytical gel-filtration chromatography; SAXS, small-angle X-ray scattering; Activity, enzymatic activity; SE-AUC, sedimentation equilibrium analytical ultracentrifugation. [d]The authors analyzed their SAXS data in three ways (fitting the entire scattering curves, Guinier analysis, and fitting the normalized forward scattered intensity); each method produced an estimated K_d value.

3.3 Results

3.3.1 Production, Purification, and Characterization of 3CL^{pro}

We used heterologous expression in *Escherichia coli* to produce a SARS-CoV-2 3CL^{pro} fusion construct composed of protease with an N-terminal GST tag and C-terminal 6×-His tag, as described by Hilgenfeld and coworkers;[129] semimature 3CL^{pro} with an autolyzed GST tag was apparent after 5 h of induction with IPTG (**Figure 3.1A**). Protein of >98% purity and with a molecular mass approximately corresponding to that of authentic 3CL^{pro} was obtained following affinity chromatography, proteolytic removal of the His tag, and anion-exchange chromatography (**Figure 3.1B**).

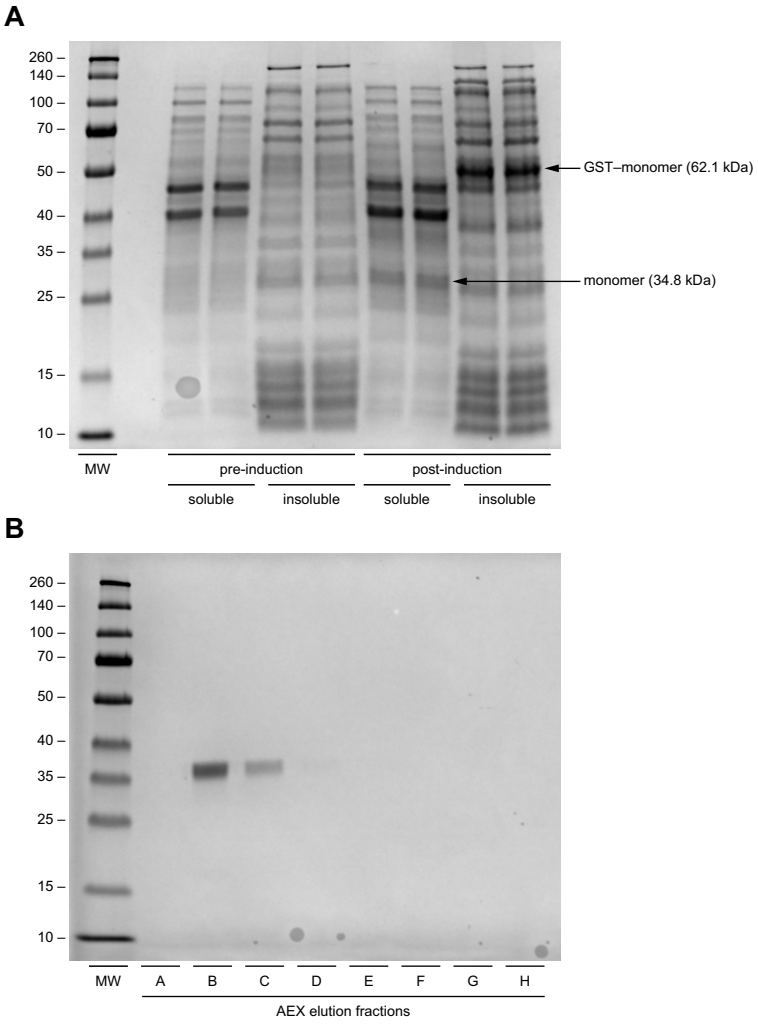


Figure 3.1. SDS–PAGE gels for the expression and purification of SARS-CoV-2 3CL^{pro}. (A) Expression gel. Expected masses account for a C terminal 6× His tag. The presence of a monomer with an autolyzed N-terminal GST tag is visible in the post-induction lysate soluble fraction. The appearance of a GST-tagged monomer in the insoluble fraction of the post-induction lysate serves as a secondary confirmation of 3CL^{pro} production. (B) Purification gel. Purity of 3CL^{pro} following the removal of the His tag with PreScission protease and anion-exchange FPLC (expected mass: 33.8 kDa for 3CL^{pro} after the removal of purification tags). Elution fractions are in the order of increasing [NaCl].

To confirm the purity and accuracy of our produced $3CL^{pro}$, we used Q−TOF mass spectrometry (**Figure 3.2A**, **B**). We observed only one protein species with an observed mass of 33,796.8 Da. This mass corresponds to that of authentic $3CL^{pro}$ (UniProt accession ID: P0DTD1, positions 3264−3569) bearing no purification tags (Δm = 5.6 ppm). We assayed the thermostability of our recombinant $3CL^{pro}$ using differential scanning fluorimetry (DSF). We observed a single thermal denaturation event at T_m = 55.2 ± 0.1 °C (mean ± SE, n = 4) (**Figure 3.2C**). Our T_m values generally agree with previous reports of 54.2 and 55.0 °C, which were reported for SARS-CoV-2 $3CL^{pro}$ with a C-terminal His tag.[126,190] $3CL^{pro}$ appears to be a stable protein at ambient and physiological temperatures, and our results suggest that a C-terminal His tag does not impair the inherent thermostability of $3CL^{pro}$.

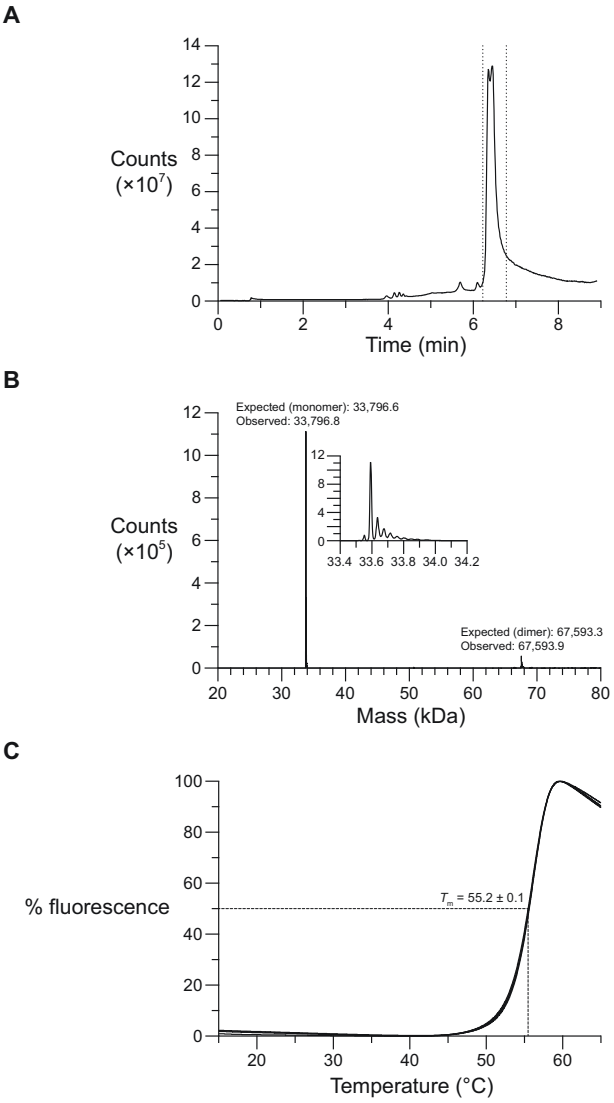


Figure 3.2. (A) Q–TOF MS total ion chromatogram of purified $3CL^{pro}$. (B) Deconvoluted mass spectrum of the bounded peak in panel A. (C) DSF curves of $3CL^{pro}$ (100 μg mL^{-1}) in 50 mM HEPES–NaOH buffer, pH 7.5, normalized to maximum fluorescence. DSF experiments were performed in quadruplicate.

3.3.2 Design, Synthesis, and Characterization of a 3CL^{pro} Substrate

We sought to choose an optimal consensus sequence for our 3CL^{pro} substrate. Following a literature review of known and putative SARS-CoV and SARS-CoV-2 3CL^{pro} peptide substrates, we chose the candidate sequence ATLQ↓SGNA, where the cleavage site is indicated by "↓". Residues preceding the scissile bond in our candidate substrate integrate those that 3CL^{pro} recognizes in its *in vivo* autocleavage from the pp1a polyprotein: AVLQ at the NSP4−3CL^{pro} interface, VTFQ at the 3CL^{pro}−NSP6 interface. Likewise, residues P1′ and P2′ in our candidate substrate recapitulate the N-terminal residues of 3CL^{pro}, again mimicking the NSP4−3CL^{pro} interface. Our candidate substrate accommodates many of the structural features observed in SARS-CoV and SARS-CoV-2 substrate-binding subsites: the strict requirement for a P1 Gln; the hydrophobic S2 and S4 subsites corresponding to a P2 Leu and P4 Ala; the shallow, solvent-exposed S3 subsite corresponding to a P3 Thr; access of the S1′ subsite occupant (corresponding to a P1′ Ser) to form a hydrogen bond with the Cys145−His41 catalytic dyad and thereby stabilize a reactive Cys145 thiolate; and the broad binding pocket of the S3′ subsite that accommodates a P3′ Asn.[188,189,203-206] To enable the detection of substrate cleavage, we appended EDANS-conjugated Glu and DABCYL-conjugated Lys to the N- and C-terminal ends of the peptide as a FRET pair; we also appended N- and C-terminal Arg to increase aqueous solubility. The structure of our substrate R− E(EDANS)−ATLQSGNA−K(DABCYL)−R is shown in **Figure 3.3A**.

The synthesis of the 3CL^{pro} substrate by solid-phase peptide synthesis was successful, as judged by analytical HPLC and MALDI−TOF mass spectrometry. A single major peak was observed by HPLC with coincident absorbance at the peptide bond, EDANS FRET donor, and

DABCYL FRET acceptor wavelengths (**Figure 3.3B**). The mass of the synthesized peptide was as expected (**Figure 3.3C**, **D**).

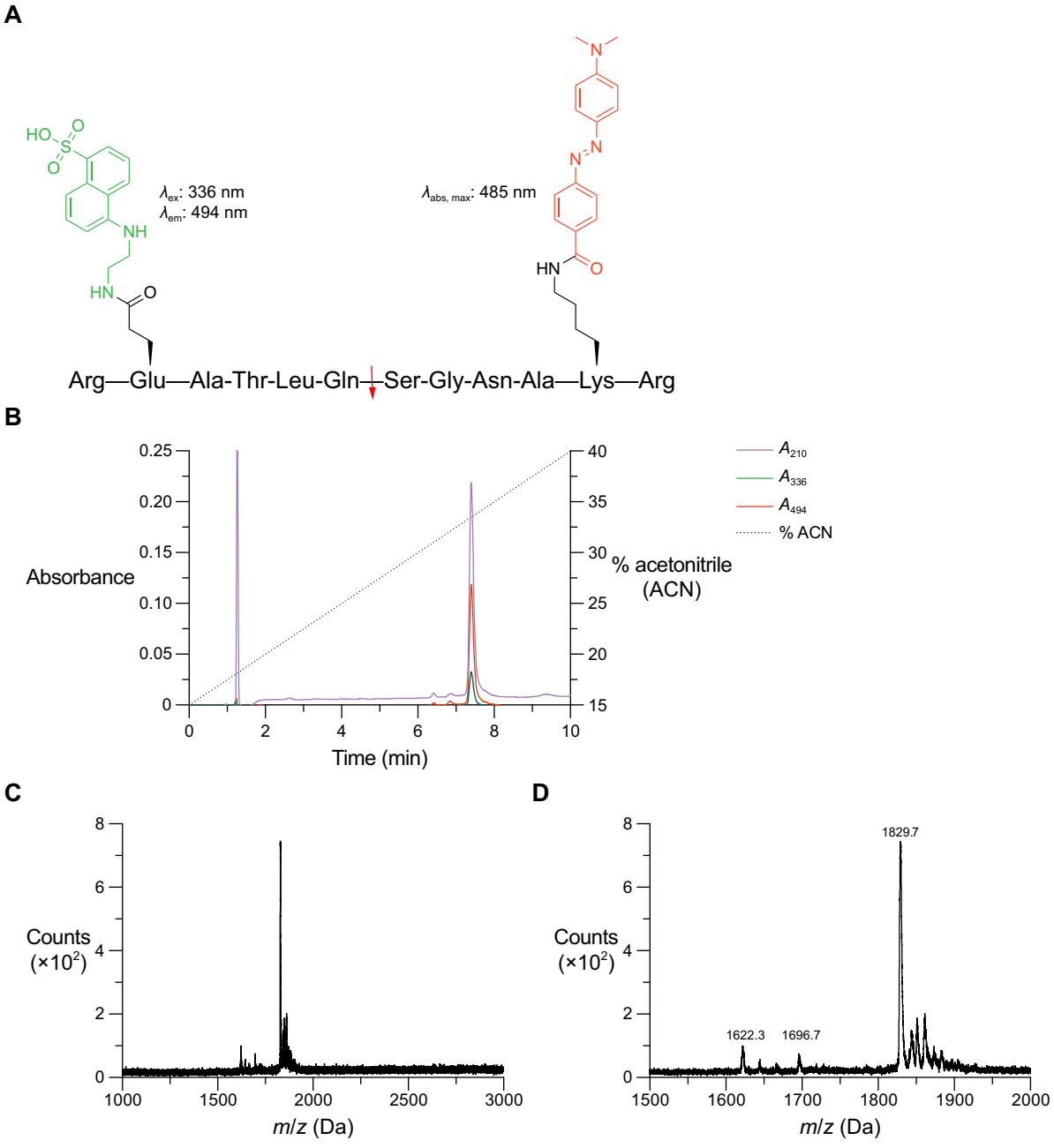

A
HO
O
S
O
NH
HN
O
λex: 336 nm
λem: 494 nm
λabs, max: 485 nm
N
N
N
HN
O
Arg—Glu—Ala-Thr-Leu-Gln-Ser-Gly-Asn-Ala—Lys—Arg
B
Absorbance
0.25
0.20
0.15
0.10
0.05
0
40
35
30
25
20
15
% acetonitrile
(ACN)
A210
A336
A494
% ACN
0
2
4
6
8
10
Time (min)
C
Counts
(×10²)
8
6
4
2
0
1000
1500
2000
2500
3000
m/z (Da)
D
Counts
(×10²)
1829.7
1622.3
1696.7
1500
1600
1700
1800
1900
2000
m/z (Da)

Figure 3.3. Characterization of the $3CL^{pro}$ substrate R–E(EDANS)–ATLQSGNA–K(DABCYL)–R. (A) Structure of the peptide (EDANS in green; DABCYL in red). Cleavage of the Gln/Ser peptide bond by $3CL^{pro}$ (red arrow) liberates the EDANS-bearing product peptide, increasing its fluorescence. (B) Analytical reversed-phase HPLC chromatogram of the synthesized peptide eluted from a C18 column with a 15–40% v/v acetonitrile gradient. Elution was monitored at λ = 210, 336, and 494 nm for the absorption of peptide bonds, the EDANS moiety, and the DABCYL moiety, respectively. Following the solvent front at t = 1 min, a single major peak with absorption at all three wavelengths is observed at $t \approx 7.5$ min. (C) Full MALDI–TOF mass spectrum, and (D) 1500–2000 Da inset for the peptide. Expected mass $[M + H]^+$, 1829.9 Da; observed mass, 1829.7 Da. Higher mass peaks correspond to salt adducts of the peptide; lower mass peaks at 1622.3 and 1696.7 Da correspond to artifactual, ionization-induced fragmentation at the EDANS and DABCYL moieties, respectively.

3.3.3 Dimerization of 3CL^{pro}

The 3CL^{pro} dimerization dissociation constant K_d was assessed from enzymological data. Only dimeric 3CL^{pro} is catalytically competent with one functional active site per dimer;[133] the maximal reaction velocity is a function of the 3CL^{pro} dimer concentration $[D]_x$:

$$V_{max} = k_{cat}[D]_x \tag{3.1}$$

with:

$$[D]_x = \frac{1}{8}\left(4E_{T,x} + K_d - \sqrt{K_d^2 + 8E_{T,x}K_d}\right) \tag{3.2}$$

where $E_{T,x}$ is the molar concentration of total 3CL^{pro} protomer. Equation 3.2 results from the definition of the dissociation constant and mass balance for $E_{T,x}$. Enzymological assay of 3CL^{pro} dimerization was necessary prior to the determination of the catalytic parameters k_{cat} and K_M in order to calculate the latter set of parameters with a known concentration of dimer.

We confirmed that our substrate is suitable for assays of 3CL^{pro} activity with observable turn-on of fluorescence in 50 mM HEPES−NaOH buffer, pH 7.5, containing DTT (1 mM) at 25 °C; with sufficient enzyme, all reactions approached completion within the observed timeframes (**Figure 3.4**).

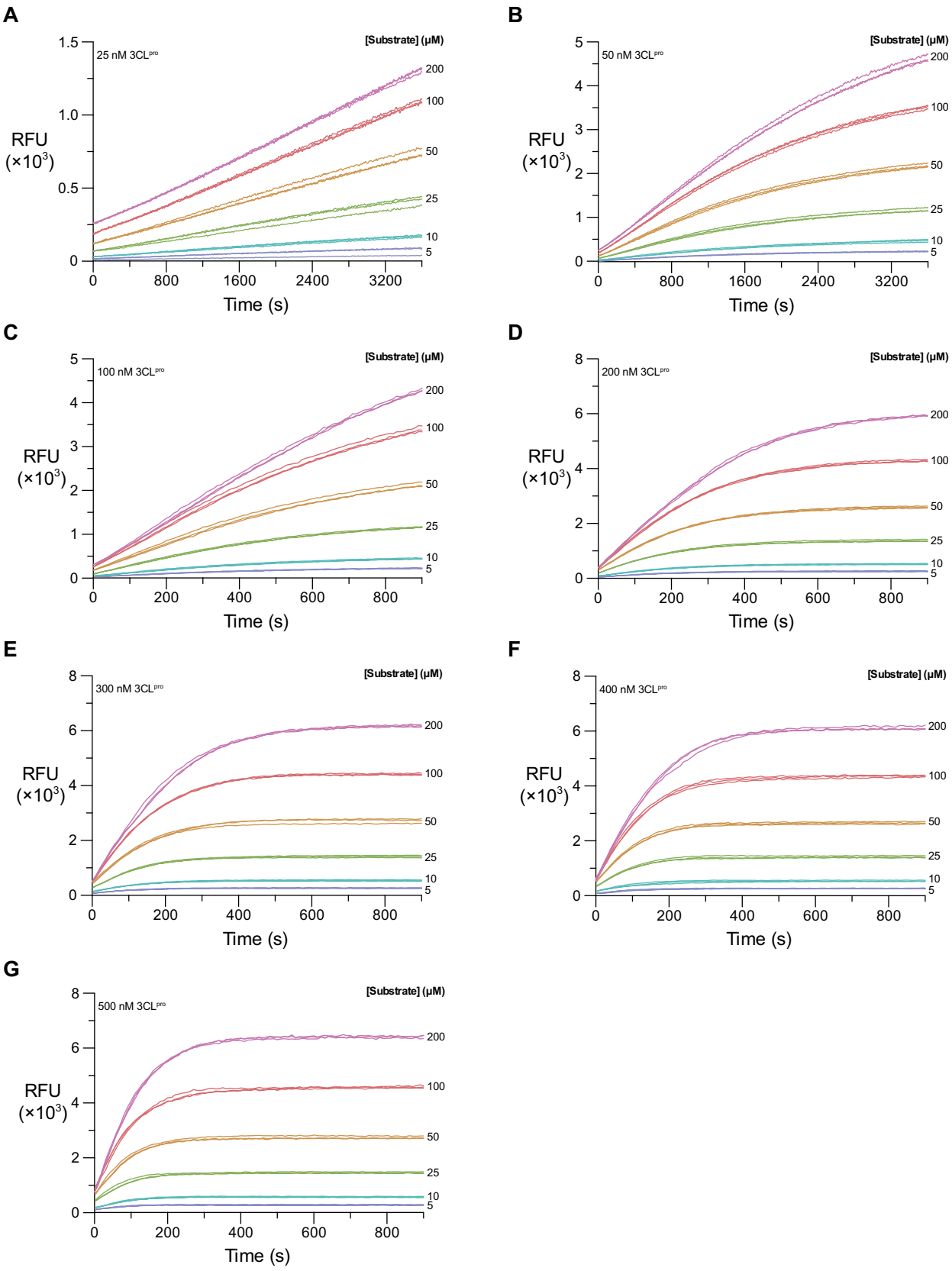

A
25 nM 3CLpro
[Substrate] (μM)
200
100
50
25
10
5
RFU
(×10^3)
Time (s)
B
50 nM 3CLpro
C
100 nM 3CLpro
D
200 nM 3CLpro
E
300 nM 3CLpro
F
400 nM 3CLpro
G
500 nM 3CLpro

Figure 3.4. Progress curves for substrate cleavage by $3CL^{pro}$ at (A) 25 nM, (B) 50 nM, (C) 100 nM, (D) 200 nM, (E) 300 nM, (F) 400 nM, and (G) 500 nM. Assays in panels A and B were monitored for 60 min; assays in panels C–G were monitored for 15 min.

The asymptotic end point fluorescence was directly proportional to substrate concentration at low concentrations, but this linear relationship broke down at higher peptide concentrations. This dichotomy is indicative of the inner filter effect, a well-documented phenomenon in FRET assays.[125,140,141] We characterized the extent of the inner filter effect in our assay by measuring the fluorescence of a fixed concentration of the R−E(EDANS)−ATLQ product (**Figure 3.5**) in the presence of varied substrate concentrations. We found that the fluorescence of the product significantly attenuates due to the inner filter effect when [substrate] > 50 μM (**Figure 3.6**).

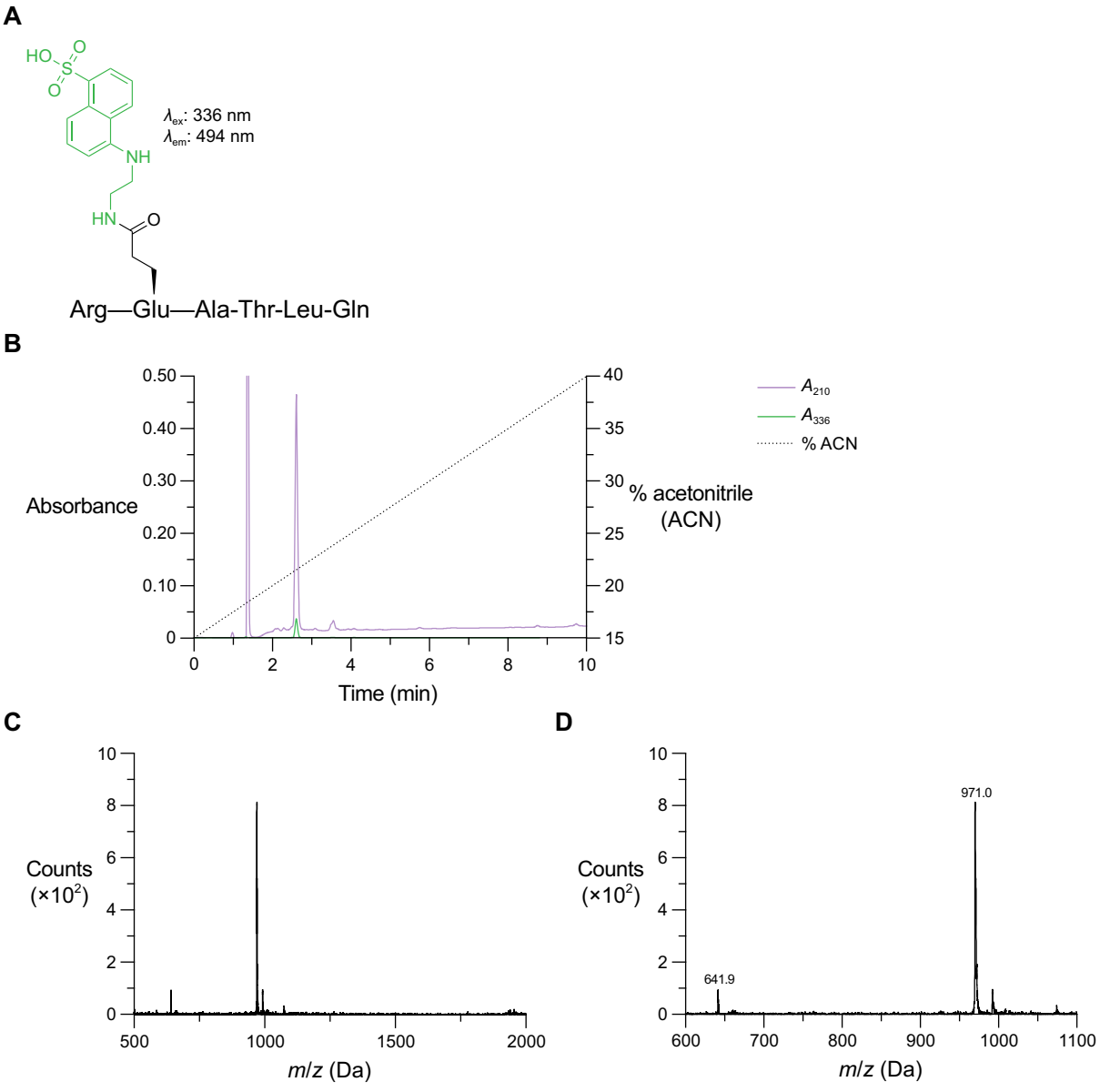

A
HO
O
S
O
O
NH
HN
O
λex: 336 nm
λem: 494 nm
Arg—Glu—Ala-Thr-Leu-Gln
B
Absorbance
0.50
0.40
0.30
0.20
0.10
0
40
35
30
25
20
15
% acetonitrile
(ACN)
0
2
4
6
8
10
Time (min)
A210
A336
% ACN
C
Counts
(×10²)
10
8
6
4
2
0
500
1000
1500
2000
m/z (Da)
D
Counts
(×10²)
971.0
641.9
600
700
800
900
1000
1100
m/z (Da)

Figure 3.5. Characterization of the $3CL^{pro}$ product R–E(EDANS)–ATLQ, used as a control peptide to characterize the inner filter effect (see **Figure 3.6**). (A) Structure of the peptide (EDANS in green). (B) Analytical reversed-phase HPLC chromatogram of the synthesized peptide eluted from a C18 column with a 15–40% v/v acetonitrile gradient. Elution was monitored at λ = 210 and 336 nm for the absorption of peptide bonds and the EDANS moiety, respectively. Following the solvent front at t = 1.4 min, a single major peak with absorption at both wavelengths is observed at $t \approx 2.6$ min. (C) Full MALDI–TOF mass spectrum, and (D) 600–1100 Da inset for the peptide. Expected mass $[M + H]^+$, 965.5 Da; observed mass, 971.0 Da. The higher mass peak corresponds to a salt adduct of the peptide; the lower mass peak at 641.9 Da corresponds to artifactual, ionization-induced fragmentation at the EDANS moiety.

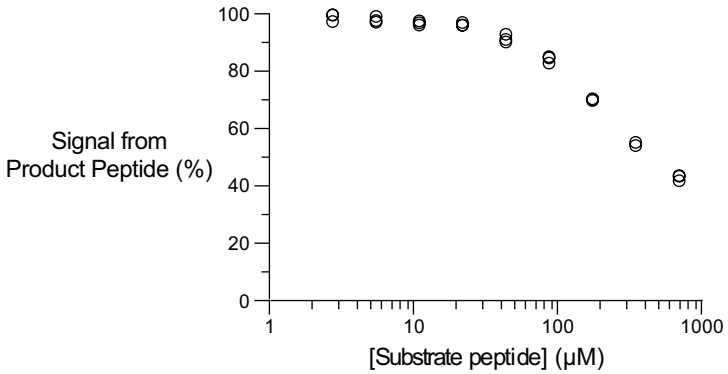


Figure 3.6. Graph showing the inner filter effect for the $3CL^{pro}$ substrate R–E(EDANS)–ATLQSGNA–K(DABCYL)–R. The EDANS-bearing product peptide R–E(EDANS)–ATLQ was synthesized and prepared at a final concentration of 10 μM in 50 mM HEPES–NaOH buffer, pH 7.5. The fluorescence intensity of the product peptide was measured in the presence of increasing concentrations of the substrate peptide; intensities were normalized to that of 10 μM product peptide in the absence of the substrate peptide. The attenuation of product fluorescence as a result of the inner filter effect becomes >10% as the total concentration of the substrate peptide, which contains the FRET-acceptor DABCYL, exceeds 50 μM.

We therefore transformed progress curve fluorescence intensity values to product concentration by equating the asymptotic end point fluorescence intensity of each progress curve to the initial substrate peptide concentration. The extent of the inner filter effect is constant over the course of a single enzymatic reaction, as the concentration of FRET donor and acceptor does not change regardless of whether they are parts of one substrate peptide or two separate product peptides.[140,141]

We assayed substrate cleavage with 25−500 nM $E_{T,x}$. For the Michaelis−Menten regime where [substrate] << K_M, the decrease in substrate concentration over time approximately follows an exponential decay with a pseudo-first-order rate constant r_x (for further details, see: Equations B.7−B.13 of Appendix B):

$$r_x = \left(\frac{k_{cat}}{K_M}\right)[D]_x \tag{3.3}$$

This rate constant is a parameter of an approximate expression for product fluorescence over time (see Equation B.12), enabling its estimation by nonlinear regression of the progress curves in **Figure 3.4**. For a given $E_{T,x}$, r_x should be constant with respect to substrate concentration; a decrease in the fitted r_x with an increase in substrate concentration indicates that the assumption [substrate] << K_M has become invalid and allows for a qualitative estimate of K_M. Plots of the fitted r_x values are shown in **Figure 3.7**.

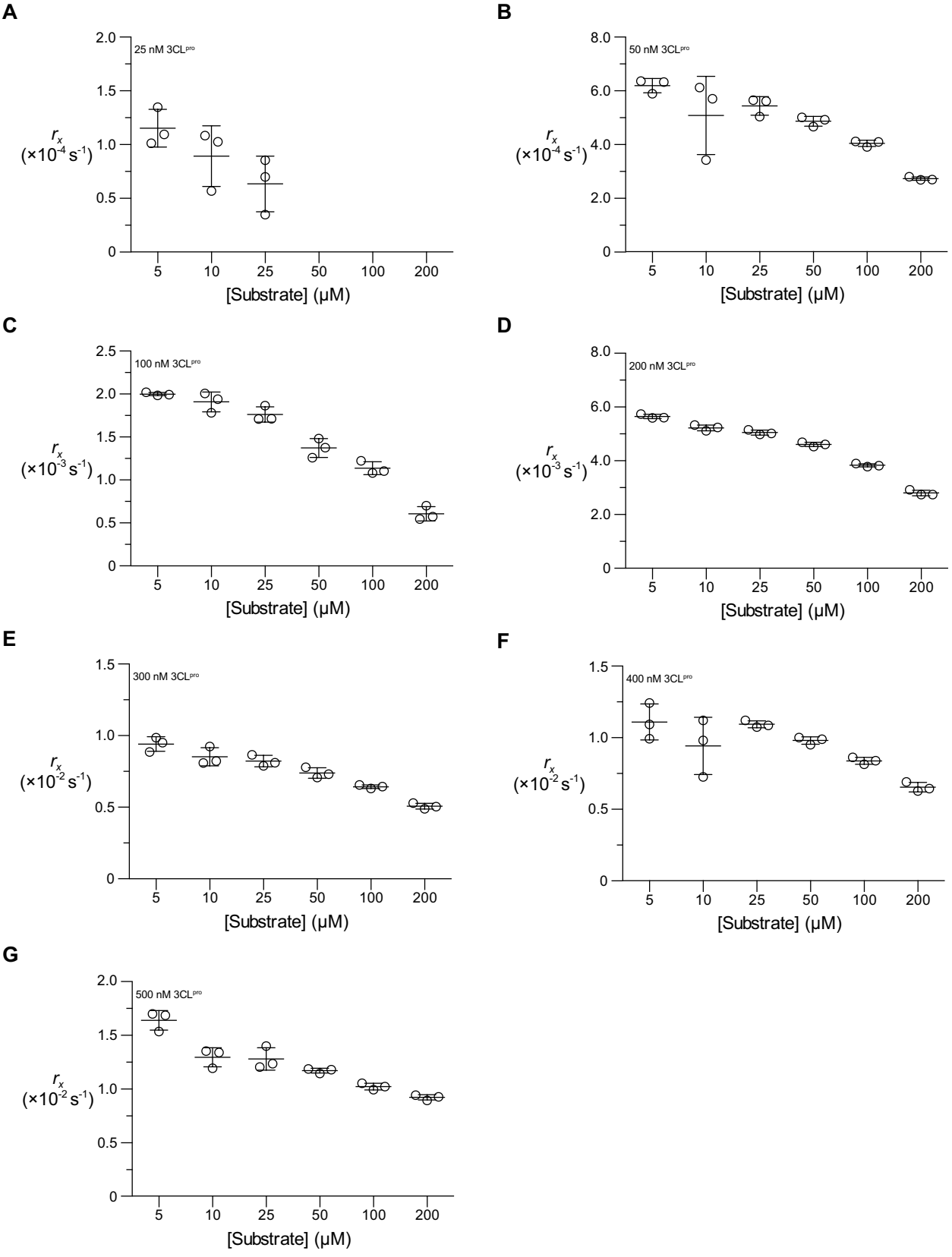

A
25 nM 3CLpro
r_x (×10^{-4} s^{-1})
2.0
1.5
1.0
0.5
0
5
10
25
50
100
200
[Substrate] (μM)
B
50 nM 3CLpro
r_x (×10^{-4} s^{-1})
8.0
6.0
4.0
2.0
0
5
10
25
50
100
200
[Substrate] (μM)
C
100 nM 3CLpro
r_x (×10^{-3} s^{-1})
2.5
2.0
1.5
1.0
0.5
0
5
10
25
50
100
200
[Substrate] (μM)
D
200 nM 3CLpro
r_x (×10^{-3} s^{-1})
8.0
6.0
4.0
2.0
0
5
10
25
50
100
200
[Substrate] (μM)
E
300 nM 3CLpro
r_x (×10^{-2} s^{-1})
1.5
1.0
0.5
0
5
10
25
50
100
200
[Substrate] (μM)
F
400 nM 3CLpro
r_x (×10^{-2} s^{-1})
1.5
1.0
0.5
0
5
10
25
50
100
200
[Substrate] (μM)
G
500 nM 3CLpro
r_x (×10^{-2} s^{-1})
2.0
1.5
1.0
0.5
0
5
10
25
50
100
200
[Substrate] (μM)

Figure 3.7. Fitted r_x values determined by nonlinear regression of the progress curves in **Figure 3.4** using the pseudo-first order approximation. Values are the mean ± SD for $3CL^{pro}$ at (A) 25 nM, (B) 50 nM, (C) 100 nM, (D) 200 nM, (E) 300 nM, (F) 400 nM, and (G) 500 nM. Nonlinear regression failed to converge for 25 nM $3CL^{pro}$ at higher substrate concentrations due to the lack of a discernable plateau fluorescence intensity in the corresponding progress curves (*e.g.*, **Figure 3.4A**), resulting in a lack of fitted r_x for substrate concentrations >25 μM.

For each substrate concentration, r_x values were normalized to the pseudo-first-order rate constant r_l corresponding to the largest $3CL^{pro}$ concentration, $E_{T,l}$. From Equations 3.2 and 3.3, this ratio is

$$\frac{r_x}{r_\mathrm{l}} = \frac{4E_{\mathrm{T},x} + K_\mathrm{d} - \sqrt{K_\mathrm{d}^2 + 8E_{\mathrm{T},x}K_\mathrm{d}}}{4E_{\mathrm{T,l}} + K_\mathrm{d} - \sqrt{K_\mathrm{d}^2 + 8E_{\mathrm{T,l}}K_\mathrm{d}}} \tag{3.4}$$

and is a function of $E_{T,x}$ with one fittable constant, K_d. Data for each substrate concentration were fitted to Equation 3.4 to derive the value of K_d (**Figure 3.8**). Within the substrate range 10−100 μM, where the pseudo-first-order approximation appears valid and raw fluorescence intensities were adequately sensitive (**Figure 3.4** and **3.7**), the fitted K_d values were highly consistent and resulted in K_d = 15.9 ± 3.8 nM, which is the mean ± SE (n = 4). Successful fitting to Equation 3.4 indicates that $3CL^{pro}$ is catalytically competent only in the dimeric state.

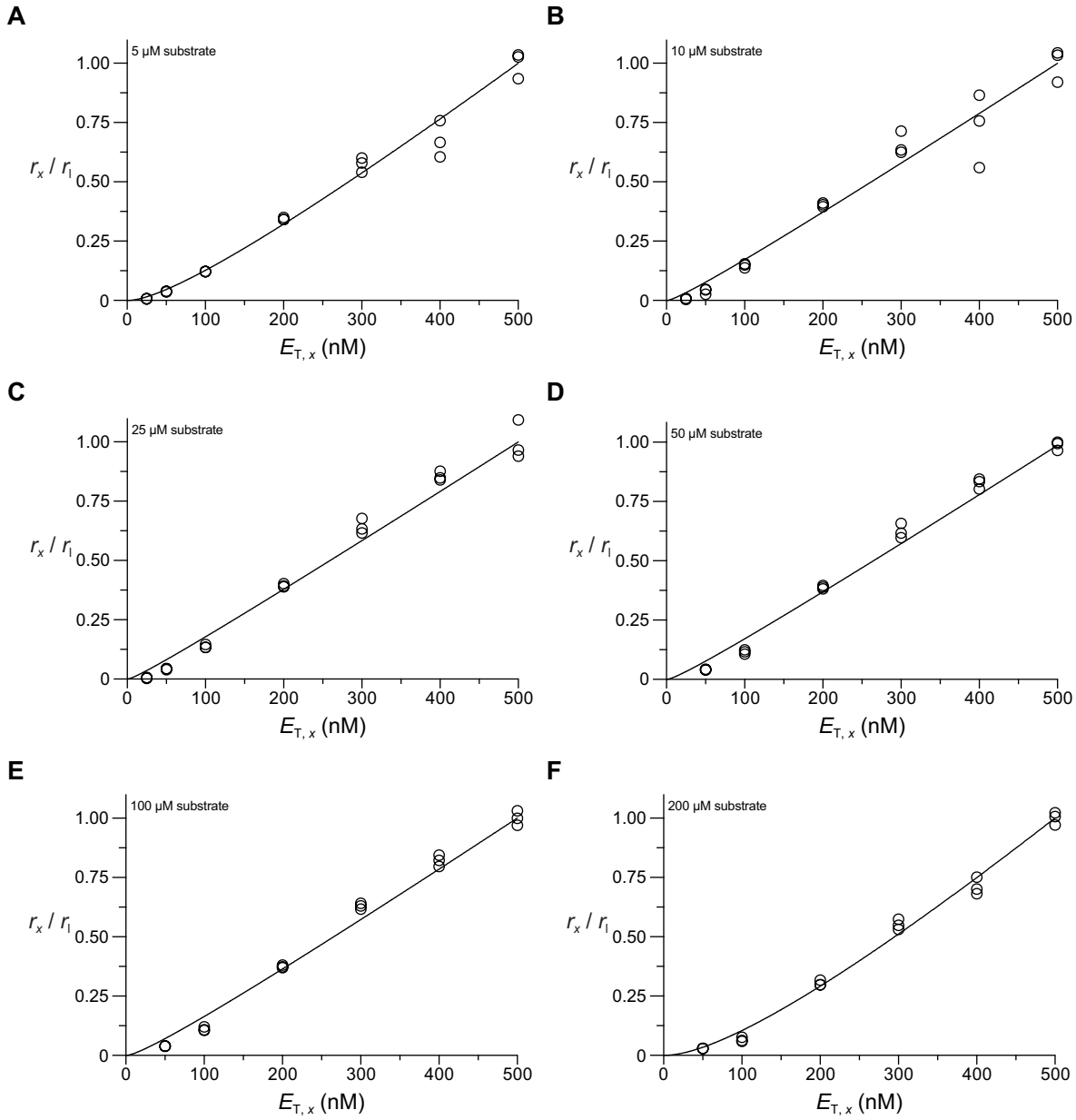

A
5 μM substrate
B
10 μM substrate
C
25 μM substrate
D
50 μM substrate
E
100 μM substrate
F
200 μM substrate
r_x / r_l
$E_{T,x}$ (nM)
0
0.25
0.50
0.75
1.00
100
200
300
400
500

Figure 3.8. Fitted r_x/r_1 plots to derive K_d from Equation 3.4. Values determined by fitting the data at each substrate concentrations are (A) K_d = 149.2 nM at 5 μM substrate, (B) K_d = 14.87 nM at 10 μM substrate, (C) K_d = 8.782 nM at 25 μM substrate, (D) K_d = 13.59 nM at 50 μM substrate, (E) K_d = 26.52 nM at 100 μM substrate, and (F) K_d = 338.4 nM at 200 μM substrate. Only fitted K_d values for 10–100 μM substrate were used to determine the mean value of K_d because 200 μM substrate appears to violate the pseudo-first order assumption that [substrate] << K_M (**Figure 3.7**), and the 5 μM substrate progress curves (**Figure 3.4**) have a relatively low change in fluorescence intensity over the course of the enzymatic reaction, which introduces ambiguity in nonlinear fits to Equation B.12.

3.3.4 Catalysis by 3CL^{pro}

We analyzed the 3CL^{pro} progress curves using both a traditional Michaelis−Menten (MM) strategy and a Bayesian Markov Chain Monte Carlo (MCMC) approach.[207] We note that traditional MM analysis requires a range of substrate concentrations of up to ~10×K_M. Although we do not know the value of K_M *a priori*, the pseudo-first-order approximation used to determine the dimerization K_d enabled us to estimate the concentration regime in which K_M exists. Based on our observation that the pseudo-first-order approximation begins breaking down as substrate concentrations exceed 100 μM, we predicted that 10×K_M is near millimolar concentration and cannot be feasibly assayed due to the magnitude of the inner filter effect at these concentrations along with limited substrate solubility at these concentrations (*cf*: **Figure 3.6**). The Bayesian MCMC method (*vide infra*) analyzes data from progress curves and thus does not require high substrate concentrations; we used MM analyses when possible to corroborate our MCMC results.

We determined Michaelis−Menten initial reaction rates from 3CL^{pro} progress curves in a semiautomated manner using ICEKAT in order to minimize the introduction of bias.[208] We restricted initial rate analyses to $E_{T,x} \leq 200$ nM, as the faster reactions at higher 3CL^{pro} concentrations resulted in insufficient data to accurately determine initial rates (*e.g.*, compare the linear region of the progress curves for high and low $E_{T,x}$ in **Figure 3.4**). MM plots for each $E_{T,x}$ are shown in **Figure 3.9** and best-fit kinetic parameters are listed in **Table 3.3**. Our MM initial rate analysis estimates that the catalytic efficiency of 3CL^{pro} with our FRET substrate is k_{cat}/K_M = (2.6 ± 0.6) × 10^4 M^{-1} s^{-1} (mean ± SE, n = 4).

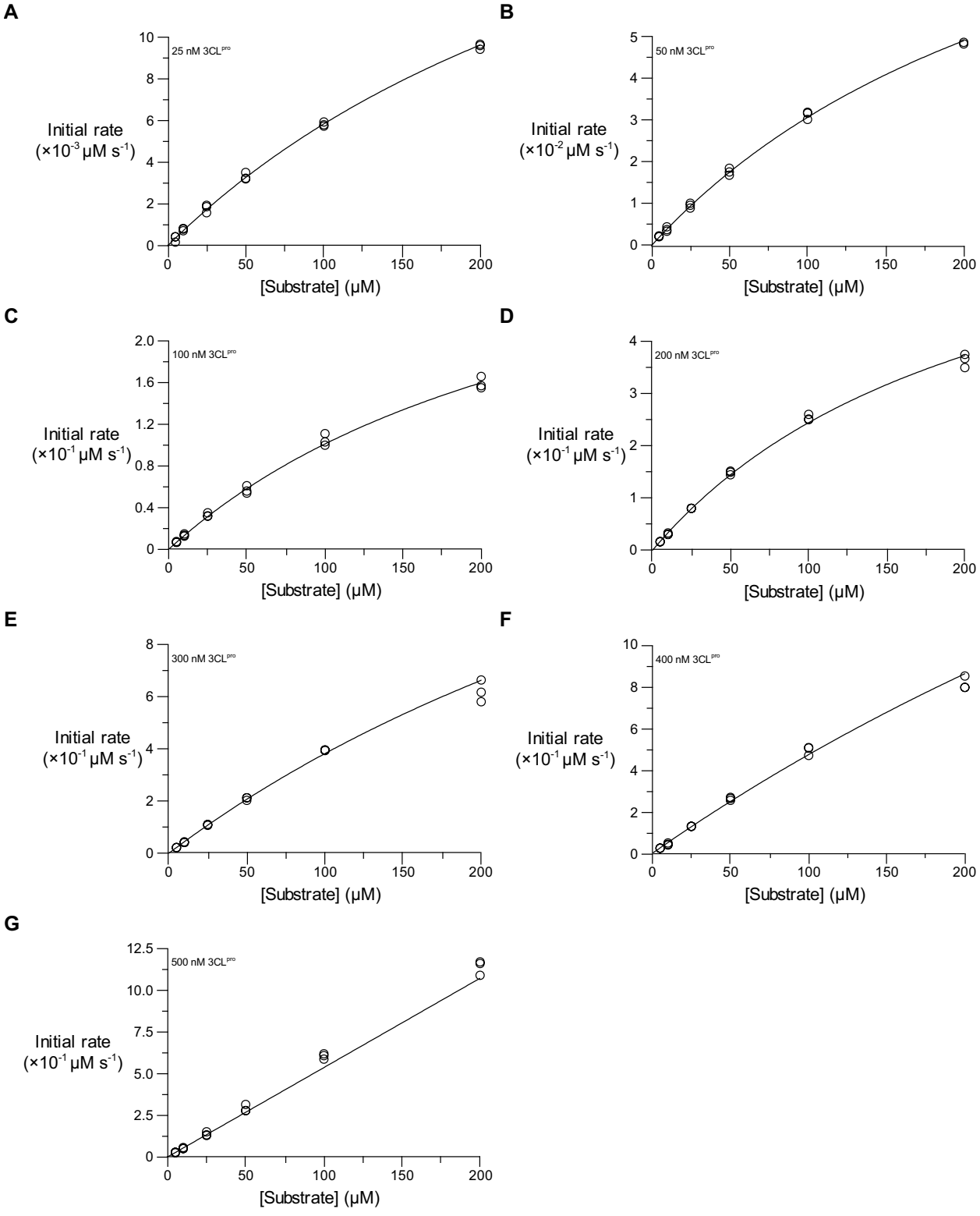

A
25 nM 3CL^pro
Initial rate
(×10^-3 μM s^-1)
[Substrate] (μM)
B
50 nM 3CL^pro
Initial rate
(×10^-2 μM s^-1)
[Substrate] (μM)
C
100 nM 3CL^pro
Initial rate
(×10^-1 μM s^-1)
[Substrate] (μM)
D
200 nM 3CL^pro
Initial rate
(×10^-1 μM s^-1)
[Substrate] (μM)
E
300 nM 3CL^pro
Initial rate
(×10^-1 μM s^-1)
[Substrate] (μM)
F
400 nM 3CL^pro
Initial rate
(×10^-1 μM s^-1)
[Substrate] (μM)
G
500 nM 3CL^pro
Initial rate
(×10^-1 μM s^-1)
[Substrate] (μM)

Figure 3.9. Michaelis–Menten curves for the cleavage of the substrate peptide by $3CL^{pro}$ at (A) 25 nM, (B) 50 nM, (C) 100 nM, (D) 200 nM, (E) 300 nM, (F) 400 nM, and (G) 500 nM. Initial rates were derived from the progress curves (**Figure 3.4**) using ICEKAT with the default setting "maximize slope magnitude."[208] Kinetic parameters are listed in **Table 3.3**.

Table 3.3. Steady-State Kinetic Parameters for Catalysis of the Cleavage of R–E(EDANS)–ATLQ↓SGNA–K(DABCYL)–R by $3CL^{pro}$.

		Michaelis–Menten analysis (ICEKAT)		Bayesian analysis (EKMCMC)	
[$3CL^{pro}$] (nM)	[Dimer] (nM)[a]	k_{cat} (s^{-1})[c]	K_M (μM)[c]	k_{cat} (s^{-1})	K_M (μM)
25	7.2	3.81	367	3.01	240
50	16.8	7.26	296	7.21	235
100	37.7	10.1	278	8.93	173
200	81.9	9.61	221	10.6	205
300	127.5	ND	ND	13.2	242
400	173.7	ND	ND	11.7	223
500	220.4	ND	ND	14.5	282
		7.7 ± 1.4[b]	291 ± 30[b]	9.8 ± 1.5[b]	229 ± 13[b]

[a]Dimer concentration calculated from total enzyme concentration using Equation 3.2 with K_d = 15.9 nM. [b]Mean ± SE (n = 4 for ICEKAT, n = 7 for EKMCMC). [c]ND, not determined.

Due to the limitations of traditional Michaelis−Menten analysis (*e.g.*, the lack of a clear plateau in the hyperbolic MM curves in **Figure 3.9**), we chose to employ the alternate analytical strategy of direct Bayesian inference from progress curves. This strategy does not necessitate that substrate concentrations greatly exceed the value of K_M and can be executed using the public-domain R computational package EKMCMC.[207,209] This strategy also maximizes the experimental efficiency and analytical rigor by utilizing full progress curves in lieu of initial rate plots derived therefrom. We independently performed a Bayesian analysis of the progress curves for each concentration of assayed $3CL^{pro}$ (*i.e.*, each plot in **Figure 3.4**). A posterior sample plot for each enzyme concentration is shown in **Figure 3.10**.

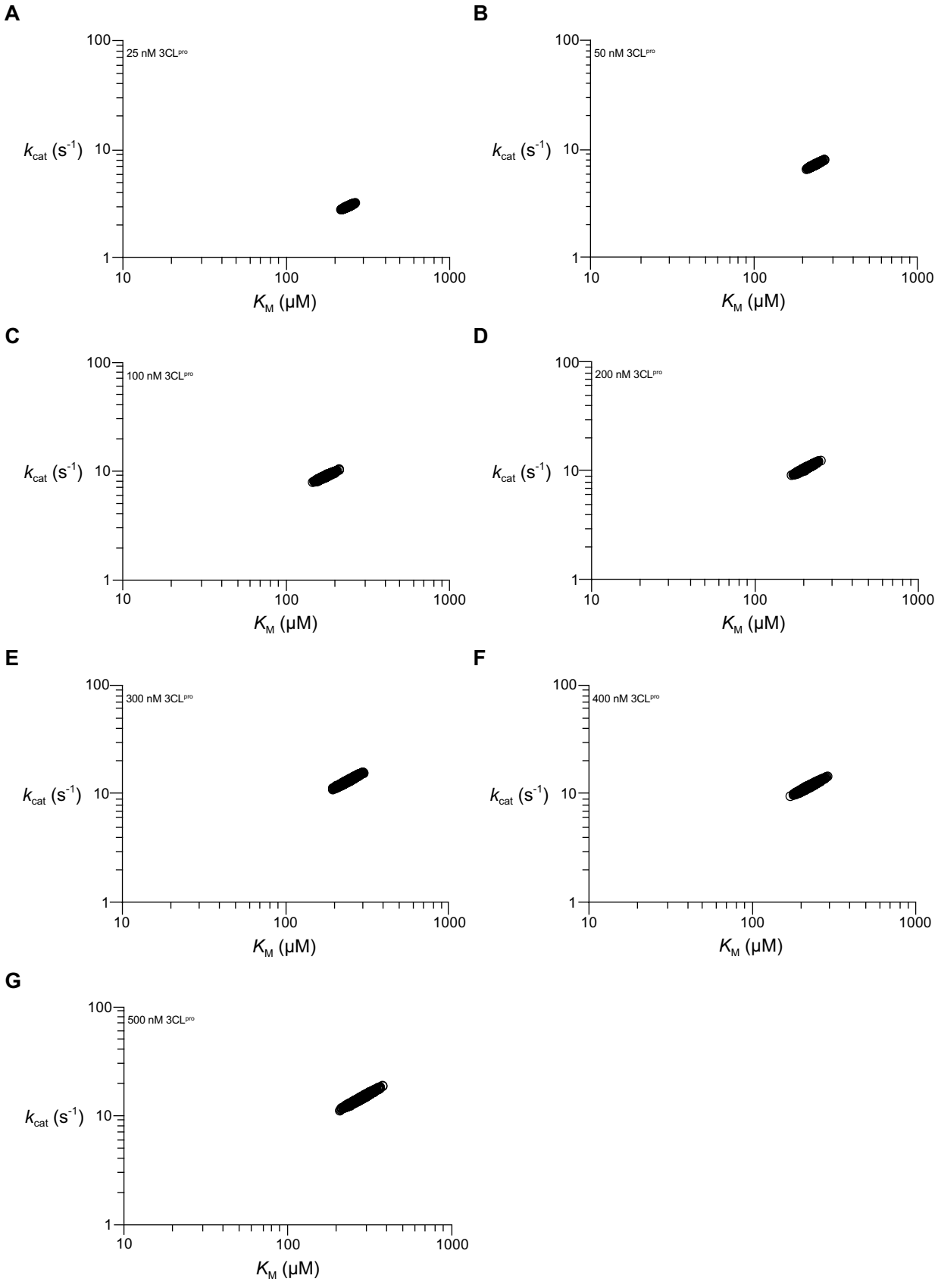

A
25 nM 3CL^pro
B
50 nM 3CL^pro
C
100 nM 3CL^pro
D
200 nM 3CL^pro
E
300 nM 3CL^pro
F
400 nM 3CL^pro
G
500 nM 3CL^pro
k_cat (s^-1)
K_M (μM)
100
10
1
10
100
1000

Figure 3.10. MCMC scatterplots of k_{cat} and K_M values for catalysis by $3CL^{pro}$, estimated by Bayesian inference (n = 3000 samples) of the progress curves in **Figure 3.4**. Data were analyzed with the EKMCMC package of Kim *et al.* using the determined $3CL^{pro}$ dimerization K_d to convert V_{max} to k_{cat}.[207,209] MCMC scatterplots are for (A) 25 nM, (B) 50 nM, (C) 100 nM, (D) 200 nM, (E) 300 nM, (F) 400 nM, and (G) 500 nM $3CL^{pro}$.

All posterior samples of k_{cat} and K_M (n = 3000 per $3CL^{pro}$ concentration) were confirmed to exhibit convergence, possess low autocorrelation between iterative estimates, and yield unimodal, bell-shaped sample distributions, as advised by the developers of the algorithm (**Figure 3.11** and **3.12**).[207,209]

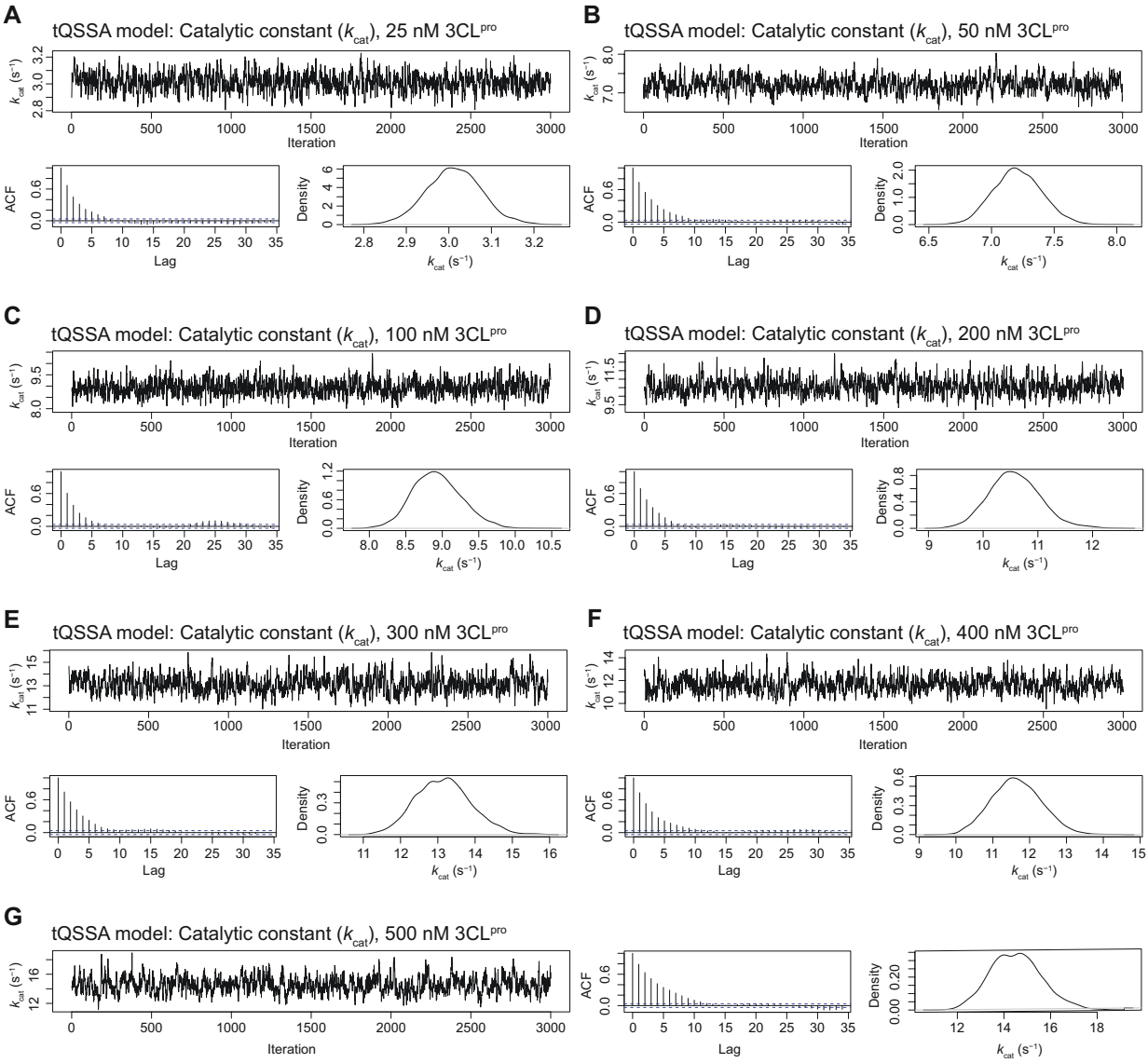

A
tQSSA model: Catalytic constant (k_{cat}), 25 nM 3CL^{pro}
B
tQSSA model: Catalytic constant (k_{cat}), 50 nM 3CL^{pro}
C
tQSSA model: Catalytic constant (k_{cat}), 100 nM 3CL^{pro}
D
tQSSA model: Catalytic constant (k_{cat}), 200 nM 3CL^{pro}
E
tQSSA model: Catalytic constant (k_{cat}), 300 nM 3CL^{pro}
F
tQSSA model: Catalytic constant (k_{cat}), 400 nM 3CL^{pro}
G
tQSSA model: Catalytic constant (k_{cat}), 500 nM 3CL^{pro}
k_{cat} (s^{-1})
Iteration
ACF
Lag
Density

Figure 3.11. Diagnostic graphs for Bayesian MCMC estimation of k_{cat} for catalysis by $3CL^{pro}$ from the progress curves in **Figure 3.4**, analyzed with the EKMCMC package of Kim *et al.* using the determined $3CL^{pro}$ dimerization K_d to convert V_{max} to k_{cat}.[207,209] Diagnostic graphs are for (A) 25 nM, (B) 50 nM, (C) 100 nM, (D) 200 nM, (E) 300 nM, (F) 400 nM, and (G) 500 nM $3CL^{pro}$. Within each sub-figure, the top graph is a trace plot showing the sampled k_{cat} values for successive iterations, the bottom-left graph is an auto-correlation function (ACF) plot showing the correlation between successive sampled k_{cat} values as a function of the distance between iterations, and the bottom-right graph is a posterior sample distribution density plot. All plots were drawn with 3000 posterior samples after removing the first 1000 samples (burn-in period) and applying a thinning rate of 1/30. All diagnostic plots are as desired: (1) the trace plots show convergent sampling, (2) the ACF plots show that successive samples rapidly become independent (ACF ≈ 0) as the lag increases, and (3) the posterior sample distribution density plots are unimodal and approximately Gaussian in shape.

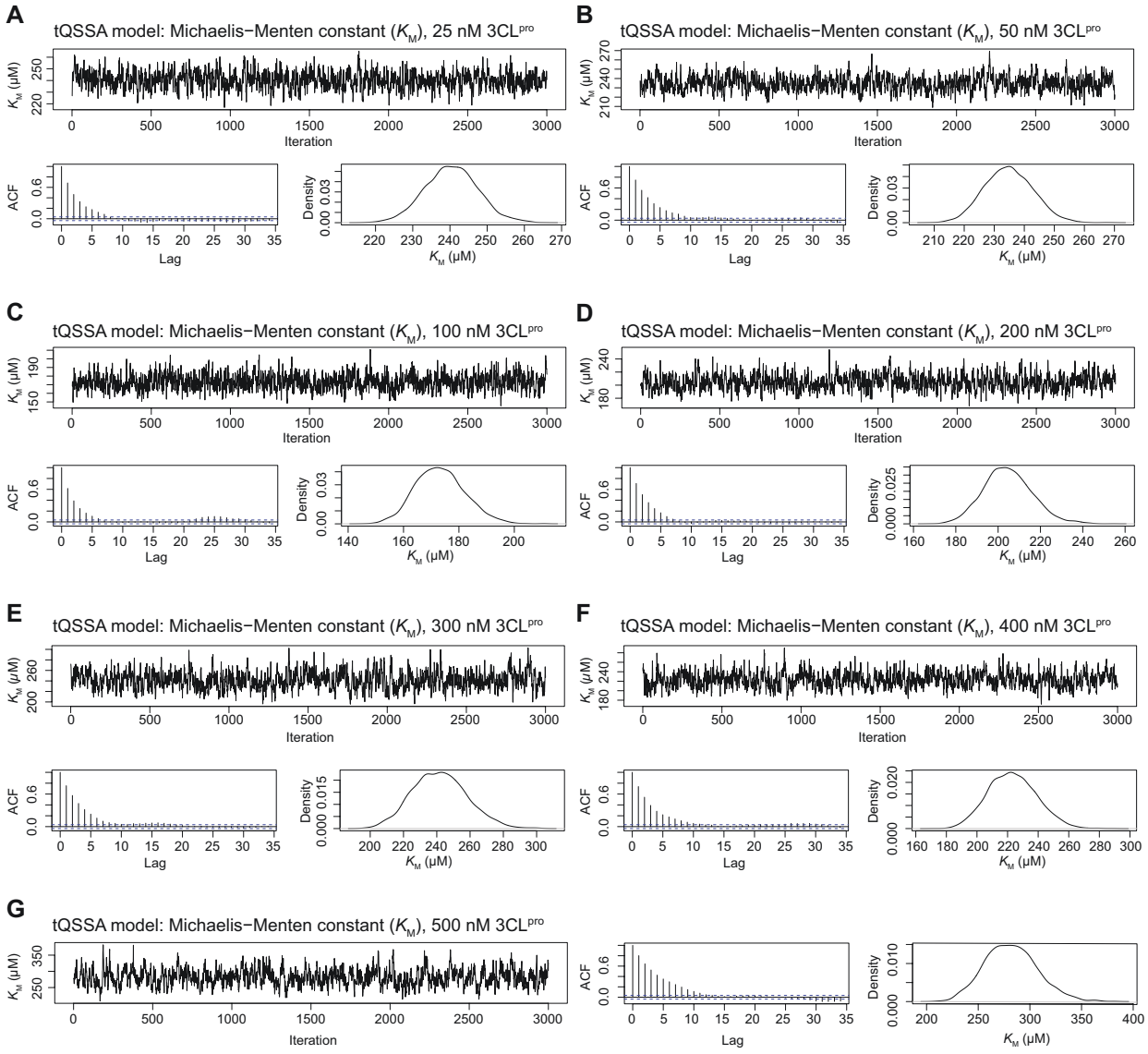

A
tQSSA model: Michaelis−Menten constant (K_M), 25 nM 3CL^{pro}
B
tQSSA model: Michaelis−Menten constant (K_M), 50 nM 3CL^{pro}
C
tQSSA model: Michaelis−Menten constant (K_M), 100 nM 3CL^{pro}
D
tQSSA model: Michaelis−Menten constant (K_M), 200 nM 3CL^{pro}
E
tQSSA model: Michaelis−Menten constant (K_M), 300 nM 3CL^{pro}
F
tQSSA model: Michaelis−Menten constant (K_M), 400 nM 3CL^{pro}
G
tQSSA model: Michaelis−Menten constant (K_M), 500 nM 3CL^{pro}
K_M (μM)
Iteration
ACF
Lag
Density
K_M (μM)

Figure 3.12. Diagnostic graphs for Bayesian MCMC estimation of K_M for catalysis by $3CL^{pro}$ from progress curves in **Figure 3.4**, analyzed with the EKMCMC package of Kim *et al*.[207,209] Diagnostic graphs are for (A) 25 nM, (B) 50 nM, (C) 100 nM, (D) 200 nM, (E) 300 nM, (F) 400 nM, and (G) 500 nM $3CL^{pro}$. Within each sub-figure, the top graph is a trace plot showing the sampled K_M values for successive iterations, the bottom-left graph is an auto-correlation function (ACF) plot showing the correlation between successively sampled K_M values as a function of the distance between iterations, and the bottom-right graph is a posterior sample distribution density plot. All plots were drawn with 3000 posterior samples after removing the first 1000 samples (burn-in period) and applying a thinning rate of 1/30. All diagnostic plots are as desired: (1) the trace plots show convergent sampling, (2) the ACF plots show that successive samples rapidly become independent (ACF ≈ 0) as the lag increases, and (3) the posterior sample distribution density plots are unimodal and approximately Gaussian in shape.

The sample mean values at each concentration of $3CL^{pro}$ are listed in **Table 3.3**; because the Bayesian analytical strategy is not limited to initial reaction rates, we were able to analyze the progress curves at all $3CL^{pro}$ concentrations. The Bayesian-estimated catalytic efficiency of $3CL^{pro}$ with our FRET substrate is $k_{cat}/K_M = (4.3 \pm 0.7) \times 10^4\ M^{-1}\ s^{-1}$ (mean ± SE, $n = 7$).

Given that the Bayesian analytical strategy is not subject to the requirements of the Michaelis−Menten initial rate strategy, as well as the fact that the MM initial rate plots only just began to approach hyperbolic asymptotes within the substrate concentrations used, we believe that the Bayesian-determined catalytic parameters are a more reliable representation of the catalytic efficiency of SARS-CoV-2 $3CL^{pro}$. Comparing the results of the Michaelis−Menten strategy and the Bayesian inference one, it is evident that the use of a more limited analytical technique alone induces a twofold underestimate of the catalytic efficiency of $3CL^{pro}$.

3.3.5 Effect of Ionic Strength on $3CL^{pro}$

Motivated by reports that catalysis by SARS-CoV $3CL^{pro}$ is highly sensitive to solution ionic strength,[125,201,210] we investigated the impact of NaCl concentration on enzymatic activity. Initially, we assayed the cleavage of 100 μM substrate by 50 nM $3CL^{pro}$ in 50 mM HEPES−NaOH buffer, pH 7.5, containing DTT (1 mM) and NaCl (0−1000 mM). All buffers were adjusted to constant pH after salt dissolution. We chose these enzyme and substrate concentrations because of the extended period of linear catalytic activity (see **Figure 3.4**), which facilitates the determination of the initial enzymatic reaction rate. As shown in **Figure 3.13**, we observed attenuation of the $3CL^{pro}$ initial reaction rate at increased ionic strength with a nearly 40% loss of enzymatic activity at 200 mM NaCl.

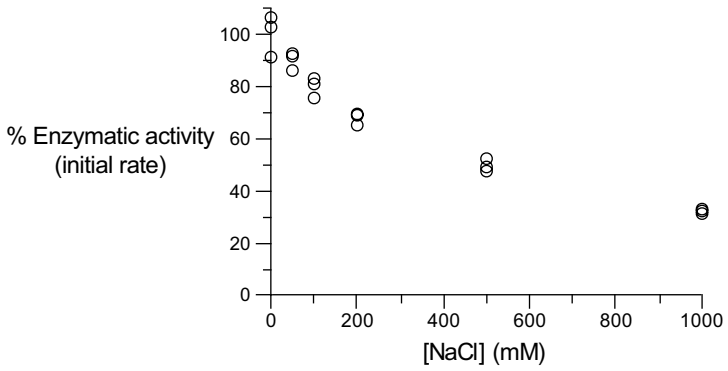


Figure 3.13. Influence of NaCl concentration on catalysis by $3CL^{pro}$ in 50 mM HEPES−NaOH buffer, pH 7.5, containing DTT (1 mM) at 25 °C. Reaction rates are normalized to those of assays run in the absence of added NaCl.

We attempted to repeat the enzymological determination of the $3CL^{pro}$ dimerization K_d for 50 mM HEPES−NaOH buffer, pH 7.5, containing NaCl (150 mM) and DTT (1 mM), as an increased K_d would result in a lower concentration of dimer at any given concentration of total $3CL^{pro}$ and would explain our observation of lower initial reaction rates in the presence of NaCl (see Equations 3.1 and 3.2). Our analytical strategy was unsuccessful, however, due to the emergence of a pronounced lag phase in the time course of the enzymatic reaction (data not shown). The length of this lag phase, which increased with lower enzyme or substrate concentrations, resulted in overestimated pseudo-first-order rate constants (r_x). Accordingly, we were unable to accurately determine the dimerization K_d by enzymatic means in the presence of added NaCl. Importantly, although the length of the lag phase is dependent on $3CL^{pro}$ and substrate concentrations, it does not appear to be as significantly influenced by the concentrations of NaCl used for the initial rates reported in **Figure 3.13**. We propose plausible causes for the emergence of the lag phase in the Discussion.

3.4 Discussion

Reliable data on the enzymology of the 3CL^{pro} protease of SARS-CoV and SARS-CoV-2 are necessary for our understanding of its catalysis and for therapeutic intervention. Values for both the catalytic efficiency (k_{cat}/K_M) and dimerization dissociation constant (K_d) vary substantially, with the latter parameter spanning nearly six orders of magnitude (see **Table 3.1** and **3.2**). The variation in K_d values is particularly concerning because the analysis of 3CL^{pro} enzymological data relies on knowing the concentration of the dimer, which manifests catalysis. At least one of the publications in our survey assumed that the analytical concentration of enzyme ($E_{T,x}$) was identical to the concentration of active catalyst ($[D]_x$).[188] This assumption is incorrect because $[D]_x$ differs from $E_{T,x}$ to an extent that depends on the value of K_d as described in Equation 3.2, which accounts for the half-site reactivity of 3CL^{pro}.[131,133,135]

Some variability in the reported values of k_{cat}/K_M is understandable. For example, there is no singular substrate that must be used for determining the 3CL^{pro} activity. The enzyme processes viral polyproteins at multiple, distinct cleavage sites and is thus somewhat promiscuous.[206] Still, nearly all data are for the cleavage of an LQ↓SG peptide bond (**Table 3.1**), as in our substrate. There is also a slight variability in solution conditions (pH, salt concentration, and temperature) that could affect enzymatic catalysis. We note that rather than determining the optimal enzyme substrate or constraining ourselves to endogenous 3CL^{pro} substrate sequences, we chose a rationally designed substrate sequence to balance catalysis with other experimental considerations that have likely limited other studies, such as peptide solubility.

Of greater concern are assay factors that would impair the dimerization of 3CL^{pro}, which would correspondingly decrease the concentration of catalytically competent enzyme and lead to erroneous normalization of enzymological data when determining the catalytic efficiency.[203,211]

The use of artificial purification tags is expected to impede dimerization as several key interprotomer interactions occur at terminal residues in 3CL^{pro}. Aside from simply assembling the dimeric state, the terminal residues of one protomer structure contribute to the active site of its sibling protomer and are critical to catalysis.[132,211] Several studies have reported that N- and C-terminal truncations of 3CL^{pro} significantly disrupt dimerization and/or catalysis,[202,212] and at least one study has shown that the introduction of exogenous terminal tags as small as two residues leads to a large decrease in catalytic activity.[194] The alteration of 3CL^{pro} termini likely disrupts key interactions, including the Ser1···Phe140′ hydrogen bonds and the Arg4···Glu290′ salt bridges (where ′ denotes the sibling protomer, and these interactions occur reciprocally between the two protomers). The loss of these interactions leads to a collapsed oxyanion hole and obviates catalytic activity by impeding stabilization of the negatively charged tetrahedral transition state during proteolysis.[132,211] The C terminus is also important for stabilizing the dimeric state, as demonstrated by several C-terminal single-point mutants that exhibited decreased catalytic activity and an apparent molecular mass corresponding to monomeric 3CL^{pro}.[213] Studies entailing assays of 3CL^{pro} with unnatural termini report some of the lowest k_{cat}/K_M values (**Table 3.1**). Surprisingly, several of the reported constructs most predisposed to dimerization include terminal tags (**Table 3.2**). This observation could be the result of variable analytical techniques used to measure K_d, some of which lack the sensitivity needed to detect nanomolar dissociation constants. For a single analytical technique employed within one study, the impact of terminal tags on dimerization is apparent. For example, Liang, Wang, and co-workers found that the addition of tags led to a ≤ 50-fold increase in the measured value of K_d.[200] We believe that the assay of *native* 3CL^{pro} is indispensable to understanding its intrinsic catalytic efficiency.

Another, albeit lesser, consideration is the concentration of salt used in the $3CL^{pro}$ assay buffers. As shown in **Figure 3.13**, there is a significant decrease in activity upon the addition of NaCl to the assay buffer, with a nearly twofold decrease in catalytic efficiency with 200 mM NaCl. Similar results were reported by Chang and co-workers for the SARS-CoV homologue, where 150 mM NaCl induced partial dissociation of the dimer.[201] Mangel and co-workers also report that NaCl disrupts SARS-CoV $3CL^{pro}$ activity, though they observed a large activity loss (80%) at only 100 mM NaCl.[210] Due to other variable assay parameters, relationships between [NaCl] and k_{cat}/K_M or K_d are difficult to distinguish in the data listed in **Table 3.1** and **3.2**, respectively, but we do note that for the SARS-CoV-2 homologue, the three lowest reported catalytic efficiencies all included NaCl in the assay buffer.

Increases in ionic strength likely diminish the free energy of Coulombic interactions, both between protomers of the dimer and between the enzyme and its bound substrate. Mangel and co-workers postulated that high ionic strength disrupts the interactions of Glu166 and a peptidic substrate.[210] Similarly, Shi and Song observed that the inclusion of NaCl in the assay buffer led to decreased catalytic activity without perturbing dimerization, suggesting that ionic strength disrupts substrate binding or catalysis itself.[213] Conversely, as previously noted, Chang and co-workers observed that NaCl disrupted the quaternary structure of $3CL^{pro}$.[201] Ferreira and Rabeh reported that SARS-CoV-2 $3CL^{pro}$ exhibits reduced thermodynamic stability when [NaCl] ≥ 100 mM without dependency on the precise salt concentration, which they attributed to the disruption of salt bridges.[126] Velazquez-Campoy, Abian, and co-workers observed a subtle decrease in the thermostability of SARS-CoV-2 $3CL^{pro}$ in the presence of 150 mM NaCl.[190] Given the aforementioned salt bridges that are responsible for stabilizing the $3CL^{pro}$ dimer, we believe that increased ionic strength would interfere with dimerization, as well as catalysis.

Because of the likelihood that NaCl disrupts catalysis either directly or through its effect on enzyme dimerization, we preemptively chose to omit NaCl in our standard $3CL^{pro}$ activity assay to optimally assess its catalytic efficiency.

The dimerization of $3CL^{pro}$ can be induced by a substrate.[214] A key residue in this process appears to be Glu166, which has been implicated in providing structure to the oxyanion hole. In the monomeric state of $3CL^{pro}$, Glu166 blocks access to the critical S1 substrate-binding subsite, but in the dimeric state, Glu166 interacts instead with Ser1′, revealing the S1 subsite and stabilizing the oxyanion hole. Intriguingly, molecular dynamics simulations suggest some conformational flexibility of Glu166 in the monomeric state, with the residue precluding access to the S1 subsite most of but critically not all of the time.[132] The series of events for substrate cleavage by $3CL^{pro}$ are often thought of as being (1) dimerization → (2) subsite revelation → (3) substrate binding → (4) catalysis, but Chou and co-workers propose that the series is (1) subsite revelation → (2) substrate binding → (3) dimerization → (4) catalysis.[214] Moreover, they suggest that Glu166 forms a hydrogen bond with Asn142, stabilizing an adjacent 3_{10}-helix to maintain a collapsed oxyanion hole. Upon substrate binding, Glu166 instead interacts with the substrate P1 glutamine and Ser1′, disrupting the Glu166···Asn142 hydrogen bond and causing conformational changes that stabilize an intact oxyanion hole.[211,214] Thus, the binding of the substrate to a monomer could stabilize Glu166 in a conformation that facilitates dimerization, forms an oxyanion hole, and enables catalysis. The lag phase that we observed in $3CL^{pro}$ activity assays with 150 mM NaCl might be a consequence of disrupted $3CL^{pro}$ dimerization, and the length of the lag phase might be variable because the extent of substrate-induced dimerization is dependent upon the substrate concentration.

3.5 Conclusions

Our analyses of 3CL^{pro} have relied on core principles of enzymology. (1) We characterize the authentic enzyme, without exogenous residues (**Figure 3.2**). (2) We designed our peptidic substrate to accommodate the structural features of the 3CL^{pro} substrate subsites and to position the pendant FRET moieties far from the scissile bond so as not to disrupt binding or catalysis. (3) We assayed 3CL^{pro} activity in the absence of NaCl, mitigating the disruption to dimerization (**Figure 3.13**). (4) We characterized the magnitude of the inner filter effect (**Figure 3.6**) and accounted for its consequences in our analyses. (5) We determined the value of K_d through enzymological assays, which are more sensitive than other methods (*e.g.*, analytical gel filtration chromatography or SDS−PAGE). (6) We determined the catalytic parameters k_{cat} and K_M for our substrate through two means: initial rate (*i.e.*, Michaelis−Menten) analysis and progress curve (*i.e.*, Bayesian Markov Chain Monte Carlo) analysis. The latter strategy maximizes data utility and unshackles enzymology from the empirical limitations of Michaelis−Menten techniques.

Our work reveals that 3CL^{pro} is a catalyst that is more capable than previously reported. Moreover, 3CL^{pro} is much more predisposed toward dimerization than is appreciated, having a K_d value in the nanomolar range. These findings define the landscape for the design of protease inhibitors for the prevention and treatment of COVID-19 infection and could facilitate efforts to combat the COVID-19 pandemic and potentially assist in the prevention of future coronavirus-based outbreaks.

3.6 Acknowledgments

E.C.W. was supported by a Graduate Research Fellowship from the NSF. This work was supported by Grants R35 GM148220 and R21 AI17166301.

A pGEX-6P-1 plasmid encoding SARS-CoV-2 $3CL^{pro}$ was a kind gift from Professor Rolf Hilgenfeld (University of Lübeck). The authors are grateful to JoLynn B. Giancola for assistance with Q–TOF and MALDI–TOF mass spectrometry, Jinyi Yang for assistance with MALDI–TOF mass spectrometry, and Dr. Nile S. Abularrage for assistance with analytical HPLC.

3.7 Experimental Details

3.7.1 Materials

All chemicals and reagents were of commercial reagent grade or better and were used without further purification.

3.7.2 Conditions

All procedures were performed in air at ambient temperature (~22 °C) and pressure (1.0 atm) unless indicated otherwise.

3.7.3 Production and Purification of SARS-CoV-2 3CL^{pro}

Authentic 3CL^{pro} was produced and purified following methods described previously.[129] Briefly, BL21-Gold(DE3) competent *Escherichia coli* were transformed with the pGEX-6P-1/3CL^{pro} plasmid by heat shock and grown overnight at 37 °C on Luria–Bertani (Miller) agar (2% w/v) containing ampicillin (100 μg mL^{-1}). A starter culture in 1× YT medium (which contained 0.8% w/v tryptone, 0.5% w/v yeast extract, 0.25% w/v NaCl, and 100 μg mL^{-1} ampicillin) was inoculated with a single transformant colony and grown overnight at 37 °C with shaking at 250 RPM. Cultures of 1× YT medium were inoculated with the starter culture to a starting OD_{600} = 0.05 and incubated at 37 °C with shaking at 250 RPM until OD_{600} = 0.8. Gene expression was induced with isopropyl β-d-1-thiogalactopyranoside (final concentration: 0.5 mM) for 5 h at 37 °C with shaking. Cultures were pelleted by centrifugation and held at −70 °C. The induction of expression was confirmed by SDS–PAGE.

Pelleted *E. coli* cells were thawed briefly and resuspended in 20 mM Tris–HCl buffer, pH 7.8, containing NaCl (150 mM) at room temperature. *E. coli* were passed through a cell disruptor

(Constant Biosystems), and the lysate was subjected to centrifugation at 38,400*g* and 4 °C for 2 h. The supernatant was loaded onto a HisTrap FF column (Cytiva) that had been pre-equilibrated with lysis buffer. The column was washed with 20 column-volumes of lysis buffer, and the target protein was eluted with a linear gradient (20 column volumes) of lysis buffer containing imidazole (0–500 mM). Fractions with target protein were simultaneously treated with PreScission protease (Cytiva) and dialyzed against 20 mM Tris–HCl buffer, pH 7.8, containing NaCl (150 mM) and DTT (1 mM) at 4 °C overnight with gentle stirring. The retentate was passed over coupled GSTtrap FF and HisTrap FF columns (Cytiva) to remove the GST tag, His tag, PreScission protease, and unprocessed $3CL^{pro}$. The flowthrough was exchanged into 20 mM Tris–HCl buffer, pH 8.0, containing DTT (1 mM), loaded on a HiTrap Q FF column (Cytiva) pre-equilibrated with the same buffer, and eluted by a linear gradient of exchange buffer with 0–500 mM NaCl over 20 column volumes. Fractions with target protein were pooled and exchanged into 20 mM Tris–HCl buffer, pH 7.8, containing NaCl (150 mM), EDTA (1 mM), DTT (1 mM). Purified protein was flash-frozen in liquid nitrogen, quantified by BCA assay, and stored at −70 °C.

3.7.4 Analysis of $3CL^{pro}$ Purity

For Q–TOF mass spectrometry of purified $3CL^{pro}$, a 1 μM solution of protein was made in aqueous acetonitrile (5% v/v) containing formic acid (0.1% v/v). A 15-μL aliquot (15 pmol of $3CL^{pro}$) was injected onto a 40 °C-thermostatted PLRP-S column (5-μm particle, 1000-Å pore; Agilent Technologies) and eluted with a gradient of aqueous acetonitrile (5–95% v/v) over 7.5 min. Eluted protein was introduced to an Agilent 6530 Q–TOF mass analyzer by ESI and analyzed in positive mode.

3.7.5 Differential Scanning Fluorimetry

The thermostability of $3CL^{pro}$ was determined with differential scanning fluorimetry. To do so, $3CL^{pro}$ was mixed with SYPRO Orange protein gel stain (Supelco) in 50 mM HEPES–NaOH buffer, pH 7.5, to a final concentration of 100 μg mL^{-1} protein, 50× SYPRO Orange (vendor stock: 5000× in DMSO). The resulting solution was then heated from 15–95 °C at 1 °C/min using a QuantStudio 7 Flex Real-Time PCR system (Applied Biosystems); fluorescence was monitored in real-time with λ_{ex} = 470 ± 15 nm and λ_{em} = 586 ± 10 nm. Data were processed with Protein Thermal Shift software (Applied Biosystems) using the Boltzmann fitting method. Values of T_m represent the temperature at which fluorescence was 50% maximal.

3.7.6 Design and Synthesis of a $3CL^{pro}$ Substrate

Aided by the homology of $3CL^{pro}$ from SARS-CoV and SARS-CoV-2, known polyprotein cleavage sites for SARS-CoV,[125] and simulated interactions in a Michaelis complex of SARS-CoV-2 $3CL^{pro}$ with a peptidic substrate,[204] a candidate substrate sequence ATLQ↓SGNA (↓, cleavage site) was chosen. The octapeptide was flanked by EDANS and DABCYL conjugated to glutamic acid and lysine, respectively, to serve as a FRET pair; the peptide was further capped by terminal arginine residues to improve solubility. The desired peptide R–E(EDANS)–ATLQ↓SGNA–K(DABCYL)–R was synthesized by solid-phase peptide synthesis. Briefly, 2-chlorotrityl chloride resin loaded with Fmoc-Arg(Pbf)-OH was deprotected and coupled to Fmoc-Lys(DABCYL)-OH using PyBOP/DIPEA as the activating agent. The substrate peptide sequence was then extended using a CEM Liberty Blue automated peptide synthesizer prior to manual coupling of the Fmoc-Glu(EDANS)-OH residue, which was achieved using PyBOP/DIPEA pre-activation. Deprotection of the N-terminal Fmoc group was then performed

before the final Fmoc-Arg(Pbf)-OH residue was coupled using PyBOP/DIPEA. The resulting peptide was deprotected and cleaved under standard SPPS conditions and purified by preparative reversed-phase HPLC. The control peptide R–E(EDANS)–ATLQ, which represents the N-terminal cleavage product of the substrate peptide, was synthesized by analogous means.

3.7.7 Analysis of Peptide Purity

For analytical HPLC of the two peptides, 50 μM peptide was dissolved in water containing trifluoroacetic acid (0.1% v/v). Peptide (2.5 nmol) was injected onto a non-thermostatted C18 column (250 × 4.6 mm, 5-μm particle, 100-Å pore; Varian) and eluted with a gradient of acetonitrile (15–40% v/v) in water over 10 min. Eluted peptide was detected with a diode array detector at λ = 210, 336, and 494 nm for peptide bond, EDANS, and DABCYL absorption, respectively.

For MALDI–TOF mass spectrometry of the two peptides, desalted peptide was spotted on an α-cyano-4-hydroxycinnamic acid matrix and analyzed with a Bruker microflex MALDI–TOF mass spectrometer in linear positive mode.

3.7.8 3CL^{pro} Kinetics Assay

Standard assays of 3CL^{pro} were performed in 50 mM HEPES–NaOH buffer, pH 7.5, containing DTT (1 mM) at 25 °C. The concentration of FRET substrate varied from 5 to 200 μM, and the concentration of 3CL^{pro} varied from 25 to 500 nM, as indicated. Substrate cleavage was initiated by the addition of an equal volume of 2× enzyme to 2× substrate. Fluorescence was monitored at λ_{ex} = 336 ± 5 nm and λ_{em} = 494 ± 10 nm over 15–60 min, as indicated. Nonlinear regression

analyses to Equations B.12 and B.14 were performed with Prism v6.0 software (GraphPad Software).

3.7.9 Michaelis–Menten Analysis of Kinetic Data

Progress curves were analyzed for initial rate Michaelis–Menten plots using the ICEKAT software of Smith and co-workers.[208] Fluorescence raw data were converted into product concentration by equating the asymptotic end point fluorescence intensity of each progress curve to the initial substrate peptide concentration. Then, ICEKAT software was used with the default setting "maximize slope magnitude" to process the data (n = 6 substrate concentrations × 3 technical replicates) independently at each concentration of $3CL^{pro}$ ($E_{T,x}$).

3.7.10 EKMCMC Analysis of Kinetic Data

Full progress curves were also analyzed with the EKMCMC R algorithm of Kim and co-workers.[207,209] Fluorescence raw data were converted into product concentrations by equating the asymptotic endpoint fluorescence intensity of each progress curve to the initial substrate peptide concentration. The EKMCMC algorithm was used to process the dataset (n = 6 substrate concentrations × 3 technical replicates) independently at each concentration of $3CL^{pro}$ ($E_{T,x}$). For the `enz_data` field within EKMCMC, the concentration of catalytically competent dimer at each $E_{T,x}$ was estimated by using the previously determined K_d value. The EKMCMC algorithm was run using the total QSSA model with an initial K_M guess of 1000 μM (`K_M_init` = 1000, `K_M_m` = 1000), a burn-in period of 1000 samples (`burn` = 1000), a thinning rate of 1/30 (`jump` = 30), and an effective 3000 MCMC iterations (`nrepeat` = 3000).

Chapter 4:
Sensitive detection of SARS-CoV-2 main protease $3CL^{pro}$ with an engineered ribonuclease zymogen

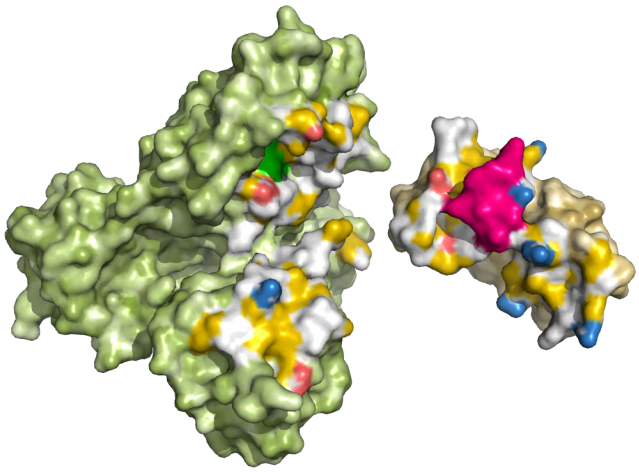
This chapter has been submitted for publication as:

Wralstad, E. C.; Raines, R. T. Sensitive detection of SARS-CoV-2 main protease $3CL^{pro}$ with an engineered ribonuclease zymogen.

Contributions: Conceptualization, E.C.W.; experimental methodology, E.C.W.; experimental investigation, E.C.W.; writing—original draft preparation, E.C.W.; writing—review and editing, R.T.R.; supervision, R.T.R.; funding acquisition, E.C.W. and R.T.R.

4.1 Abstract

Alongside vaccines and antiviral therapeutics, diagnostic tools are a crucial aid to combat the COVID-19 pandemic caused by the etiological agent SARS-CoV-2. All common assays for infection rely on the detection of viral sub-components, including structural proteins of the virion or fragments of the viral genome. Selective pressure imposed by human intervention of COVID-19 can, however, induce viral mutations that decrease the sensitivity of diagnostic assays based on biomolecular structure, leading to an increase in false-negative results. In comparison, mutations are unlikely to alter the *function* of viral proteins, and viral machinery is under less selective pressure from vaccines and therapeutics. Toward this end, we used intein-mediated *cis*-splicing to create a circular ribonuclease zymogen that is activated by the SARS-CoV-2 main protease, $3CL^{pro}$. Zymogen activation by $3CL^{pro}$ leads to a >300-fold increase in ribonucleolytic activity, which can be detected with a highly sensitive fluorogenic substrate. This coupled assay can detect low nanomolar concentrations of $3CL^{pro}$ within a timeframe comparable to that of common antigen-detection protocols. More generally, the concept of detecting a protease by activating a ribonuclease could be the basis of diagnostic tools for other indications.



4.2 Introduction

The COVID-19 pandemic caused by the betacoronavirus SARS-CoV-2 has been a public health crisis of extraordinary proportions. Success in combatting SARS-CoV-2 has been achieved by the development of preventive vaccines, therapeutic interventions, and diagnostic tools.[215] Still, the landscape has shifted continually. Several therapeutic agents employed in the treatment of COVID-19, including the neutralizing monoclonal antibodies produced by Lilly, Regeneron, GlaxoSmithKline, and others; the adenosine nucleoside analog Remdesivir; and the combination protease inhibitor formulation Paxlovid have exhibited reduced efficacy as a consequence of mutations in the SARS-CoV-2 genome.[187,216-218] Similarly, rapid antigen detection testing (RADT), which often targets the SARS-CoV-2 spike and nucleocapsid proteins and serves as an indispensable front-line tool to mitigate viral transmission, has suffered from reduced sensitivity.[137,138] For example, the appearance of the Omicron SARS-CoV-2 variant led to a days-long lag between the detection of viral infection by RADT versus PCR.[139] Even PCR-based testing, which is the gold standard, has lost sensitivity due to "gene dropout" in the spike and nucleocapsid proteins.[219-221]

RADT, PCR, and other diagnostic tools rely on the structure of a viral molecule. We sought to develop a diagnostic strategy that relies instead on molecular *function*. Such a function-based approach would mitigate the risk of mutation-based sensitivity loss because viral evolution to evade detection is unlikely to impair its replisome.

An obvious basis for a diagnostic assay relying on function is the catalytic cleavage of a fluorogenic substrate by a viral enzyme.[222-224] Such an assay is likely, however, to lack sensitivity because each turnover yields only one fluorescent molecule. We reasoned that the sensitivity

could be amplified if the viral enzyme instead activated a zymogen that then catalyzed the cleavage of a fluorogenic substrate (**Figure 4.1**).

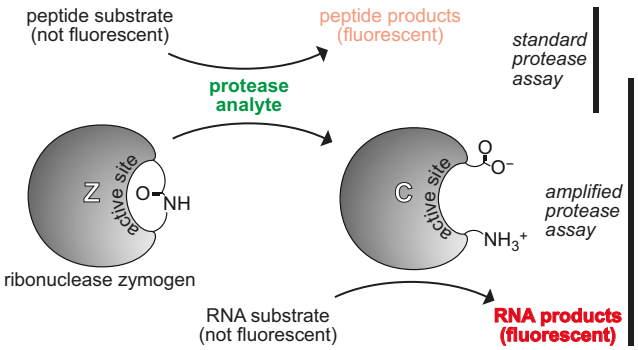


Figure 4.1. Standard and amplified assays for a protein analyte. In a standard assay, a protease cleaves a fluorogenic substrate. In an amplified assay (as in this work), a protease cleaves a peptide bond in the linker of a ribonuclease zymogen (*Z*), producing a cleaved zymogen (*C*) with an uncovered active site. Zymogen cleavage is detected by assaying ribonucleolytic activity with a fluorogenic substrate.

Moreover, we suspected that using two different types of enzymes (*e.g.*, a protease and a nuclease) would avoid complications from crosstalk between the enzymes and their substrates and products.

As a viral analyte, we choose the main protease $3CL^{pro}$. SARS-CoV-2 relies on this protease to process its two polyproteins at 11 cleavage sites, each flanked by a glutamine (P1=Gln in the Schechter-Berger nomenclature) and small, uncharged amino acid (P1′=Gly, Ala, or Ser).[125,204] Recently, we showed that $3CL^{pro}$ cleaves the consensus substrate sequence ATLQ↓SGNA (↓, cleavage site) with a high value of $k_{cat}/K_M = (4.3 \pm 0.7)\times10^4$ M^{-1} s^{-1}.[180]

As a zymogen, we choose one engineered from human ribonuclease 1 (RNase 1). The N and C termini of RNase 1 and its homologues flank the active site. Accordingly, connecting the native termini with a peptidic linker occludes the active site.[94-97] The linker of our zymogen contains the sequence: ATLQSGNA, as its cleavage by $3CL^{pro}$ would uncover the active site. Zymogen activation could be detected with a hypersensitive assay that is based on a fluorogenic ribonuclease substrate.[35,36]

Here, we design and prepare a ribonuclease zymogen and demonstrate that it is specifically cleaved by $3CL^{pro}$ within its inactivating linker. Using this assay, we can detect low nanomolar concentrations of $3CL^{pro}$ within a timeframe comparable to current RADT. For larger $3CL^{pro}$ concentrations, detection of the viral protease occurs within several minutes. Our strategy could inspire next-generation diagnostics that mitigate the sensitivity losses that can result from viral evolution.

4.3 Results

4.3.1 Modeling the zymogen and its interaction with 3CL^{pro}

We computationally predicted the structure of our 3CL^{pro}-directed, RNase 1-based zymogen with the AlphaFold2-based algorithm ColabFold and *in silico* permutations of the zymogen sequence. We designed those permutations in order to accommodate the software by installing false termini.[225] We assessed the predictions based on their structural similarity to native RNase 1 as well as the proximity of their false termini; four permuted sequences generated sensible zymogen models and are shown in **Figure 4.2**. The permuted sequences, individual model structures, and AlphaFold2 quality metrics are shown in **Figure 4.3**.

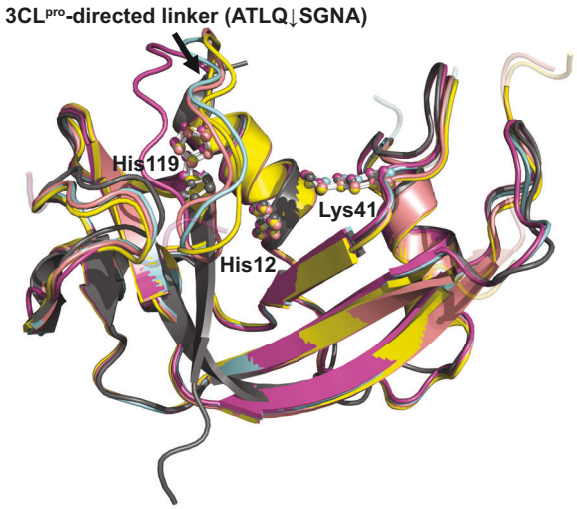
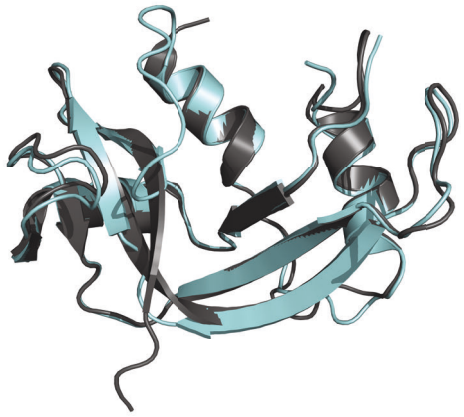


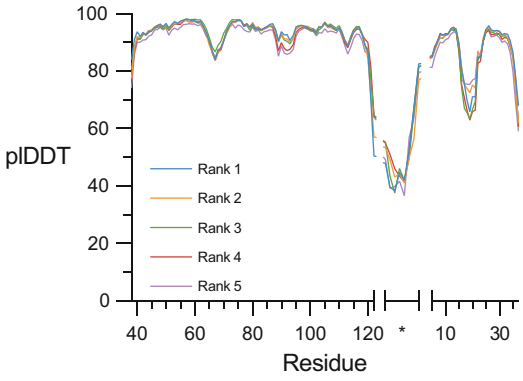
Figure 4.2. AlphaFold2-generated zymogen models based on sequences permuted around Q37|R38, G68|Q69, N88|R89, and R89|R90 (in color) aligned to wild-type RNase 1 (in gray), cartoon form with active site residues in ball-and-stick form.

A

```
  40          50          60          70          80          90          100         110         120
GRC KPVNTFVHEP LVDVQNVCFQ EKVTCKNGQG NCYKSNSSMH ITDCRLTNGS RYPNCAYRTS PKERHIIVAC EGSPYVPVHF

122 ******** 5    10          20          30       37
 DA ATLQSGNA AKKFQR QHMDSDSSPS SSSTYCNQMM RRRNMTQ
```

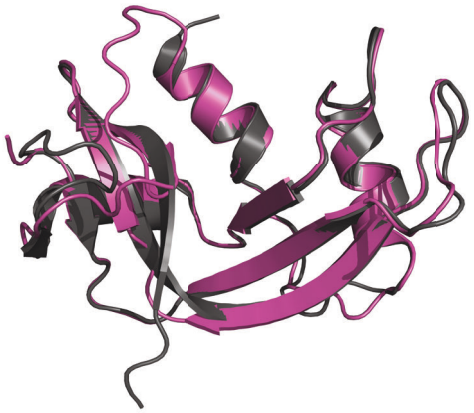


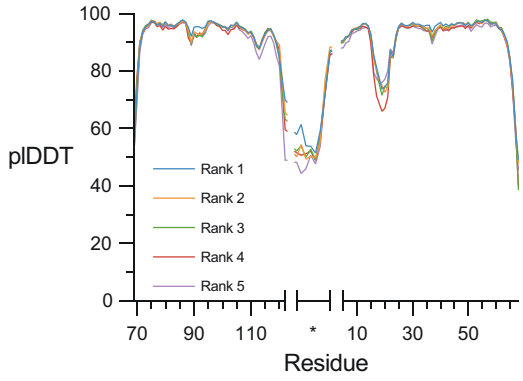


B

```
  70          80          90          100         110         120 122 ******** 5    10
QG NCYKSNSSMH ITDCRLTNGS RYPNCAYRTS PKERHIIVAC EGSPYVPVHF   DA ATLQSGNA AKKFQR

          20          30          40          50          60        68
QHMDSDSSPS SSSTYCNQMM RRRNMTQGRC KPVNTFVHEP LVDVQNVCFQ EKVTCKNG
```





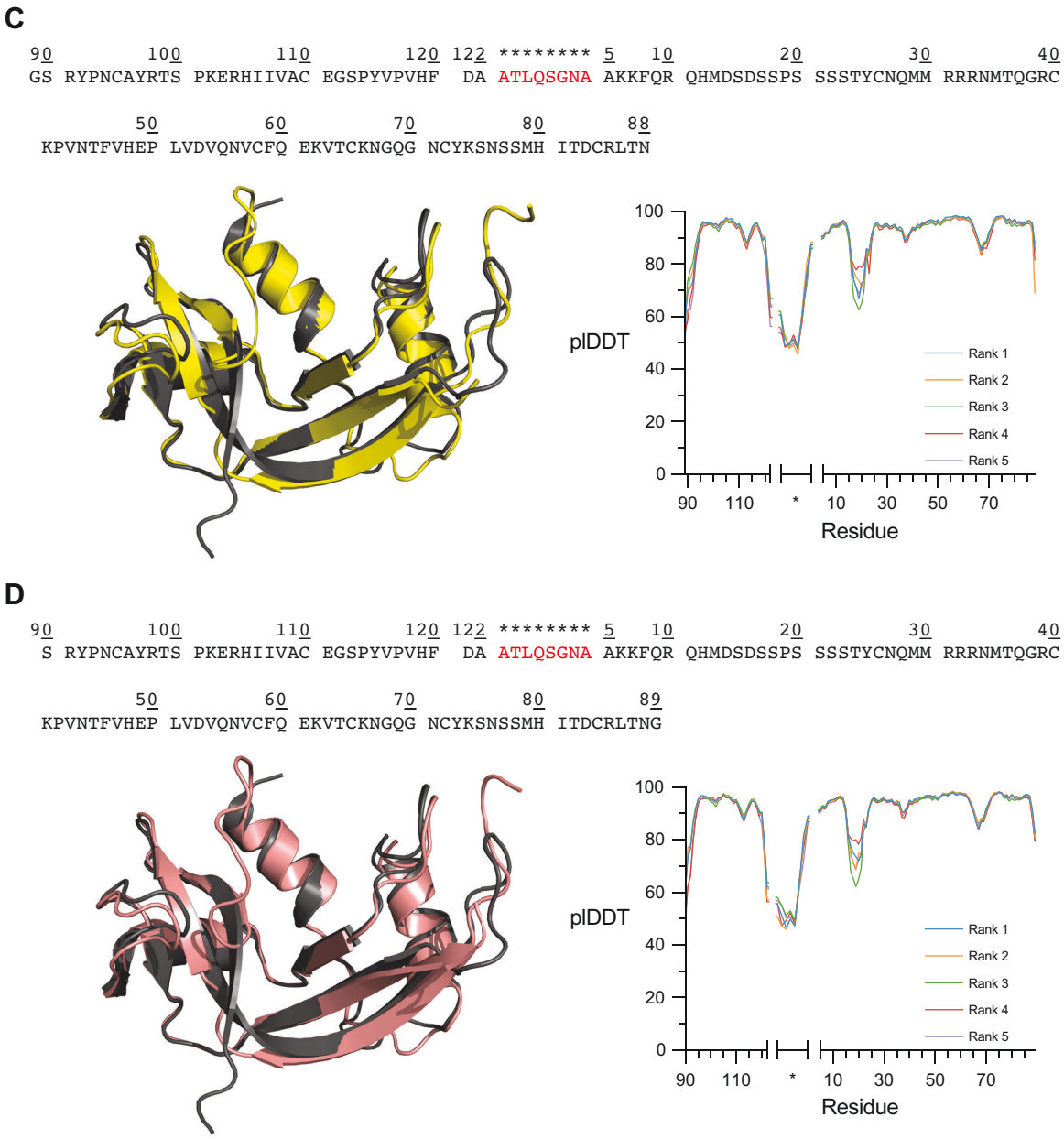


Figure 4.3. Zymogen structure predictions generated using AlphaFold2 with protein sequences circularly permuted around (A) Q37|R38, (B) G68|Q69, (C) N88|R89, and (D) R89|R90. Each subpanel shows: (top) the permuted protein sequence with the linker in red font and denoted by *; (left) the top-ranked prediction model structurally aligned to wild-type RNase 1 (in gray); and (right) plDDT scores for the five models produced from a given permutant sequence.

The backbone alignment of the models to wild-type RNase 1 reveals a high degree of structural identity with RMSD ≤ 2.5 Å. The permutant-dependent false termini are 4.1–20.4 Å apart (C-terminal carbonyl carbon to N-terminal amino nitrogen). The Q37|R38-permuted sequence yielded the shortest terminal distance and was thus a model most reminiscent of the circular zymogen, which has a peptide bond between the false termini.

Catalysis of RNA cleavage by RNase 1 and its homologues relies primarily on the side chains of three residues: His12, Lys41, and His119 (**Figure 4.4**).[22]

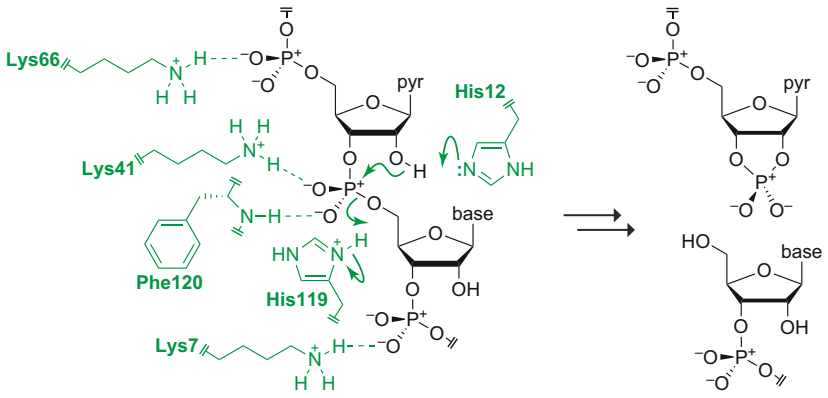


Figure 4.4. Putative mechanism of catalysis of RNA cleavage by RNase 1.

Based on the AlphaFold2-generated models, we predict that His12 and Lys41 are not significantly perturbed in the zymogen relative to RNase 1, but His119 is dislocated, as shown in **Figure 4.5**.

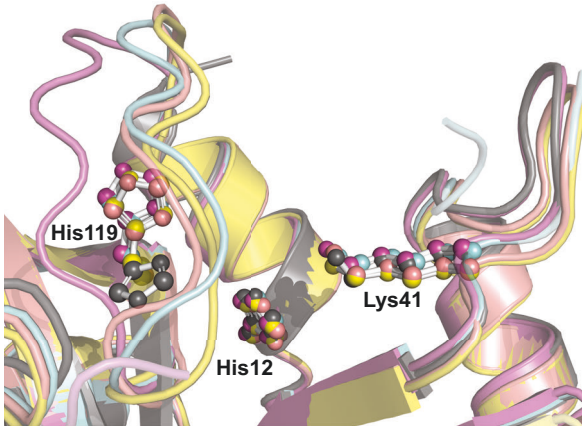


Figure 4.5. Active sites of the AlphaFold2-generated zymogen models permuted around Q37|R38, G68|Q69, N88|R89, and R89|R90 (in color) compared to that of wild-type RNase 1 (in gray), cartoon form with active site residues in ball-and-stick form.

This dislocation consistently appears between models: across four permuted zymogen sequences and five ranked models per sequence, the imidazole ring of His119 is rotated 123 ± 2° (mean ± SD) out of alignment from its native position (**Figure 4.6**).

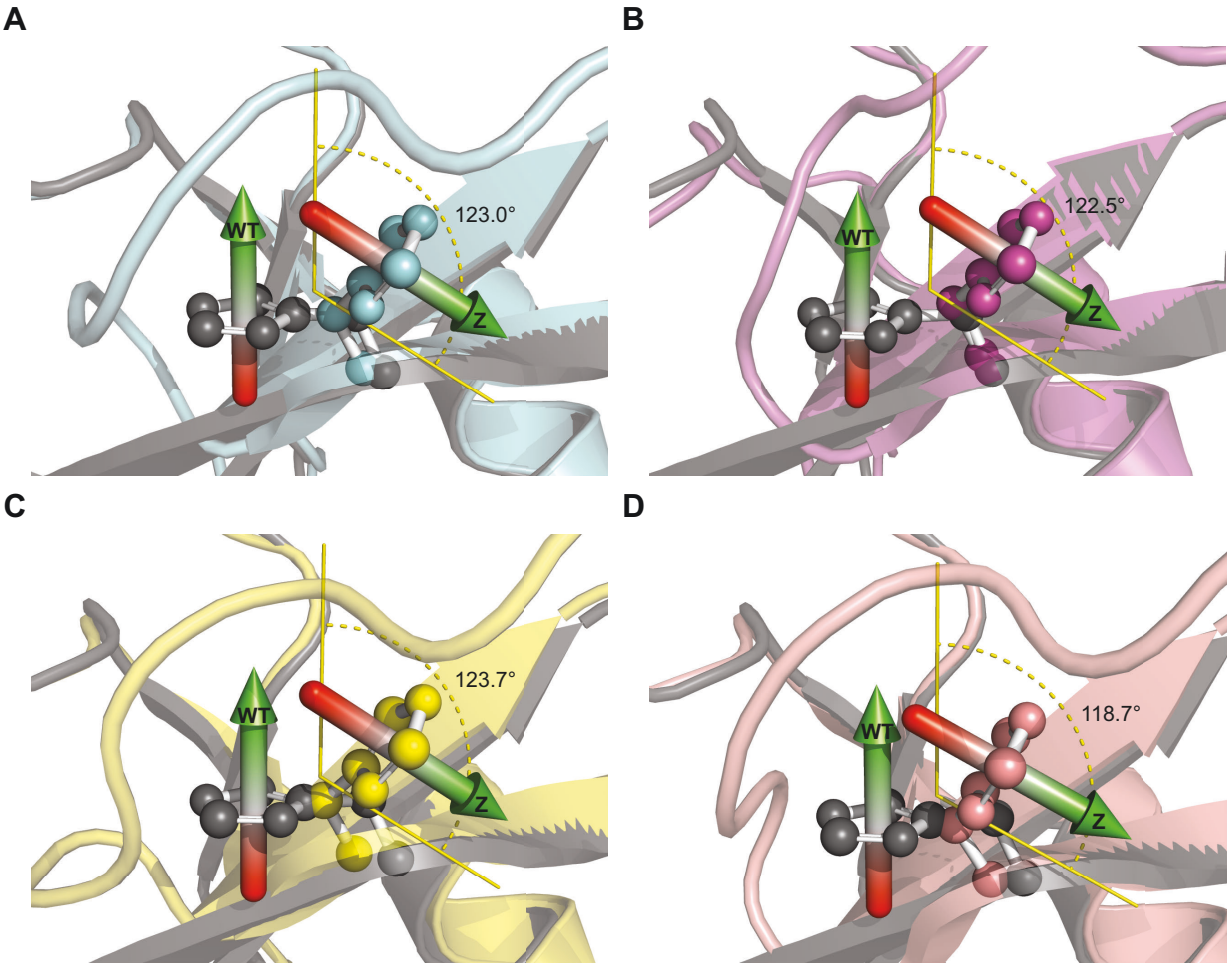


Figure 4.6. Dislocation angle of His119 in the AlphaFold2-generated zymogen structure predictions relative to wild-type RNase 1. Arrows represent the vector normal to the plane of the His119 imidazole ring in the wild-type protein, *WT*, and in the zymogen, *Z*. Panels represent the top-ranked model prediction generated with protein sequences circularly permuted around (A) Q37|R38, (B) G68|Q69, (C) N88|R89, and (D) R89|R90.

Docking the tetranucleotide substrate analog d(ATAA) to the zymogen also reveals a significant steric clash between the zymogen linker and nucleotides on the 3′ side of the scissile phosphodiester bond (**Figure 4.7**). We predict that active-site distortion and steric hindrance can work in tandem to inactivate the enzymatic activity of our engineered zymogen, as occurs in natural zymogens.[226]

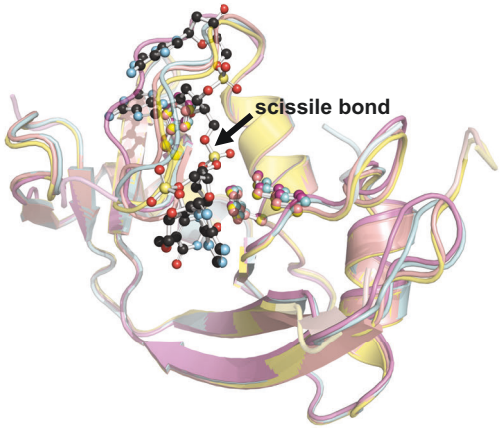
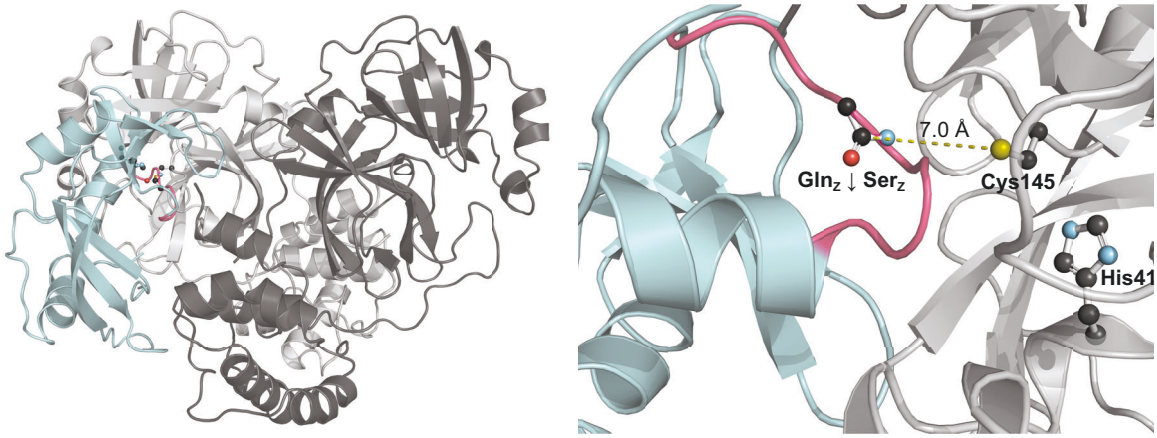


Figure 4.7. Docking of the ribonuclease substrate analog d(ATAA) with the AlphaFold2-generated zymogen models permuted around Q37|R38, G68|Q69, N88|R89, and R89|R90 (cartoon form with active site residues in ball-and-stick form, in color). Substrate is in ball-and-stick form/CPK color scheme and is derived from PDB ID 1RCN.[227]

To model interactions between the zymogen and 3CL^{pro}, we submitted the optimal zymogen model (Q37|R38: 1.08-Å RMSD to RNase 1, 4.8-Å terminal gap) and the structure of SARS-CoV 3CL^{pro} (PDB ID: 1UJ1[135]) to the ClusPro protein–protein docking server.[228] We chose to use the 3CL^{pro} homologue structure in lieu of a structure from SARS-CoV-2 because the former captures the physiologically relevant protease dimer within the crystallographic asymmetric unit rather than generating a dimer by symmetry operations; the two active sites of dimeric 3CL^{pro} are reported to be conformationally nonidentical, and an accurate representation of the two component protomers is necessary for relevant zymogen docking.[131,133,135] Because SARS-CoV and SARS-CoV-2 3CL^{pro} share 96% sequence identity, our docking results are likely to be valid for SARS-CoV-2 3CL^{pro}.

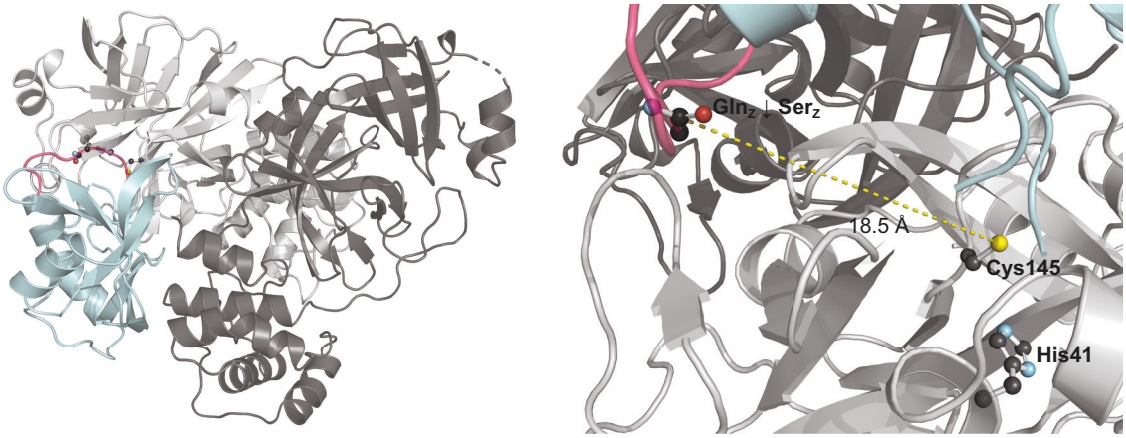
Surface modeling with ClusPro revealed significant topological complementarity between the zymogen and 3CL^{pro} (see Digital Supplementary Material file **wralstad-evansw-phd-chem-2024-supplemental_1.mp4**, which depicts surface-modeled protein–protein docking of the zymogen and 3CL^{pro}). Specifically, ClusPro yields a high-ranking model (3 of 30) in which the zymogen linker is inserted into one of the 3CL^{pro} active sites (**Figure 4.8A, left**). In the zymogen linker, the main-chain carbonyl carbon atom of Gln_Z (C) is the intended electrophile. In the high-ranking model, this carbon is within 7 Å of the sidechain sulfur atom (S^{γ}) of Cys145, which is the active-site nucleophile (**Figure 4.8A, right**). Notably, 6 of 30 ClusPro docked structures have the scissile bond of the zymogen near the active site of 3CL^{pro}, with an average $S^{\gamma}\cdots C$ distance of 12.3 Å and a range of 7.0–18.5 Å (**Figure 4.8**).

A



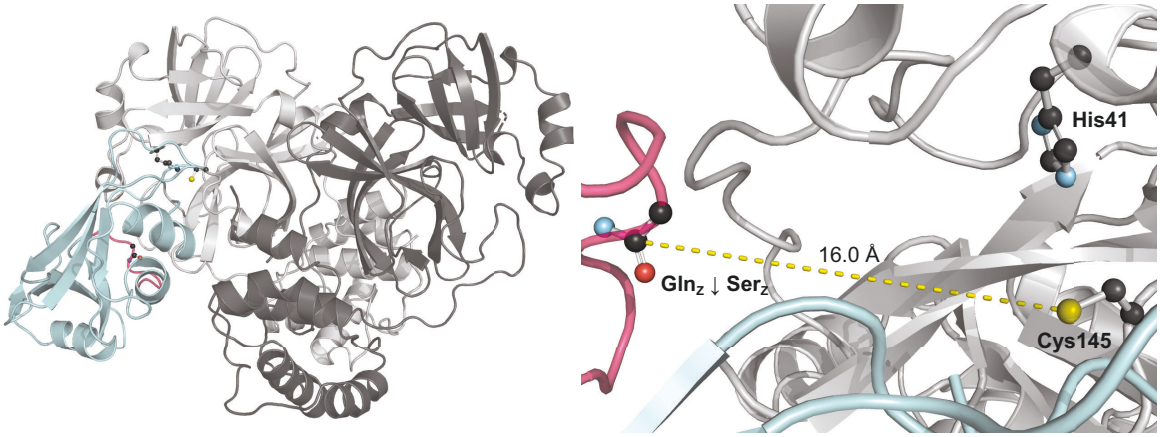
Model 3 of 30
59 members; center score: −706.4; lowest score: −808.7

B



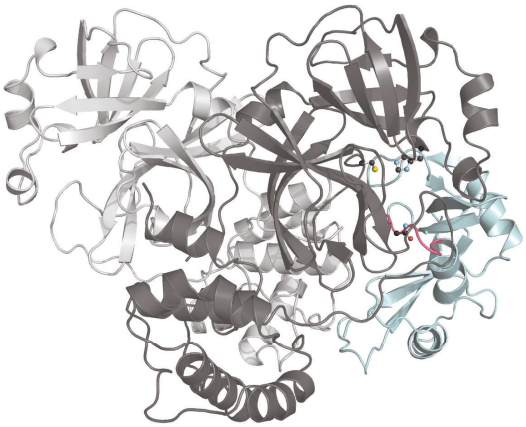
Model 10 of 30
27 members; center score: −664.2; lowest score: −821.0

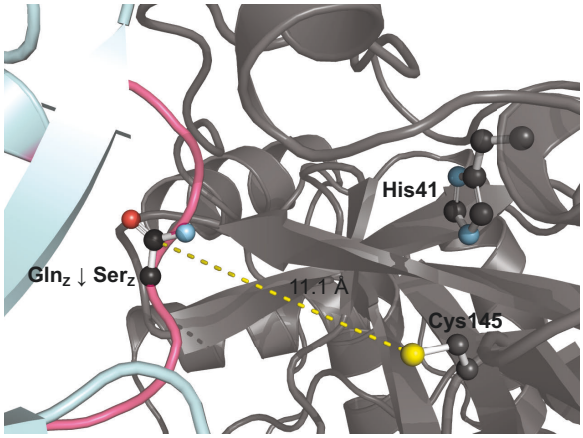
C



Model 12 of 30
25 members; center score: −699.0; lowest score: −755.7

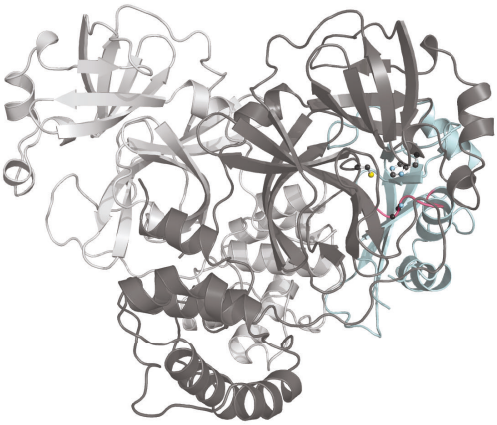
D

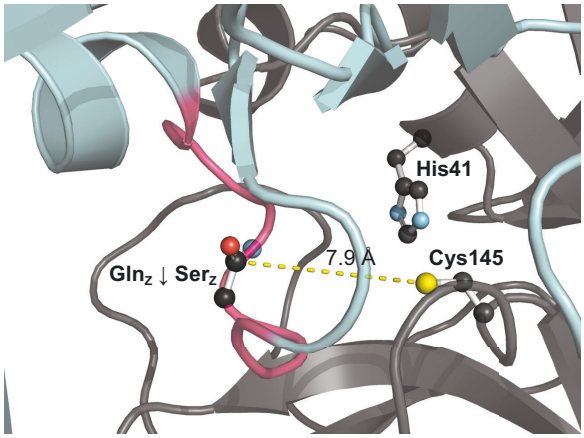




Model 20 of 30
15 members; center score: −680.2; lowest score: −718.7

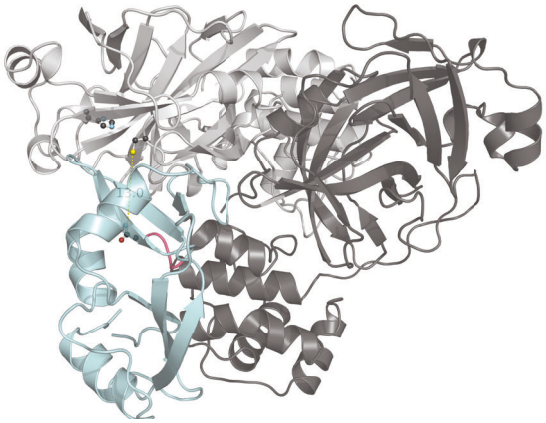
E

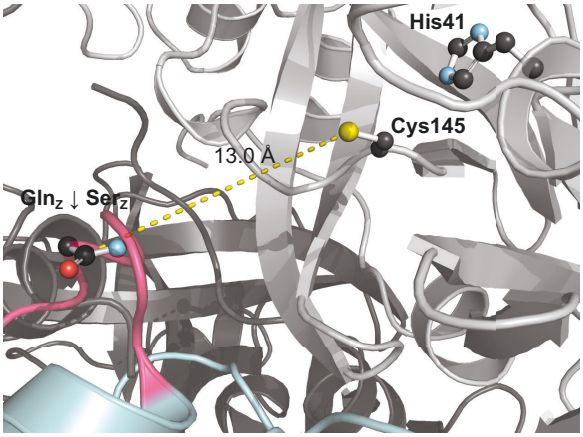




Model 22 of 30
12 members; center score: −681.5; lowest score: −773.9

F





Model 30 of 30
6 members; center score: −665.3; lowest score: −700.4

Figure 4.8. Images of ClusPro-predicted zymogen–3CL^{pro} models (A–F) in which the zymogen linker is proximal to the protease active site. Each subpanel shows: (left) the global interaction of the 3CL^{pro} protomers (cartoon, distinct shades of gray) and the zymogen (cartoon, teal with linker in magenta) with atoms relevant to linker cleavage in ball-and-stick form (CPK color scheme); (right) the location of the scissile amide bond of the zymogen linker ($Gln_Z \downarrow Ser_Z$) in relation to the 3CL^{pro} catalytic residue Cys145.

4.3.2 Characterizing the nature of zymogen cleavage by 3CL^{pro}

We used mass spectrometry to observe zymogen cleavage by 3CL^{pro} because of the characteristic +18 Da shift upon amide bond hydrolysis. The waterfall plot in **Figure 4.9** shows that over 50% of the zymogen is cleaved within 5 minutes and that cleavage is effectively complete within 1 hour.

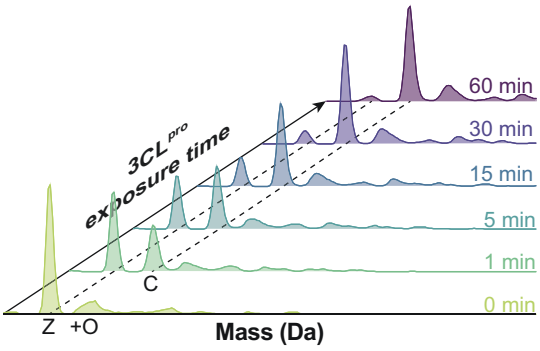


Figure 4.9. Waterfall plot of deconvoluted mass spectrometry data for zymogen (*Z*) conversion to cleaved zymogen (*C*) by $3CL^{pro}$; peak area is proportional to percent of total species. +O indicates a +16 Da peak frequently seen in RNase samples due to methionine oxidation.

Heat denaturation experiments revealed that the zymogen is considerably more thermostable upon linker cleavage (ΔT_m = +21.0 ± 0.1 °C) (**Figure 4.10**), as expected from the release of the conformational strain incurred by cyclization.[97]

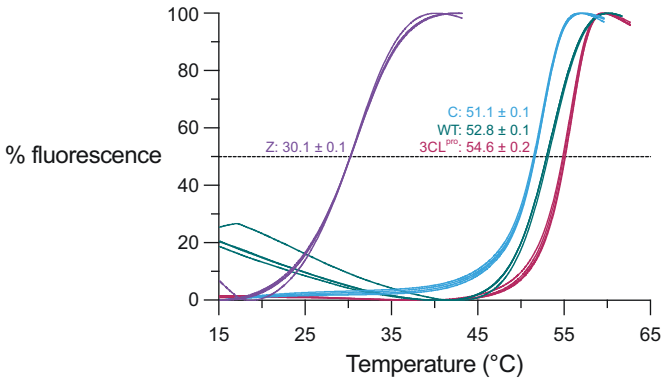


Figure 4.10. Thermal denaturation curves for the zymogen (*Z*) and cleaved zymogen (*C*), with 3CL^{pro} and wild-type RNase 1 (*WT*) shown for comparison. Thermal denaturation temperatures, T_m, are reported as mean ± SE (n = 4).

The zymogen also appears resistant to off-target proteolysis by $3CL^{pro}$; even after incubating the zymogen with $3CL^{pro}$ for 3 hours, proteolytic degradation products are not visible by SDS–PAGE (**Figure 4.11**), and mass spectrometry reveals only the +18 Da adduct corresponding to one proteolytic event.

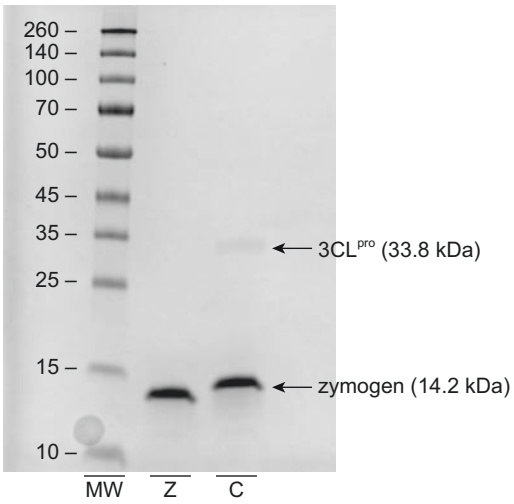


Figure 4.11. SDS–PAGE of the zymogen before (*Z*) and after (*C*) $3CL^{pro}$ treatment for 3 hours. *MW*, molecular weight standards (masses in kDa).

Using the tetranucleotide 6-FAM–dArU$(dA)_2$–6-TAMRA as a fluorogenic ribonuclease substrate, we determined that the uncleaved zymogen has a catalytic efficiency of $(9.0 \pm 0.1) \times 10^3$ M^{-1} s^{-1}, which is 0.04% that of wild-type RNase 1. Meanwhile, the fully cleaved zymogen has a k_{cat}/K_M value of $(2.9 \pm 0.1) \times 10^6$ M^{-1} s^{-1}; thus, the release of cyclic strain restores activity to 14% of that of the wild-type enzyme and leads to a 321-fold increase in activity relative to the cyclized protein.

4.3.3 Detecting 3CL^{pro} by the signal amplification of zymogen activation

After demonstrating that the zymogen is specifically cleaved in its linker, we characterized whether the zymogen can be used to detect 3CL^{pro} and distinguish samples harboring low levels of protease from samples that are protease-free. As expected, ribonucleolytic activity is not observed in the 15-minute observation period at any of the employed zymogen concentrations if 3CL^{pro} is not present (**Figure 4.12**).

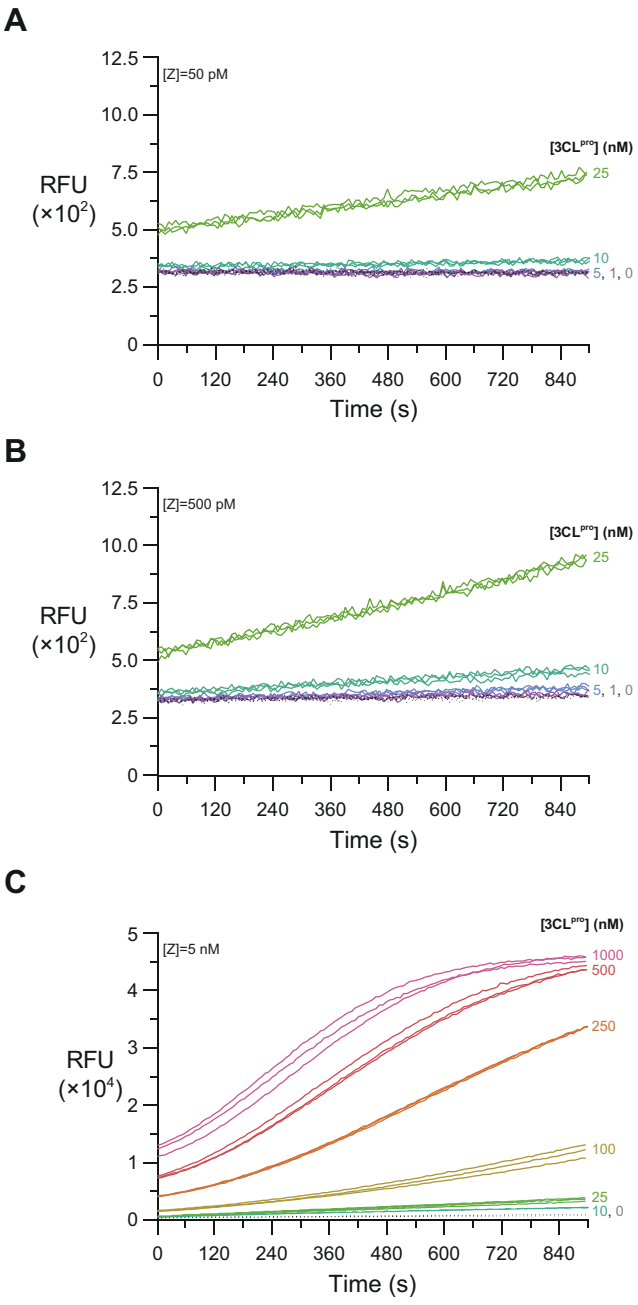
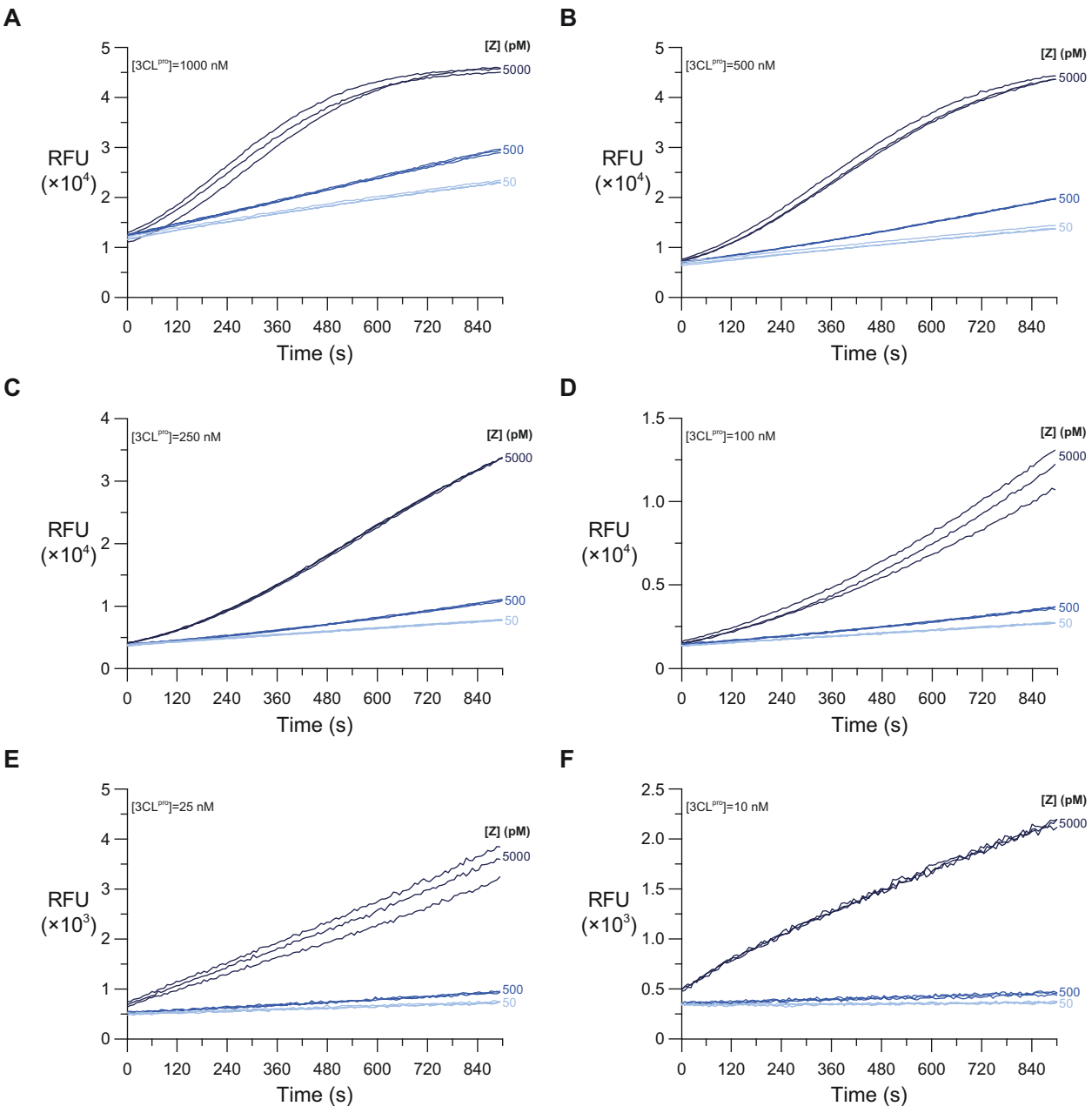


Figure 4.12. Detection of 3CL^{pro} by zymogen activation with the coupled enzyme activity assay. Fluorescence intensity is shown for turnover of the RNA substrate upon 3CL^{pro}-mediated cleavage of (A) 50 pM, (B) 500 pM, and (C) 5 nM zymogen. 3CL^{pro} concentration (n = 3 replicates) is indicated by line color. Progress curves are also shown as a function of 3CL^{pro} concentration in **Figure 4.13**.

With as little as 50 pM zymogen, we can detect the presence of ≥10 nM $3CL^{pro}$ by its activation of the zymogen and the subsequent turnover of our ribonuclease substrate (**Figure 4.12A**). Increasing the zymogen concentration to 500 pM increases the detection sensitivity to ≥5 nM $3CL^{pro}$ (**Figure 4.12B**). The twofold increase in $3CL^{pro}$ detection sensitivity with a tenfold increase in zymogen concentration suggests that the inherent catalytic activity of $3CL^{pro}$, rather than the rate of substrate turnover by activated zymogen, limits detection. A further increase in zymogen concentration to 5 nM enables expeditious detection of $3CL^{pro}$; at protease concentrations ≥250 nM, over half of the ribonuclease substrate was turned over within the observation period, and the same benchmark was reached within 5 minutes at 1000 nM $3CL^{pro}$ (**Figure 4.12C**). All timecourse data are shown in **Figure 4.13**.


A
[3CL^{pro}]=1000 nM
[Z] (pM)
5000
500
50
RFU
(×10^4)
Time (s)
B
[3CL^{pro}]=500 nM
[Z] (pM)
5000
500
50
RFU
(×10^4)
Time (s)
C
[3CL^{pro}]=250 nM
[Z] (pM)
5000
500
50
RFU
(×10^4)
Time (s)
D
[3CL^{pro}]=100 nM
[Z] (pM)
5000
500
50
RFU
(×10^4)
Time (s)
E
[3CL^{pro}]=25 nM
[Z] (pM)
5000
500
50
RFU
(×10^3)
Time (s)
F
[3CL^{pro}]=10 nM
[Z] (pM)
5000
500
50
RFU
(×10^3)
Time (s)

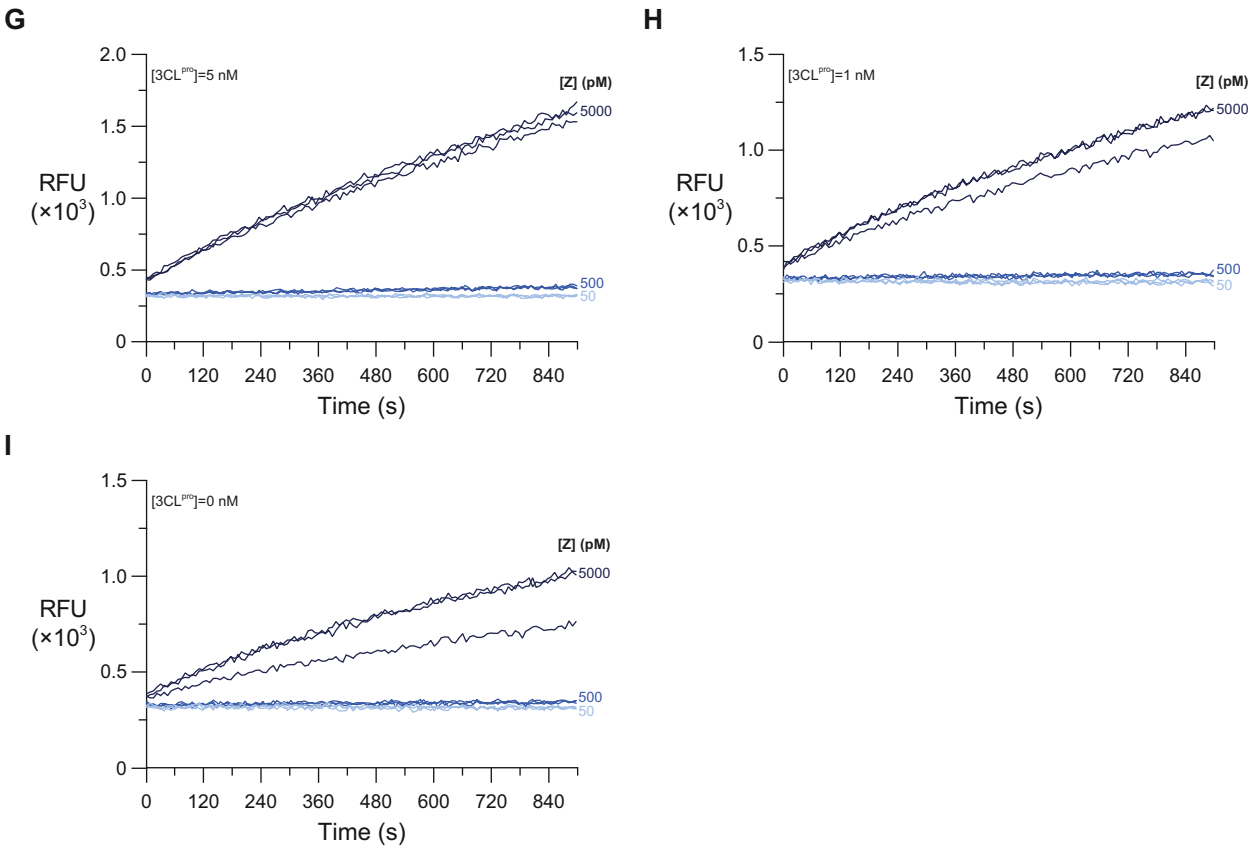


Figure 4.13. Graphs of the real-time detection of $3CL^{pro}$ via zymogen activation and subsequent ribonucleolysis. Each panel depicts zymogen activity at one $3CL^{pro}$ concentration: (A) 1000 nM; (B) 500 nM, (C) 250 nM, (D) 100 nM, (E) 25 nM, (F) 10 nM, (G) 5 nM, (H) 1 nM, and (I) 0 nM. All reactions were conducted with 200 nM substrate in 100 mM Tris-HCl buffer, pH 7.5, containing NaCl (10 mM) at 25 °C. Progress curves shown in **Figure 4.12** are from these data.

Using the timecourse data for 500 pM and 5 nM zymogen at [$3CL^{pro}$] ≥ 100 nM, we modeled the rate of substrate turnover in our coupled reaction detection system **(Figure 4.14**).

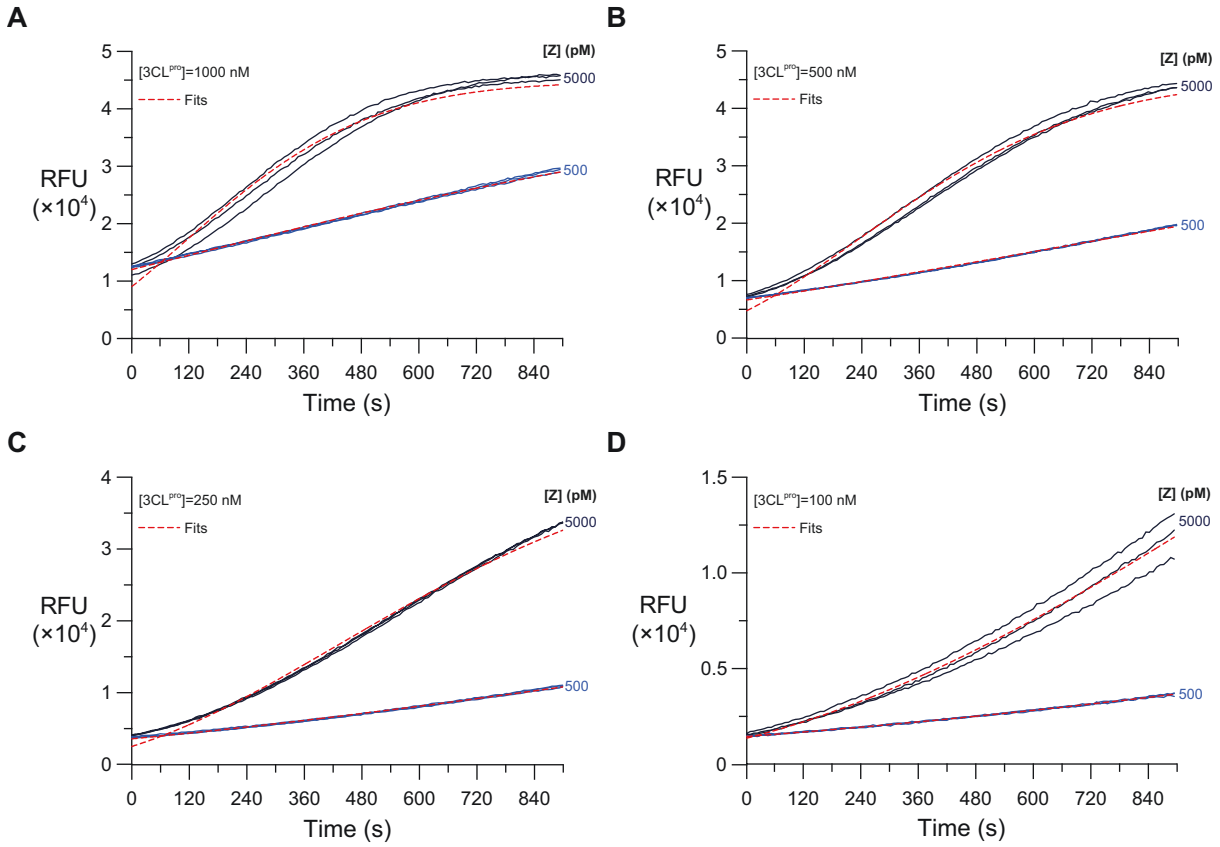


Figure 4.14. Graphs of the fitting of data from **Figure 4.13** for one [3CL^{pro}]: (A) 1000 nM; (B) 500 nM, (C) 250 nM, and (D) 100 nM. Catalytic efficiencies were determined by nonlinear regression of each timecourse dataset to Equation C.14; each fit is indicated by a red dashed line.

We determined that the k_{cat}/K_M value for the cleavage of the zymogen linker by $3CL^{pro}$ is $(1.2 \pm 0.1) \times 10^3$ M^{-1} s^{-1} (mean ± SE) (see Appendix C for a complete derivation of the equations governing the reaction kinetics). For comparison, we recently reported that the same protease substrate sequence installed in the fluorogenic dodecapeptide R–E(EDANS)–ATLQSGNA–K(DABCYL)–R is cleaved by $3CL^{pro}$ with a k_{cat}/K_M value of $(4.3 \pm 0.7) \times 10^4$ M^{-1} s^{-1}.[180]

4.4 Discussion

Our zymogen design follows a strategy that we used previously to generate ribonuclease zymogens activated by other viral proteases: terminal truncation of RNase 1 in conjunction with protein cyclization with a peptidic linker to induce conformational strain and steric occlusion.[97] Our computational modeling corroborates the efficacy of this approach to diminish ribonucleolytic activity. As shown in **Figure 4.5** and **4.6**, AlphaFold2 predicts that a key active site residue, His119, is significantly dislocated upon cyclization, which should deter catalysis of RNA cleavage (**Figure 4.4**). The linker also introduces significant steric bulk in the vicinity of the active site; **Figure 4.7** suggests that even a substrate with just two nucleotides on the 3′ side of the scissile phosphodiester bond clashes with the linker. Thus, conformational distortion and steric occlusion appear to work in tandem to achieve the >300-fold inactivation of ribonucleolytic activity and result in a k_{cat}/K_M value that is comparable to that of other RNase zymogens.[94,95,97] Given that the thermostability of the zymogen returns nearly to that of wild-type RNase 1 upon cleavage (**Figure 4.10**), we hypothesize that structural relaxation is responsible for the ability of the cleaved zymogen to achieve nearly wild-type ribonucleolytic activity and is a testament to the remarkable conformational stability of ribonucleases.[226,229] The newly unconstrained residues of the cleaved linker could occlude the active site and lead to a ribonucleolytic activity that is less than that of wild-type RNase 1.

The activation of the zymogen is both expeditious and specific to the linker (**Figure 4.9**, **4.11**), as desired for our context. $3CL^{pro}$ cleaves a peptide bond following a glutamine residue.[206] Although seven endogenous glutamine residues exist in our zymogen as possible off-target $3CL^{pro}$ proteolysis sites, each is in a context that is incompatible with the substrate specificity of $3CL^{pro}$ (see **Table 4.1**).

Table 4.1. Location of potential off-target 3CL^{pro} cleavage sites in the zymogen and identified 3CL^{pro} substrate incompatibilities.

Gln[a]	Incompatibility with 3CL^{pro} substrate specificity[b]
9	P4=Lys6 (*large*), P1′=Arg10 (*bulky, cationic*)
11	P4=Phe8 (*bulky*), P2=Arg10 (*cationic*), P1′=His12 (*bulky*)
28	P4=Tyr25 (*bulky*), P1′=Met29 (*bulky*)
37	P2=Thr36 (*hydrophilic*)
55	P1′=Asn56 (*bulky*)
60	P1′=Glu61 (*anionic*)
69	P4=Lys66 (*bulky*)

[a]Using native RNase 1 residue numbering, the zymogen sequence is:

```
5   10          20          30          40          50          60          70
AKKFQR QHMDSDSSPS SSSTYCNQMM RRRNMTQGRC KPVNTFVHEP LVDVQNVCFQ EKVTCKNGQG
        80          90         100         110         120 122 ********
NCYKSNSSMH ITDCRLTNGS RYPNCAYRTS PKERHIIVAC EGSPYVPVHF  DA ATLQSGNA
```

The linker is in red font and denoted by *.

[b]Using Schechter-Berger nomenclature, 3CL^{pro} strictly requires glutamine at P1 and preferentially binds hydrophobic residues at P2; small residues at P4; and small, uncharged residues at P1′.[203-206]

We believe that the high degree of topological complementarity between the zymogen and $3CL^{pro}$ (see Digital Supplementary Material file **wralstad-evansw-phd-chem-2024-supplemental_1.mp4**), along with the relative flexibility of the zymogen linker and its ability to weave into the active site of $3CL^{pro}$ (**Figure 4.8**), explain the appreciable rate of $3CL^{pro}$ cleavage of the zymogen linker (which is ~3% that when cleaving the same eight-residue sequence in an untethered peptide).

Our coupled system (**Figure 4.1**) enables detection of even low concentrations of $3CL^{pro}$ within a timeframe comparable to RADT (**Figure 4.12**). It is estimated that an infected individual harbors up to 10^9–10^{11} SARS-CoV-2 virion particles in total and that patient samples contain $>10^6$ viral copies per mL, even during early infection.[230-232] Though the number of $3CL^{pro}$ copies produced per virion has not been reported, we anticipate that the importance of $3CL^{pro}$ to the SARS-CoV-2 replisome complex means that infected patient samples harbor zymogen-detectable concentrations of $3CL^{pro}$. In support of this hypothesis, immunofluorescence and immunoblotting assays have detected $3CL^{pro}$ within infected cells as soon as six hours post-infection.[233,234] Although nanomolar concentrations of $3CL^{pro}$ can be detected directly with an assay for proteolytic activity (*e.g.*, with a fluorogenic peptide substrate), we recently reported that detection of 25 nM $3CL^{pro}$ requires high micromolar concentrations of peptide substrate and assay times >30 minutes.[180] By contrast, all reagents in our zymogen assay are nanomolar or below, and we easily detect $3CL^{pro}$ within 15 minutes.

Critically, because the linker sequence in our zymogen is similar to endogenous $3CL^{pro}$ polyprotein recognition sequences, we believe that our detection strategy is resistant to viral mutation-based sensitivity loss. In addition, $3CL^{pro}$ is not under the same magnitude of selective pressure as other diagnostic targets, such as the spike and nucleocapsid proteins. These attributes

are intrinsic advantages of a function-based diagnostic strategy and should mitigate the retooling that has been required of other SARS-CoV-2 diagnostics upon emergence of variants.[235,236]

4.5 Conclusions

We have demonstrated that an engineered ribonuclease zymogen can be used for the sensitive detection of a viral protease. Our computational analyses and biophysical experiments support the hypothesis that conformational strain impairs ribonucleolytic activity in a manner that is reversible upon linker cleavage. The signal amplification that results from this coupled enzyme assay enables us to detect even low nanomolar concentrations of $3CL^{pro}$ in a timeframe comparable to rapid antigen detection testing. Future efforts will focus on adapting this proof-of-concept to a point-of-care or at-home diagnostic tool, including immobilizing the zymogen for lateral flow assays and developing alternative ribonuclease substrates that would enable facile detection of zymogen activation without fluorimetry.

4.6 Acknowledgments

E.C.W. was supported by a Graduate Research Fellowship from the NSF. This work was supported by Grants R01 CA073808 and R21 AI171663 (NIH).

4.7 Experimental Details

4.7.1 Modeling the zymogen and its interactions

AlphaFold2 does not accept circular proteins as an input. Accordingly, the sequence of the circular zymogen was cleaved in four locations (**Figure 4.3**), and each of these sequences was submitted to AlphaFold2 for structural modeling.[225] The suitability of each permutant model was evaluated based on its plDDT scores and structural backbone alignment to the progenitor RNase 1 structure (PDB ID: 1Z7XZ).[174]

To model the steric clash between the zymogen and an RNA substrate, PyMOL was used to structurally align the zymogen models and the empirical RNase 1 structure to the structure of bovine RNase A that had been co-crystallized with the substrate analog d(ApTpApApG) (PDB ID: 1RCN).[227]

The ClusPro protein-protein docking server was used to predict the interactions of the Q37|R38 permutant with SAS-CoV $3CL^{pro}$ (PDB ID: 1UJ1).[135,228,237-239]

4.7.2 Preparing the zymogen and protease

The $3CL^{pro}$-directed ribonuclease zymogen was prepared by methods similar to those described previously.[97] Briefly, Gibson assembly was used to construct an expression plasmid that combined five gene fragments 5′ to 3′ along the sense strand: (1) the *Npu*C fragment of the split intein from *Nostoc punctiforme* PCC73102; (2) residues 58–122 of mature *Homo sapiens* RNase 1; (3) residues ATLQSGNA to serve as a cyclizing linker for RNase 1 with recognition by $3CL^{pro}$; (4) residues 5–57 of RNase 1; and (5) the *Npu*N fragment of the split intein.[115,120]

Electrocompetent BL21(DE3) *E. coli* were transformed with the assembled plasmid and positive selection using 200 μg mL^{-1} ampicillin. Cultures were grown to OD_{600} = 1.80 in Terrific

Broth at 37 °C with shaking at 250 RPM, and zymogen expression was induced by the addition of isopropyl ß-D-1-thiogalactopyranoside (to 1 mM) and growth for 4 hours. Cultures were pelleted by centrifugation, and the pelleted cells were stored at –70 °C.

Pelleted *E. coli* cells were thawed and resuspended in 20 mM Tris-HCl buffer, pH 7.6, containing EDTA (10 mM) at room temperature. The cells were lysed by sonication with fifteen rounds of 10-s sonication at 30% amplitude, followed by ≥ 60 s on ice. The lysate was subjected to centrifugation, and inclusion bodies were denatured in 20 mM Tris-HCl buffer, pH 8.0, containing guanidinium-HCl (7 M), EDTA (10 mM), and DTT (100 mM) at room temperature. The solubilized protein was diluted tenfold with 20 mM acetic acid, and the resulting suspension was clarified by centrifugation. The supernatant was dialyzed against 20 mM acetic acid overnight at 4 °C to further clarify. Following centrifugation, the supernatant was dropwise added to a refolding buffer of 100 mM Tris-HCl, pH 7.8, containing EDTA (10 mM), L-arginine (0.5 M), reduced glutathione (1.0 mM), and oxidized glutathione (0.2 mM). The zymogen was allowed to refold at 4 °C without agitation for 5 days. The refolded zymogen was purified by gel-filtration and cation-exchange fast protein liquid chromatography. All chromatography buffers were treated with DEPC prior to use, except for Tris, which was added from ribonuclease-free stocks. Purified protein was buffer-exchanged into 50 mM sodium acetate buffer, pH 5.0, flash-frozen in liquid nitrogen, quantified by BCA assay, and stored at –70 °C. Protein purity was verified by SDS–PAGE, and protein identity was confirmed by mass spectrometry (*vide infra*).

Authentic SARS-CoV-2 $3CL^{pro}$ was prepared as we recently reported.[180]

4.7.3 Monitoring the timecourse of zymogen cleavage by 3CL^{pro}

Zymogen (25 μM) and 3CL^{pro} (1 μM) were mixed in 50 mM HEPES-NaOH buffer, pH 7.5, and the resulting solution was incubated at 20 °C without agitation. At each timepoint, an aliquot of the reaction mixture was acidified to ~0.1% v/v formic acid to halt proteolysis, as 3CL^{pro} has negligible activity at pH <5.[193,199] High-resolution mass spectrometry of an 8 pmol sample was performed using ESI mass spectrometry on an Agilent 6530C Accurate-Mass Q–TOF mass spectrometer equipped with a PLRP-S column (Agilent Technologies, 2.1 × 50 mm, 5-μm particle, 1000-Å pore). A gradient of 5–95% v/v acetonitrile (0.1% v/v formic acid) in water (0.1% v/v formic acid) over 7 min was used for all samples. Before Q–TOF LC-MS analysis, all samples were passed through a Spin-X Centrifugal Tube Filter (0.22-μm, cellulose acetate membrane) from R&D Systems (Minneapolis, MN).

4.7.4 Determining the zymogen's stability upon cleavage by 3CL^{pro}

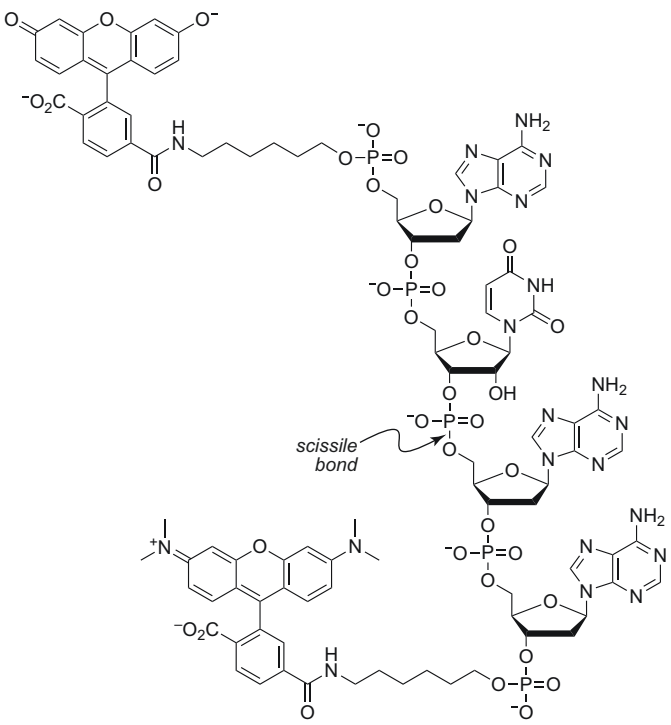
Differential scanning fluorimetry (DSF) was used to determine the thermostability of the zymogen before and after linker cleavage. Zymogen was cleaved as in the timecourse monitoring (*vide supra*) apart from using higher protein concentrations (94.1 μM zymogen, 3.76 μM 3CL^{pro}) to achieve final concentrations amenable to DSF. After 1 hour, SYPRO Orange protein gel stain (Supelco) was added to achieve 1 mg mL^{-1} protein and 50× SYPRO Orange (vendor stock: 5000× in DMSO) in 50 mM HEPES-NaOH buffer, pH 7.5. Aliquots of uncleaved zymogen, wild-type RNase 1, and 3CL^{pro} were similarly prepared with SYPRO Orange. The resulting samples were heated from 15–95 °C at 1 °C/min using a QuantStudio 7 Flex Real-Time PCR system (Applied Biosystems); fluorescence was monitored in real-time with λ_{ex} = 470 ± 15 nm and λ_{em} = 586 ± 10 nm. Data were processed with Protein Thermal Shift software (Applied

Biosystems) using the Boltzmann fitting method. Values of T_m represent the temperature at which fluorescence reached 50% of its maximum.

To characterize the potential for off-target proteolytic degradation, the zymogen was incubated with $3CL^{pro}$ in 50 mM HEPES-NaOH buffer, pH 7.5, for 3 hours (94.1 μM zymogen, 3.76 μM $3CL^{pro}$). Aliquots of cleaved and uncleaved zymogen were diluted to 0.4 mg mL^{-1} with 50 mM HEPES-NaOH buffer, pH 7.5, mixed 1:1 with 2× Laemmli sample buffer (Bio-Rad) containing 2-mercaptoethanol (Sigma), incubated at 98 °C for 5 min, cooled, and subjected to electrophoresis (2 μg load) at 4 °C on a precast polyacrylamide/Tris-glycine SDS–PAGE gel (Bio-Rad) alongside a molecular weight standard (Thermo Scientific). The gel was stained with Bio-Safe Coomassie G-250 (Bio-Rad), destained in ultrapure water, and imaged with an Amersham 600 gel imager (Cytiva). Image editing in ImageJ consisted only of cropping and brightness/contrast adjustment to maximize band sensitivity.

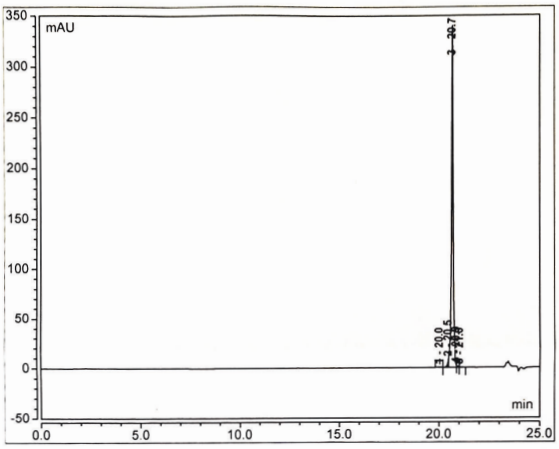
4.7.5 Characterizing the ribonucleolytic activity of the zymogen

The tetranucleotide ribonuclease substrate 6-FAM–dA<u>r</u>U$(dA)_2$–6-TAMRA (**Figure 4.15**) was obtained from Microsynth AG and had a mass of 2307.5 Da (expected, 2312.7 Da) according to MALDI–TOF mass spectrometry.



Oligo-ID 04778720 **Microsynth**

Order-ID:	821713	*Injection Volume:*	**5 µl**
Chromatography:	**Reverse Phase HPLC**	*Wavelength:*	**260 nm**
Buffer system:	**TEAA; 50°C**	*Recording Time:*	**03.04.2023 15:37**

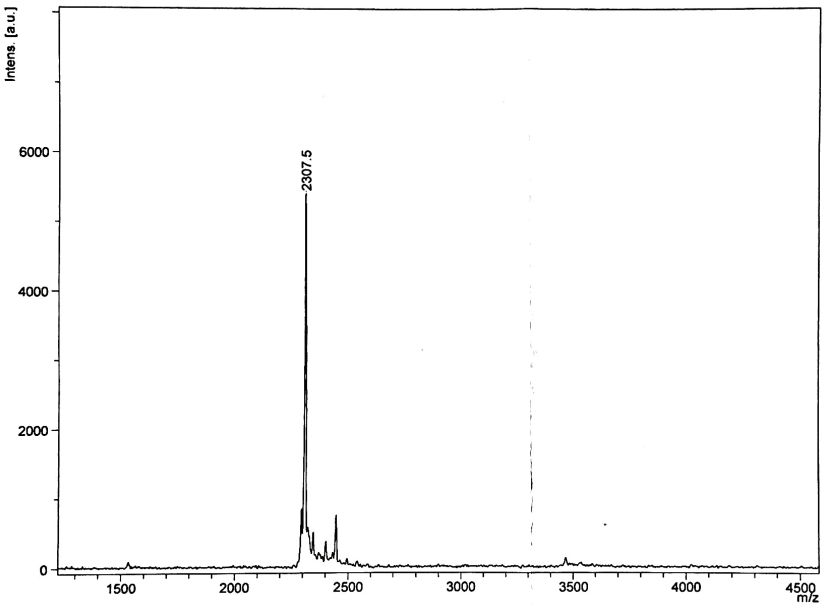


No.	Ret.Time min	Peak Number	Peak Height mAU	Peak Area mAU*min	Relative Area %
1	20.03	*1*	0.9	0.1	0.28
2	20.47	*2*	8.3	0.4	1.17
3	20.68	*3*	332.7	31.9	97.90
4	20.87	*4*	1.8	0.1	0.41
5	21.04	*5*	0.6	0.1	0.24
Total:			344.3	32.6	100.0

Oligo ID # 4778720

Theoretical Mass [g/mol]: 2312.66

Acceptance Criteria: Mw ± 0.5%



Microsynth

Figure 4.15. Structure of 6-FAM–dArU(dA)$_2$–6-TAMRA, which is the fluorogenic substrate for assays of ribonucleolytic activity and has only one scissile phosphodiester bond. The cleavage of that bond increases the fluorescence of the FAM moiety with λ_{ex} = 493 ± 5 nm and λ_{em} = 515 ± 5 nm. Insets: HPLC chromatogram (top) and MALDI–TOF mass spectrum (bottom) of synthetic 6-FAM–dArU(dA)$_2$–6-TAMRA.

Catalysis of RNA hydrolysis was assessed by mixing the uncleaved zymogen (12.5 nM) with 6-FAM–dArU(dA)$_2$–6-TAMRA (200 nM) in ribonuclease-free 100 mM Tris-HCl buffer, pH 7.5, containing NaCl (10 mM). Ultrapure water was DEPC-treated before the addition of Tris, due to the reactivity of its primary amino group with DEPC. Substrate fluorescence was measured at λ_{ex} = 493 ± 5 nm and λ_{em} = 515 ± 5 nm with a Spark multimode plate reader (Tecan) before the addition of zymogen to determine the background fluorescence, I_0. After adding zymogen, the reaction mixture was incubated at 25 °C and continuously monitored for 4 minutes. Wild-type RNase 1 was then added to a concentration of ~25 μM to cleave all of the substrate and determine I_{max}. The value of k_{cat}/K_M was calculated with the equation:

$$\frac{k_{\mathrm{cat}}}{K_{\mathrm{M}}} = \frac{{}^{\Delta I}/_{\Delta t}}{(I_{\mathrm{max}} - I_0)[\mathrm{E}]} \tag{4.1}$$

where $\Delta I/\Delta t$ is the slope of the fluorescence versus time data collected during the 4-minute observation period, and [E] is the concentration of enzyme.

Catalysis of RNA cleavage by the activated zymogen was determined by first treating the zymogen with 3CL^{pro} as in the DSF assay (*vide supra*). The ribonucleolytic activity of the cleaved zymogen was then assayed analogously to the uncleaved zymogen, except for the use of only 12.5 pM of the activated zymogen.

Real-time detection of 3CL^{pro} by zymogen cleavage and subsequent turn-on of ribonucleolytic activity was accomplished by mixing variable zymogen concentrations (5000, 500, or 50 pM) with variable 3CL^{pro} concentrations (1000, 500, 250, 100, 25, 10, 5, 1, or 0 nM) and 200 nM substrate as before, now with continuous monitoring for 15 minutes. All other reaction conditions were consistent with the assay of the intact and activated zymogen. To determine the catalytic efficiency of zymogen linker cleavage by 3CL^{pro}, timecourse data were

fitted by nonlinear regression to Equation C.14. Nonlinear regression was restricted to datasets from 500 and 5000 pM zymogen with [$3CL^{pro}$] ≥ 100 nM, where the curvature of the timecourse data was sufficient to allow for convergence of the nonlinear regression algorithm.

Appendix A:

Semisynthesis of Human Ribonuclease–S

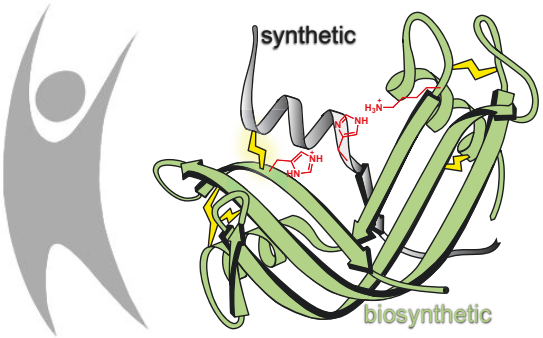
Reprinted with permission from ref. 240.

Sayers, J.; Wralstad, E. C.; Raines, R. T. Semisynthesis of human ribonuclease-S. *Bioconjugate Chem.* **2021**, *32*, 82-87.

Contributions: Conceptualization, J.S.; experimental methodology, J.S. and E.C.W.; experimental investigation, J.S. and E.C.W.; writing—original draft preparation, J.S. and E.C.W.; writing—review and editing, R.T.R.; supervision, R.T.R.; funding acquisition, J.S., E.C.W., and R.T.R.

A.1 Abstract

Since its conception, the ribonuclease S complex (RNase S) has led to historic discoveries in protein chemistry, enzymology, and related fields. Derived by the proteolytic cleavage of a single peptide bond in bovine pancreatic ribonuclease (RNase A), RNase S serves as a convenient and reliable model system for incorporating unlimited functionality into an enzyme. Applications of the RNase S system in biomedicine and biotechnology have, however, been hindered by two shortcomings: (1) the bovine-derived enzyme could elicit an immune response in humans, and (2) the complex is susceptible to dissociation. Here, we have addressed both limitations in the first semisynthesis of an RNase S conjugate derived from human pancreatic ribonuclease and stabilized by a covalent interfragment cross-link. We anticipate that this strategy will enable unprecedented applications of the "RNase–S" system.



A.2 Introduction

Semisynthesis allows access to precisely modified variants of otherwise native enzymes and other proteins. By combining recombinant DNA expression and organic synthesis, noncanonical residues can be incorporated for the precise interrogation and manipulation of protein structure and function.[241] The first enzyme to be prepared by merging biosynthesis and chemical synthesis was ribonuclease S (RNase S). RNase S is a noncovalent complex that forms upon the association of two fragments of bovine pancreatic ribonuclease (RNase A; UniProtKB Q9BEC3) that are generated upon proteolysis with subtilisin. Alone, the S-peptide (1–20) and S-protein (21–124) fragments are catalytically inactive, but they associate spontaneously to reconstitute the enzymic active site and restore ribonucleolytic activity (**Figure A.1**).[29,103,242,243]

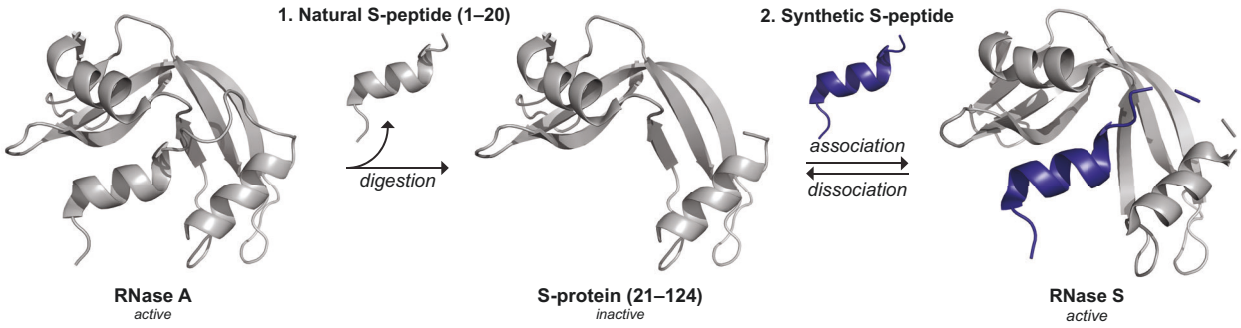


Figure A.1. Traditional production of semisynthetic RNase S. Step 1: Proteolytic digestion of RNase A to cleave S-peptide (residues 1–20) and yield catalytically inactive S-protein (residues 21–124). Step 2: Reversible association of S-protein and a synthetic S-peptide variant to produce catalytically active, semisynthetic RNase S. Images were produced with PyMOL software and PDB entries 1JVT[244] and 1RNU.[104]

A truncated variant of S-peptide (residues 1–15) also forms an active complex.[245] By preparation of synthetic variants of S-peptide, a wide array of chemical modifications have been incorporated into RNase S. The ensuing data have enhanced our understanding of protein structure, protein folding, protein–protein interactions, and enzymology.[29,242,243,246] S-peptide can even serve as an affinity tag for the purification and detection of recombinant proteins.[108-110,247] In recent studies, semisynthetic RNase S variants have been used to stabilize α-helices via peptide stapling,[248,249] control peptide–protein binding with photoswitchable peptides,[250] and validate L-tellurienylalanine as a phenylalanine isostere.[251]

RNase 1 (UniProtKB P07998), which is the human homologue of RNase A, has demonstrable utility in biomedical contexts. In particular, an engineered variant of RNase 1 exhibits specific toxicity for cancer cells and is in clinical trials as a chemotherapeutic agent.[252,253] RNase 1 also has considerable antiviral activity,[44] and manifests antibacterial activity in the presence of a pore-forming peptide.[254] To date, however, RNase S has not been employed in the clinic. Indeed, such applications are discouraged by the potential immunogenicity of the bovine-derived complex in humans.[255] Hence, we reasoned that an RNase S of human origin could have an expanded scope.

Traditionally, S-protein is generated by treating intact, folded RNase A with subtilisin, a bacterial serine protease.[2] Proteolytic cleavage occurs within a surface loop bordering the active site (residues 15–25), primarily between residues Ala20 and Ser21. Because the bovine and human pancreatic ribonucleases have 68% sequence identity, we inferred that the same strategy could be employed to generate an S-protein fragment from RNase 1. To the contrary, however, treating RNase 1 with subtilisin under similar conditions does not result in the formation of the corresponding S-protein and S-peptide fragments.[256] This resistance arises because the key

surface loop (residues 15–25) of RNase 1 is sterically incompatible with the subtilisin active site due to a conformational change induced by Pro19.[256,257] Moreover, S-protein does not fold properly in the absence of S-peptide,[258] and thus cannot be produced simply by heterologous expression.

There have been unsuccessful attempts to express the native RNase 1 21–127 fragment in *Escherichia coli*. Some success has been achieved through the digestion of an engineered chimeric bovine/human ribonuclease.[255,259] As an alternative strategy, we imagined that enterokinase, which is a highly specific protease, could be employed to produce the S-protein fragment directly from recombinant RNase 1. Specifically, we envisaged that the enterokinase recognition sequence (DDDDK↓) could be installed via site-directed mutagenesis at the boundary of the S-peptide and S-protein fragments.[20] Proteolytic cleavage, which occurs at the C-terminus of the recognition site, would release native S-protein along with S-peptide fused to the enterokinase recognition sequence, DDDDK.

In addition to circumventing issues associated with immunogenicity, enhancing the stability of the RNase S complex is crucial to enabling therapeutic applications. Despite being highly specific, the noncovalent interactions between the S-protein and S-peptide fragments render the complex prone to dissociation (K_d = 49 nM), particularly at low concentrations.[260] Additionally, RNase S has greater conformational entropy than does intact RNase A, resulting in decreased thermodynamic stability and higher susceptibility toward denaturation and proteolysis.[260,261] Using a strategy employed to enhance the conformational stability of intact ribonucleases,[262,263] we previously described bovine S-protein and S-peptide variants that were able to form a disulfide cross-link (Cys^4–Cys^{118}) at the fragment interface.[20] By restricting conformational flexibility, this additional disulfide bond was found to improve thermodynamic stability and

increase resistance to proteolysis. Crucially, the disulfide cross-link was not highly detrimental to the ribonucleolytic activity of the enzyme. Given the similar three-dimensional structures of RNase A and RNase 1, we hypothesized that a cross-linking strategy could also stabilize an RNase S complex derived from RNase 1. Toward this end, we report herein a novel semisynthetic strategy for the preparation of human RNase–S, where the "–" highlights the covalent cross-link between the two fragments.

A.3 Results & Discussion

We designed the requisite variant of RNase 1 with an enterokinase cleavage site inserted between residues Ser20 and Ser21 (**Figure A.2**).

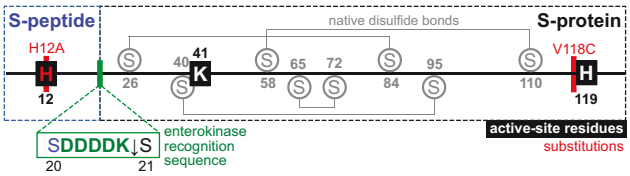


Figure A.2. Map of the amino acid sequence of the RNase 1 variant. The locations of the enzymic active site residues, native disulfide bonds, substitutions, and the inserted enterokinase recognition sequence are shown.

As a precaution against contaminant activity arising from undigested enzyme, we replaced His12, which is a key active site residue within S-peptide, with alanine. This substitution reduces catalytic activity by $>10^4$-fold without causing structural perturbations.[40,264-266] We introduced an additional substitution, Val118Cys, to serve as a conjugation handle for disulfide cross-linking with synthetic S-peptide variants. The ensuing RNase 1 variant was produced and purified as described previously (**Figure A.3**, **A.4**).[20]

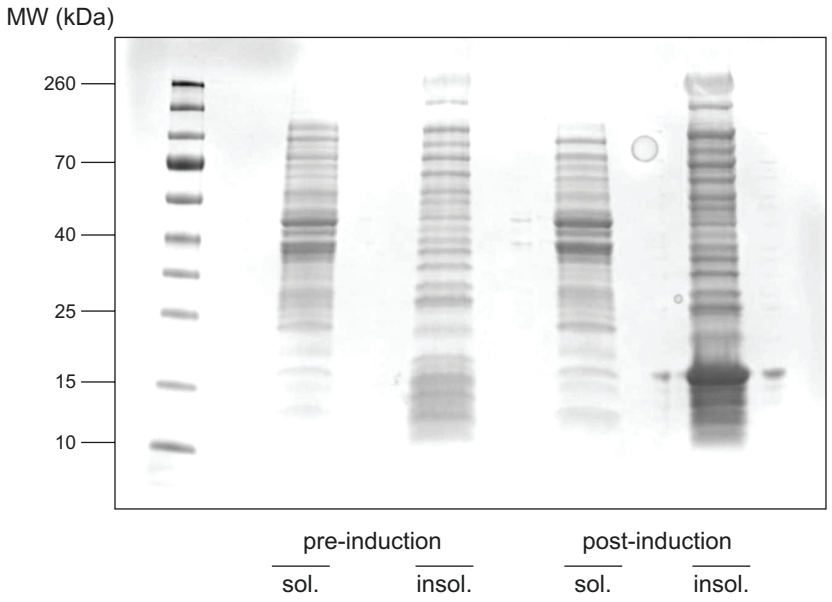


Figure A.3. SDS–PAGE analysis of the production of the RNase 1 variant, stained with Coomassie. Expression was induced by the addition of IPTG.

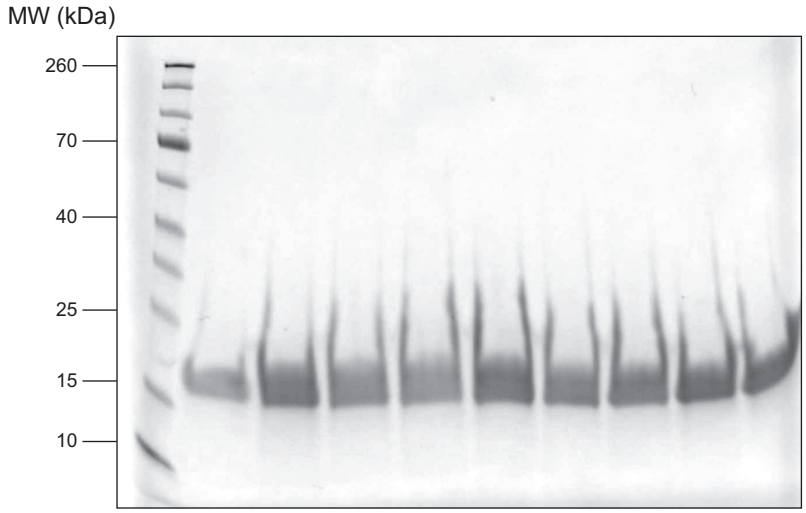


Figure A.4. SDS–PAGE analysis of fractions collected after cation-exchange chromatography of the RNase 1 variant, stained with Coomassie.

Following purification, we protected the unpaired thiol of Cys118 by reaction with 5,5′-dithio-(2-nitrobenzoic acid) (DTNB) to form a mixed disulfide with 2-nitro-5-thiobenzoic acid (NTB) as a safeguard against adventitious oxidation during storage. To produce the S-protein fragment, the RNase 1 variant (1 mg mL^{-1} in 20 mM Tris-HCl buffer, pH 7.4, containing NaCl (50 mM) and $CaCl_2$ (2 mM)) was treated with enterokinase (10 U per 1 mg RNase 1), resulting in efficient, sequence-specific cleavage of the protein in <6 h (**Figure A.5**).

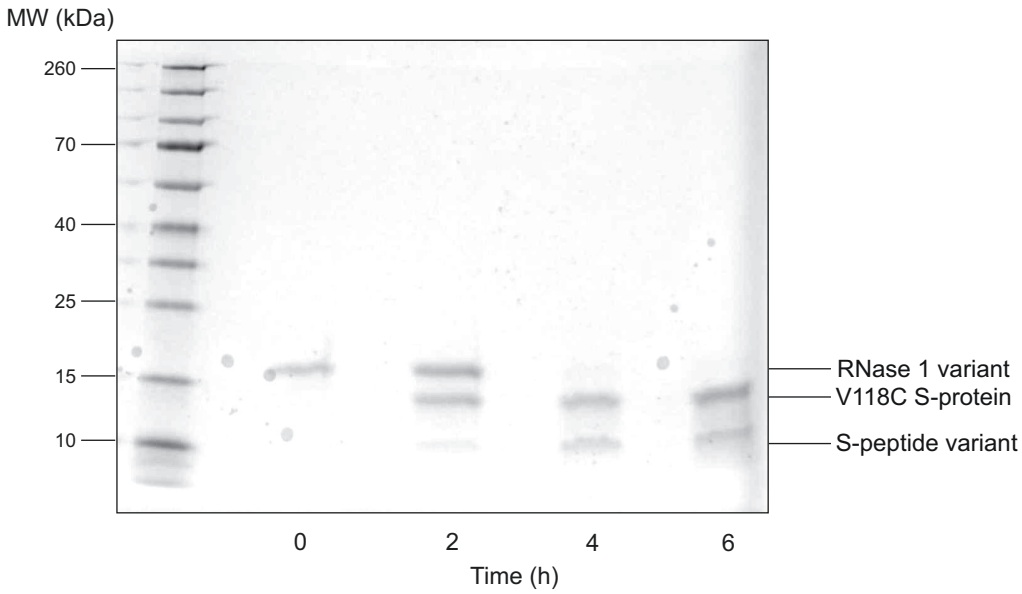


Figure A.5. SDS–PAGE analysis of the digestion of the RNase 1 variant with enterokinase (10 U per 1 mg) to yield V118C S-protein and a variant peptide.

The H12A S-peptide fragment bearing the C-terminal enterokinase recognition sequence (DDDDK) was then discarded. We separated the fragments with semipreparative reversed-phase HPLC, and isolated purified V118C(NTB) S-protein (**Figure A.6**, **A.7**).

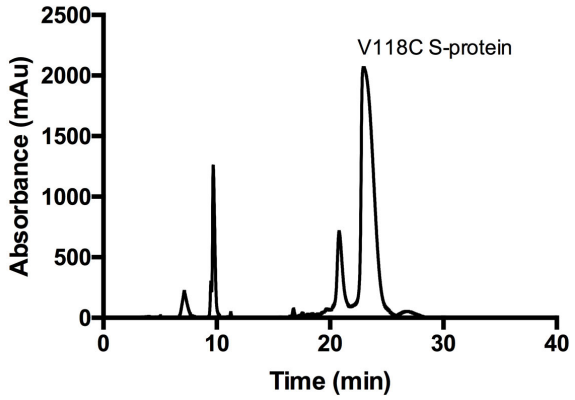


Figure A.6. Preparative HPLC trace of crude digestion reaction. V118C(NTB) S-protein retention time, 22.97 min.

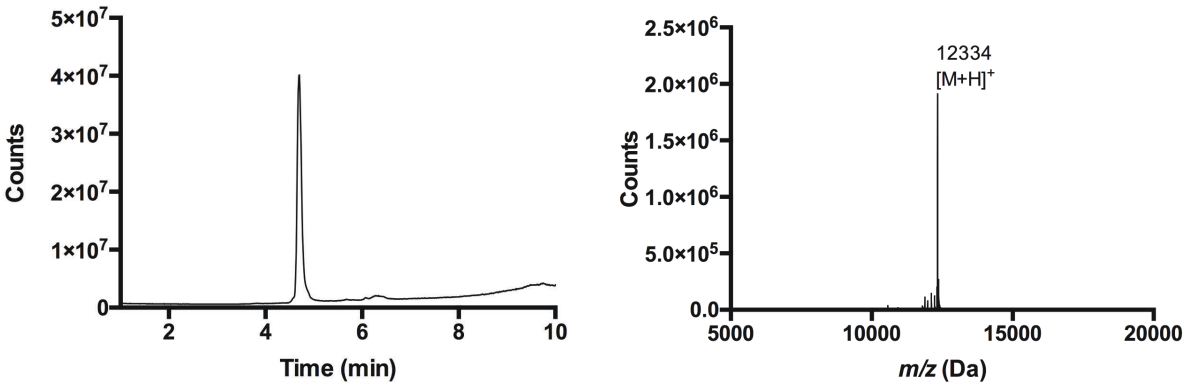


Figure A.7. Q–TOF mass spectrometry total ion chromatogram (left) and deconvoluted mass spectrum (right) for purified V118C(NTB) S-protein. Retention time, 4.70 min (5–95% v/v acetonitrile over 10 min). Expected mass, 12,333.9 Da.

To conjugate V118C S-protein with synthetic S-peptide (residues 1–20 of the RNase 1 sequence), we envisioned the formation of a stabilizing disulfide bond between Cys118 and a cysteine residue incorporated into a synthetic S-peptide. A *de novo* designed disulfide bridge between residues 4 and 118 of intact RNase 1 has been reported to enhance the thermodynamic stability of the intact protein.[263] In this strategy, however, replacing Arg4 with a cysteine residue eliminates a charged residue on the periphery of phosphoryl group-binding subsites[33,267] and could diminish ribonucleolytic activity (**Figure A.8**).[268]

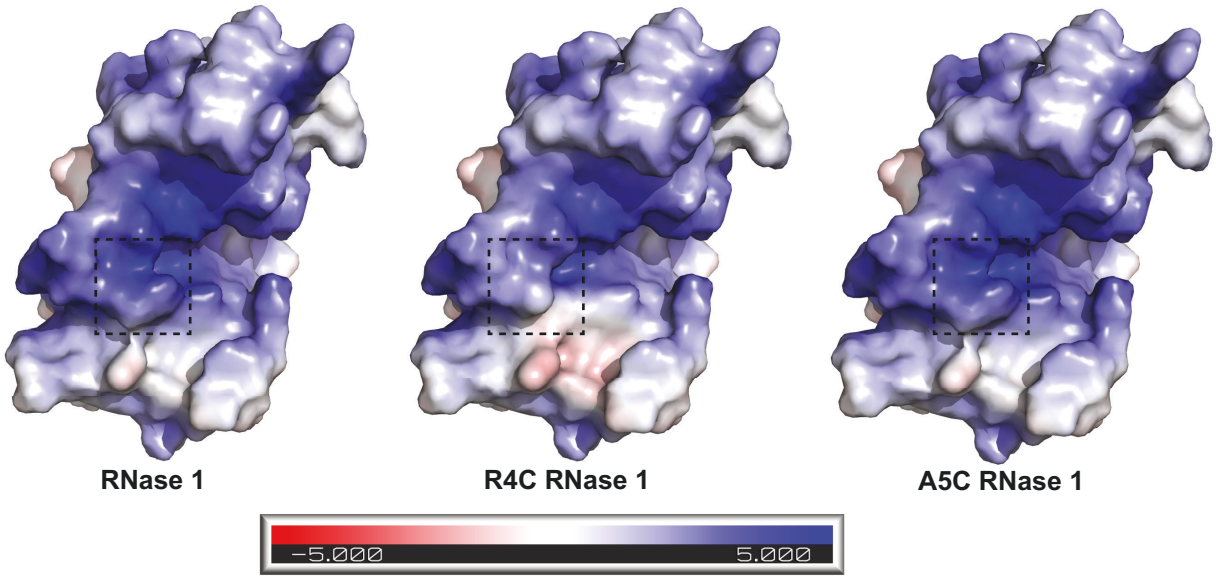


Figure A.8. Electrostatic potential molecular surfaces of RNase 1 (PDB entry 1Z7X) calculated with the Adaptive Poisson–Boltzmann Solver of PyMOL software from Schrödinger (New York, NY). Left, unmodified RNase 1 sequence. Middle, R4C RNase 1 variant. Right, A5C RNase 1 variant. The positive electrostatic potential in the region highlighted by the dashed box is reduced substantially by the R4C substitution but less so by the A5C substitution.

To avoid this pitfall, we examined the three-dimensional structure of RNase 1 (PDB entry 1Z7X[174]) to identify an alternative site for a cross-link within the S-peptide sequence. We found that the C^{α}–C^{α} distances from residue 118 to Arg4 and Ala5 are 5.3 and 4.8 Å, respectively (**Figure A.9**).

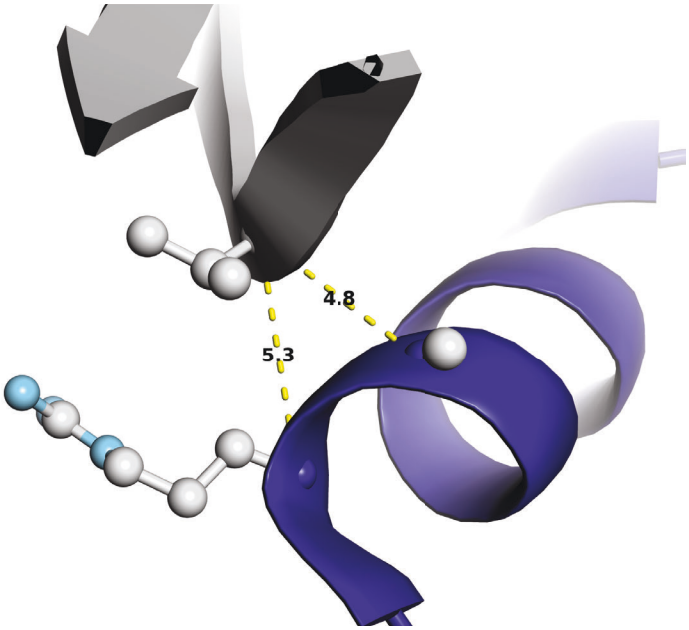


Figure A.9. Image of C^{α}–C^{α} distances from Val118 to Arg4 (5.3 Å) and Ala5 (4.8 Å) as measured in PDB entry 1Z7X with PyMOL software.

Since the mean C^{α}–C^{α} distance in a cystine residue is 5.6 Å,[269] we hypothesized that an A5C substitution would enable facile disulfide cross-linking without a significant alteration of conformation or charge.

We assembled human RNase–S by a convergent scheme (**Figure A.10**).

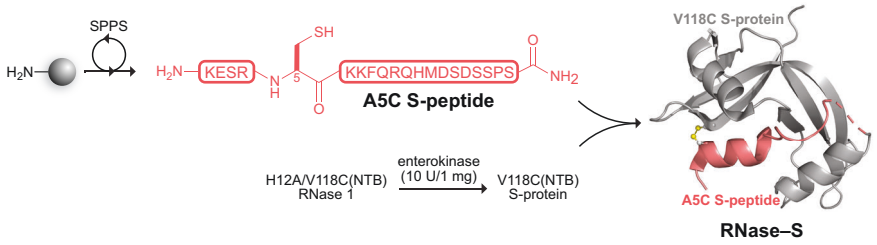


Figure A.10. Scheme for the semisynthesis of human RNase–S. Image was produced with PyMOL software and PDB entry 1Z7X.[174]

Specifically, we prepared A5C S-peptide by standard Fmoc-strategy solid-phase peptide synthesis and purified the peptide by preparative reversed-phase HPLC, obtaining a high overall yield (**Figure A.11**).

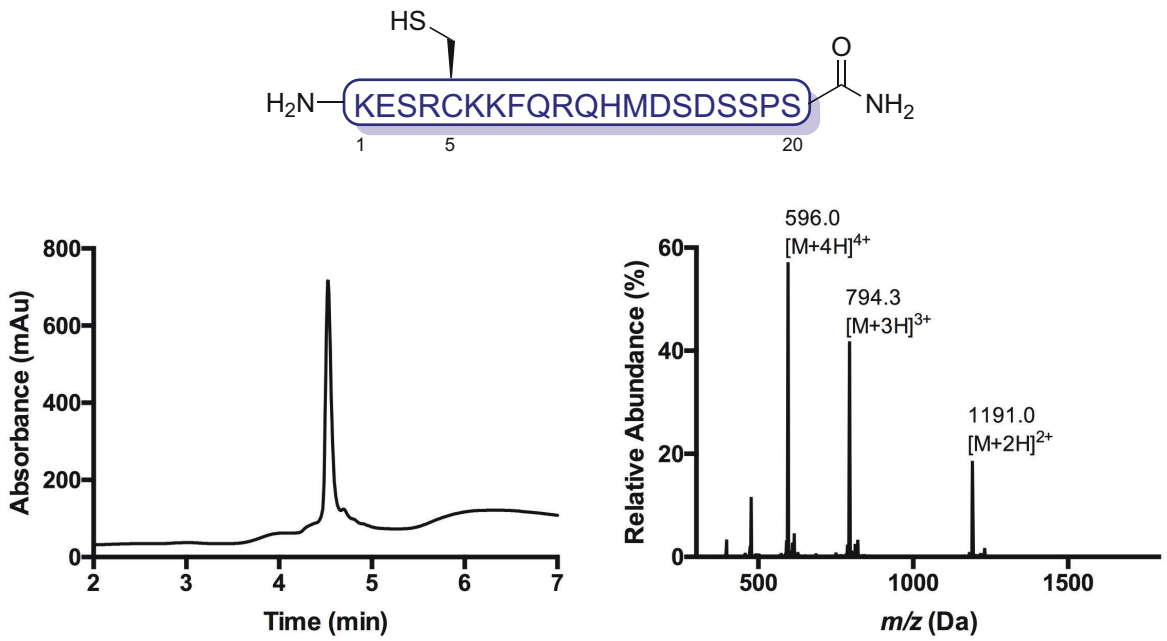


Figure A.11. LC/MS analysis of purified A5C S-peptide. Retention time, 4.53 min (0–10% v/v acetonitrile over 7 min, λ = 210 nm). *m*/*z* expected $[M + 2H]^{2+}$: 1191.3, $[M + 3H]^{3+}$: 794.5, $[M + 4H]^{4+}$: 596.2. *m*/*z* found (ESI^+) $[M + 2H]^{2+}$: 1191.0, $[M + 3H]^{3+}$: 794.3, $[M + 4H]^{4+}$: 596.0.

Immediately prior to conjugation, we prepared a solution of the lyophilized peptide (420 μM in phosphate-buffered saline, pH 8.0, containing EDTA (40 mM)) and treated the solution with immobilized tris(carboxyethyl)phosphine to ensure complete reduction of the peptide. After 30 min at 37 °C, we added the A5C S-peptide solution to a solution of NTB-protected V118C S-protein (8.1 μM in 0.2 M sodium acetate buffer, pH 5.5), and agitated the resulting solution gently at 37 °C. The solution gradually turned slightly yellow, indicative of the release of the NTB^{2-} dianion, which provided a convenient visual indication that the conjugation reaction was proceeding. After only 1 h, we found the formation of disulfide-linked RNase–S to be complete by MALDI–TOF mass spectrometry. The conjugate was separated from residual excess A5C S-peptide by cation-exchange chromatography and judged to be pure by SDS–PAGE (**Figure A.12**).and Q–TOF mass spectrometry (**Figure A.13**).

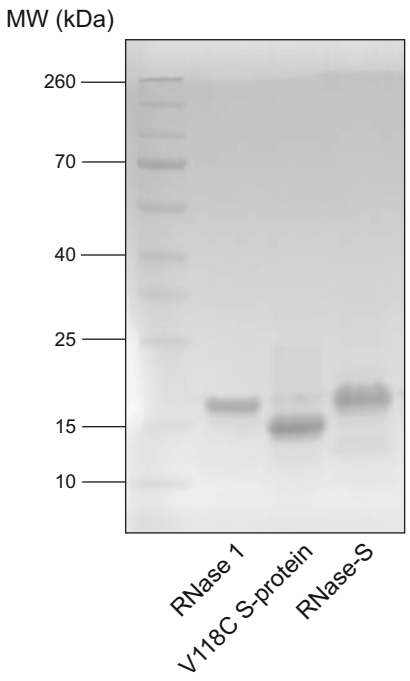


Figure A.12. SDS–PAGE analysis of RNase–S conjugate under non-reducing conditions with Coomassie staining. Lane 1: molecular weight marker. Lane 2: wild-type RNase 1. Lane 3: V118C(NTB) S-protein. Lane 4: RNase–S conjugate.

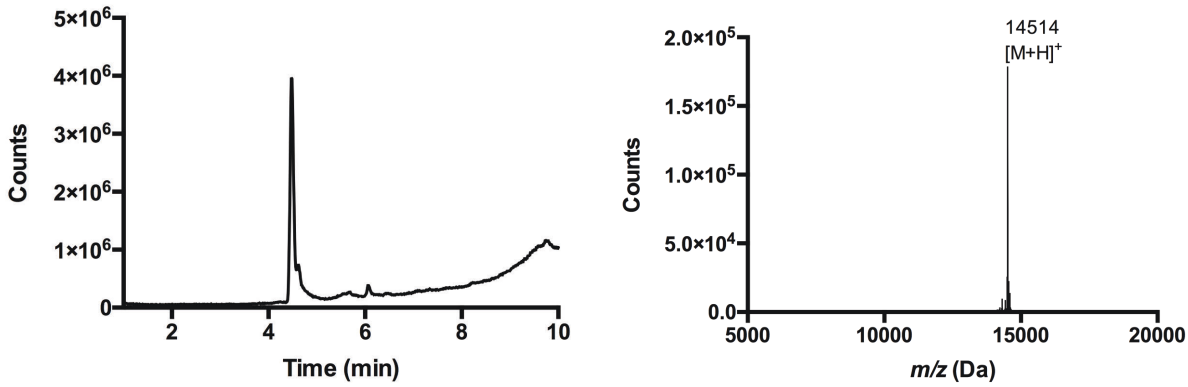


Figure A.13. Q–TOF mass spectrometry total ion chromatogram (left) and deconvoluted mass spectrum (right) for purified RNase–S conjugate. Retention time, 4.47 min (5–95% v/v acetonitrile over 10 min). Expected mass, 14,515.3 Da.

Next, we subjected the purified RNase–S conjugate to a fluorogenic assay of ribonucleolytic activity to confirm that the active site had been reconstituted correctly. In this assay, the fluorescence of the substrate, 6-FAM–dArU$(dA)_2$–6-TAMRA, increases by 200-fold upon cleavage of its single RNA residue by a ribonuclease.[35] Gratifyingly, the semisynthetic RNase–S conjugate displayed robust activity, with $k_{cat}/K_M = 3.6 \times 10^5\ M^{-1}s^{-1}$ (**Figure A.14**).

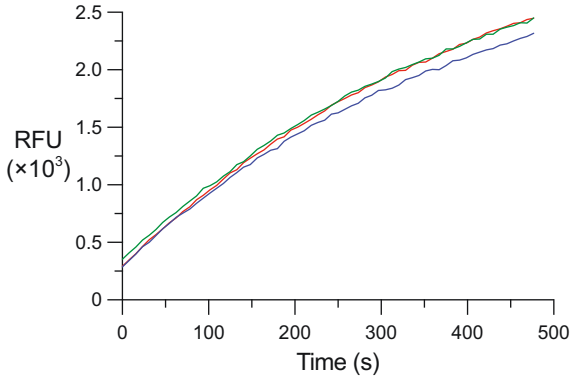


Figure A.14. Graph showing the ribonucleolytic activity of RNase–S (5 nM) with a fluorogenic substrate, 6-FAM–dArU(dA)$_2$–6-TAMRA (200 nM), in 0.10 M DEPC-treated,[270] OVS-free[271] MES-NaOH buffer, pH 6.0, containing NaCl (0.10 M); λ_{ex} 493 nm, λ_{em} 515 nm. The assay was performed in triplicate.

This value is substantial, though somewhat lower than that of wild-type RNase 1 under the same conditions ($k_{cat}/K_M = 9.5 \times 10^6\ M^{-1}s^{-1}$). Notably, these assays were conducted with an enzymic concentration (5 nM) at which the noncovalent bovine RNase S complex is largely dissociated and thus inactive.[260]

Then, we used differential scanning fluorimetry[272] to assess the thermostability of the RNase–S conjugate. We found that RNase–S has a T_m value of (44.5 ± 0.3) °C (**Figure A.15**).

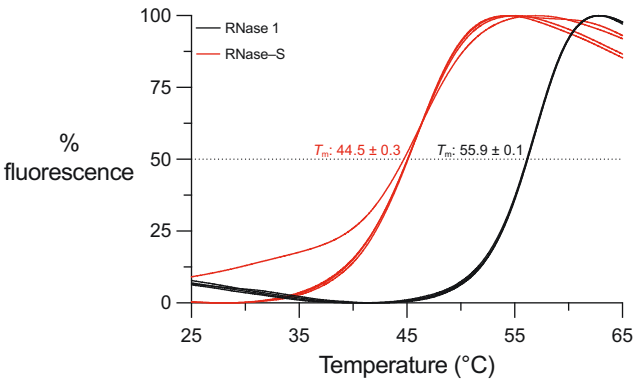


Figure A.15. Heat denaturation profiles of human RNase–S and human RNase 1 obtained by differential scanning fluorimetry using SYPRO Orange, λ_{ex} = 470 ± 15 nm, λ_{em} = 586 ± 10 nm. Assays were performed in quadruplicate.

The data were indicative of a two-state transition, that is, the absence of an intermediate between the native and denatured states. As a comparator, we also assessed the thermostability of wild-type RNase 1. We found a T_m value of (55.9 ± 0.1) °C, which is indistinguishable from one that we reported previously.[273] Thus, RNase–S has lower conformational stability than does RNase 1, but its T_m value is well above physiological temperature.

In contrast to the four native disulfide bonds, Cys5–Cys118 is likely to be more solvent accessible and might, therefore, be more susceptible to reduction in the folded conformation.[263] To gauge the stability of the 5–118 cross-link in a reducing environment like the cytosol, we treated RNase–S with glutathione (1 mM, 10:1 GSH/GSSG) at 37 °C and monitored reduction by SDS–PAGE. After 1 h, we could detect some dissociated S-peptide from reduction of the Cys5–Cys118 bond. Still, intact RNase–S survived for up to 4 h (**Figure A.16**).

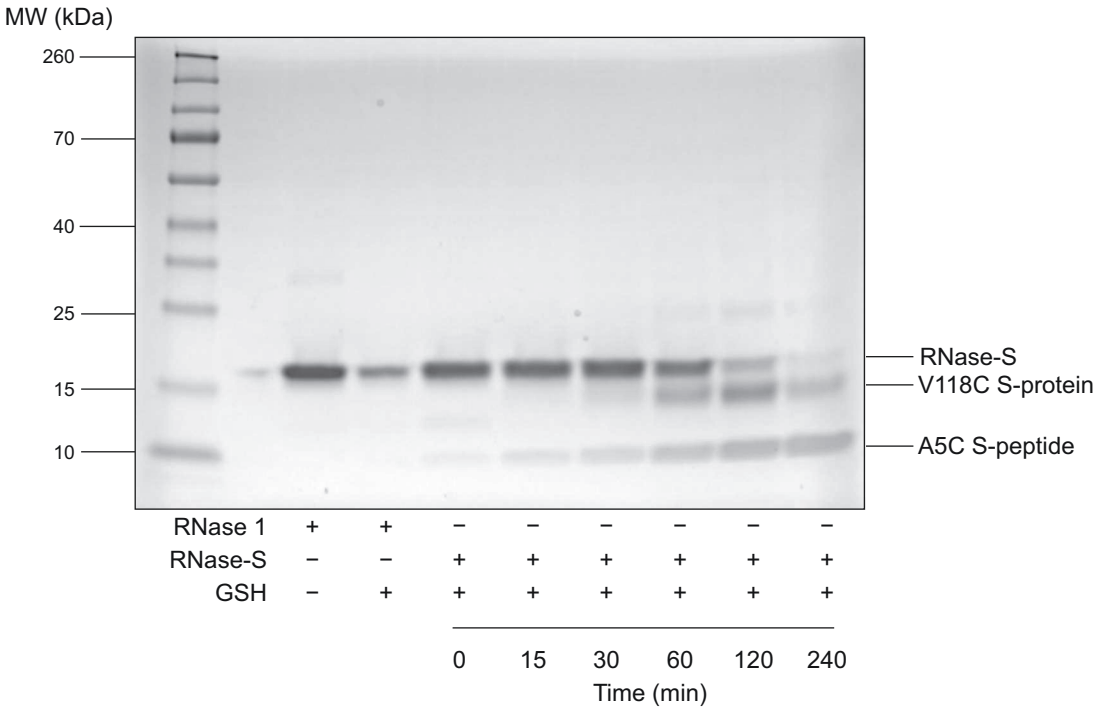


Figure A.16. SDS–PAGE analysis of RNase–S conjugate incubated at 37 °C with glutathione (1 mM, 10:1 GSH/GSSG). Lane 1: molecular weight marker. Lane 3, wild-type RNase 1. Lane 4: RNase 1 plus glutathione. Lanes 5–10: RNase–S plus glutathione.

To prolong the half-life of the RNase–S conjugate even further, we attempted to generate a thioether–cross-linked variant. Specifically, we dehydrated A5C S-peptide to form A5Dha S-peptide (**Figure A.17**, **A.18**), which we linked to V118C S-protein by a conjugate addition reaction.

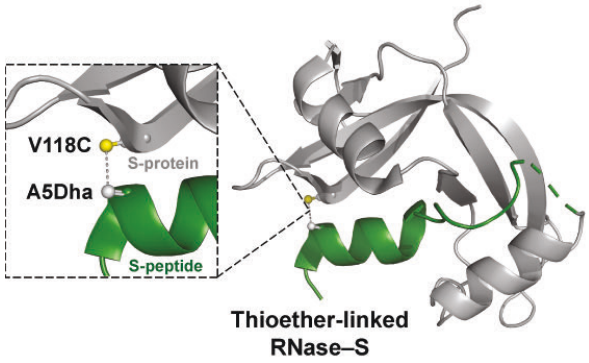


Figure A.17. Image of the envisioned thioether crosslink between V118C S-protein and A5Dha S-peptide as generated with PyMOL software and adapted from PDB entry 1RNU.[104]

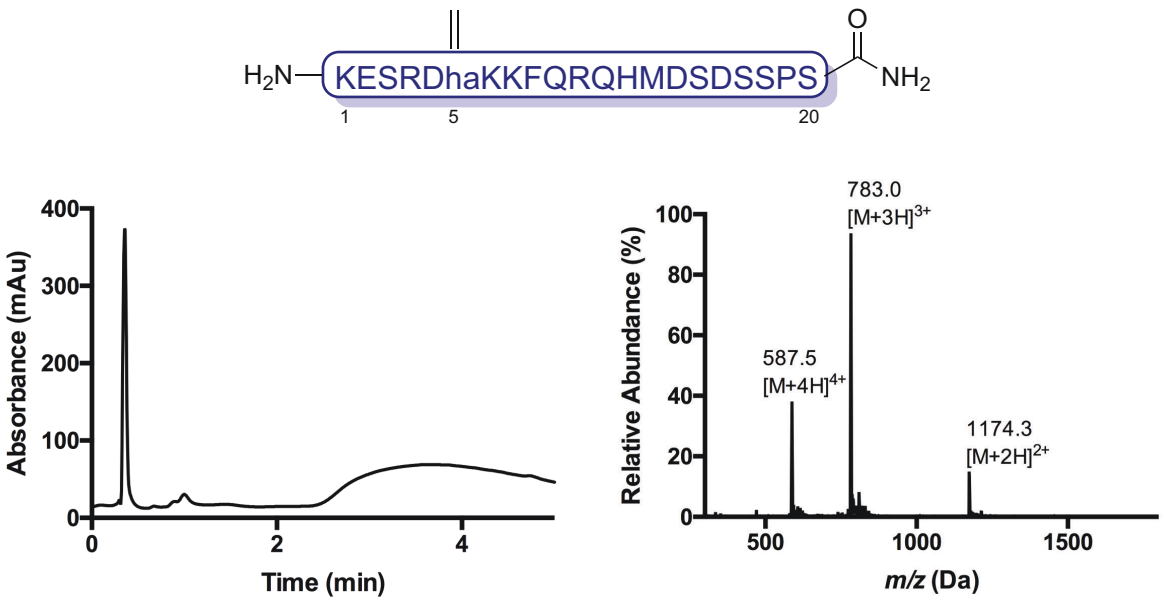


Figure A.18. LC/MS analysis of purified A5Dha S-peptide. Retention time, 0.36 min (5–95% v/v acetonitrile over 7 min, λ = 210 nm). *m*/*z* expected $[M + 2H]^{2+}$: 1174.3, $[M + 3H]^{3+}$: 783.2, $[M + 4H]^{4+}$: 587.6. *m*/*z* found (ESI^+) $[M + 2H]^{2+}$: 1174.3, $[M + 3H]^{3+}$: 783.0, $[M + 4H]^{4+}$: 587.5.

After 24 h, we did observe the formation of some of the desired thioether-linked RNase–S (**Figure A.19**).

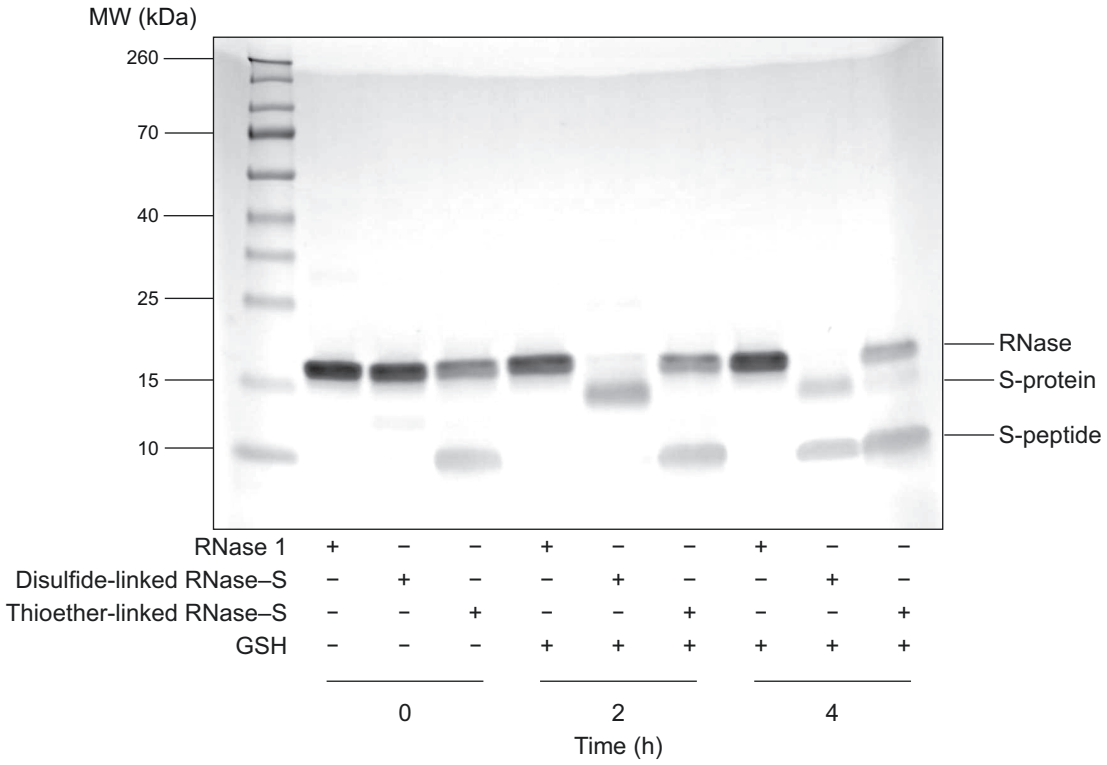


Figure A.19. SDS–PAGE analysis of RNase 1, disulfide-linked RNase–S and thioether-linked RNase–S incubated at 37 °C with glutathione (1 mM, 10:1 GSH/GSSG).

Unfortunately, the requisite pH of this reaction (≥7.5) led to dimerization of V118C S-protein and compromised the yield of the desired product. We therefore proceeded with the highly efficient disulfide-cross-linking strategy, which proceeds in nearly quantitative yield within 1 h.

Finally, we realized that this semisynthetic strategy offers convenient access to active human ribonucleases bearing homogeneous, noncanonical modifications. To demonstrate this capability, we prepared exemplar S-peptide variants bearing N-terminal alkyne (**Figure A.20**), biotin (**Figure A.21**), and fluorescein moieties (**Figure A.22**) and conjugated the modified peptides to V118C S-protein as before.

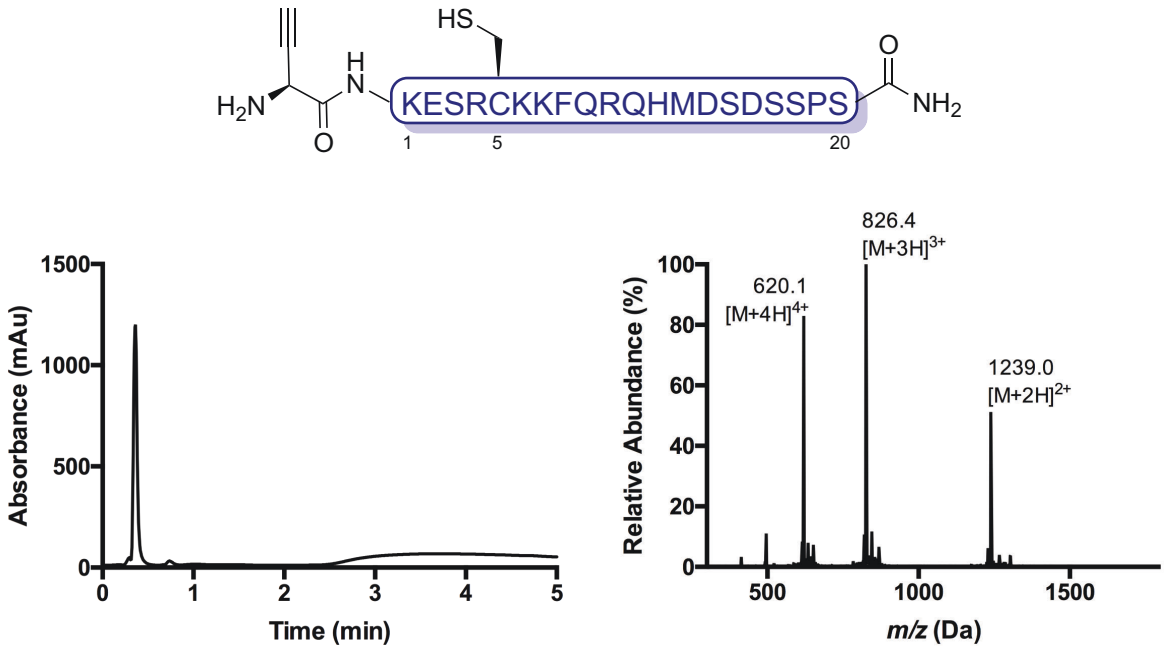


Figure A.20. LC/MS analysis of purified alkyne–A5C S-peptide. Retention time, 4.66 min (0–10% v/v acetonitrile over 7 min, λ = 210 nm). *m*/*z* expected (Met13 oxidized) $[M + 2H]^{2+}$: 1239.6, $[M + 3H]^{3+}$: 826.7, $[M + 4H]^{4+}$: 620.3. *m*/*z* found (ESI^{+}) $[M + 2H]^{2+}$: 1239.0, $[M + 3H]^{3+}$: 826.4, $[M + 4H]^{4+}$: 620.1.

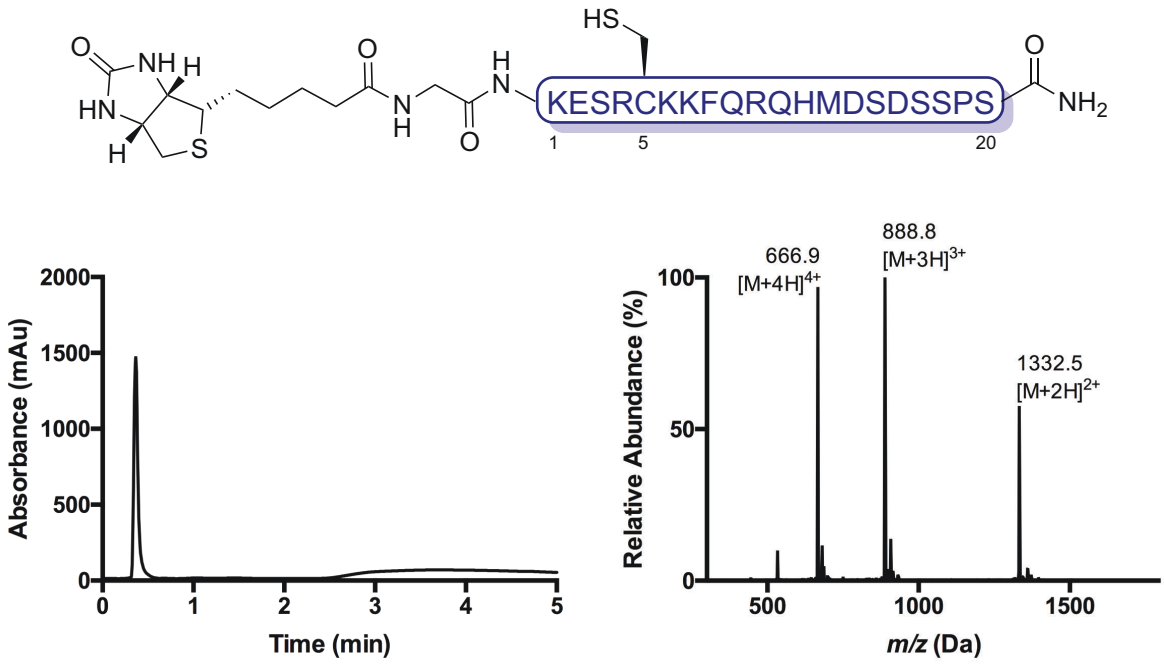


Figure A.21. LC/MS analysis of purified biotin–A5C S-peptide. Retention time, 5.16 min (0–10% v/v acetonitrile over 7 min, λ = 210 nm). *m*/*z* expected $[M + 2H]^{2+}$: 1333.0, $[M + 3H]^{3+}$: 889.0, $[M + 4H]^{4+}$: 667.0. *m*/*z* found (ESI^+) $[M + 2H]^{2+}$: 1332.5, $[M + 3H]^{3+}$: 888.8, $[M + 4H]^{4+}$: 666.9.

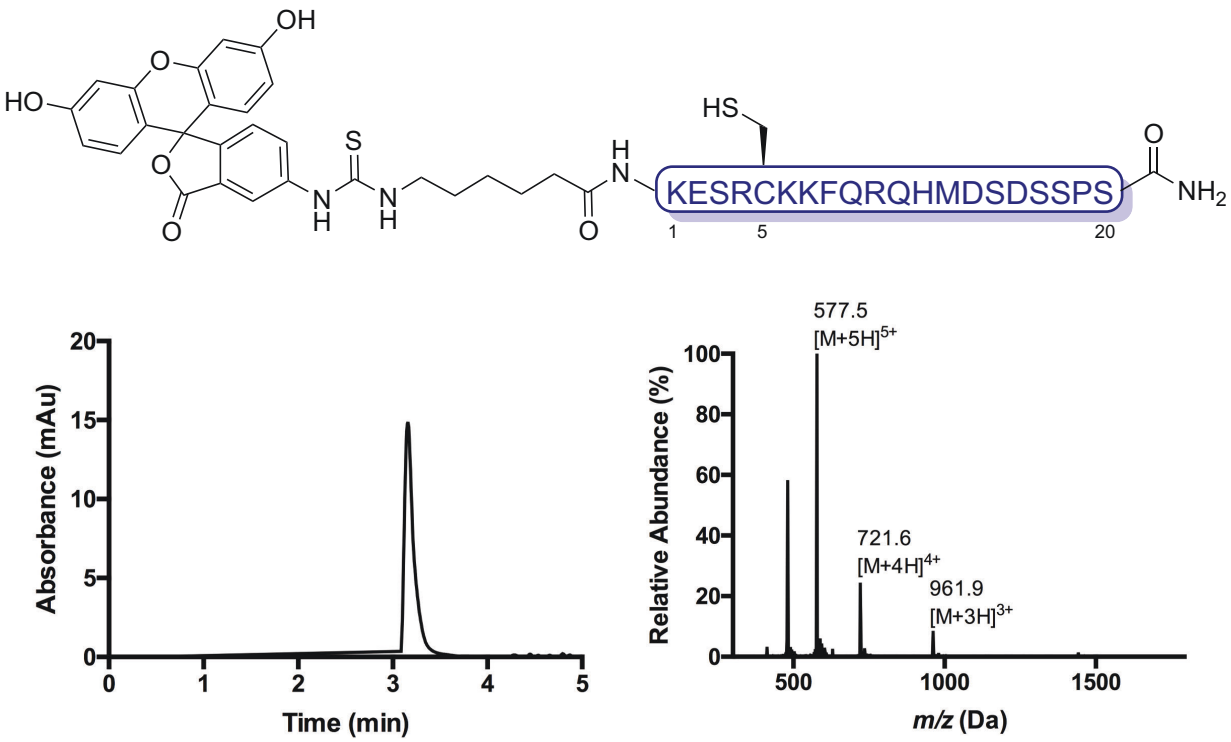


Figure A.22. LC/MS analysis of purified fluorescein–A5C S-peptide. Retention time, 3.16 min (5–95% v/v acetonitrile over 7 min, λ = 280 nm). *m*/*z* expected $[M + 3H]^{3+}$: 962.1, $[M + 4H]^{4+}$: 721.8, $[M + 5H]^{5+}$: 577.6. *m*/*z* found (ESI^+) $[M + 3H]^{3+}$: 961.9, $[M + 4H]^{4+}$: 721.6, $[M + 5H]^{5+}$: 577.5.

We isolated the resulting homogeneously tagged RNase–S variants by cation-exchange chromatography and characterized them by Q–TOF mass spectrometry (**Figure A.23**, **A.24**, and **A.25**).

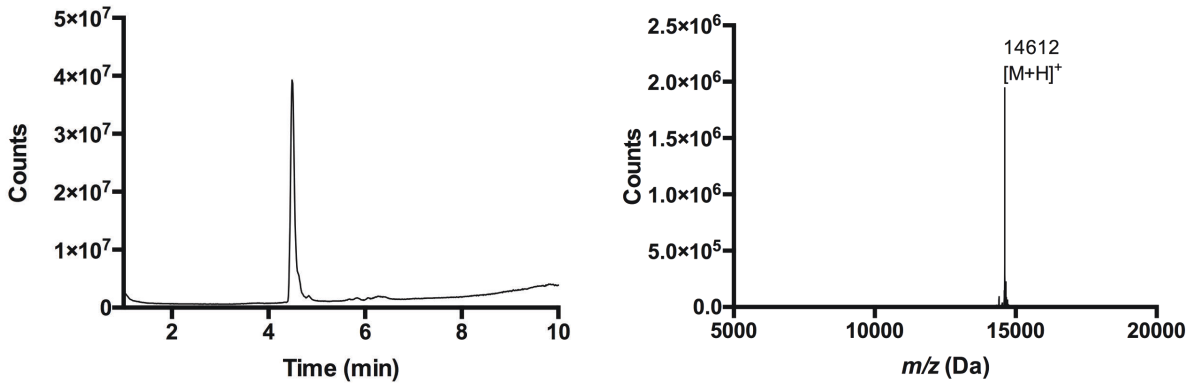


Figure A.23. Q–TOF mass spectrometry total ion chromatogram (left) and deconvoluted mass spectrum (right) for purified alkyne–RNase–S conjugate. Retention time, 4.49 min (5–95% v/v acetonitrile over 10 min). Expected mass, 14,611.9 Da.

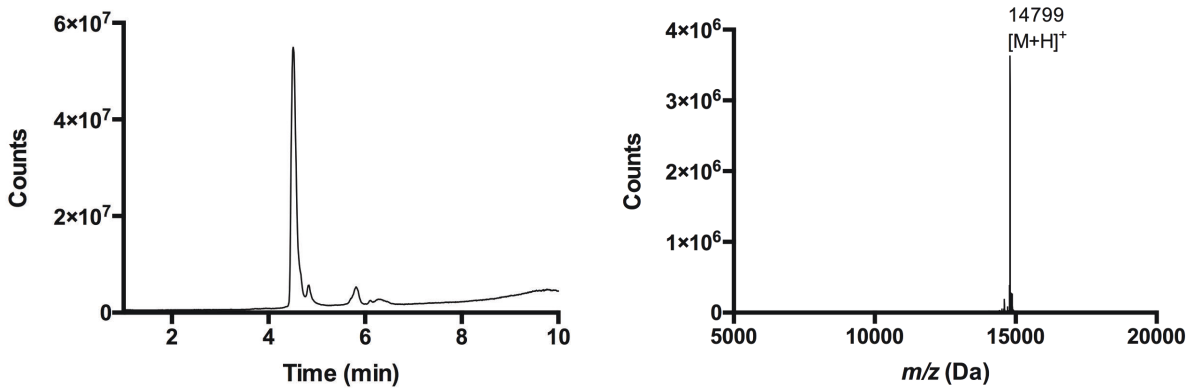


Figure A.24. Q–TOF mass spectrometry total ion chromatogram (left) and deconvoluted mass spectrum (right) for purified biotin–RNase–S conjugate. Retention time, 4.52 min (5–95% v/v acetonitrile over 10 min). Expected mass, 14,798.7 Da.

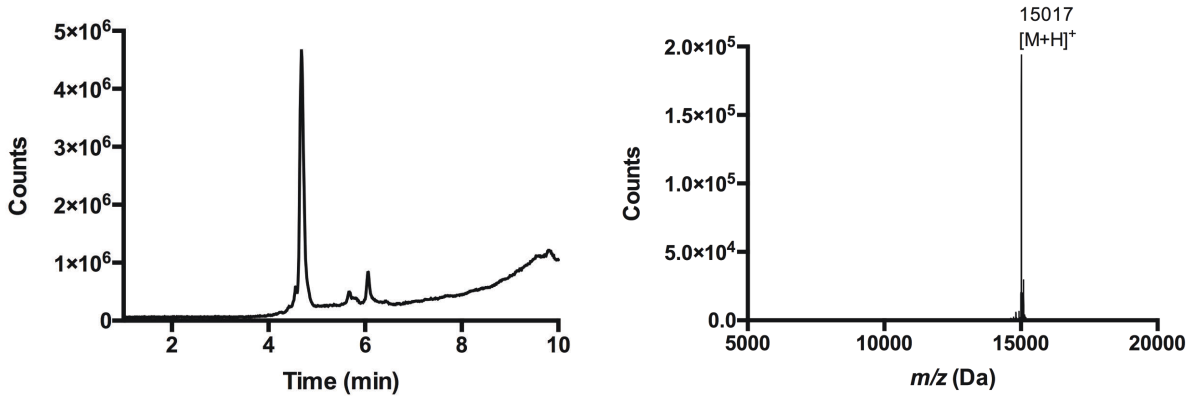


Figure A.25. Q–TOF mass spectrometry total ion chromatogram (left) and deconvoluted mass spectrum (right) for purified fluorescein–RNase–S conjugate. Retention time, 4.68 min (5–95% v/v acetonitrile over 10 min). Expected mass, 15,018.0 Da.

We detected alkyne–RNase–S after a Cu(I)-catalyzed alkyne–azide cycloaddition (CuAAC) with 5-TAMRA–azide and SDS–PAGE (**Figure A.26**), biotin–RNase–S with an antibiotin immunoblot (**Figure A.27**), and fluorescein–RNase–S with fluorescence imaging following SDS–PAGE (**Figure A.28**).

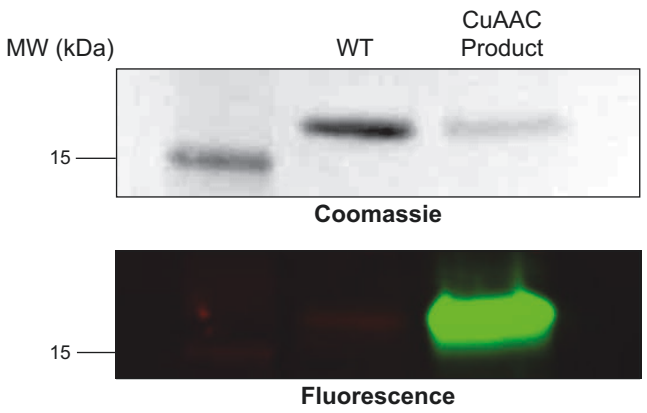


Figure A.26. Analysis of CuAAC reaction of alkyne–RNase–S and Azide-Fluor 545 by SDS–PAGE with Coomassie staining and fluorescence imaging. Lane 1: molecular weight marker. Lane 2: wild-type RNase 1 (*WT*). Lane 3: product of the CuAAC reaction.

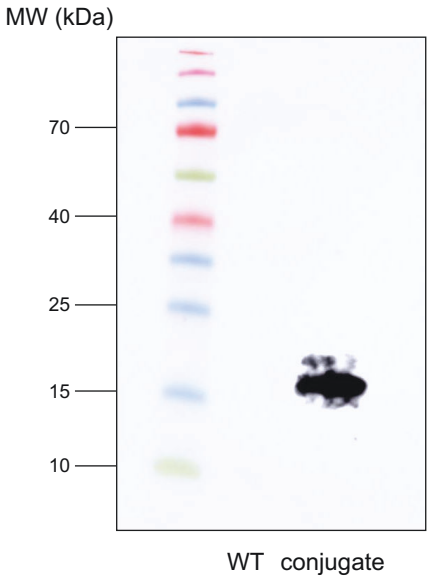


Figure A.27. Immunoblot of biotin–RNase–S conjugate. Lane 1: molecular weight marker. Lane 2: wild-type RNase 1 (*WT*). Lane 3: biotin–RNase–S conjugate.

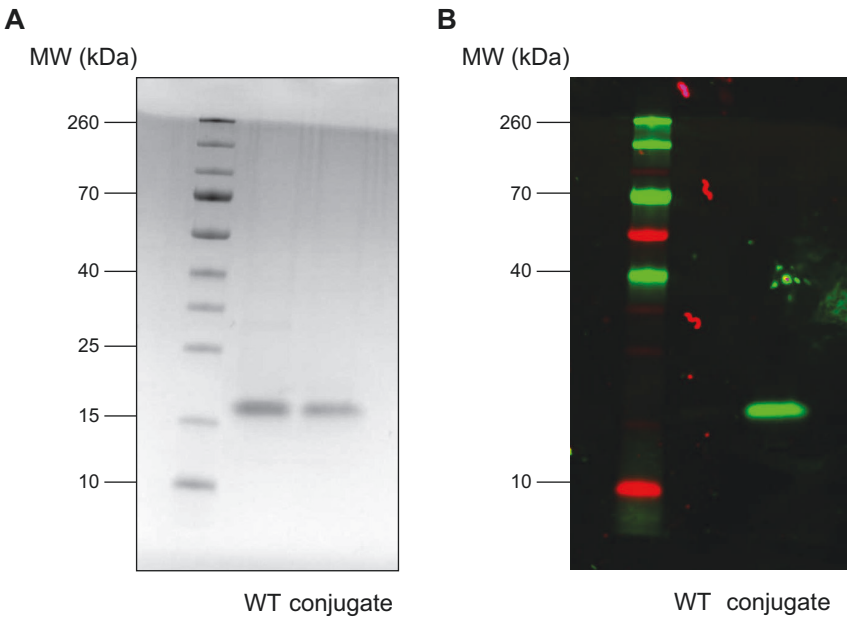


Figure A.28. Analysis of fluorescein–RNase–S conjugate by SDS–PAGE with Coomassie staining (A) and fluorescence imaging (B). Lane 1: molecular weight marker. Lane 2: wild-type RNase 1 (*WT*). Lane 3: fluorescein–RNase–S conjugate.

All three conjugates had high, nearly indistinguishable ribonucleolytic activity (**Table A.1**, **Figure A.29**).

Table A.1. Values of k_{cat}/K_M for catalysis by ribonucleases.

Ribonuclease	k_{cat}/K_M ($M^{-1}s^{-1}$)
RNase 1	$(9.47 \pm 0.33) \times 10^6$
RNase–S	$(3.63 \pm 0.15) \times 10^5$
Alkyne–RNase–S	$(2.23 \pm 0.05) \times 10^4$
Biotin–RNase–S	$(2.08 \pm 0.21) \times 10^4$
Fluorescein–RNase–S	$(2.74 \pm 0.10) \times 10^4$

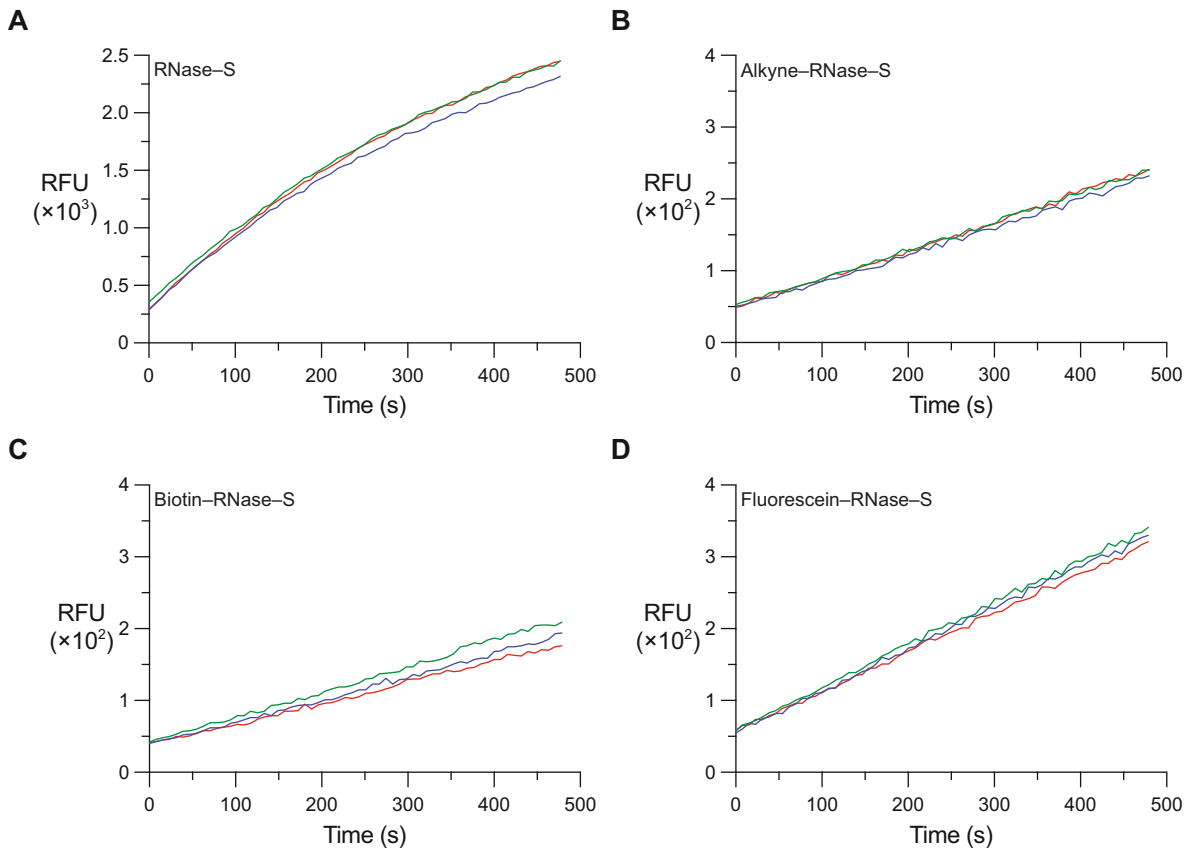


Figure A.29. Graph showing the ribonucleolytic activity of RNase–S and its conjugates (5 nM) with a fluorogenic substrate, 6-FAM–dArU(dA)$_2$–6-TAMRA (200 nM), in 0.10 M MES–NaOH buffer, pH 6.0, containing NaCl (0.10 M). λ_{ex} = 493 nm, λ_{em} = 515 nm. Assays were performed in triplicate. Data for RNase–S are also shown in **Figure A.14**.

Notably, with the requisite fragments in hand, each RNase–S conjugate could be assembled, purified, and used for an application in less than one day. These examples highlight the efficiency and flexibility of our semisynthetic strategy, which offers unprecedented opportunities to produce an unlimited variety of precisely modified variants of an active human ribonuclease.

A.4 Conclusions

Sixty years after the description of bovine RNase S, we employed recombinant DNA technology and a sequence-specific protease to achieve the first preparation of its human homologue. Strategic incorporation of an unpaired cysteine residue into the S-protein sequence enabled disulfide cross-linking with a chemically synthesized S-peptide fragment to produce an RNase–S conjugate. Conveniently, the disulfide conjugation reaction proceeds rapidly in mild conditions and without any exogenous reagents. High ribonucleolytic activity confirmed the successful reconstitution of the enzymic active site. The semisynthetic nature of this strategy enables synthetic S-peptide variants bearing a plethora of modifications to be incorporated into an active catalyst of RNA cleavage. We anticipate that this highly efficient semisynthetic strategy will find broad utility in preparing modified human RNase–S variants that enable novel applications in biotechnology and biomedicine.

A.5 Acknowledgments

E.C.W. was supported by a Graduate Research Fellowship from the NSF. This work was supported by Grant R01 CA073808 (NIH).

A.6 Experimental Details

A.6.1 Materials

The BL21(DE3) strain of *Escherichia coli* was from Novagen (Madison, WI). Ribonuclease A (RNase A) was product number R6513 from Sigma–Aldrich (St. Louis, MO). A fluorogenic ribonuclease substrate (6-FAM–dArU$(dA)_2$–6-TAMRA) and DNA oligonucleotides were from Integrated DNA Technologies (Coralville, IA). Phosphate-buffered saline (PBS) contained Na_2HPO_4 (10 mM), KH_2PO_4 (1.8 mM), NaCl (137 mM), and KCl (2.7 mM) at pH 7.4. Immobilized TCEP was product 77712 from Thermo Fischer Scientific (Waltham, MA). A Spectra Multicolor Broad Range Protein Ladder from Thermo Fisher Scientific was used as a molecular mass standard during SDS–PAGE. Gels were stained with a solution of Bio-Safe™ Coomassie Brilliant Blue G-250 from Bio-Rad (Hercules, CA). SYPRO Orange was from Sigma–Aldrich. All other chemicals and reagents were of commercial reagent grade or better and were used without further purification.

A.6.2 Instrumentation

Aqueous solutions were made with water that was generated with a Milli-Q® IQ 7000 water purification system from Millipore Sigma and had resistivity ≥18 $M\Omega \cdot cm^{-1}$ at 25 °C. SDS–PAGE analyses were performed with Any kD™ Mini-PROTEAN® TGX™ Precast Gels in a Mini-PROTEAN Tetra cell. Gels were imaged with an Amersham Imager 600 from GE Life Sciences and images were processed with ImageJ software. FPLC chromatography was performed with an ÄKTA Pure system from GE Healthcare Life Sciences (Piscataway, NJ), and the results were analyzed with the UNICORN Control System. HiTrap SP-HP and HiLoad® 26/600 Superdex® 75 prep grade columns for protein purification were from GE Healthcare Life

Sciences. Protein concentrations were determined by using either a bicinchonic acid (BCA) assay kit from Thermo Fisher Scientific and an infinite M1000 microplate reader from Tecan (Zürich, Switzerland) or by absorbance at 280 nm with a DS-11 UV–Vis spectrophotometer from DeNovix (Wilmington, DE). Mass spectrometry of synthetic peptides was performed with a nominal mass 6125B mass spectrometer attached to a 1260 Infinity LC from Agilent Technologies (Santa Clara, CA). The intact molecular masses of RNase–S conjugates were determined by ESI mass spectrometry with a 6530 Accurate-Mass Q–TOF LC/MS equipped with a PLRP-S column (1000-Å pore size, 5-μM particle size, 50-mm length, and 2.1-mm ID) from Agilent Technologies. Synthetic peptides were purified with a 1260 Infinity Preparative LC System from Agilent Technologies equipped with a Nucleosil C18 column (100 Å, 5 μm, 250 × 21 mm) from Macherey–Nagel (Düren, Germany).

A.6.3 Conditions

All procedures were performed in air at ambient temperature (~22 °C) and pressure (1.0 atm) unless indicated otherwise.

A.6.4 S–peptide preparation

A.6.4.A Preloading Rink Amide ProTide resin

Rink Amide ProTide resin was initially washed with dichloromethane (DCM) (5×) and dimethylformamide (DMF) (5×), followed by removal of the Fmoc protecting group by treatment with 20% v/v 4-methyl piperidine in DMF (2 × 5 min). The resin was washed with DMF (5×), DCM (5×), and DMF (5×). To a solution of the first residue to be coupled (Fmoc-AA-OH, 4 equiv) in DMF (final concentration 0.1 M) was added PyBOP (4 equiv) and DIPEA (8 equiv). After 2 min of preactivation, the mixture was added to the resin and incubated with agitation for 2 h. The resin was washed with DMF (5×), DCM (5×), and DMF (5×). Any unreacted amino groups remaining on the resin were capped by treatment with a solution of 10% v/v acetic anhydride in pyridine (2 × 3 min) and then washed with DMF (5×), DCM (5×), and DMF (5×). The resin was subsequently submitted to iterative peptide assembly (Fmoc-SPPS).

A.6.4.B Preloading Cl-TCP(Cl) ProTide resin

Cl-TCP(Cl) ProTide resin was swollen in DCM for 30 min then washed with DCM (5×) and DMF (5×). A solution of Fmoc-AA-OH (0.2 M) in DMF was added to the resin. To a 0.125 M KI solution in DMF was added DIPEA (1.0 M final concentration), and the resulting solution was added to the resin. The resin was agitated overnight. The resin was washed with DMF (5×) and DCM (5×) and was treated with a solution of 17:2:1 DCM/methanol/*i*Pr_2NEt for 30 min to cap any unreacted 2-chlorotrityl chloride resin handles. Finally, the resin was washed with DCM (5×) and DMF (5×) and subsequently submitted to iterative peptide assembly (Fmoc-SPPS).

A.6.4.C Estimating the load of the first amino acid

The resin was treated with 20% v/v 4-methyl piperidine in DMF (2 × 3 min). The combined filtrate was diluted to 10 mL with DMF. A 10-μL aliquot of the solution was further diluted 200-fold with DMF, and the UV absorbance of the piperidine–fulvene adduct was measured (λ = 301 nm, ε = 7800 $M^{-1}cm^{-1}$). The amount of amino acid loaded onto the resin could then be estimated.

A.6.4.D General iterative peptide assembly (Fmoc-SPPS)

Deprotection. The resin was treated with 20% v/v 4-methyl piperidine in DMF (2 × 3 min) and washed with DMF (5×), DCM (5×), and DMF (5×).

General amino acid coupling. A preactivated solution of Fmoc-AA-OH (4 equiv), HATU (4 equiv) and DIPEA (8 equiv) in DMF (final concentration: 0.10 M) was added to the resin. After 2 h of agitation, the resin was washed with DMF (5×), DCM (5×), and DMF (5×).

Capping. A solution of 10% v/v acetic anhydride in pyridine was added to the resin (2 × 3 min). The resin was washed with DMF (5×), DCM (5×), and DMF (5×).

Global deprotection and resin cleavage. A solution of 90:5:2.5:2.5 trifluoroacetic acid/triisopropylsilane (iPr$_3$SiH)/thioanisole/H_2O was added to the resin. After 2 h of agitation, the resin was washed with trifluoroacetic acid (3 × 2 mL).

Work-up. The cleavage filtrate was concentrated to dryness under a stream of nitrogen gas. The residue was suspended in diethyl ether and cooled to 0 °C. The precipitate was pelleted by centrifugation at 4000 RPM for 5 min, the supernatant was decanted, and the pellet was dried under a gentle stream of nitrogen gas. The resulting white powder was stored at −20 °C until it could be purified via reversed-phase preparative HPLC.

A.6.4.E Automated SPPS

Microwave-assisted peptide synthesis was performed at 90 °C on a Liberty Blue® automated microwave peptide synthesiser from CEM (Matthews, NC). The general coupling method was employed as follows: 2-min coupling (90 °C), 1-min deprotection (90 °C), and 1-min associated washes and liquid handling. A fourfold excess of Fmoc-AA-OH, HOBt, and DIC was used, as is standard.

A.6.4.F A5C S-peptide synthesis

A5C S-Peptide was prepared on Rink Amide ProTide resin (116 mg, 0.43 mmol g^{-1}, 50 μmol) via automated Fmoc-SPPS as described in **Section A.6.4.E**. Following complete elongation of the sequence, the peptide was subjected to global deprotection/resin cleavage and precipitated from cold diethyl ether (**Section A.6.4.D**) to produce a white powder. The crude material was purified via reversed-phase preparative HPLC (10–40% v/v acetonitrile over 20 min, 0.1% v/v trifluoroacetic acid) to afford A5C S-peptide as a white solid after lyophilization (19 mg, 7.98 μmol, 12% yield).

A.6.4.G Alkyne–A5C S-peptide synthesis

Alkyne-tagged A5C S-peptide was prepared on Rink Amide ProTide resin (77 mg, 0.43 mmol g^{-1}, 33 μmol). The sequence was elongated to Lys1 via automated Fmoc-SPPS as described in **Section A.6.4.E**. Fmoc-propargyl-Gly-OH (22 mg, 66 μmol, 2 equiv) was coupled to the N terminus according to the general amino acid coupling procedure described in **Section A.6.4.D**. Following deprotection of the N-terminal Fmoc protecting group, the crude peptide was deprotected, cleaved from the resin and precipitated from cold diethyl ether (**Section A.6.4.D**).

Finally, the crude material was purified via reversed-phase preparative HPLC (10–40% v/v acetonitrile over 20 min, 0.1% v/v trifluoroacetic acid) to afford alkyne–A5C S-peptide as a white solid after lyophilization (8 mg, 3.22 μmol, 10% yield).

A.6.4.H Biotin–A5C S-peptide synthesis

Biotin-tagged A5C S-peptide was prepared on Rink Amide ProTide resin (77 mg, 0.43 mmol g^{-1}, 33 μmol). The sequence was elongated via automated Fmoc-SPPS as described in **Section A.6.4.E**, with an additional Gly residue added to the N terminus. Following Fmoc-deprotection, (+)-biotin *N*-hydroxysuccinimide ester (23 mg, 66 μmol, 2 equiv) was coupled as a solution in DMF with DIPEA (12 μL, 66 μmol, 2 equiv) for 2 h. The crude peptide was deprotected, cleaved from the resin, and precipitated from cold diethyl ether (**Section A.6.4.D**) before purification via reversed-phase preparative HPLC (10–40% v/v acetonitrile over 20 min, 0.1% v/v trifluoroacetic acid) to afford biotin–A5C S-peptide as a white solid after lyophilization (11 mg, 4.13 μmol, 12% yield).

A.6.4.I Fluorescein–A5C S-peptide synthesis

Fluorescein-tagged A5C S-peptide was prepared on Rink Amide ProTide resin (116 mg, 0.43 mmol g^{-1}, 50 μmol). The sequence was elongated via automated Fmoc-SPPS as described in **Section A.6.4.E**, with an additional 6-aminohexanoic acid residue added to the N terminus. Inclusion of this non-α amino acid spacer was necessary to avoid acid-catalyzed thiohydantoin formation via the Edman degradation pathway.[274] Following Fmoc-deprotection, fluorescein isothiocyanate (98 mg, 250 μmol, 5 equiv) was coupled as a solution in DMF with DIPEA (45 μL, 250 μmol, 5 equiv) for 2 h. The crude peptide was deprotected, cleaved from the resin, and

precipitated from cold diethyl ether **Section A.6.4.D**) before purification via reversed-phase preparative HPLC (10–40% v/v acetonitrile over 20 min, 0.1% v/v trifluoroacetic acid) to afford fluorescein–A5C S-peptide as a white solid after lyophilization (13 mg, 4.50 μmol, 9% yield).

A.6.4.J A5Dha S-peptide synthesis

To a solution of A5C S-peptide (20 mg, 8.40 μmol) in 50 mM sodium phosphate buffer, pH 8.0, containing EDTA (20 mM), was added 2,5-dibromohexanediamide (20 mg, 66.2 μmol, 7.9 equiv).[275] The reaction mixture was agitated overnight at 37 °C. Purification was performed via reversed-phase preparative HPLC (10–30% v/v acetonitrile over 40 min, 0.1% v/v trifluoroacetic acid) to afford A5Dha S-peptide as a white solid after lyophilization (7.05 mg, 3.00 μmol, 35% yield).

A.6.5 S–protein preparation

A.6.5.A Preparation of the RNase 1 variant

Synthetic cDNA encoding the H12A/S20_S21insDDDDK/V118C RNase 1 variant flanked by regions of homology near the T7 promoter and terminator of the pET22b vector were obtained from Integrated DNA Technologies. Linear pET22b was prepared by PCR using primers that complement the DNA encoding RNase 1:

Forward: `5′-ACTGAGCTCGTGGTGGTGG-3′`

Reverse: `3′-CCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACATATG-5′`

All gene and plasmid fragments were combined with Gibson assembly for expression in *E. coli*. The corresponding amino acid sequence for the RNase 1 variant is as follows (changes in **red**). Note that the variant is missing the inconsequential C-terminal threonine residue (Thr128), which was absent from both the initial report of the amino acid sequence[276] and the crystal structure.[174]

```
        10         20         30         40         50         60         70
KESRAKKFQR QAMDSDSSPS DDDDKSSSTY CNQMMRRRNM TQGRCKPVNT FVHEPLVDVQ NVCFQEKVTC
        80         90        100        110        120        130
KNGQGNCYKS NSSMHITDCR LTNGSRYPNC AYRTSPKERH IIVACEGSPY VPCHFDASVE DS_
```

The RNase 1 variant was produced and purified essentially as described previously.[178] Briefly, transformed BL21(DE3) cells were grown at 37 °C with shaking in Terrific Broth containing ampicillin (400 μg mL^{-1}) until OD_{600} = 1.8–2.2. Gene expression was induced by the addition of

0.5 mM isopropyl-β-d-thiogalactopyranoside (IPTG). After 3–4 h, cells were harvested by centrifugation and lysed at 19.0 kPSI with a benchtop cell disruptor (Constant Systems). After centrifugation at 10,500 RPM for 45 min, the resulting inclusion bodies were dissolved in 20 mM Tris-HCl buffer, pH 8.0, containing guanidine–HCl (7 M), EDTA (10 mM), and DTT (100 mM). This solution was diluted tenfold by the slow addition of degassed 20 mM acetic acid, then subjected to centrifugation at 10,500 RPM for 45 min. The resulting supernatant was then dialyzed using 3.5-kDa MWCO tubing (Spectrum Labs) against 16 L of 20 mM acetic acid overnight. After centrifugation at 10,500 RPM for 40 min, the retentate was added dropwise to refolding buffer, which was 100 mM Tris-HCl buffer, pH 7.8, containing NaCl (100 mM), reduced glutathione (1.0 mM), and oxidized glutathione (0.2 mM), then allowed to refold at 4 °C for 5 days. The pH of the solution was adjusted to 5.0, and the resulting solution was concentrated to 10 mL with an Amicon® Stirred Cell concentrator (EMD Millipore) with Hydrostart® 10-kDa filters (Sartorius). The resulting protein solution was purified by chromatography on a HiLoad® 26/600 Superdex® 75 pg gel-filtration column from GE Healthcare Life Sciences with 50 mM sodium acetate buffer, pH 5.0, containing NaCl (100 mM) and sodium azide (0.05% w/v).

After gel-filtration chromatography, the fractions containing ribonuclease were combined, and the pH of the resulting solution was increased by adding 2.0 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 a fourfold molar excess to the protein. After incubation for 10 min at 4 °C, the pH of the solution was adjusted to 5.0 by the addition of 3.0 M sodium acetate, and the resulting solution was incubated at 4 °C overnight. The protected protein was purified further by passage through a HiTrap® SP cation-

exchange column from GE Healthcare Life Sciences with 50 mM sodium acetate buffer, pH 5.0, and a gradient of 0.0–1.0 M NaCl over 10 min.

A.6.5.B Preparation of V118C(NTB) S-protein

$CaCl_2$ (2 mM final concentration) was added to the RNase 1 variant (10 mL of a 1 mg mL^{-1} solution) in 20 mM Tris-HCl buffer, pH 7.4, containing NaCl (50 mM). Enterokinase (light chain) was product number P8070S from New England Biolabs. Enterokinase (10 U per 1 mg of protein) was added, and the reaction mixture was incubated with gentle agitation. Aliquots were removed over 24 h and subjected to SDS–PAGE analysis.

The V118C(NTB) S-protein fragment was isolated by semi-preparative reversed-phase HPLC with a C4 column and two-step linear gradient (Step 1: 20–50% v/v solvent B over 3 min; Step 2: 50–100% v/v solvent B over 25 min. Solvent A is 50 mM sodium phosphate buffer, pH 2.7. Solvent B is 40% v/v solvent A and 60% v/v acetonitrile). Fractions containing V118C(NTB) S-protein were pooled and concentrated under a stream of nitrogen gas to remove acetonitrile. The concentration of V118C(NTB) S-protein was determined by absorbance at 280 nm using $\varepsilon = 7950\ M^{-1}cm^{-1}$. The identity of each peak on the HPLC trace was confirmed by Q–TOF mass spectrometry.

A.6.6 Disulfide conjugation to prepare RNase–S constructs

Lyophilized A5C S-peptide, alkyne–A5C S-peptide, biotin–A5C S-peptide, or fluorescein–A5C S-peptide (10 equiv) was dissolved in PBS, pH 8.0, containing EDTA (40 mM), to a final concentration of 1 mg mL^{-1}. An equivalent volume of immobilized TCEP was drained using a fritted syringe, and then the S-peptide solution was taken up and allowed to reduce for 30 min at 37 °C. Then, the reduced S-peptide solution was ejected directly into a solution of NTB-protected V118C S-protein (0.1 mg mL^{-1} in sodium acetate buffer, pH 5.5). The conjugation reaction was allowed to proceed for 1 h with gentle agitation at 37 °C. The solution gradually turns yellow due to removal of the NTB protecting group on Cys118. Complete conversion to the disulfide-linked RNase–S conjugate was confirmed by MALDI–TOF mass spectrometry. The solution was concentrated to ≤500 µL using a Vivaspin® 20 centrifugal concentrator (5000 MWCO PES) and then purified by passage through a HiTrap® SP cation-exchange column (GE Healthcare Life Sciences) with 50 mM sodium acetate buffer, pH 5.0, and a gradient of 0.0–0.8 M NaCl over 80 min. The conjugated product elutes at approximately 0.5 M NaCl.

A.6.7 Thioether conjugation to prepare RNase–S

NTB-protected V118C S-protein (0.3 mg mL^{-1} in 50 mM sodium phosphate buffer, pH 2.7) was deprotected for 1 h at 37 °C with immobilized TCEP with gentle agitation. After this time, the yellow solution was desalted and exchanged into 50 mM Tris-HCl buffer, pH 7.0, containing NaCl (50 mM) and EDTA (10 mM) using a PD MiniTrap G-25 column (GE Healthcare Life Sciences). The solution of deprotected V118C S-protein was immediately combined with A5Dha S-peptide (12 equiv, 0.4 mg mL^{-1}) in 50 mM Tris-HCl buffer, pH 7.0, containing NaCl (50 mM) and EDTA (10 mM). The conjugation reaction was allowed to proceed overnight with gentle

agitation. The solution was concentrated to ≤500 μL using a Vivaspin® 20 centrifugal concentrator (5000 MWCO PES) and then purified by passage through a HiTrap® SP cation-exchange column (GE Healthcare Life Sciences) with 50 mM sodium acetate buffer, pH 5.0, and a gradient of 0.0–0.8 M NaCl over 80 min. The thioether-linked RNase–S product elutes at approximately 0.5 M NaCl.

It should be noted that a substantial amount of S-protein dimerization was observed under the conditions required for thioether conjugation. We were, however, able to isolate a small amount of the desired thioether-linked RNase–S conjugate. To confirm that the thioether-crosslink enhanced the stability of the conjugate in a reducing environment, we incubated thioether-linked RNase–S with glutathione and monitored the conjugate stability over 4 h at 37 °C.

A.6.8 Detection of RNase–S constructs

A.6.8.A Detection of Alkyne–RNase–S

Alkyne–RNase–S (1 mg mL^{-1} in PBS) was reacted with 5-carboxytetramethylrhodamine-azide (Azide-Fluor 545; 5-TAMRA–azide; Sigma–Aldrich product number 760757) in a CuAAC reaction. The reaction was performed by the addition of $CuSO_4$ (to 1 mM from a 100× stock solution in water), Tris((1-benzyl-4-triazolyl)methyl)amine (to 0.1 mM from a 100× stock solution in DMSO), Azide-fluor 545 (to 0.1 mM from a 100× stock solution in DMSO), and sodium ascorbate (to 2 mM from a 200× stock solution in water). The reaction mixture was agitated for 1 h before analysis by SDS–PAGE.

A.6.8.B Detection of Biotin–RNase–S

The biotin–RNase–S conjugate was subjected to SDS–PAGE under non-reducing conditions and then transferred onto a PVDF membrane using an iBlot® Gel Transfer Device. After protein transfer, the membrane was washed with Tris-buffered saline supplemented with Tween® 20 (0.1% v/v) (TBST, 5 × 5 min). The membrane was blocked with TBST containing non-fat dry milk (5% w/v) with gentle agitation at 4 °C overnight and then washed with TBST (5 × 5 min). The membrane was incubated with anti-biotin HRP-linked antibody, which was product number 7075S from Cell Signaling Technologies (Danvers, MA) and was diluted 1:1500 in TBST containing bovine serum albumin (3% w/v), with gentle agitation for 1 h. After incubation, the membrane was washed with TBST (5 × 5 min) and the blot was then developed using SuperSignal West Pico PLUS Chemiluminescent Substrate according to the manufacturer's protocol (Thermo Scientific).

A.6.9 Assays of ribonucleolytic activity

Assays were carried out in 0.10 M DEPC-treated,[270] OVS-free[271] MES-NaOH buffer, pH 6.0, containing NaCl (0.10 M), 6-FAM–dArU(dA)$_2$–6-TAMRA (200 nM), and a ribonuclease (5 nM).[35] Fluorescence intensity was measured with a M1000 microplate reader (Tecan) by monitoring at λ_{ex} = 493 ± 5 nm and λ_{em} = 515 ± 5 nm. Assays were performed in triplicate in a flat, black, 96-well plate (Corning). The value of k_{cat}/K_M was calculated with the equation:

$$\frac{k_{\mathrm{cat}}}{K_{\mathrm{M}}} = \frac{{}^{\Delta I}/_{\Delta t}}{(I_{\mathrm{max}} - I_0)[\mathrm{E}]} \tag{A.1}$$

where $\Delta I/\Delta t$ is the slope of the fluorescence vs. time data collected during the observation period and [E] is the concentration of enzyme. The background intensity of the substrate (I_0) was determined by measuring fluorescence intensity prior to the addition of enzyme. Product

intensity (I_{max}) was determined by measuring fluorescence intensity after the addition of enzyme to a final concentration of 5 μM, which was sufficient to cleave all of the substrate.

A.6.10 Determining protein thermostability by differential scanning fluorimetry

Differential scanning fluorimetry (DSF) was used to determine the thermostability of RNase–S in relation to RNase 1. SYPRO Orange protein gel stain (Supelco) was added to RNase–S and RNase 1 to achieve 1 mg mL^{-1} protein and 50× SYPRO Orange (vendor stock: 5000× in DMSO) in Dulbecco's phosphate-buffered saline (Gibco). The resulting samples were heated from 15–95 °C at 1 °C/min using a QuantStudio 7 Flex Real-Time PCR system (Applied Biosystems); fluorescence was monitored in real-time with λ_{ex} = 470 ± 15 nm and λ_{em} = 586 ± 10 nm. Data were processed with Protein Thermal Shift software (Applied Biosystems) using the Boltzmann fitting method. Values of T_m represent the temperature at which fluorescence reached 50% of its maximum.

Appendix B:
Primer: Determining the K_d of an Obligate Homodimeric Enzyme with Half-Site Reactivity from Progress Curves

Reprinted with permission from ref. 180.

Wralstad, E. C.; Sayers, J.; Raines, R. T. Bayesian inference elucidates the catalytic competency of the SARS-CoV-2 main protease $3CL^{pro}$. *Anal. Chem.* **2023**, *95*, 14981-14989.

B.1 Michaelis-Menten kinetics in terms of dimer

For the dimerization of two enzyme monomers M to form the dimeric species D, $2\mathrm{M} \rightleftharpoons \mathrm{D}$ and:

$$K_\mathrm{d} = \frac{[\mathrm{M}]^2}{[\mathrm{D}]} \tag{B.1}$$

$$E_\mathrm{T} = [\mathrm{M}] + 2[\mathrm{D}] \tag{B.2}$$

where E_T is the analytical enzyme concentration on a per-protomer basis.

For a homodimeric enzyme in which only the dimer exhibits catalytic competency and the enzyme has half-site reactivity (i.e., protomer active sites alternate catalytic function), the maximal enzyme reaction rate is

$$V_\mathrm{max} = k_\mathrm{cat}[\mathrm{D}] \tag{B.3}$$

From Equation B.3, the Michaelis-Menten equation can be written as:

$$\begin{aligned} v &= \frac{V_\mathrm{max}[\mathrm{S}]}{K_\mathrm{M} + [\mathrm{S}]} \\ &= \frac{k_\mathrm{cat}[\mathrm{D}][\mathrm{S}]}{K_\mathrm{M} + [\mathrm{S}]} \end{aligned} \tag{B.4}$$

where [S] is the free substrate concentration, k_cat is the enzyme's turnover rate, and K_M is the Michaelis constant. It is assumed that the total substrate concentration $[\mathrm{S}]_0 >> E_\mathrm{T}$, and thus $[\mathrm{S}] \approx [\mathrm{S}]_0$ (i.e., the amount of substrate bound in the substrate-enzyme complex is negligible at any given moment). We seek to recast Equation B.4 in terms of the total enzyme concentration E_T and the enzyme dissociation constant K_d.

B.2 Quantifying dimer concentration

Equation B.1 can be rewritten as

$$[\mathrm{M}]^2 - K_\mathrm{d}[\mathrm{D}] = 0$$

Inserting Equation B.2 results in:

$$(E_{\mathrm{T}} - 2[\mathrm{D}])^2 - K_{\mathrm{d}}[\mathrm{D}] = 0$$

which, upon expansion, is

$$4[\mathrm{D}]^2 - (4E_{\mathrm{T}} + K_{\mathrm{d}})[\mathrm{D}] + E_{\mathrm{T}}^2 = 0$$

Solving this quadratic expression in terms of [D] leads to:

$$[\mathrm{D}] = \frac{1}{8}\left(4E_{\mathrm{T}} + K_{\mathrm{d}} \pm \sqrt{K_{\mathrm{d}}^2 + 8E_{\mathrm{T}}K_{\mathrm{d}}}\right)$$

From mass balance (Equation B.2), we know that $[\mathrm{D}] \leq \frac{1}{2}E_{\mathrm{T}}$, which is the first term in the above expression for [D]. Thus, the additive solution for [D] is not physically realizable, and therefore the concentration of dimer is equal to:

$$[\mathrm{D}] = \frac{1}{8}\left(4E_{\mathrm{T}} + K_{\mathrm{d}} - \sqrt{K_{\mathrm{d}}^2 + 8E_{\mathrm{T}}K_{\mathrm{d}}}\right) \tag{B.5}$$

B.3 Reframing the Michaelis-Menten equation

With an analytical expression for [D] in hand, we can rewrite Equation B.4:

$$v = \frac{1}{8}\left(4E_{\mathrm{T}} + K_{\mathrm{d}} - \sqrt{K_{\mathrm{d}}^2 + 8E_{\mathrm{T}}K_{\mathrm{d}}}\right)\left(\frac{k_{\mathrm{cat}}[\mathrm{S}]}{K_{\mathrm{M}} + [\mathrm{S}]}\right) \tag{B.6}$$

The second term in parentheses depends only on intrinsic enzyme properties (k_{cat}, K_{M}) and the design of the assay ([S]), and it is entirely independent of E_{T}.

For the first term in parentheses, we consider two asymptotic conditions. In the first case of negligble enzyme, $E_{\mathrm{T}} << K_{\mathrm{d}} \rightarrow 4E_{\mathrm{T}} + K_{\mathrm{d}} \approx K_{\mathrm{d}}$ and $K_{\mathrm{d}}^2 + 8K_{\mathrm{d}}E_{\mathrm{T}} \approx K_{\mathrm{d}}^2$. As a result, v tends to zero, as is expected under conditions where nearly all of the enzyme is in the inactive monomeric state.

In the second case of abundant enzyme, $E_{\mathrm{T}} >> K_{\mathrm{d}}$, the parenthetical expression tends to $\frac{1}{2}E_{\mathrm{T}}$, as expected under conditions where all of the enzyme is driven to dimerize.

B.4 Pseudo-first order conditions

Focusing on the Michaelis-Menten equation in the form described by Equation B.4, with [D] described by Equation B.5, we consider the case where $[\mathrm{S}] << K_{\mathrm{M}}$. Under these conditions, the Michaelis-Menten equation simplifies to:

$$v = -\frac{\mathrm{d}[\mathrm{S}]}{\mathrm{d}t} = \left(\frac{k_{\mathrm{cat}}}{K_{\mathrm{M}}}\right)[\mathrm{D}][\mathrm{S}]$$

Integrating this equation to solve for [S] results in:

$$[\mathrm{S}] = Ce^{-\left(\frac{k_{\mathrm{cat}}}{K_{\mathrm{M}}}\right)[\mathrm{D}]t} \tag{B.7}$$

where C is a constant of integration that is solved for upon considering the precise format of the enzyme assay (see below).

B.5 Fluorescence-based assay monitoring

For the catalyzed reaction S $\rightarrow$ P, product concentration [P] is related to substrate concentration by $[\mathrm{P}] = [\mathrm{S}]_0 - [\mathrm{S}]$ (again, it is assumed that $E_\mathrm{T} << [\mathrm{S}]_0$, and therefore the concentration of enzyme-bound substrate may be neglected). Substituting Equation B.7 into this expression results in:

$$[\mathrm{P}] = [\mathrm{S}]_0 - Ce^{-\left(\frac{k_\mathrm{cat}}{K_\mathrm{M}}\right)[\mathrm{D}]t}$$

As expected, $\lim_{t\rightarrow\infty}[\mathrm{P}] = [\mathrm{S}]_0$. Invoking the initial condition that $[\mathrm{P}] = 0$ at $t = 0$ results in $C = [\mathrm{S}]_0$, and therefore

$$[\mathrm{P}] = [\mathrm{S}]_0\left(1 - e^{-\left(\frac{k_\mathrm{cat}}{K_\mathrm{M}}\right)[\mathrm{D}]t}\right) \tag{B.8}$$

For an enzyme assay in which a non-fluorescent substrate is converted to a fluorescent product (such as a FRET cleavage assay), the fluorescence intensity F is related to [P] by

$$F = F_0 + \alpha[\mathrm{P}] \tag{B.9}$$

where F_0 is the background fluorescence intensity and α is a proportionality constant. Given that the maximum fluorescence intensity F_max will be observed when all substrate has been converted to product (i.e., $[\mathrm{P}] = [\mathrm{S}]_0$),

$$F_\mathrm{max} = F_0 + \alpha[\mathrm{S}]_0 \tag{B.10}$$

Rearranging Equations B.9 and B.10 to isolate [P] and $[\mathrm{S}]_0$, respectively, then inserting the expressions into Equation B.8 results in:

$$F = F_0 + (F_\mathrm{max} - F_0)\left(1 - e^{-\left(\frac{k_\mathrm{cat}}{K_\mathrm{M}}\right)[\mathrm{D}]t}\right) \tag{B.11}$$

B.6 Determining $K_\mathbf{d}$ and $\frac{k_\mathbf{cat}}{K_\mathbf{M}}$

To start, we assay the homodimeric enzyme under a range of substrate and enzyme concentrations. The substrate concentration should span orders of magnitude with the goal of assaying concentrations less than the (unknown) K_M, and the enzyme concentration E_T should be significantly less than $[\mathrm{S}]_0$. Then, the fluorescence vs time data are fitted by nonlinear regression to a modified form of Equation B.11:

$$F = F_0 + (F_{\max} - F_0)\left(1 - e^{-r_x t}\right) \tag{B.12}$$

where

$$r_x = \left(\frac{k_\mathrm{cat}}{K_\mathrm{M}}\right)[\mathrm{D}]_x \tag{B.13}$$

and $[\mathrm{D}]_x$ is the concentration of enzyme dimer for a given analytical enzyme concentration $E_\mathrm{T} = E_{\mathrm{T},x}$ (x denoting that several concentrations are assayed in the complete dataset). Note that r_x is independent of $[\mathrm{S}]_0$, and for a single $E_{\mathrm{T},x}$ assayed across a range of $[\mathrm{S}]_0$, a consistent fitted r_x should be determined as long as the condition $[\mathrm{S}]_0 << K_\mathrm{M}$ is valid. In fact, by examining at what $[\mathrm{S}]_0$ the fitted r_x appears to begin deviating, we can qualitatively judge at what $[\mathrm{S}]_0$ the pseudo-first order approximation begins breaking down and therefore in what approximate regime the value of the unknown K_M must be.

Focusing on a single $[\mathrm{S}]_0$ for which the pseudo-first order approximation is valid, we determine the fitted r_x for all assayed $E_{\mathrm{T},x}$. Then, we let r_1 represent the fitted r_x for the largest enzyme concentration assayed, $E_{\mathrm{T},1}$. We normalize the fitted r_x values to r_1:

$$\begin{aligned}\frac{r_x}{r_1} &= \frac{\left(\frac{k_\mathrm{cat}}{K_\mathrm{M}}\right)[\mathrm{D}]_x}{\left(\frac{k_\mathrm{cat}}{K_\mathrm{M}}\right)[\mathrm{D}]_1} \\ &= \frac{[\mathrm{D}]_x}{[\mathrm{D}]_1}\end{aligned}$$

Inserting Equation B.5 into the above equation yields:

$$\frac{r_x}{r_1} = \frac{4E_{T,x} + K_d - \sqrt{K_d^2 + 8K_d E_{T,x}}}{4E_{T,1} + K_d - \sqrt{K_d^2 + 8K_d E_{T,1}}} \tag{B.14}$$

A plot of r_x/r_1 vs $E_{T,x}$ is a function with one fittable parameter, K_d (recall that $E_{T,1}$ is simply the largest enzyme concentration used and is therefore an empirically-defined constant). Thus, nonlinear regression of the data to Equation B.14 will determine K_d.

B.7 Summary

We assay an obligate homodimeric enzyme with half-site reactivity using a FRET pair-tagged substrate to produce a fluorescent product under a range of substrate concentrations $[S]_0$ and analytical enzyme concentrations $E_{T,x}$. We fit the data of each assay (i.e., one $[S]_0$ and $E_{T,x}$) to Equation B.12:

$$F = F_0 + (F_{max} - F_0)\left(1 - e^{-r_x t}\right) \tag{B.15}$$

This yields a collection of fitted r_x.

For a given $E_{T,x}$, we assess the fitted r_x across the assayed substrate concentrations and qualitatively decide where the assumption $[S]_0 << K_M$ breaks down. We restrict further analysis to the $[S]_0$ below this cutoff. Within each $[S]_0$ for which the pseudo-first order approximation appears valid, we normalize the fitted r_x to r_1, which is simply the fitted r_x at the largest enzyme concentration $E_{T,1}$. Plotting $\frac{r_x}{r_1}$ against $E_{T,x}$ and fitting the transformed data to Equation B.14

$$\frac{r_x}{r_1} = \frac{4E_{T,x} + K_d - \sqrt{K_d^2 + 8K_d E_{T,x}}}{4E_{T,1} + K_d - \sqrt{K_d^2 + 8K_d E_{T,1}}} \tag{B.16}$$

will yield K_d.

Appendix C:
Primer: Ribonucleolytic Activity of the Uncleaved and Cleaved Zymogen

This primer has been included as Supporting Information for:

Wralstad, E. C.; Raines, R. T. Sensitive detection of SARS-CoV-2 main protease $3CL^{pro}$ with an engineered ribonuclease zymogen.

C.1 Defining the reaction system

We have a ribonuclease (RNase) zymogen of concentration $[\mathrm{Z}]$ whose active site is occluded by a peptidic linker that is recognized and cleaved by $3\mathrm{CL}^{\mathrm{pro}}$. We let U represent the zymogen in its unactivated, pre-cleaved state and A represent the activated zymogen after linker cleavage. If Z represents the total population of zymogen regardless of its cleavage state, then by mass balance,

$$[\mathrm{U}] + [\mathrm{A}] = [\mathrm{Z}] \tag{C.1}$$

Both the unactivated and activated zymogen are capable of cleaving a generic RNA substrate S to yield a product P, albeit with different catalytic efficiencies:

$$\mathrm{U} + \mathrm{S} \rightarrow \mathrm{U} + \mathrm{P} \tag{C.2}$$

$$\mathrm{A} + \mathrm{S} \rightarrow \mathrm{A} + \mathrm{P} \tag{C.3}$$

The unactivated zymogen hydrolyzes the substrate with catalytic efficiency $\left(k_{\mathrm{cat}}/K_{\mathrm{M}}\right)_{\mathrm{U}}$, and the activated zymogen correspondingly hydrolyzes with efficiency $\left(k_{\mathrm{cat}}/K_{\mathrm{M}}\right)_{\mathrm{A}}$.

Letting R represent $3\mathrm{CL}^{\mathrm{pro}}$, the reaction equation for zymogen cleavage is:

$$\mathrm{R} + \mathrm{U} \rightarrow \mathrm{R} + \mathrm{A} \tag{C.4}$$

The protease likewise has a catalytic efficiency $\left(k_{\mathrm{cat}}/K_{\mathrm{M}}\right)_{\mathrm{R}}$ for cleavage of the zymogen linker. In the context of the protease cleavage reaction, the unactivated and activated zymogen are substrate and product, respectively, rather than catalysts.

C.2 Modeling the time evolution of zymogen activation

Proteolytic activation of the zymogen follows standard Michaelis-Menten kinetics, and the initial rate v_{R} of zymogen activation represented by chemical reaction C.4 is expressed by

$$v_{\mathrm{R}} = \frac{k_{\mathrm{cat,R}}[\mathrm{R}][\mathrm{U}]}{K_{\mathrm{M,R}} + [\mathrm{U}]} \tag{C.5}$$

where $k_{\mathrm{cat,R}}$ is the rate constant and $K_{\mathrm{M,R}}$ is the Michaelis constant for proteolytic cleavage. We make our first simplifying assumption, the validity of which will be justified shortly:

Assumption 1: the concentration of unactivated zymogen is always much less than $K_{\mathrm{M,R}}$

By invoking $[\mathrm{U}] << K_{\mathrm{M,R}}$, equation C.5 can be approximated and solved for $[\mathrm{U}]$ as a function of time t:

$$v_{\mathrm{R}} = -\frac{\mathrm{d}[\mathrm{U}]}{\mathrm{d}t} = \left(\frac{k_{\mathrm{cat}}}{K_{\mathrm{M}}}\right)_{\mathrm{R}} [\mathrm{R}][\mathrm{U}] = r[\mathrm{U}]$$

where we define the pseudo-first order rate constant $r = \left(\frac{k_{\mathrm{cat}}}{K_{\mathrm{M}}}\right)_{\mathrm{R}} [\mathrm{R}]$. Then:

$$\begin{aligned}
\frac{\mathrm{d}[\mathrm{U}]}{[\mathrm{U}]} &= -r\,\mathrm{d}t \\
\int \frac{1}{[\mathrm{U}]}\,\mathrm{d}[\mathrm{U}] &= -r \int \mathrm{d}t \\
\ln[\mathrm{U}] + B &= -rt + C
\end{aligned}$$

After consolidating the constants of integration B and C and cleaning up the equation, we are left with:

$$[\mathrm{U}] = De^{-rt}$$

for some constant D. With equation C.1, we can express the time-dependent concentration of activated zymogen by

$$[\mathrm{A}] = [\mathrm{Z}] - De^{-rt}$$

Applying the initial condition that $[\mathrm{A}] = 0$ at $t = 0$, we find that

$$[\mathrm{U}] = [\mathrm{Z}]e^{-rt} \tag{C.6}$$

$$[\mathrm{A}] = [\mathrm{Z}]\left(1 - e^{-rt}\right) \tag{C.7}$$

We note that the assumption $[\mathrm{U}] << K_{\mathrm{M,R}}$ is dependent on the protease of interest and the identity of the zymogen linker. In our 3CL$^{\mathrm{pro}}$-directed zymogen, we have installed the linker sequence ATLQSGNA. Our lab recently characterized 3CL$^{\mathrm{pro}}$ cleavage of the peptidic substrate R–E(EDANS)–ATLQSGNA–K(DABCYL)–R and found K_{M}=229 μM (at 25 °C in 50 mM HEPES, pH 7.5). It is likely that the zymogen linker (installed within a protein) is a poorer substrate than the analogous free peptide, thus $K_{\mathrm{M,R}}$ is believed to be higher for cleavage of the zymogen linker. The zymogen is utilized at concentrations well below this (e.g., high picomolar to low nanomolar), as it is itself being used as a catalyst in the cleavage of an RNA substrate. As a result, the approximation $[\mathrm{U}] << K_{\mathrm{M,R}}$ should be valid.

C.3 Characterizing ribonucleolytic activity over time

Both the unactivated and activated zymogen follow Michaelis-Menten kinetics in their cleavage of an RNA substrate (chemical reactions C.2 and C.3):

$$v_{\mathrm{U}} = \frac{k_{\mathrm{cat,U}}[\mathrm{U}][\mathrm{S}]}{K_{\mathrm{M,U}} + [\mathrm{S}]}$$

$$v_{\mathrm{A}} = \frac{k_{\mathrm{cat,A}}[\mathrm{A}][\mathrm{S}]}{K_{\mathrm{M,A}} + [\mathrm{S}]}$$

We now make our second simplifying assumption:

Assumption 2: the concentration of RNA substrate is always much less than both $K_{M,U}$ and $K_{M,A}$

This assumption can be rendered valid by the design of the zymogen activity assay, namely the initial substrate concentration $[S]_0$ that we choose to run the assay at. The K_M for cleavage of the short oligonucleotide substrate dArU$(dA)_2$ by wild-type RNase A (the bovine homologue of the zymogen's progenitor human RNase 1, 68% sequence identity) is ~22 μM. Typical ribonucleolytic activity assays are conducted with substrate in the mid- to high-nanomolar concentration range; thus, this assumption is valid.

Applying Assumption 2 results in:

$$
\begin{aligned}
v_U &= \left(\frac{k_{cat}}{K_M}\right)_U [U][S] = c_U[U][S] \\
v_A &= \left(\frac{k_{cat}}{K_M}\right)_A [A][S] = c_A[A][S] \\
v_T &= -\frac{d[S]}{dt} = v_U + v_A = (c_U[U] + c_A[A])[S]
\end{aligned}
$$

For notational simplicity, we denote the catalytic efficiencies of the unactivated and activated zymogen by $c_U = \left(\frac{k_{cat}}{K_M}\right)_U$ and $c_A = \left(\frac{k_{cat}}{K_M}\right)_A$, respectively. As [U] and [A] are defined by equations C.6 and C.7, we can express the rate of RNA turnover as:

$$
\frac{d[S]}{dt} = -[Z]\left(c_U e^{-rt} + c_A\left(1 - e^{-rt}\right)\right)[S] \tag{C.8}
$$

Note that if $[R] = 0$ (*i.e.*, no protease is present to cleave the zymogen), then equation C.8 simplifies to the general expression for pseudo-first order enzyme kinetics, as would be expected when the zymogen remains unactivated at all times. Rearranging terms and integrating results in:

$$\int \frac{1}{[\mathrm{S}]}\,\mathrm{d}[\mathrm{S}] = -[\mathrm{Z}]c_\mathrm{U}\int e^{-rt}\,\mathrm{d}t - [\mathrm{Z}]c_\mathrm{A}\int \left(1-e^{-rt}\right)\,\mathrm{d}t$$
$$\ln[\mathrm{S}] + J = \left(\frac{[\mathrm{Z}]}{r}\right)c_\mathrm{U}e^{-rt} - [\mathrm{Z}]c_\mathrm{A}t - \left(\frac{[\mathrm{Z}]}{r}\right)c_\mathrm{A}e^{-rt} + K$$

for some constants of integration J and K. Note that this expression for $[\mathrm{S}]$ requires that $r \neq 0$, which is to say that the expression is only applicable when the protease is present (for if $[\mathrm{R}] = 0$, then $r = 0$). After consolidating the constants of integration and cleaning up the equation, we are left with

$$[\mathrm{S}] = L\,\exp\left(\frac{[\mathrm{Z}]}{r}\left(c_\mathrm{U}-c_\mathrm{A}\right)e^{-rt} - [\mathrm{Z}]c_\mathrm{A}t\right) \tag{C.9}$$

for some constant L. If we apply the initial condition that at $[\mathrm{S}] = [\mathrm{S}]_0$ at $t = 0$, we learn that

$$L = [\mathrm{S}]_0\,\exp\left(-\frac{[\mathrm{Z}]}{r}\left(c_\mathrm{U}-c_\mathrm{A}\right)\right)$$

Inserting this into equation C.9 results in:

$$\begin{aligned}
[\mathrm{S}] &= [\mathrm{S}]_0\,\exp\left(-\frac{[\mathrm{Z}]}{r}\left(c_\mathrm{U}-c_\mathrm{A}\right)\right)\exp\left(\frac{[\mathrm{Z}]}{r}\left(c_\mathrm{U}-c_\mathrm{A}\right)e^{-rt} - [\mathrm{Z}]c_\mathrm{A}t\right)\\
&= [\mathrm{S}]_0\,\exp\left(\frac{[\mathrm{Z}]}{r}\left(c_\mathrm{U}-c_\mathrm{A}\right)\left(e^{-rt}-1\right) - [\mathrm{Z}]c_\mathrm{A}t\right)\\
&= [\mathrm{S}]_0\,\exp\left(\frac{[\mathrm{Z}]}{r}\left(c_\mathrm{A}-c_\mathrm{U}\right)\left(1-e^{-rt}\right) - [\mathrm{Z}]c_\mathrm{A}t\right)
\end{aligned}$$

Finally, if we make a common assumption of enzyme assays that the substrate concentration greatly exceeds that of the enzyme (i.e., $[\mathrm{S}] >> [\mathrm{Z}]$), we can neglect the concentration of zymogen-bound substrate and say that the product concentration is $[\mathrm{P}] = [\mathrm{S}]_0 - [\mathrm{S}]$. Thus,

$$[\mathrm{P}] = [\mathrm{S}]_0\left(1-\exp\left(\frac{[\mathrm{Z}]}{r}\left(c_\mathrm{A}-c_\mathrm{U}\right)\left(1-e^{-rt}\right) - [\mathrm{Z}]c_\mathrm{A}t\right)\right) \tag{C.10}$$

C.4 Monitoring ribonuclease activity by fluorescence

We monitor ribonucleolytic activity using a FRET-tagged, chimeric oligonucleotide substrate 6-FAM–dArU(dA)$_2$–6-TAMRA in which the core tetranucleotide has one scissile bond (rU↓dA); substrate cleavage liberates the donor FAM from the quencher TAMRA, inducing fluorescence and enabling monitoring product formation. The relationship between fluorescence intensity F and product concentration [P] is linear:

$$\begin{aligned} F &= \alpha[\mathrm{P}] + F_0 \\ \Rightarrow [\mathrm{P}] &= \alpha^{-1}(F - F_0) \end{aligned}$$

where α is a proportionality constant (dictated by experimental parameters such as instrument gain) and F_0 is the background fluorescence intensity. Given that the maximum concentration of product $[\mathrm{P}]_{\max} = [\mathrm{S}]_0$, there is a corresponding maximum fluorescence intensity:

$$\begin{aligned} F_{\max} &= \alpha[\mathrm{P}]_{\max} + F_0 = \alpha[\mathrm{S}]_0 + F_0 \\ \Rightarrow [\mathrm{S}]_0 &= \alpha^{-1}(F_{\max} - F_0) \end{aligned}$$

Inserting the above expressions into equation C.10 allows us to model product fluorescence as a function of time:

$$F = F_0 + (F_{\max} - F_0)\left(1 - \exp\left(\frac{[\mathrm{Z}]}{r}(c_{\mathrm{A}} - c_{\mathrm{U}})\left(1 - e^{-rt}\right) - [\mathrm{Z}]c_{\mathrm{A}}t\right)\right) \tag{C.11}$$

C.5 Summary

We have an RNase zymogen of concentration $[\mathrm{Z}]$ whose active site is occluded by a peptidic linker that is recognized and cleaved by $3\mathrm{CL}^{\mathrm{pro}}$. If at $t = 0$ we mix the zymogen with $3\mathrm{CL}^{\mathrm{pro}}$, then the concentration of uncleaved/unactivated and cleaved/activated zymogen at any time t is determined by

$$[\mathrm{U}] = [\mathrm{Z}]e^{-rt} \tag{C.12}$$

$$[\mathrm{A}] = [\mathrm{Z}]\left(1 - e^{-rt}\right) \tag{C.13}$$

where $r = \left(\frac{k_{\mathrm{cat}}}{K_{\mathrm{M}}}\right)_{\mathrm{R}} [\mathrm{R}]$.

If, at the same time that we mix the zymogen and $3\mathrm{CL}^{\mathrm{pro}}$, we mix in the oligonucleotide RNase substrate 6-FAM–dArU$(\mathrm{dA})_2$–6-TAMRA, then the fluorescence intensity of the product is governed by

$$F = F_0 + (F_{\max} - F_0)\left(1 - \exp\left(\frac{[\mathrm{Z}]}{r}(c_{\mathrm{A}} - c_{\mathrm{U}})\left(1 - e^{-rt}\right) - [\mathrm{Z}]c_{\mathrm{A}}t\right)\right) \tag{C.14}$$

REFERENCES

(1) Numata, K. How to define and study structural proteins as biopolymer materials. *Polym. J. (Tokyo, Jpn.)* **2020**, *52*, 1043-1056.

(2) Nair, A.; Chauhan, P.; Saha, B.; Kubatzky, K. F. Conceptual evolution of cell signaling. *Int. J. Mol. Sci.* **2019**, *20*, 3292.

(3) Porto de Souza Vandenberghe, L.; Karp, S. G.; Binder Pagnoncelli, M. G.; von Linsingen Tavares, M.; Libardi Junior, N.; Valladares Diestra, K.; Viesser, J. A.; Soccol, C. R. Classification of enzymes and catalytic properties. In *Biomass, Biofuels, Biochemicals*, Singh, S. P., Pandey, A., Singhania, R. R., Larroche, C., Li, Z. Eds.; Elsevier, 2020; pp 11-30.

(4) Keller, M. A.; Piedrafita, G.; Ralser, M. The widespread role of non-enzymatic reactions in cellular metabolism. *Curr. Opin. Biotechnol.* **2015**, *34*, 153-161.

(5) Doig, A. J. Frozen, but no accident – why the 20 standard amino acids were selected. *FEBS J.* **2017**, *284*, 1296-1305.

(6) Zwanzig, R.; Szabo, A.; Bagchi, B. Levinthal's paradox. *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 20-22.

(7) Uversky, V. N. Posttranslational Modification. In *Brenner's Encyclopedia of Genetics*, 2 ed.; Maloy, S., Hughes, K. Eds.; Academic Press, 2013; pp 425-430.

(8) Walsh, C. *Posttranslational Modification of Proteins: Expanding Nature's Inventory*; Roberts and Co. Publishers, 2006.

(9) Cecarini, V.; Gee, J.; Fioretti, E.; Amici, M.; Angeletti, M.; Eleuteri, A. M.; Keller, J. N. Protein oxidation and cellular homeostasis: emphasis on metabolism. *Biochim. Biophys. Acta* **2007**, *1773*, 93-104.

(10) Engqvist, M. K. M.; Rabe, K. S. Applications of protein engineering and directed evolution in plant research. *Plant Physiol.* **2019**, *179*, 907-917.

(11) Tobin, H. P.; Richards, H. D.; Callender, A. R.; Wilson, J. C. Protein engineering: a new frontier for biological therapeutics. *Curr. Drug Metab.* **2014**, *15*, 743-756.

(12) Wen, F.; Nair, N. U.; Zhao, H. Protein engineering in designing tailored enzymes and microorganisms for biofuels production. *Curr. Opin. Biotechnol.* **2009**, *20*, 412-419.

(13) Liu, R.; Liang, L.; Lacerda, M. P.; Freed, E. F.; Eckert, C. A. Advances in protein engineering and its application in synthetic biology. In *New Frontiers and Applications of Synthetic Biology*, Singh, V. Ed.; Academic Press, 2022; pp 147-158.

(14) Leland, P. A.; Schultz, L. W.; Kim, B.-M.; Raines, R. T. Ribonuclease A variants with potent cytotoxic activity. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 10407-10412.

(15) Bretscher, L. E.; Abel, R. L.; Raines, R. T. A Ribonuclease A variant with low catalytic activity but high cytotoxicity. *J. Biol. Chem.* **2000**, *275*, 9893-9896.

(16) Leland, P. A.; Staniszewski, K. E.; Kim, B.-M.; Raines, R. T. Endowing human pancreatic ribonuclease with toxicity for cancer cells. *J. Biol. Chem.* **2001**, *276*, 43095-43102.

(17) Hoang, T. T.; Tanrikulu, C.; Vatland, Q. A.; Hoang, T. M.; Raines, R. T. A human Ribonuclease variant and ERK-pathway inhibitors exhibit highly synergistic toxicity for cancer cells. *Mol. Cancer Ther.* **2018**, *17*, 2622-2632.

(18) Mix, K. A.; Lomax, J. E.; Raines, R. T. Cytosolic delivery of proteins by bioreversible esterification. *J. Am. Chem. Soc.* **2017**, *139*, 14396-14398.

(19) Ressler, V. T.; Mix, K. A.; Raines, R. T. Esterification delivers a functional enzyme into a human cell. *ACS Chem. Biol.* **2019**, *14*, 599-602.
(20) Watkins, R. W.; Arnold, U.; Raines, R. T. Ribonuclease S redux. *Chem. Commun.* **2011**, *47*, 973-975.
(21) Jacob, F. o. *The Possible and the Actual*; University of Washington Press, 1982.
(22) Raines, R. T. Ribonuclease A. *Chem. Rev.* **1998**, *98*, 1045-1065.
(23) Jones, W. The action of boiled pancreas extract on yeast nucleic acid. *Am. J. Physiol.* **1920**, *52*, 203-207.
(24) Kunitz, M. Isolation from beef pancreas of a crystalline protein possessing ribonuclease activity. *Science* **1939**, *90*, 112-113.
(25) Findlay, D.; Herries, D. G.; Mathias, A. P.; Rabin, B. R.; Ross, C. A. The active site and mechanism of action of bovine pancreatic Ribonuclease. *Nature* **1961**, *190*, 781-784.
(26) Anfinsen, C. B. Principles that govern the folding of protein chains. *Science* **1973**, *181*, 223-230.
(27) Moore, S.; Stein, W. H. Chemical structures of pancreatic ribonuclease and deoxyribonuclease. *Science* **1973**, *180*, 458-464.
(28) Merrifield, B. Solid phase synthesis. *Science* **1986**, *232*, 341-347.
(29) Richards, F. M. Linderstrøm-Lang and the Carlsberg laboratory: the view of a postdoctoral fellow in 1954. *Protein Sci.* **1992**, *1*, 1721-1730.
(30) Deavin, A.; Mathias, A. P.; Rabin, B. R. Mechanism of action of bovine pancreatic ribonuclease. *Biochem. J.* **1966**, *101*, 14C-16C. (acccessed 10/9/2023).
(31) Nogués, M. V.; Vilanova, M.; Cuchillo, C. M. Bovine pancreatic ribonuclease A as a model of an enzyme with multiple substrate binding sites. *Biochim. Biophys. Acta, Protein Struct. Mol. Enzymol.* **1995**, *1253*, 16-24.
(32) DelCardayré, S. B.; Raines, R. T. Structural determinants of enzymatic processivity. *Biochemistry* **1994**, *33*, 6031-6037.
(33) Cuchillo, C. M.; Nogués, M. V.; Raines, R. T. Bovine pancreatic ribonuclease: fifty years of the first enzymatic reaction mechanism. *Biochemistry* **2011**, *50*, 7835-7841.
(34) Park, C.; Raines, R. T. Catalysis by Ribonuclease A is limited by the rate of substrate association. *Biochemistry* **2003**, *42*, 3509-3518.
(35) Kelemen, B. R.; Klink, T. A.; Behlke, M. A.; Eubanks, S. R.; Leland, P. A.; Raines, R. T. Hypersensitive substrate for ribonucleases. *Nucleic Acids Res.* **1999**, *27*, 3696-3701.
(36) Park, C.; Kelemen, B. R.; Klink, T. A.; Sweeney, R. Y.; Behlke, M. A.; Eubanks, S. R.; Raines, R. T. Fast, facile, hypersensitive assays for ribonucleolytic activity. In *Methods in Enzymology*, Vol. 341; Academic Press, 2001; pp 81-94.
(37) Park, C.; Raines, R. T. Quantitative analysis of the effect of salt concentration on enzymatic catalysis. *J. Am. Chem. Soc.* **2001**, *123*, 11472-11479.
(38) Fisher, B. M.; Ha, J.-H.; Raines, R. T. Coulombic forces in protein-RNA interactions: binding and cleavage by Ribonuclease A and variants at Lys7, Arg10, and Lys66. *Biochemistry* **1998**, *37*, 12121-12132.
(39) Wlodawer, A.; Svensson, L. A.; Sjoelin, L.; Gilliland, G. L. Structure of phosphate-free ribonuclease A refined at 1.26 Å. *Biochemistry* **1988**, *27*, 2705-2717.
(40) Lomax, J. E.; Eller, C. H.; Raines, R. T. Comparative functional analysis of ribonuclease 1 homologs: molecular insights into evolving vertebrate physiology. *Biochem. J.* **2017**, *474*, 2219-2233.

(41) Beintema, J. J.; Kleineidam, R. G. The ribonuclease A superfamily: general discussion. *Cell. Mol. Life Sci.* **1998**, *54*, 825-832.
(42) Koczera, P.; Martin, L.; Marx, G.; Schuerholz, T. The Ribonuclease A superfamily in humans: canonical RNases as the buttress of innate immunity. *Int. J. Mol. Sci.* **2016**, *17*, 1278.
(43) Prats-Ejarque, G.; Lu, L.; Salazar, V. A.; Moussaoui, M.; Boix, E. Evolutionary trends in RNA base selectivity within the RNase A superfamily. *Front. Pharmacol.* **2019**, *10*, 1170.
(44) Lu, L.; Li, J.; Moussaoui, M.; Boix, E. Immune modulation by human secreted RNases at the extracellular space. *Front. Immunol.* **2018**, *9*, 1012.
(45) Weil, L.; Buchert, A. R.; Maher, J. Photoöxidation of crystalline lysozyme in the presence of methylene blue and its relation to enzymatic activity. *Arch. Biochem. Biophys.* **1952**, *40*, 245-252.
(46) Weil, L.; Gordon, W. G.; Buchert, A. R. Photoöxidation of amino acids in the presence of methylene blue. *Arch. Biochem. Biophys.* **1951**, *33*, 90-109.
(47) Weil, L.; James, S.; Buchert, A. R. Photooxidation of crystalline chymotrypsin in the presence of methylene blue. *Arch. Biochem. Biophys.* **1953**, *46*, 266-278.
(48) Weil, L.; Seibles, T. S. Photooxidation of crystalline ribonuclease in the presence of methylene blue. *Arch. Biochem. Biophys.* **1955**, *54*, 368-377.
(49) Weil, L.; Buchert, A. R. Photoöxidation of crystalline β-lactoglobulin in the presence of methylene blue. *Arch. Biochem. Biophys.* **1951**, *34*, 1-15.
(50) Gundlach, H. G.; Stein, W. H.; Moore, S. The nature of the amino acid residues involved in the inactivation of ribonuclease by iodoacetate. *J. Biol. Chem.* **1959**, *234*, 1754-1760.
(51) Gundlach, H. G.; Moore, S.; Stein, W. H. The reaction of iodoacetate with methionine. *J. Biol. Chem.* **1959**, *234*, 1761-1764.
(52) Neumann, N. P.; Moore, S.; Stein, W. H. Modification of the methionine residues in ribonuclease. *Biochemistry* **1962**, *1*, 68-75.
(53) Fruchter, R. G.; Crestfield, A. M. The specific alkylation by iodoacetamide of Histidine-12 in the active site of Ribonuclease. *J. Biol. Chem.* **1967**, *242*, 5807-5812.
(54) Brosnan, J. T.; Brosnan, M. E. The sulfur-containing amino acids: an overview. *J. Nutr.* **2006**, *136*, 1636S-1640S.
(55) Lucas-Lenard, J.; Lipmann, F. Protein biosynthesis. *Annu. Rev. Biochem.* **1971**, *40*, 409-448.
(56) Echols, N.; Harrison, P.; Balasubramanian, S.; Luscombe, N. M.; Bertone, P.; Zhang, Z.; Gerstein, M. Comprehensive analysis of amino acid and nucleotide composition in eukaryotic genomes, comparing genes and pseudogenes. *Nucleic Acids Res.* **2002**, *30*, 2515-2523.
(57) Tsuji, J.; Nydza, R.; Wolcott, E.; Mannor, E.; Moran, B.; Hesson, G.; Arvidson, T.; Howe, K.; Hayes, R.; Ramirez, M.; Way, M. The frequencies of amino acids encoded by genomes that utilize standard and nonstandard genetic codes. *Bios* **2010**, *81*, 22-31.
(58) Xiao, Q.; Zhang, F.; Nacev, B. A.; Liu, J. O.; Pei, D. Protein N-terminal processing: substrate specificity of *Escherichia coli* and human methionine aminopeptidases. *Biochemistry* **2010**, *49*, 5588-5599.
(59) Kaleta, C.; Schäuble, S.; Rinas, U.; Schuster, S. Metabolic costs of amino acid and protein production in *Escherichia coli*. *Biotechnol. J.* **2013**, *8*, 1105-1114.
(60) Aledo, J. C. Methionine in proteins: The Cinderella of the proteinogenic amino acids. *Protein Sci.* **2019**, *28*, 1785-1796.

(61) Levine, R. L.; Mosoni, L.; Berlett, B. S.; Stadtman, E. R. Methionine residues as endogenous antioxidants in proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1996**, *93*, 15036-15040.
(62) Brot, N.; Weissbach, H. Biochemistry and physiological role of methionine sulfoxide residues in proteins. *Arch. Biochem. Biophys.* **1983**, *223*, 271-281.
(63) Stadtman, E. R. Cyclic oxidation and reduction of methionine residues of proteins in antioxidant defense and cellular regulation. *Arch. Biochem. Biophys.* **2004**, *423*, 2-5.
(64) Levine, R. L.; Moskovitz, J.; Stadtman, E. R. Oxidation of methionine in proteins : roles in antioxidant defense and cellular regulation. *IUBMB Life* **2000**, *50*, 301-307.
(65) Schöneich, C. Methionine oxidation by reactive oxygen species: reaction mechanisms and relevance to Alzheimer's disease. *Biochim. Biophys. Acta* **2005**, *1703*, 111-119.
(66) Kim, N. H.; Jeong, M. S.; Choi, S. Y.; Kang, J. H. Oxidative modification of Cytochrome C by hydrogen peroxide. *Mol. Cells* **2006**, *22*, 220-227.
(67) Ghesquière, B.; Jonckheere, V.; Colaert, N.; Van Durme, J.; Timmerman, E.; Goethals, M.; Schymkowitz, J.; Rousseau, F.; Vandekerckhove, J.; Gevaert, K. Redox proteomics of protein-bound methionine oxidation. *Mol. Cell. Proteomics* **2011**, *10*, M110.006866.
(68) Bergès, J.; de Oliveira, P.; Fourré, I.; Houée-Levin, C. The one-electron reduction potential of methionine-containing peptides depends on the sequence. *J. Phys. Chem. B* **2012**, *116*, 9352-9362.
(69) Xu, K.; Uversky, V. N.; Xue, B. Local flexibility facilitates oxidization of buried methionine residues. *Protein Pept. Lett.* **2012**, *19*, 688-697.
(70) Varadarajan, S.; Yatin, S.; Aksenova, M.; Butterfield, D. A. Review: Alzheimer's amyloid β-peptide-associated free radical oxidative stress and neurotoxicity. *J. Struct. Biol.* **2000**, *130*, 184-208.
(71) Maltsev, A. S.; Chen, J.; Levine, R. L.; Bax, A. Site-specific interaction between α-Synuclein and membranes probed by NMR-observed methionine oxidation rates. *J. Am. Chem. Soc.* **2013**, *135*, 2943-2946.
(72) Carp, H.; Miller, F.; Hoidal, J. R.; Janoff, A. Potential mechanism of emphysema: alpha 1-proteinase inhibitor recovered from lungs of cigarette smokers contains oxidized methionine and has decreased elastase inhibitory capacity. *Proc. Natl. Acad. Sci. U. S. A.* **1982**, *79*, 2041-2045.
(73) Lei, K.-F.; Wang, Y.-F.; Zhu, X.-Q.; Lu, P.-C.; Sun, B.-S.; Jia, H.-L.; Ren, N.; Ye, Q.-H.; Sun, H.-C.; Wang, L.; et al. Identification of MSRA gene on chromosome 8p as a candidate metastasis suppressor for human hepatitis B virus-positive hepatocellular carcinoma. *BMC Cancer* **2007**, *7*, 172.
(74) De Luca, A.; Sanna, F.; Sallese, M.; Ruggiero, C.; Grossi, M.; Sacchetta, P.; Rossi, C.; De Laurenzi, V.; Di Ilio, C.; Favaloro, B. Methionine sulfoxide reductase A down-regulation in human breast cancer cells results in a more aggressive phenotype. *Proc. Natl. Acad. Sci. U. S. A.* **2010**, *107*, 18628-18633.
(75) Ahmed, Z. M.; Yousaf, R.; Lee, B. C.; Khan, S. N.; Lee, S.; Lee, K.; Husnain, T.; Rehman, A. U.; Bonneux, S.; Ansar, M.; et al. Functional null mutations of MSRB3 encoding Methionine Sulfoxide Reductase are associated with human deafness DFNB74. *Am. J. Hum. Genet.* **2011**, *88*, 19-29.
(76) Lim, J. M.; Kim, G.; Levine, R. L. Methionine in proteins: it's not just for protein initiation anymore. *Neurochem. Res.* **2019**, *44*, 247-257.

(77) Newberry, R. W.; Raines, R. T. Secondary forces in protein folding. *ACS Chem. Biol.* **2019**, *14*, 1677-1686.
(78) Meyer, E. A.; Castellano, R. K.; Diederich, F. Interactions with aromatic rings in chemical and biological recognition. *Angew. Chem., Int. Ed.* **2003**, *42*, 1210-1250.
(79) Valley, C. C.; Cembran, A.; Perlmutter, J. D.; Lewis, A. K.; Labello, N. P.; Gao, J.; Sachs, J. N. The methionine-aromatic motif plays a unique role in stabilizing protein structure. *J. Biol. Chem.* **2012**, *287*, 34979-34991.
(80) Salonen, L. M.; Ellermann, M.; Diederich, F. Aromatic rings in chemical and biological recognition: energetics and structures. *Angew. Chem., Int. Ed.* **2011**, *50*, 4808-4842.
(81) Pal, D.; Chakrabarti, P. Non-hydrogen bond interactions involving the methionine sulfur atom. *J. Biomol. Struct. Dyn.* **2001**, *19*, 115-128.
(82) Aledo, J. C.; Cantón, F. R.; Veredas, F. J. Sulphur atoms from methionines interacting with aromatic residues are less prone to oxidation. *Sci. Rep.* **2015**, *5*, 16955.
(83) Michod, R. E.; Viossat, Y.; Solari, C. A.; Hurand, M.; Nedelcu, A. M. Life-history evolution and the origin of multicellularity. *J. Theor. Biol.* **2006**, *239*, 257-272.
(84) Niu, D. K.; Chen, J.-K. Evolutionary advantages of cell specialization: save and protect DNA. *J. Theor. Biol.* **1997**, *187*, 39-43.
(85) Kay, J.; Kassell, B. The autoactivation of trypsinogen. *J. Biol. Chem.* **1971**, *246*, 6661-6665. (acccessed 2023/11/03).
(86) Stroud, R. M.; Kossiakoff, A. A.; Chambers, J. L. Mechanisms of zymogen activation. *Annu. Rev. Biophys. Bioeng.* **1977**, *6*, 177-193.
(87) Neurath, H. The activation of zymogens. In *Advances in Protein Chemistry*, Anfinsen, C. B., Anson, M. L., Bailey, K., Edsall, J. T. Eds.; Vol. 12; Academic Press, 1957; pp 319-386.
(88) Shi, Y. Caspase activation, inhibition, and reactivation: a mechanistic view. *Protein Sci.* **2004**, *13*, 1979-1987.
(89) Nemerson, Y.; Furie, B.; Jackson, C. M. Zymogens and cofactors of blood coagulation. *Crit. Rev. Biochem.* **1980**, *9*, 45-85.
(90) Ivanciu, L.; Toso, R.; Margaritis, P.; Pavani, G.; Kim, H.; Schlachterman, A.; Liu, J.-H.; Clerin, V.; Pittman, D. D.; Rose-Miranda, R.; et al. A zymogen-like factor Xa variant corrects the coagulation defect in hemophilia. *Nat. Biotechnol.* **2011**, *29*, 1028-1033.
(91) Ng, D. Y. W.; Arzt, M.; Wu, Y.; Kuan, S. L.; Lamla, M.; Weil, T. Constructing hybrid protein zymogens through protective dendritic assembly. *Angew. Chem., Int. Ed.* **2014**, *53*, 324-328.
(92) Shapira, A.; Gal-Tanamy, M.; Nahary, L.; Litvak-Greenfeld, D.; Zemel, R.; Tur-Kaspa, R.; Benhar, I. Engineered toxins "zymoxins" are activated by the HCV NS3 protease by removal of an inhibitory protein domain. *PLoS ONE* **2011**, *6*, e15916.
(93) Yu, Y.; Lutz, S. Circular permutation: a different way to engineer enzyme structure and function. *Curr. Trends Biotechnol.* **2011**, *29*, 18-25. (acccessed 2023/10/14).
(94) Plainkum, P.; Fuchs, S. M.; Wiyakrutta, S.; Raines, R. T. Creation of a zymogen. *Nat. Struct. Biol.* **2003**, *10*, 115-119.
(95) Johnson, R. J.; Lin, S. R.; Raines, R. T. A ribonuclease zymogen activated by the NS3 protease of the hepatitis C virus. *FEBS J.* **2006**, *273*, 5457-5465.
(96) Turcotte, R. F.; Raines, R. T. Design and characterization of an HIV-specific Ribonuclease zymogen. *AIDS Res. Hum. Retroviruses* **2008**, *24*, 1357-1363.
(97) Windsor, I. W.; Graff, C. J.; Raines, R. T. Circular zymogens of human Ribonuclease 1. *Protein Sci.* **2019**, *28*, 1713-1719.

(98) Windsor, I. W.; Dudley, D. M.; O'Connor, D. H.; Raines, R. T. Ribonuclease zymogen induces cytotoxicity upon HIV-1 infection. *AIDS Res. Ther.* **2021**, *18*, 77.
(99) Fothergill-Gilmore, L. A. Protein fragmentation. In *Protein Biotechnology: Isolation, Characterization, and Stabilization*, Franks, F. Ed.; Humana Press, 1993; pp 237-249.
(100) Vlasak, J.; Ionescu, R. Fragmentation of monoclonal antibodies. *mAbs* **2011**, *3*, 253-263.
(101) Sierra, S.; Toneatti, R.; González-Maeso, J. Class A GPCR oligomerization: reasons of controversy. In *GPCRs*, Jastrzebska, B., Park, P. S. H. Eds.; Academic Press, 2020; pp 121-140.
(102) Michnick, S. W.; Ear, P. H.; Manderson, E. N.; Remy, I.; Stefan, E. Universal strategies in research and drug discovery based on protein-fragment complementation assays. *Nat. Rev. Drug Discovery* **2007**, *6*, 569-582.
(103) Richards, F. M.; Vithayathil, P. J. The preparation of subtilisin-modified ribonuclease and the separation of the peptide and protein components. *J. Biol. Chem.* **1959**, *234*, 1459-1465.
(104) Kim, E. E.; Varadarajan, R.; Wyckoff, H. W.; Richards, F. M. Refinement of the crystal structure of ribonuclease S. Comparison with and between the various ribonuclease A structures. *Biochemistry* **1992**, *31*, 12304-12314.
(105) Nadig, G.; Ratnaparkhi, G. S.; Varadarajan, R.; Vishveshwara, S. Dynamics of ribonuclease A and ribonuclease S: computational and experimental studies. *Protein Sci.* **1996**, *5*, 2104-2114.
(106) Thomson, J.; Ratnaparkhi, G. S.; Varadarajan, R.; Sturtevant, J. M.; Richards, F. M. Thermodynamic and structural consequences of changing a sulfur atom to a methylene group in the M13Nle mutation in ribonuclease-S. *Biochemistry* **1994**, *33*, 8587-8593.
(107) Dunn, B. M.; Dibello, C.; Kirk, K. L.; Cohen, L. A.; Chaiken, I. M. Synthesis, purification, and properties of a semisynthetic ribonuclease S incorporating 4-fluoro-L-histidine at position 12. *J. Biol. Chem.* **1974**, *249*, 6295-6301.
(108) Kim, J. S.; Raines, R. T. Ribonuclease S-peptide as a carrier in fusion proteins. *Protein Sci.* **1993**, *2*, 348-356.
(109) Kim, J.-S.; Raines, R. T. Peptide tags for a dual affinity fusion system. *Anal. Biochem.* **1994**, *219*, 165-166.
(110) Raines, R. T.; McCormick, M.; Van Oosbree, T. R.; Mierendorf, R. C. The S·tag fusion system for protein purification. In *Methods in Enzymology*, Vol. 326; Academic Press, 2000; pp 362-376.
(111) Wildemann, D.; Schiene-Fischer, C.; Aumüller, T.; Bachmann, A.; Kiefhaber, T.; Lücke, C.; Fischer, G. A nearly isosteric photosensitive amide-backbone substitution allows enzyme activity switching in ribonuclease S. *J. Am. Chem. Soc.* **2007**, *129*, 4910-4918.
(112) Backer, M. V.; Gaynutdinov, T. I.; Aloise, R.; Przekop, K.; Backer, J. M. Engineering S-protein fragments of bovine ribonuclease A for targeted drug delivery. *Protein Expression Purif.* **2002**, *26*, 455-461.
(113) Gaidamakova, E. K.; Backer, M. V.; Backer, J. M. Molecular vehicle for target-mediated delivery of therapeutics and diagnostics. *J. Controlled Release* **2001**, *74*, 341-347.
(114) Liang, Y.; Furukawa, H.; Sakamoto, K.; Inaba, H.; Matsuura, K. Anticancer activity of reconstituted ribonuclease S-decorated artificial viral capsid. *ChemBioChem* **2022**, *23*.
(115) Shah, N. H.; Muir, T. W. Inteins: nature's gift to protein chemists. *Chem. Sci.* **2014**, *5*, 446-461.

(116) Gogarten, J. P.; Senejani, A. G.; Zhaxybayeva, O.; Olendzenski, L.; Hilario, E. Inteins: structure, function, and evolution. *Annu. Rev. Microbiol.* **2002**, *56*, 263-287.
(117) Wang, H.; Wang, L.; Zhong, B.; Dai, Z. Protein splicing of inteins: a powerful tool in synthetic biology. *Front. Bioeng. Biotechnol.* **2022**, *10*, Mini Review.
(118) Shah, N. H.; Eryilmaz, E.; Cowburn, D.; Muir, T. W. Extein residues play an intimitate role in the rate-limiting step of protein trans-splicing. *J. Am. Chem. Soc.* **2013**, *135*, 5839-5847.
(119) Shah, N. H.; Eryilmaz, E.; Cowburn, D.; Muir, T. W. Naturally split inteins assemble through a "capture and collapse" mechanism. *J. Am. Chem. Soc.* **2013**, *135*, 18673-18681.
(120) Zettler, J.; Schütz, V.; Mootz, H. D. The naturally split Npu DnaE intein exhibits an extraordinarily high rate in the protein trans-splicing reaction. *FEBS Lett.* **2009**, *583*, 909-914.
(121) Shah, N. H.; Dann, G. P.; Vila-Perelló, M.; Liu, Z.; Muir, T. W. Ultrafast protein splicing is common among cyanobacterial split inteins: implications for protein engineering. *J. Am. Chem. Soc.* **2012**, *134*, 11338-11341.
(122) Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. A novel coronavirus from patients with pneumonia in China, 2019. *N. Engl. J. Med.* **2020**, *382*, 727-733.
(123) Zhou, P.; Yang, X.-L.; Wang, X.-G.; Hu, B.; Zhang, L.; Zhang, W.; Si, H.-R.; Zhu, Y.; Li, B.; Huang, C.-L.; et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **2020**, *579*, 270-273.
(124) Sternberg, A.; Naujokat, C. Structural features of coronavirus SARS-CoV-2 spike protein: targets for vaccination. *Life Sci.* **2020**, *257*, 118056-118056.
(125) Grum-Tokars, V.; Ratia, K.; Begaye, A.; Baker, S. C.; Mesecar, A. D. Evaluating the 3C-like protease activity of SARS-Coronavirus: recommendations for standardized assays for drug discovery. *Virus Res.* **2008**, *133*, 63-73.
(126) Ferreira, J. C.; Rabeh, W. M. Biochemical and biophysical characterization of the main protease, 3-chymotrypsin-like protease (3CLpro) from the novel coronavirus SARS-CoV 2. *Sci. Rep.* **2020**, *10*, 22200-22200.
(127) de Vries, M.; Mohamed, A. S.; Prescott, R. A.; Valero-Jimenez, A. M.; Desvignes, L.; O'Connor, R.; Steppan, C.; Devlin, J. C.; Ivanova, E.; Herrera, A.; et al. A comparative analysis of SARS-CoV-2 antivirals characterizes $3CL^{pro}$ inhibitor PF-00835231 as a potential new treatment for COVID-19. *J. Virol.* **2021**, *95*, e01819-01820.
(128) Morse, J. S.; Lalonde, T.; Xu, S.; Liu, W. R. Learning from the past: possible urgent prevention and treatment options for severe acute respiratory infections caused by 2019-nCoV. *ChemBioChem* **2020**, *21*, 730-738.
(129) Zhang, L.; Lin, D.; Sun, X.; Curth, U.; Drosten, C.; Sauerhering, L.; Becker, S.; Rox, K.; Hilgenfeld, R. Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. *Science* **2020**, *368*, 409-412.
(130) Jin, Z.; Du, X.; Xu, Y.; Deng, Y.; Liu, M.; Zhao, Y.; Zhang, B.; Li, X.; Zhang, L.; Peng, C.; et al. Structure of M^{pro} from COVID-19 virus and discovery of its inhibitors. *Nature* **2020**, *582*, 289-293.
(131) Durdagi, S.; Dağ, Ç.; Dogan, B.; Yigin, M.; Avsar, T.; Buyukdag, C.; Erol, I.; Ertem, F. B.; Calis, S.; Yildirim, G.; et al. Near-physiological-temperature serial crystallography

reveals conformations of SARS-CoV-2 main protease active site for improved drug repurposing. *Structure* **2021**, *29*, 1382-1396.e1386.
(132) Tan, J.; Verschueren, K. H. G.; Anand, K.; Shen, J.; Yang, M.; Xu, Y.; Rao, Z.; Bigalke, J.; Heisen, B.; Mesters, J. R.; et al. pH-dependent conformational flexibility of the SARS-CoV main proteinase (M^{pro}) dimer: molecular dynamics simulations and multiple x-ray structure analyses. *J. Mol. Biol.* **2005**, *354*, 25-40.
(133) Chen, H.; Wei, P.; Huang, C.; Tan, L.; Liu, Y.; Lai, L. Only one protomer is active in the dimer of SARS 3C-like proteinase. *J. Biol. Chem.* **2006**, *281*, 13894-13898.
(134) Amamuddy, O. S.; Verkhivker, G. M.; Bishop, Ö. T. Impact of early pandemic stage mutations on molecular dynamics of SARS-CoV-2 M^{pro}. *J. Chem. Inf. Model.* **2020**, *60*, 5080-5102.
(135) Yang, H.; Yang, M.; Ding, Y.; Liu, Y.; Lou, Z.; Zhou, Z.; Sun, L.; Mo, L.; Ye, S.; Pang, H.; et al. The crystal structures of severe acute respiratory syndrome virus main protease and its complex with an inhibitor. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 13190-13195.
(136) Abbasi, M.; Tvakoli, N.; Bagheri Faradonbeh, S.; Bakhshayeshi, A. Cost-effectiveness analysis of rapid test compared to polymerase chain reaction (PCR) in patients with Acute Respiratory Syndrome. *Med. J. Islam. Repub. Iran* **2022**, *36*, 276-282, Original Research. (acccessed 2022).
(137) de Michelena, P.; Torres, I.; Ramos-García, Á.; Gozalbes, V.; Ruiz, N.; Sanmartín, A.; Botija, P.; Poujois, S.; Huntley, D.; Albert, E.; Navarro, D. Real-life performance of a COVID-19 rapid antigen detection test targeting the SARS-CoV-2 nucleoprotein for diagnosis of COVID-19 due to the Omicron variant. *J. Infect.* **2022**, *84*, e64-e66.
(138) Mahmoudinobar, F.; Britton, D.; Montclare, J. K. Protein-based lateral flow assays for COVID-19 detection. *Protein Eng., Des. Sel.* **2021**, *34*, 1-10.
(139) Adamson, B.; Sikka, R.; Wyllie, A. L.; Premsrirut, P. Discordant SARS-CoV-2 PCR and rapid antigen test results when infectious: a December 2021 occupational case series. *medRxiv* **2022**. DOI: 10.1101/2022.01.04.22268770.
(140) Liu, Y.; Kati, W.; Chen, C.-M.; Tripathi, R.; Molla, A.; Kohlbrenner, W. Use of a fluorescence plate reader for measuring kinetic parameters with inner filter effect correction. *Anal. Biochem.* **1999**, *267*, 331-335.
(141) Palmier, M. O.; Van Doren, S. R. Rapid determination of enzyme kinetics from fluorescence: overcoming the inner filter effect. *Anal. Biochem.* **2007**, *371*, 43-51.
(142) Benham, A. M. Protein secretion and the endoplasmic reticulum. *Cold Spring Harbor Perspect. Biol.* **2012**, *4*, a012872-a012872.
(143) Schöneich, C. Mechanisms of metal-catalyzed oxidation of histidine to 2-oxo-histidine in peptides and proteins. *J. Pharm. Biomed. Anal.* **2000**, *21*, 1093-1097.
(144) Jackson, D. Y.; Burnier, J.; Quan, C.; Stanley, M.; Tom, J.; Wells, J. A. A designed peptide ligase for total synthesis of Ribonuclease A with unnatural catalytic residues. *Science* **1994**, *266*, 243-247.
(145) Uchida, K. Histidine and lysine as targets of oxidative modification. *Amino Acids* **2003**, *25*, 249-257.
(146) Ottaviano, F. G.; Handy, D. E.; Loscalzo, J. Redox regulation in the extracellular environment. *Circ. J.* **2008**, *72*, 1-16.
(147) Weickmann, J. L.; Olson, E. M.; Glitz, D. G. Immunological assay of pancreatic ribonuclease in serum as an indicator of pancreatic cancer. *Cancer Res.* **1984**, *44*, 1682-1687. From NLM.

(148) Brock, J. W. C.; Ames, J. M.; Thorpe, S. R.; Baynes, J. W. Formation of methionine sulfoxide during glycoxidation and lipoxidation of ribonuclease A. *Arch. Biochem. Biophys.* **2007**, *457*, 170-176.
(149) Holm, L.; Laiho, A.; Törönen, P.; Salgado, M. DALI shines a light on remote homologs: one hundred discoveries. *Protein Sci.* **2023**, *32*.
(150) Marino, S. M.; Gladyshev, V. N. Cysteine function governs Its conservation and degeneration and restricts its utilization on protein surfaces. *J. Mol. Biol.* **2010**, *404*, 902-916.
(151) Dereeper, A.; Guignon, V.; Blanc, G.; Audic, S.; Buffet, S.; Chevenet, F.; Dufayard, J. F.; Guindon, S.; Lefort, V.; Lescot, M.; et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res.* **2008**, *36*, W465-W469.
(152) Klink, T. A.; Woycechowsky, K. J.; Taylor, K. M.; Raines, R. T. Contribution of disulfide bonds to the conformational stability and catalytic activity of ribonuclease A. *Eur. J. Biochem.* **2000**, *267*, 566-572.
(153) Smith, B. D.; Raines, R. T. Genetic selection for critical residues in Ribonucleases. *J. Mol. Biol.* **2006**, *362*, 459-478.
(154) Crooks, G. E.; Hon, G.; Chandonia, J.-M.; Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **2004**, *14*, 1188-1190.
(155) Paiardini, A.; Bossa, F.; Pascarella, S. CAMPO, SCR_FIND and CHC_FIND: a suite of web tools for computational structural biology. *Nucleic Acids Res.* **2005**, *33*, W50-W55.
(156) Lin, S.; Yang, X.; Jia, S.; Weeks, A. M.; Hornsby, M.; Lee, P. S.; Nichiporuk, R. V.; Iavarone, A. T.; Wells, J. A.; Dean Toste, F.; Chang, C. J. Redox-based reagents for chemoselective methionine bioconjugation. *Science* **2017**, *355*, 597-602.
(157) Lagerwerf, F. M.; van de Weert, M.; Heerma, W.; Haverkamp, J. Identification of oxidized methionine in peptides. *Rapid Commun. Mass Spectrom.* **1996**, *10*, 1905-1910.
(158) Barker, D. G.; Bruton, C. J. The fate of norleucine as a replacement for methionine in protein synthesis. *J. Mol. Biol.* **1979**, *133*, 217-231.
(159) van Hest, J. C. M.; Kiick, K. L.; Tirrell, D. A. Efficient incorporation of unsaturated methionine analogues into proteins in vivo. *J. Am. Chem. Soc.* **2000**, *122*, 1282-1288.
(160) Cirino, P. C.; Tang, Y.; Takahashi, K.; Tirrell, D. A.; Arnold, F. H. Global incorporation of norleucine in place of methionine in cytochrome P450 BM-3 heme domain increases peroxygenase activity. *Biotechnol. Bioeng.* **2003**, *83*, 729-734.
(161) Hartrampf, N.; Saebi, A.; Poskus, M.; Gates, Z. P.; Callahan, A. J.; Cowfer, A. E.; Hanna, S.; Antilla, S.; Schissel, C. K.; Quartararo, A. J.; et al. Synthesis of proteins by automated flow chemistry. *Science* **2020**, *368*, 980-987.
(162) Uchida, K.; Kawakishi, S. Reaction of a histidyl residue analog with hydrogen peroxide in the presence of copper(II) oon. *J. Agric. Food Chem.* **1990**, *38*, 660-664.
(163) Collin, F. Chemical basis of reactive oxygen species reactivity and involvement in neurodegenerative diseases. *Int. J. Mol. Sci.* **2019**, *20*.
(164) Stark, G. R.; Stein, W. H. Alkylation of the methionine residues of Ribonuclease in 8 M urea. *J. Biol. Chem.* **1964**, *239*, 3755-3761.
(165) Ribeiro, A. J. M.; Tyzack, J. D.; Borkakoti, N.; Holliday, G. L.; Thornton, J. M. A global analysis of function and conservation of catalytic residues in enzymes. *J. Biol. Chem.* **2020**, *295*, 314-324.

(166) Schafer, F. Q.; Buettner, G. R. Redox environment of the cell as viewed through the redox state of the glutathione disulfide/glutathione couple. *Free Radical Biol. Med.* **2001**, *30*, 1191-1212.
(167) Korndörfer, I. P.; Brueckner, F.; Skerra, A. The crystal structure of the human (S100A8/S100A9) 2 heterotetramer, Calprotectin, illustrates how conformational changes of interacting α-Helices can determine specific association of two EF-hand proteins. *J. Mol. Biol.* **2007**, *370*, 887-898.
(168) Sroussi, H. Y.; Köhler, G. A.; Agabian, N.; Villines, D.; Palefsky, J. M. Substitution of methionine 63 or 83 in S100A9 and cysteine 42 in S100A8 abrogate the antifungal activities of S100A8/A9: potential role for oxidative regulation. *FEMS Immunol. Med. Microbiol.* **2009**, *55*, 55-61.
(169) Samanta, U.; Bahnson, B. J. Crystal structure of human plasma platelet-activating factor acetylhydrolase. *J. Biol. Chem.* **2008**, *283*, 31617-31624.
(170) Tjoelker, L. W.; Wilder, C.; Eberhardt, C.; Stafforinit, D. M.; Dietsch, G.; Schimpf, B.; Hooper, S.; Trong, H. L.; Cousens, L. S.; Zimmerman, G. A.; et al. Anti-inflammatory properties of a platelet-activating factor acetylhydrolase. *Nature* **1995**, *374*, 549-553.
(171) Tew, D. G.; Southan, C.; Rice, S. Q. J.; Lawrence, G. M. P.; Li, H.; Boyd, H. F.; Moores, K.; Gloger, I. S.; Macphee, C. H. Purification, properties, sequencing, and cloning of a lipoprotein-associated, serine-dependent phospholipase involved in the oxidative modification of low-density lipoproteins. *Arterioscler., Thromb., Vasc. Biol.* **1996**, *16*, 591-599.
(172) Gaboriaud, C.; Serre, L.; Guy-Crotte, O.; Forest, E.; Fontecilla-Camps, J.-C. Crystal structure of human trypsin 1: unexpected phosphorylation of Tyr151. *J. Mol. Biol.* **1996**, *259*, 995-1010.
(173) Orabi, E. A.; English, A. M. Predicting structural and energetic changes in Met–aromatic motifs on methionine oxidation to the sulfoxide and sulfone. *Phys. Chem. Chem. Phys.* **2018**, *20*, 23132-23141.
(174) Johnson, R. J.; McCoy, J. G.; Bingman, C. A.; Phillips Jr, G. N.; Raines, R. T. Inhibition of human pancreatic ribonuclease by the human ribonuclease inhibitor protein. *J. Mol. Biol.* **2007**, *368*, 434-449.
(175) Edgar, R. C. Muscle5: high-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat. Commun.* **2022**, *13*, 6968.
(176) Janson, G.; Paiardini, A. PyMod 3: a complete suite for structural bioinformatics in PyMOL. *Bioinformatics* **2020**, *37*, 1471-1472. (acccessed 9/21/2023).
(177) Willard, L.; Ranjan, A.; Zhang, H.; Monzavi, H.; Boyko, R. F.; Sykes, B. D.; Wishart, D. S. VADAR: a web server for quantitative evaluation of protein structure quality. *Nucleic Acids Res.* **2003**, *31*, 3316-3319.
(178) Rutkoski, T. J.; Kurten, E. L.; Mitchell, J. C.; Raines, R. T. Disruption of shape-complementarity markers to create cytotoxic variants of Ribonuclease A. *J. Mol. Biol.* **2005**, *354*, 41-54.
(179) Nandhini, K. P.; Alhassan, M.; Veale, C. G. L.; Albericio, F.; De La Torre, B. G. Methionine-containing peptides: avoiding secondary reactions in the final global deprotection. *ACS Omega* **2023**, *8*, 15631-15637.
(180) Wralstad, E. C.; Sayers, J.; Raines, R. T. Bayesian inference elucidates the catalytic competency of the SARS-CoV-2 main protease $3CL^{pro}$. *Anal. Chem.* **2023**, *95*, 14981-14989.

(181) Krammer, F. SARS-CoV-2 vaccines in development. *Nature* **2020**, *586*, 516-527.
(182) Hoffman, R. L.; Kania, R. S.; Brothers, M. A.; Davies, J. F.; Ferre, R. A.; Gajiwala, K. S.; He, M.; Hogan, R. J.; Kozminski, K.; Li, L. Y.; et al. Discovery of ketone-based covalent inhibitors of coronavirus 3CL proteases for the potential therapeutic treatment of COVID-19. *J. Med. Chem.* **2020**, *63*, 12725-12747.
(183) Hwang, Y.-C.; Lu, R.-M.; Su, S.-C.; Chiang, P.-Y.; Ko, S.-H.; Ke, F.-Y.; Liang, K.-H.; Hsieh, T.-Y.; Wu, H.-C. Monoclonal antibodies for COVID-19 therapy and SARS-CoV-2 detection. *J. Biomed. Sci. (London, U. K.)* **2022**, *29*, 1-1.
(184) Weisblum, Y.; Schmidt, F.; Zhang, F.; DaSilva, J.; Poston, D.; Lorenzi, J. C. C.; Muecksch, F.; Rutkowska, M.; Hoffmann, H.-H.; Michailidis, E.; et al. Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants. *eLife* **2020**, *9*, e61312-e61312.
(185) Li, R.; Liu, J.; Zhang, H. The challenge of emerging SARS-CoV-2 mutants to vaccine development. *J. Genet. Genomics* **2021**, *48*, 102-106.
(186) Gupta, R. K.; Topol, E. J. COVID-19 vaccine breakthrough infections. *Science* **2021**, *374*, 1561-1562.
(187) Planas, D.; Saunders, N.; Maes, P.; Guivel-Benhassine, F.; Planchais, C.; Buchrieser, J.; Bolland, W.-H.; Porrot, F.; Staropoli, I.; Lemoine, F.; et al. Considerable escape of SARS-CoV-2 Omicron to antibody neutralization. *Nature* **2022**, *602*, 671-675.
(188) Lee, J.; Worrall, L. J.; Vuckovic, M.; Rosell, F. I.; Gentile, F.; Ton, A.-T.; Caveney, N. A.; Ban, F.; Cherkasov, A.; Paetzel, M.; Strynadka, N. C. J. Crystallographic structure of wild-type SARS-CoV-2 main protease acyl-enzyme intermediate with physiological C-terminal autoprocessing site. *Nat. Commun.* **2020**, *11*, 5877-5877.
(189) MacDonald, E. A.; Frey, G.; Namchuk, M. N.; Harrison, S. C.; Hinshaw, S. M.; Windsor, I. W. Recognition of divergent viral substrates by the SARS-CoV-2 main protease. *ACS Infect. Dis.* **2021**, *7*, 2591-2595.
(190) Abian, O.; Ortega-Alarcon, D.; Jimenez-Alesanco, A.; Ceballos-Laita, L.; Vega, S.; Reyburn, H. T.; Rizzuti, B.; Velazquez-Campoy, A. Structural stability of SARS-CoV-2 3CLpro and identification of quercetin as an inhibitor by experimental screening. *Int. J. Biol. Macromol.* **2020**, *164*, 1693-1703.
(191) Graziano, V.; McGrath, W. J.; Yang, L.; Mangel, W. F. SARS CoV main proteinase: the monomer–dimer equilibrium dissociation constant. *Biochemistry* **2006**, *45*, 14632-14641.
(192) Kuo, C.-J.; Chi, Y.-H.; Hsu, J. T. A.; Liang, P.-H. Characterization of SARS main protease and inhibitor assay using a fluorogenic substrate. *Biochem. Biophys. Res. Commun.* **2004**, *318*, 862-867.
(193) Solowiej, J.; Thomson, J. A.; Ryan, K.; Luo, C.; He, M.; Lou, J.; Murray, B. W. Steady-state and pre-steady-state kinetic evaluation of Severe Acute Respiratory Syndrome coronavirus (SARS-CoV) $3CL^{pro}$ cysteine protease: development of an ion-pair model for catalysis. *Biochemistry* **2008**, *47*, 2617-2630.
(194) Xue, X.; Yang, H.; Shen, W.; Zhao, Q.; Li, J.; Yang, K.; Chen, C.; Jin, Y.; Bartlam, M.; Rao, Z. Production of authentic SARS-CoV M^{pro} with enhanced activity: application as a novel tag-cleavage endopeptidase for protein overproduction. *J. Mol. Biol.* **2007**, *366*, 965-975.
(195) Barrila, J.; Gabelli, S. B.; Bacha, U.; Amzel, L. M.; Freire, E. Mutation of Asn28 disrupts the dimerization and enzymatic activity of SARS $3CL^{pro}$. *Biochemistry* **2010**, *49*, 4308-4317.

(196) Yang, H.; Xie, W.; Xue, X.; Yang, K.; Ma, J.; Liang, W.; Zhao, Q.; Zhou, Z.; Pei, D.; Ziebuhr, J.; et al. Design of wide-spectrum inhibitors targeting coronavirus main proteases. *PLoS Biol.* **2005**, *3*, e324-e324.
(197) Wang, H.; He, S.; Deng, W.; Zhang, Y.; Li, G.; Sun, J.; Zhao, W.; Guo, Y.; Yin, Z.; Li, D.; Shang, L. Comprehensive insights into the catalytic mechanism of Middle East Respiratory Syndrome 3C-like protease and Severe Acute Respiratory Syndrome 3C-like protease. *ACS Catal.* **2020**, *10*, 5871-5890.
(198) Fan, K.; Wei, P.; Feng, Q.; Chen, S.; Huang, C.; Ma, L.; Lai, B.; Pei, J.; Liu, Y.; Chen, J.; Lai, L. Biosynthesis, purification, and substrate specificity of Severe Acute Respiratory Syndrome coronavirus 3C-like proteinase. *J. Biol. Chem.* **2004**, *279*, 1637-1642.
(199) Huang, C.; Wei, P.; Fan, K.; Liu, Y.; Lai, L. 3C-like proteinase from SARS coronavirus catalyzes substrate hydrolysis by a general base mechanism. *Biochemistry* **2004**, *43*, 4568-4574.
(200) Hsu, M.-F.; Kuo, C.-J.; Chang, K.-T.; Chang, H.-C.; Chou, C.-C.; Ko, T.-P.; Shr, H.-L.; Chang, G.-G.; Wang, A. H. J.; Liang, P.-H. Mechanism of the maturation process of SARS-CoV 3CL protease. *J. Biol. Chem.* **2005**, *280*, 31257-31266.
(201) Chou, C. Y.; Chang, H.-C.; Hsu, W.-C.; Lin, T.-Z.; Lin, C.-H.; Chang, G.-G. Quaternary structure of the Severe Acute Respiratory Syndrome (SARS) coronavirus main protease. *Biochemistry* **2004**, *43*, 14958-14970.
(202) Chen, S.; Chen, L.; Tan, J.; Chen, J.; Du, L.; Sun, T.; Shen, J.; Chen, K.; Jiang, H.; Shen, X. Severe Acute Respiratory Syndrome Coronavirus 3C-like proteinase N terminus is indispensable for proteolytic activity but not for enzyme dimerization. *J. Biol. Chem.* **2005**, *280*, 164-173.
(203) Anand, K.; Ziebuhr, J.; Wadhwani, P.; Mesters, J. R.; Hilgenfeld, R. Coronavirus main proteinase ($3CL^{pro}$) structure: basis for design of anti-SARS drugs. *Science* **2003**, *300*, 1763-1767.
(204) Ramos-Guzmán, C. A.; Ruiz-Pernía, J. J.; Tuñón, I. Unraveling the SARS-CoV-2 main protease mechanism using multiscale methods. *ACS Catal.* **2020**, *10*, 12544-12554.
(205) Kneller, D. W.; Phillips, G.; Weiss, K. L.; Pant, S.; Zhang, Q.; O'Neill, H. M.; Coates, L.; Kovalevsky, A. Unusual zwitterionic catalytic site of SARS-CoV-2 main protease revealed by neutron crystallography. *J. Biol. Chem.* **2020**, *295*, 17365-17373.
(206) Koudelka, T.; Boger, J.; Henkel, A.; Schönherr, R.; Krantz, S.; Fuchs, S.; Rodríguez, E.; Redecke, L.; Tholey, A. N-Terminomics for the identification of in vitro substrates and cleavage site specificity of the SARS-CoV-2 main protease. *Proteomics* **2021**, *21*, e2000246-e2000246.
(207) Choi, B.; Rempala, G. A.; Kim, J. K. Beyond the Michaelis-Menten equation: accurate and efficient estimation of enzyme kinetic parameters. *Sci. Rep.* **2017**, *7*, 17018-17018.
(208) Olp, M. D.; Kalous, K. S.; Smith, B. C. ICEKAT: an interactive online tool for calculating initial rates from continuous enzyme kinetic traces. *BMC Bioinf.* **2020**, *21*, 186-186.
(209) Hong, H.; Choi, B.; Kim, J. K. Beyond the Michaelis–Menten: Bayesian inference for enzyme kinetic analysis. In *Computational Methods for Estimating the Kinetic Parameters of Biological Systems*, Vanhaelen, Q. Ed.; Humana, New York, NY, 2022; pp 47-64.
(210) Graziano, V.; McGrath, W. J.; DeGruccio, A. M.; Dunn, J. J.; Mangel, W. F. Enzymatic activity of the SARS coronavirus main proteinase dimer. *FEBS Lett.* **2006**, *580*, 2577-2583.

(211) Suárez, D.; Díaz, N. SARS-CoV-2 main protease: a molecular dynamics study. *J. Chem. Inf. Model.* **2020**, *60*, 5815-5831.
(212) Goyal, B.; Goyal, D. Targeting the dimerization of the main protease of coronaviruses: a potential broad-spectrum therapeutic strategy. *ACS Comb. Sci.* **2020**, *22*, 297-305.
(213) Shi, J.; Song, J. The catalysis of the SARS 3C-like protease is under extensive regulation by its extra domain. *FEBS J.* **2006**, *273*, 1035-1045.
(214) Cheng, S.-C.; Chang, G.-G.; Chou, C.-Y. Mutation of Glu-166 blocks the substrate-induced dimerization of SARS coronavirus main protease. *Biophys. J.* **2010**, *98*, 1327-1336.
(215) Yuan, Y.; Jiao, B.; Qu, L.; Yang, D.; Liu, R. The development of COVID-19 treatment. *Front. Immunol.* **2023**, *14*, 1125246-1125246.
(216) Gandhi, S.; Klein, J.; Robertson, A. J.; Peña-Hernández, M. A.; Lin, M. J.; Roychoudhury, P.; Lu, P.; Fournier, J.; Ferguson, D.; Mohamed Bakhash, S. A. K.; et al. De novo emergence of a remdesivir resistance mutation during treatment of persistent SARS-CoV-2 infection in an immunocompromised patient: a case report. *Nat. Commun.* **2022**, *13*, 1-8.
(217) Rockett, R.; Basile, K.; Maddocks, S.; Fong, W.; Agius, J. E.; Johnson-Mackinnon, J.; Arnott, A.; Chandra, S.; Gall, M.; Draper, J.; et al. Resistance mutations in SARS-CoV-2 Delta variant after Sotrovimab use. *N. Engl. J. Med.* **2022**, *386*, 1477-1479.
(218) Service, R. F. Bad news for Paxlovid? Resistance may be coming. *Science* **2022**, *377*, 138-139.
(219) Ippoliti, C.; De Maio, F.; Santarelli, G.; Marchetti, S.; Vella, A.; Santangelo, R.; Sanguinetti, M.; Posteraro, B. Rapid detection of the Omicron (B.1.1.529) SARS-CoV-2 variant using a COVID-19 diagnostic PCR assay. *Microbiol. Spectrum* **2022**, *10*, e0099022-e0099022.
(220) Isabel, S.; Abdulnoor, M.; Boissinot, K.; Isabel, M. R.; de Borja, R.; Zuzarte, P. C.; Sjaarda, C. P.; Barker, K. R.; Sheth, P. M.; Matukas, L. M.; et al. Emergence of a mutation in the nucleocapsid gene of SARS-CoV-2 interferes with PCR detection in Canada. *Sci. Rep.* **2022**, *12*, 10867-10867.
(221) Robles-Escajeda, E.; Mohl, J. E.; Contreras, L.; Betancourt, A. P.; Mancera, B. M.; Kirken, R. A.; Rodriguez, G. Rapid shift from SARS-CoV-2 Delta to Omicron sub-variants within a dynamic Southern U.S. borderplex. *Viruses* **2023**, *15*, 658.
(222) Chyan, W.; Raines, R. T. Enzyme-activated fluorogenic probes for live-cell and in vivo imaging. *ACS Chem. Biol.* **2018**, *13*, 1810-1823.
(223) Singh, H.; Tiwari, K.; Tiwari, R.; Pramanik, S. K.; Das, A. Small molecule as fluorescent probes for monitoring intracellular enzymatic transformations. *Chem. Rev.* **2019**, *119*, 11718-11760.
(224) Wu, X.; Wang, R.; Kwon, N.; Ma, H.; Yoon, J. Activatable fluorescent probes for *in situ* imaging of enzymes. *Chem. Soc. Rev.* **2022**, *51*, 450-463.
(225) Mirdita, M.; Schütze, K.; Moriwaki, Y.; Heo, L.; Ovchinnikov, S.; Steinegger, M. ColabFold: making protein folding accessible to all. *Nat. Methods* **2022**, *19*, 679-682.
(226) Khan, A. R.; James, M. N. G. Molecular mechanisms for the conversion of zymogens to active proteolytic enzymes. *Protein Sci.* **1998**, *7*, 815-836.
(227) Fontecilla-Camps, J. C.; de Llorens, R.; le Du, M. H.; Cuchillo, C. M. Crystal structure of ribonuclease A•d(ApTpApApG) complex. Direct evidence for extended substrate recognition. *J. Biol. Chem.* **1994**, *269*, 21526-21531.

(228) Kozakov, D.; Hall, D. R.; Xia, B.; Porter, K. A.; Padhorny, D.; Yueh, C.; Beglov, D.; Vajda, S. The ClusPro web server for protein–protein docking. *Nat. Protoc.* **2017**, *12*, 255-278.
(229) Vilà, R.; Benito, A.; Ribó, M.; Vilanova, M. Mapping the stability custers in bovine pancreatic ribonuclease A. *Biopolymers* **2009**, *91*, 1038-1047.
(230) Pan, Y.; Zhang, D.; Yang, P.; Poon, L. L. M.; Wang, Q. Viral load of SARS-CoV-2 in clinical samples. *Lancet Infect. Dis.* **2020**, *20*, 411-412.
(231) Puhach, O.; Meyer, B.; Eckerle, I. SARS-CoV-2 viral load and shedding kinetics. *Nat. Rev. Microbiol.* **2023**, *21*, 147-161.
(232) Sender, R.; Bar-On, Y. M.; Gleizer, S.; Bernshtein, B.; Flamholz, A.; Phillips, R.; Milo, R. The total number and mass of SARS-CoV-2 virions. *Proc. Natl. Acad. Sci. U. S. A.* **2021**, *118*, e2024815118-e2024815118.
(233) O'Brien, A.; Chen, D.-Y.; Hackbart, M.; Close, B. J.; O'Brien, T. E.; Saeed, M.; Baker, S. C. Detecting SARS-CoV-2 3CLpro expression and activity using a polyclonal antiserum and a luciferase-based biosensor. *Virology* **2021**, *556*, 73-78.
(234) Shi, F.-S.; Yu, Y.; Li, Y.-L.; Cui, L.; Zhao, Z.; Wang, M.; Wang, B.; Zhang, R.; Huang, Y.-W. Expression profile and localization of SARS-CoV-2 nonstructural replicase proteins in infected cells. *Microbiol. Spectrum* **2022**, *10*, e0074422.
(235) Alquraan, L.; Alzoubi, K. H.; Rababa'h, S. Y. Mutations of SARS-CoV-2 and their impact on disease diagnosis and severity. *Inform. Med. Unlocked* **2023**, *39*, 101256-101256.
(236) Ascoli, C. A. Could mutations of SARS-CoV-2 suppress diagnostic detection? *Nat. Biotechnol.* **2021**, *39*, 274-275.
(237) Desta, I. T.; Porter, K. A.; Xia, B.; Kozakov, D.; Vajda, S. Performance and its limits in rigid body protein-protein docking. *Structure* **2020**, *28*, 1071-1081.e1073.
(238) Kozakov, D.; Beglov, D.; Bohnuud, T.; Mottarella, S. E.; Xia, B.; Hall, D. R.; Vajda, S. How good is automated protein docking? *Proteins: Struct., Funct., Bioinf.* **2013**, *81*, 2159-2166.
(239) Vajda, S.; Yueh, C.; Beglov, D.; Bohnuud, T.; Mottarella, S. E.; Xia, B.; Hall, D. R.; Kozakov, D. New additions to the ClusPro server motivated by CAPRI. *Proteins: Struct., Funct., Bioinf.* **2017**, *85*, 435-444.
(240) Sayers, J.; Wralstad, E. C.; Raines, R. T. Semisynthesis of human ribonuclease-S. *Bioconjugate Chem.* **2021**, *32*, 82-87.
(241) Thompson, R. E.; Muir, T. W. Chemoenzymatic semisynthesis of proteins. *Chem. Rev.* **2020**, *120*, 3051-3126.
(242) Richards, F. M. Whatever happened to the fun? An autobiographical investigation. *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 1-25.
(243) Richards, F. M.; Wyckoff, H. W. Bovine Pancreatic Ribonuclease. In *The Enzymes*, Boyer, P. D. Ed.; Vol. 4; Academic Press, 1971; pp 647-806.
(244) Vitagliano, L.; Merlino, A.; Zagari, A.; Mazzarella, L. Reversible substrate-induced domain motions in ribonuclease A. *Proteins: Struct., Funct., Bioinf.* **2002**, *46*, 97-104.
(245) Taylor, H. C.; Richardson, D. C.; Richardson, J. S.; Wlodawer, A.; Komoriya, A.; Chaiken, I. M. “Active” conformation of an inactive semi-synthetic ribonuclease-S. *J. Mol. Biol.* **1981**, *149*, 313-317.
(246) Genz, M.; Sträter, N. Posttranslational incorporation of noncanonical amino acids in the RNase S system by semisynthetic protein assembly. In *Protein Design*, Springer New York, 2014; pp 71-87.

(247) Karpeisky, M. Y.; Senchenko, V. N.; Dianova, M. V.; Kanevsky, V. Y. Formation and properties of S-protein complex with S-peptide-containing fusion protein. *FEBS Lett.* **1994**, *339*, 209-212.
(248) Grison, C. M.; Burslem, G. M.; Miles, J. A.; Pilsl, L. K. A.; Yeo, D. J.; Imani, Z.; Warriner, S. L.; Webb, M. E.; Wilson, A. J. Double quick, double click reversible peptide "stapling". *Chem. Sci.* **2017**, *8*, 5166-5171.
(249) Assem, N.; Ferreira, D. J.; Wolan, D. W.; Dawson, P. E. Acetone-linked peptides: a convergent approach for peptide macrocyclization and labeling. *Angew. Chem., Int. Ed.* **2015**, *54*, 8665-8668.
(250) Jankovic, B.; Gulzar, A.; Zanobini, C.; Bozovic, O.; Wolf, S.; Stock, G.; Hamm, P. Photocontrolling protein–peptide interactions: from minimal perturbation to complete unbinding. *J. Am. Chem. Soc.* **2019**, *141*, 10702-10710.
(251) Vurgun, N.; Nitz, M. Validation of L-tellurienylalanine as a phenylalanine isostere. *ChemBioChem* **2020**, *21*, 1136-1139.
(252) Strong, L. E.; Kink, J. A.; Pensinger, D.; Mei, B.; Shahan, M.; Raines, R. T. Efficacy of ribonuclease QBI-139 in combination with standard of care therapies. *Cancer Res.* **2012**, *72* (8_Supplement), 1838-1838. (acccessed 10/27/2023).
(253) Strong, L. E.; Kink, J. A.; Mei, B.; Shahan, M. N.; Raines, R. T. First-in-human phase I clinical trial of QBI-139, a human ribonuclease variant, in solid tumors. *J. Clin. Oncol.* **2012**, *30* (15_suppl), TPS3113-TPS3113.
(254) Eller, C. H.; Raines, R. T. Antimicrobial synergy of a ribonuclease and a peptide secreted by human cells. *ACS Infect. Dis.* **2020**, *6*, 3083-3088.
(255) Backer, M. V.; Gaynutdinov, T. I.; Gorshkova, I. I.; Crouch, R. J.; Hu, T.; Aloise, R.; Arab, M.; Przekop, K.; Backer, J. M. Humanized docking system for assembly of targeting drug delivery complexes. *J. Controlled Release* **2003**, *89*, 499-511.
(256) Pous, J.; Mallorquí-Fernández, G.; Peracaula, R.; Terzyan, S. S.; Futami, J.; Tada, H.; Yamada, H.; Seno, M.; De Llorens, R.; Gomis-Rüth, F. X.; Coll, M. Three-dimensional structure of human RNase 1ΔN7 at 1.9 Å resolution. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2001**, *57*, 498-505.
(257) Gupta, V.; Muyldermans, S.; Wyns, L.; Salunke, D. M. The crystal structure of recombinant rat pancreatic RNase A. *Proteins: Struct., Funct., Bioinf.* **1999**, *35*, 1-12.
(258) Kato, I.; Anfinsen, C. B. On the stabilization of ribonuclease S-protein by ribonuclease S-peptide. *J. Biol. Chem.* **1969**, *244*, 1004-1007.
(259) Gaynutdinov, T. I.; Myshkin, E.; Backer, J. M.; Backer, M. V. Chimeric ribonuclease as a source of human adapter protein for targeted drug delivery. *Protein Eng., Des. Sel.* **2003**, *16*, 771-775. (acccessed 10/27/2023).
(260) Connelly, P. R.; Varadarajan, R.; Sturtevant, J. M.; Richards, F. M. Thermodynamics of protein-peptide interactions in the ribonuclease S system studied by titration calorimetry. *Biochemistry* **1990**, *29*, 6108-6114.
(261) Catanzano, F.; Giancola, C.; Graziano, G.; Barone, G. Temperature-induced denaturation of ribonuclease S: a thermodynamic study. *Biochemistry* **1996**, *35*, 13378-13385.
(262) Klink, T. A.; Raines, R. T. Conformational stability is a determinant of ribonuclease A cytotoxicity. *J. Biol. Chem.* **2000**, *275*, 17463-17467.
(263) Futami, J.; Tada, H.; Seno, M.; Ishikami, S.; Yamada, H. Stabilization of human RNase 1 by introduction of a disulfide bond between residues 4 and 118. *J. Biochem.* **2000**, *128*, 245-250.

(264) Eller, C. H.; Lomax, J. E.; Raines, R. T. Bovine brain ribonuclease is the functional homolog of human ribonuclease 1. *J. Biol. Chem.* **2014**, *289*, 25996-26006.
(265) Thomas, S. P.; Kim, E.; Kim, J.-S.; Raines, R. T. Knockout of the ribonuclease inhibitor gene leaves human cells vulnerable to secretory ribonucleases. *Biochemistry* **2016**, *55*, 6359-6362.
(266) Thompson, J. E.; Raines, R. T. Value of general acid−base catalysis to ribonuclease A. *J. Am. Chem. Soc.* **1994**, *116*, 5467-5468.
(267) Parés, X.; Nogués, M. V.; de Llorens, R.; Cuchillo, C. M. Structure and function of ribonuclease A binding subsites. *Essays Biochem.* **1991**, *26*, 89-103. From NLM.
(268) Sorrentino, S.; Naddeo, M.; Russo, A.; D'Alessio, G. Degradation of double-stranded RNA by human pancreatic ribonuclease: crucial role of noncatalytic basic amino acid residues. *Biochemistry* **2003**, *42*, 10182-10190.
(269) Schmidt, B.; Ho, L.; Hogg, P. J. Allosteric disulfide bonds. *Biochemistry* **2006**, *45*, 7429-7433.
(270) Green, M. R.; Sambrook, J. How to win the battle with RNase. *Cold Spring Harb. Protoc.* **2019**, *2019*, pdb.top101857.
(271) Smith, B. D.; Soellner, M. B.; Raines, R. T. Potent inhibition of ribonuclease A by oligo(vinylsulfonic acid). *J. Biol. Chem.* **2003**, *278*, 20934-20938.
(272) Niesen, F. H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat. Protoc.* **2007**, *2*, 2212-2221.
(273) Ressler, V. T.; Raines, R. T. Consequences of the endogenous N-glycosylation of human ribonuclease 1. *Biochemistry* **2019**, *58*, 987-996.
(274) Jullian, M.; Hernandez, A.; Maurras, A.; Puget, K.; Amblard, M.; Martinez, J.; Subra, G. N-terminus FITC labeling of peptides on solid support: the truth behind the spacer. *Tetrahedron Lett.* **2009**, *50*, 260-263.
(275) Chalker, J. M.; Gunnoo, S. B.; Boutureira, O.; Gerstberger, S. C.; Fernández-González, M.; Bernardes, G. J. L.; Griffin, L.; Hailu, H.; Schofield, C. J.; Davis, B. G. Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chem. Sci.* **2011**, *2*, 1666.
(276) Beintema, J. J.; Wietzes, P.; Weickmann, J. L.; Glitz, D. G. The amino acid sequence of human pancreatic ribonuclease. *Anal. Biochem.* **1984**, *136*, 48-64.