BIOCONJUGATION:

LINKAGE STABILITY AND NOVEL METHODS

by

Jeet Kalia

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Biochemistry)

at the

UNIVERSITY OF WISCONSIN–MADISON

2008

A dissertation entitled

Bioconjugation: Linkage Stability and Novel Methods

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

by

Jeet Kalia

Date of Final Oral Examination:

June 25, 2008

Month & Year Degree to be awarded: **December** **May** **August** 2008

****

Approval Signatures of Dissertation Committee

Signature, Dean of Graduate School

Martin Cadwallader/EH

Committee's Page. This page is not to be hand-written except for the signatures

Dedicated to Maa and Baba.

BIOCONJUGATION:

LINKAGE STABILTY AND NOVEL METHODS

Jeet Kalia

Under the supervision of Professor Ronald T. Raines

At the University of Wisconsin–Madison

"Bioconjugation" refers to covalent derivatization of proteins, DNA, RNA, and carbohydrates. Bioconjugation enables installment of moieties that endow biomolecules with desirable properties. For example, introduction of certain functional groups within biomolecules provides sites for highly chemoselective modification leading to their immobilization on surfaces and labeling with small molecules.

Bioconjugation linkages must possess high stability in aqueous solutions. Chapter 2 investigates the hydrolytic stability of oximes and hydrazones—commonly utilized carbon–nitrogen double bond-containing linkages. Isostructural hydrazones and an oxime were synthesized and their hydrolytic rate constants were obtained by using ^{1}H NMR spectroscopy. Oximes were observed to possess much greater hydrolytic stability as compared to hydrazones. The study informs on the proper use of carbon–nitrogen double bonds for bioconjugation, and lends insight into the mechanism of hydrolysis.

Maleimides are commonly utilized for thiol-mediated bioconjugation, especially for site-specific derivatization of proteins. Chapter 3 discusses the hydrolytic susceptibility of the imido group in a maleimide conjugate, and proposes a method for minimizing the resultant heterogeneity by utilizing catalysts that drive the hydrolysis to completion.

The remarkable stability of amides makes them highly desirable for bioconjugation. We have developed a general method for the site-specific modification of proteins via amide linkages. In this method, proteins are first labeled on their C-termini with the azido functional group by utilizing intein chemistry. Subsequently, a chemoselective amide bond-forming reaction of the azido group—the Staudinger ligation—is performed on the azido–proteins. Using this approach, we have immobilized the ribonuclease A protein rapidly and site-specifically via amide linkages (Chapters 4 and 5). Furthermore, we have utilized this method to generate azido–single-chain antibodies, and improved the scope and the efficiency of this method by integrating it with yeast surface display, to enable the efficient production of both prokaryotic and eukaryotic azido–proteins (Appendix I).

Acknowledgements

I have been incredibly fortunate to have two impeccable role models, who are also my best friends—my mother and father. Every day, I get inspired by their tremendous honesty, integrity, determination and humility. This thesis would have been impossible without their support and encouragement. The journey through graduate school was bumpy for me at times. Apart from the research, living far away from home was also a huge challenge. On many occasions I relied completely on my mother's never-ending confidence in me, and my father's incredible insight, even on scientific matters completely unrelated to his field of work, to bail me out of tough situations. Their lives epitomize the philosophy of "high thoughts and simple living". I aspire to be like them. It is impossible to put my gratitude for them into words.

The last year of graduate school was an especially stressful time for me due to a long postdoctoral search. During this time, I met Pearl Wichaidit. Pearl has been like a pillar of support for me throughout my last year in graduate school. Spending time with her has been my most enjoyable time during graduate school, outside the laboratory. In her company, I truly felt at home—the first time I felt this way in the United States. Pearl's selflessness and genuineness are extremely rare qualities. I hope some of her remarkable qualities have brushed on to me. She has been an incredible confidante and a true friend to me. My last year in graduate school would have been extremely tough without her. I look forward to more days of watching CSI-Miami, eating at great restaurants, and playing tennis with her.

The members of the Raines lab, past and present, have been a constant source of knowledge, support, and advice. I am especially thankful to Dr. Brian Miller, who set a high standard for me to emulate, both in science and in scientific integrity. Sharing a bay with him in the lab during my first two years in graduate school was immensely helpful to me. He continues to be a great mentor and friend. Dr. Sunil Chandran taught me a lot of the tricks of organic synthesis when I joined the lab. I have also learnt a great deal from Dr. Eddie Myers and Dr. Frank Kotch—two postdocs in the lab. The long conversations on science I had with Eddie had a profound impact on my development as a scientist. I thank him for sharing his ideas with me and for his friendship. I also thank Steve Fuchs, a senior graduate student, for his advice on career and science. He continues to be a good friend.

I was very fortunate to have shared the lab with some great colleagues. Among them are Matt Shoulders, Tom Rutkoski, Bryan Smith, Eugene Lee, Eric Benedict, and Amit Choudhary. I enjoyed working with them on the bench, going for lunches especially to Indian buffets, and discussing various aspects of life and science with them. I am grateful to them for helping me improve my manuscripts, and for their opinions and advice on career and science. My classmates, Annie Tam and Jeremy Johnson were always available with ideas and suggestions. I am also grateful to Cindy Chao for being such a good bay-mate, and for her friendship. I am going to miss stealing her food. Nitin Agarwal and Aaron Lowe were great collaborators in the Chemical Engineering department. I am especially thankful to Nitin for working around my unpredictable schedule during my postdoc interview process.

Finally, I would like to thank Professor Ronald T. Raines for allowing me to join his laboratory. I continue to be amazed at the diversity of Ron's research interests, and his enthusiasm for science. Ron is a highly creative and talented scientist, and he works really hard. I thank him for providing me independence on all my projects. He also provided me the opportunity to attend several conferences, including the International Summer School in Spetses, Greece in 2005. I learnt how to convey scientific ideas clearly from him. I thank him for maintaining a group of intelligent and hard-working scientists who make the lab a wonderful and stimulating place to work.

Table of Contents

List of Tables

List of Schemes

List of Figures

List of Abbreviations

ε	extinction coefficient
6-FAM	6-carboxyfluorescein
6-TAMRA	6-carboxytetramethylrhodamine
Ac	acetate
Ala	alanine
amu	atomic mass units
Boc	*tert*-butoxycarbonyl
cDNA	complementary deoxyribonucleic acid
CEA	carcinoembryonic antigen
ddH_2O	distilled deionized water
DIEA	*N,N*-diisopropylethylamine
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DSS	sodium 2,2-dimethyl-2-silapentane-5-sulfonate
DTT	dithiothreitol
EDC	1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride
EDTA	ethylenediaminetetraacetic acid
EGFR	epidermal growth factor receptor
ESI	electrospray ionization

FPLC	fast protein liquid chromatography
FRET	Förster resonance energy transfer
Gly	glycine
GSH	reduced glutathione
GSSG	oxidized glutathione
h	hour
HCl	hydrochloric acid
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HPLC	high-performance liquid chromatography
HRMS	high-resolution mass spectrometry
IPTG	isopropyl-1-thio-β-D-galactopyranoside
k_{cat}	first-order enzymatic rate constant
K_d	equilibrium dissociation constant
kDa	kilodalton
K_m	Michaelis constant
LB	Luria–Bertani
MALDI–TOF	matrix assisted laser desorption ionization–time of flight
MES	2-(*N*-morpholino)-ethanesulfonic acid
Met	methionine
MOPS	3-(*N*-morpholino)-propanesulfonic acid
min	minute
MRI	magnetic resonance imaging

MW	molecular weight
NaCl	sodium chloride
NHS	*N*-hydroxysuccinimide
NMR	nuclear magnetic resonance
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethyleneglycol
pK_a	log of the acid dissociation constant
PyBOP	benzotriazole-1-yl-oxy-tris-pyrrolidino-phosphonium hexafluorophosphate
R^1	alkyl group
R^2	alkyl group
R^3	alkyl group
R^4	alkyl group
RI	ribonuclease inhibitor
RNA	ribonucleic acid
RNase A	unglycosylated bovine pancreatic ribonuclease
s	second
SAM	self-assembled monolayer
scFv	single-chain antibody
SDS–PAGE	sodium dodecyl sulfate–polyacrylamide gel electrophoresis

SPR	surface plasmon resonance
t	time
$t_{1/2}$	half-life
*t*Bu	*tert*-butyl
TEA	*N,N,N*-triethylamine
TFA	trifluoroacetic acid
THF	tetrahydrofuran
TLC	thin–layer chromatography
Tris	2-amino-2-(hydroxymethyl)-1,3-propanediol
UV	ultraviolet

Chapter One

Introduction

1.1 Scope

Biomolecules can be conjugated to small molecules and surfaces either covalently or non-covalently. Examples of non-covalent bioconjugation include adsorption of proteins onto surfaces (Butler *et al.*, 1992; Duburcq *et al.*, 2004), immobilization of carbohydrates via electrostatic interactions (Shipp and Hsieh-Wilson, 2007), and immobilization of proteins by affinity tag-mediated complex formation (Paborsky *et al.*, 1996; Zhu *et al.*, 2001). Non-covalent interactions are considerably weaker than covalent interactions. Therefore, covalent bioconjugation (Figure 1.1) is the preferable form of bioconjugation, and is the focus of this thesis.

1.2 Motivation for Bioconjugation

Discovery of biological interactions

Our understanding of biological phenomena is at a rudimentary level despite centuries of research. The enormous complexity and diversity of life forms presents a veritable challenge to scientists attempting to unlock the mysteries of biological processes. The discovery that genes contain the information required to generate proteins—molecules that orchestrate a vast number of biological processes—provided a universal axiom within which discovery-based investigations could be performed (Crick, 1958; Crick, 1970). Therefore, deciphering the genetic composition of various organisms was a logical next step towards understanding biology. The resultant genome sequencing projects have yielded a wealth of information. The initial enthusiasm over the obtainment of genetic information about various organisms was,

however, mitigated by the realization that the utility of this information is limited without the knowledge of the functions of proteins encoded by genes. Moreover, elucidation of the functions of other biomolecules such as RNA and carbohydrates is imperative. Proteins, DNA, RNA, and carbohydrates regulate and perform biological functions by binding to ligands. Therefore, discovery of natural ligands of biomolecules is crucial to understanding biology.

A promising approach for ligand discovery involves appending biomolecules of interest with synthetic small molecules that can function as probes for reporting ligand binding (Rup and O'Hara, 2007). Examples of such probes include fluorescent molecules (Perez-Howard *et al.*, 1995; Zhang *et al.*, 2008), biotin (Ducoux *et al.*, 2001; Jannatipour *et al.*, 2001), and NMR probes (Pellecchia *et al.*, 2002). The ability to rapidly screen large numbers of potential ligands for biomolecules is highly desirable. Such approaches are termed "high-throughput" approaches. A particularly promising high-throughput approach involves introduction of nonnatural functional groups into biomolecules, followed by site-specific immobilization on surfaces via a chemical reaction that occurs exclusively at the appended functional group (Figure 1.1B). In a biological system, such a chemoselective functional group is referred to as a "bioorthogonal" functional group. The immobilized biomolecule is subsequently treated with various molecules to identify ligands. DNA microarrays (Brown and Botstein, 1999; Hoheisel, 2006; Lipshutz *et al.*, 1999) and protein microarrays (Kingsmore, 2006; Schena, 2004; Wilson and Nock, 2003) are examples of this approach.

Biochemical assays

Small molecules appended to biomolecules are often utilized as probes for rigorous biochemical assays. For example, Förster resonance energy transfer (FRET)-based approaches generate signals sensitive to molecular conformational changes in the 1–10 nm range (Jares-Erijman and Jovin, 2003). A typical FRET experiment entails attachment of a pair of fluorescent molecules to different regions of a biomolecule. One of these fluorophores serves as a "donor" by transferring energy nonradiatively to the other fluorophore which functions as an "acceptor". Subsequently, the acceptor emits radiation at its characteristic emission frequency, thereby reporting on the distance between the donor and acceptor. FRET has been utilized to characterize protein folding (Schuler and Eaton, 2008), RNA folding (Bokinsky and Zhuang, 2005; Smith *et al.*, 2008), and biochemical reactions (Kersteen *et al.*, 2005; Newman and Zhang, 2008). Modern single-molecule fluorescence approaches have elevated FRET-based approaches to an unprecedented level of specificity (Cornish and Ha, 2007; Ha *et al.*, 1996).

Non-fluorescent small molecules are also employed as mechanistic probes. For example, a biotin molecule was attached to a K^+ ion channel, and the conformational changes accompanying channel opening were mapped by measuring the accessibility of biotin to avidin (Jiang *et al.*, 2003). In another example, a nitrile group was introduced into an enzyme as a vibrational probe, and its stretching frequency was utilized as a sensitive reporter for the electrostatic environment within the enzymic active site (Sigala *et al.*, 2007).

Diagnostic applications

Qualitative and quantitative detection of analytes in clinical samples is crucial for the early diagnosis of disease. The enormous complexity and heterogeneity of clinical samples renders detection of individual molecules extremely challenging. Chromatographic purification of analytes prior to analysis is time-consuming and labor-intensive, and hence impractical. Traditionally, clinical chemistry and immunological methods have been utilized for the analysis of clinical samples.

Clinical chemistry exploits an intrinsic physical or chemical property of the analyte to generate a unique signal, thus circumventing analyte purification. Examples of this approach include spectrophotometric detection of metal ions and chromogenic and fluorogenic substrate-based assays for characterizing enzymes of interest (Burtis and Ashwood, 1994). Clinical chemistry approaches are limited to special cases because many analytes lack a unique signal-generating property. Moreover, clinical chemistry approaches are often not sensitive enough to be useful in clinical samples. In contrast, immunological approaches are much more sensitive—antibodies are utilized to detect analytes by using sensitive immunoassays (Price and Newman, 1997). The high specificity of antibody–antigen interactions avoids sample purification. Moreover, since antibodies can be generated against almost any analyte, this method is widely applicable.

Traditional diagnostic methods discussed above require significant biochemical experimental protocols that are time-consuming and require specialized laboratory equipment. Therefore, the applicability of such approaches is limited. There is an

urgent need to develop reusable biosensors for economical and rapid detection of analytes that would be usable in locations far removed from a laboratory setting, for example, in the office of a medical doctor or in a remote geographical location. Most biosensors consist of biomolecules attached to surfaces via robust bioconjugation linkages. For example, a commercially available glucose sensor has been developed in which glucose oxidase is immobilized to an electrode surface. The immobilized enzyme converts glucose into hydrogen peroxide, which is recorded as a digital signal. This device is utilized commercially to monitor glucose levels in diabetes patients (Heller and Feldman, in press). Some biosensor applications employ optical techniques such as surface plasmon resonance (SPR) to detect binding of analytes to biomolecules immobilized on a surface. SPR is used to measure binding of ligands, and yields accurate binding constant values (Aslan *et al.*, 2005; Hartmann-Petersen and Gordon, 2005). The BIAcore™ instrument utilized for SPR experiments is, however, expensive. A more practical and highly sensitive novel detection method based on the orientational behavior of liquid crystals on nanostructured surfaces is showing immense promise (Bertics *et al.*, 2007; Gupta *et al.*, 1998; Lowe *et al.*, 2008; Luk *et al.*, 2004).

In vivo imaging

The diagnostic methods discussed above are limited to cases wherein the nature of the disease allows for the preparation of clinical samples. In many cases, sample preparation is unfeasible, and the diagnosis needs to be performed directly inside the

body. Methods such as magnetic resonance imaging (MRI) and radioimaging are employed in such situations.

Contrast agents are used to improve signal-sensitivity in MRI. Gadolinium complexes are effective contrast agents (Allen *et al.*, 2006; Caravan, 2006). Antibodies conjugated to gadolinium complexes have been utilized for *in vivo* targeting (Shreve and Aisen, 1986). Other contrast agents such as magnetite have also been conjugated to antibodies for similar applications (Tiefenauer *et al.*, 1993).

Radioimaging is another powerful method for *in vivo* imaging. Isotopes of iodine (^{123}I and ^{131}I) are commonly used radionuclides. Iodine is especially convenient because it can be readily introduced into proteins (Holohan *et al.*, 1973), but the observation of *in vivo* deiodination has raised doubts over the use of iodine. Metal nuclides ^{111}In and ^{99m}Tc are useful alternatives. The facile introduction of metals into proteins by attachment of organic chelating agents such as EDTA is an advantage of this method (Parker, 1990).

PEGylation

Conjugation of polyethyleneglycol (PEG) molecules to proteins is a well-established technique. Commonly referred to as "PEGylation", attachment of PEGs endows proteins with many favorable characteristics such as enhanced water solubility, reduced immunogenicity, improved circulating half-life *in vivo*, enhanced proteolytic resistance, reduced toxicity, and improved thermal and mechanical

stability. PEGylation has been reviewed extensively (Chapman, 2002; Roberts *et al.*, 2002; Zalipsky, 1995), and will not be discussed in detail here.

Industrial applications

Immobilized enzymes are used as catalysts in industry (Chibata *et al.*, 1979). The first commercial application of immobilized enzymes was the resolution of amino acids by an aminocyclase (Tosa *et al.*, 1967). Applications in the food industry include use of fumarase to catalyze the isomerization of fumaric acid to malic acid. The pharmaceutical industry employs immobilized enzymes for the synthesis of drugs. For example, immobilized penicillin amidase is used in the preparation of 6-aminopenicillanic acid (Lagerlof *et al.*, 1976). Applications of bioconjugation are also prevalent in the chemical industry. One prominent example is the utilization of immobilized nitrile hydratase for the production of acrylamide from acrylonitrile (Kobayashi *et al.*, 1992).

1.3. Bioconjugation Linkages

Traditional strategies for covalent bioconjugation preclude control over the regiochemistry of reactions, producing heterogeneous reaction products. Poor control over the site of modification often results in loss of the biological function of the target biomolecule (Cha *et al.*, 2005). In contrast, novel methods of bioconjugation are highly site-specific and cause minimal perturbation of the biologically active form of the biomolecule. Moreover, biomolecules immobilized site-specifically can

possess higher ligand binding ability (Du and Saavedra, 2003; Luk *et al.*, 2004; Peluso *et al.*, 2003). Therefore, site-specific bioconjugation is preferable to random bioconjugation. Common linkages for site-specific bioconjugation are discussed below.

Linkages generated upon reaction with thiols

Thiols are potent nucleophiles in aqueous solutions at neutral pH. Therefore, derivatization of proteins via the thiolate group of a cysteine residue is a popular method of bioconjugation. An advantage of bioconjugation via cysteines is the low frequency of occurence of cysteines as compared to other amino acids in proteins, thus enabling highly regioselective bioconjugation (McCaldon and Argos, 1988). In many cases, bioconjugation can be performed via an artificially engineered cysteine residue.

Typical thiol-reactive functional groups include iodoacetamides, maleimides, and disulfides (Figure 1.2). Iodoacetamides (Figure 1.2A) have been classically utilized for determining the presence of free cysteines in proteins (Gurd, 1972). More recently, iodoacetamides have been used extensively for labeling proteins with fluorophores, PEGylation, and protein immobilization (Aslam and Dent, 1998). Maleimides are also commonly used electrophiles for thiol-mediated bioconjugation (Aslam and Dent, 1998; Lundblad, 2005). Thiolates undergo a Michael addition reaction with maleimides to form succinimidyl thioethers (Figure 1.2B). An undesirable aspect of maleimide conjugates is the hydrolytic susceptibility of their

imido groups, resulting in heterogeneity. This complicating aspect of maleimide chemistry is discussed in Chapter 3.

The thiol-selectivity of maleimides and iodoacetamides is compromised at high concentrations of the reagents—nucleophilic side chains of amino acid residues such as histidines and lysines can be covalently modified. In contrast, disulfide reagents (Figure 1.2C) are much more selective to thiols. Still, disulfides are susceptible to reduction by biological reducing agents such as glutathione. Therefore, the use of disulfides is limited to *in vitro* applications such as protein crosslinking (Armstrong *et al.*, 2006; Li *et al.*, 2008) and immobilization of peptides and proteins (Yeo *et al.*, 2004).

Linkages containing carbon–nitrogen double bonds

The facile synthesis of carbon–nitrogen double bonds via condensation of nitrogen bases with aldehydes and ketones in aqueous solutions at neutral pH renders them attractive for bioconjugation. Hydrazones are generated when the nitrogen base is a hydrazine. Oximes are formed when the nitrogen base is an alkoxyamine (Figure 1.3). Hydrazones and oximes are significantly more stable than are imines—the products of condensation of amines with aldehydes and ketones. Consequently, hydrazones and oximes are preferred to imines for bioconjugation.

Carbohydrates are especially amenable to modification with carbon–nitrogen double bonds as their hydroxyl groups can be readily converted into aldehydes, and then treated with acylhydrazine-appended small molecules to form acylhydrazones

(Wilchek and Bayer, 1987). Carbohydrates immobilized via oxime linkages have been used to generate carbohydrate microarrays (Gama *et al.*, 2006).

There are numerous examples in the literature of oxime and hydrazone conjugates of oligonucleotides (Zatsepin *et al.*, 2005). For example, acylhydrazone linkages were utilized for the immobilization of aldehydic oligonucleotides on surfaces displaying acylhydrazines (Kremsky *et al.*, 1987). Additionally, peptide nucleic acid–peptide conjugates have been generated using oxime conjugation (Neuner *et al.*, 2003).

Peptide microarrays generated by immobilizing peptides via acylhydrazone linkages were utilized for the sensitive detection of antibodies in blood samples (Duburcq *et al.*, 2004; Melnyk *et al.*, 2002). Peptides and small molecules were immobilized via oxime linkages onto glass slides displaying aldehydes, and the resulting microarray was used for protein binding and cell adhesion assays (Falsey *et al.*, 2001). Peptide fragments bearing aminoxy functional groups were incubated with a poly-aldehydic template to generate large protein-like molecules containing multiple oxime linkages (Rose, 1994; Rose *et al.*, 1996). Kent and co-workers utilized the chemoselectivity of oxime formation to assemble a transcription factor-related protein that is not readily accessible by recombinant DNA technology (Canne *et al.*, 1995). A conceptually related approach was used by Melnyk and coworkers to synthesize glycodendrimers appended with an antigen (Grandjean *et al.*, 2000; Grandjean *et al.*, 2002).

The relative hydrolytic stability of hydrazones and oximes is discussed in detail in Chapter 2. Oximes were observed to be significantly more stable than hydrazones.

Therefore, oximes are preferable to hydrazones for stable bioconjugation. Chapter 2 also elucidates the mechanistic basis underlying the high stability of oximes.

Linkages generated by cycloaddition

The recent discovery of the enormous rate acceleration of Huisgen 1,3-dipolar azide–alkyne cycloaddition by Cu(I) has established cycloaddition as one of the most useful reactions of bioconjugation (Huisgen, 1963; Kolb and Sharpless, 2003; Moses and Moorhouse, 2007). Also referred to as the "click reaction", Cu(I)-catalyzed azide–alkyne cycloaddition results in the formation of a triazole linkage in aqueous solutions (Figure 1.4A). This reaction has been utilized for proteomics applications (Speers *et al.*, 2003; Speers and Cravatt, 2005), protein and peptide immobilization (Gauchet *et al.*, 2006; Sun *et al.*, 2006), and labeling with small molecules (Kalia and Raines, 2006; Macpherson *et al.*, 2007; Peschke *et al.*, 2007; Zhou and Fahrni, 2004). Furthermore, click chemistry has been used for immobilizing carbohydrates onto glass slides (Sun *et al.*, 2006), high-density functionalization of modified DNA (Gierlich *et al.*, 2006; Gramlich *et al.*, 2008), and conjugating virus particles with fluorescent molecules (Wang *et al.*, 2003b). Chapter 4 describes a general method for installing the requisite azido group at the C-terminus of any protein. The resulting azido–protein was treated with alkynyl fluorescein to perform Cu(I)-catalyzed Huisgen 1,3-dipolar azide–alkyne cycloaddition, thereby labeling the protein site-specifically with fluorescein. Appendix I describes a method of site-specific immobilization of azido-proteins on alkyne-displaying chips via this reaction.

The Cu(I)-catalyzed version of click chemistry discussed above often results in cytotoxicity and protein precipitation due to Cu(I) present in the reaction (Kalia and Raines, 2006; Speers *et al.*, 2003). Moreover, the reaction rates are slow, precluding its use for studying cellular processes. To overcome these drawbacks, Bertozzi and co-workers have introduced a copper-free version of Huisgen's 1,3-dipolar azide–alkyne cycloaddition reaction that utilizes a ring-strained cyclooctyne group, enabling the reaction to proceed rapidly without a catalyst (Agard *et al.*, 2006; Agard *et al.*, 2004; Baskin *et al.*, 2007) (Figure 1.4B). The recent use of copper-free click chemistry for *in vivo* imaging underlines its enormous potential for bioconjugation (Laughlin *et al.*, 2008).

Another cycloaddition reaction—the Diels-Alder reaction—between a diene on a peptide and a dienophile on a glass surface has been utilized for peptide immobilization (Houseman *et al.*, 2002b). A similar approach was employed for immobilizing carbohydrates onto glass slides displaying hydroquinone functional groups (Houseman and Mrksich, 2002).

Amide linkages

Amide bonds have a half life of ~500 years in aqueous solutions (Radzicka and Wolfenden, 1996). The remarkable stability of amides is attractive for bioconjugation. Random introduction of amide linkages in biomolecules is trivial. For example, a protein can be treated with small molecules and surfaces displaying activated esters to generate amides at lysine side chains of the protein (Lahiri *et al.*,

1999; Luk *et al.*, 2004). Yet, site-specific generation of amides is challenging. Native chemical ligation and Staudinger ligation are two promising approaches for generating amide linkages site-specifically in a protein (Nilsson *et al.*, 2005).

In native chemical ligation, an N-terminal cysteine residue reacts with a thioester to undergo transthioesterification followed by a rapid *S*→*N* acyl transfer to form an amide (Figure 1.5). This reaction is a powerful tool for peptide ligation (Dawson and Kent, 2000; Dawson *et al.*, 1994; Kent, 2003). Expressed protein ligation is an extension of native chemical ligation (Evans and Xu, 2002; Muir, 2003; Muralidharan and Muir, 2006). In this method, a target protein is expressed as a fusion protein with an intein—a protein subunit that catalyzes the formation of a thioester at the C-terminus of the target protein. Protein–intein fusion proteins are treated with peptides containing an N-terminal cysteine residue to effect native chemical ligation (Figure 1.6). Surfaces displaying cysteines can be treated with protein–intein thioesters to perform site-specific protein immobilization via amide bonds (Camarero *et al.*, 2004). Using a similar approach, fluorescent molecules were site-specifically conjugated to proteins (Wood *et al.*, 2004). Furthermore, proteins have been biotinylated using expressed protein ligation, and utilized for high-throughput proteomics (Lue *et al.*, 2004). An undesirable aspect of native chemical ligation and expressed protein ligation is the introduction of a residual thiol at the site of bioconjugation, which can be a focal point for undesirable side reactions (Friedman, 1999; Raines, 1997; Terrettax *et al.*, 2002). Chemical desulfurization approaches (Yan and Dawson, 2001) provide a solution to the above problem, but they cannot be utilized if the protein

consists of any other cysteine residues. Nitrogen nucleophiles such as hydrazines and hydroxylamine react with protein–intein thioesters without installing a residual reactive group. Chapter 4 discusses the reactivity of nitrogen nucleophiles with protein–intein thioesters in detail.

The Staudinger ligation provides a solution to the cysteine limitation (Köhn and Breinbauer, 2004; Nilsson *et al.*, 2005). This reaction is based on the Staudinger reduction, in which an azide is reduced to an amine by a phosphine (Gololobov and Kasukhin, 1992; Staudinger and Meyer, 1919). Staudinger ligation utilizes a phosphine that serves as an acyl donor—the phosphorus first attacks the azide forming an iminophosphorane which is subsequently acylated with the concomitant liberation of nitrogen gas to form an amidophosphonium salt that eventually hydrolyzes to yield the amide (Lin *et al.*, 2005; Soellner *et al.*, 2006). One version of Staudinger ligation leaves a phosphine oxide in the amide product (Figure 7A) (Lin *et al.*, 2005; Luchansky *et al.*, 2003; Saxon and Bertozzi, 2000). Another version—the "traceless" Staudinger ligation—utilizes a phosphinothioester that yields an acyclic amidophosphonium salt, resulting in a final amide product that lacks the phosphine oxide moiety (Figure 7B) (Nilsson *et al.*, 2000; Nilsson *et al.*, 2001; Soellner *et al.*, 2002; Soellner *et al.*, 2006). Since the amide generated is a native peptide bond, this reaction enables total synthesis of proteins (Nilsson *et al.*, 2003; Nilsson *et al.*, 2005).

Staudinger ligation is extensively used for bioconjugation. For example an azido group was installed into a protein by utilizing azidohomoalanine-activated methionyl-tRNA synthetase of *E. coli* expressed in methionine-depleted cultures, and subjected

to Staudinger ligation with a peptide (Kiick *et al.*, 2002). Site-specific labeling of DNA by fluorescent molecules has also been performed by Staudinger ligation (Wang *et al.*, 2003a). The Staudinger ligation has been utilized for *N*-glycopeptide synthesis (Yi *et al.*, 2004). Furthermore, azido-labeled carbohydrates introduced on cell surfaces have been modified by Staudinger ligation with biotin-phosphane and FLAG-phosphane, enabling quantitative measurements by flow cytometry (Saxon and Bertozzi, 2000; Saxon *et al.*, 2002). In addition, Staudinger ligation has been utilized for rapid and site-specific immobilization of peptides and proteins (Gauchet *et al.*, 2006; Kalia *et al.*, 2007; Köhn *et al.*, 2003; Soellner *et al.*, 2003). Immobilization of azido–proteins by traceless Staudinger ligation is described in Chapter 5.

1.4. Conclusions

Applications of bioconjugation range from fundamental science to industrial and clinical research. Choosing the optimal bioconjugation linkage for a particular application is crucial. It is imperative to ensure that the linkage is stable during the course of its use. Therefore, analyses of the stability of commonly used bioconjugation linkages are required. In addition to possessing high stability, an optimal bioconjugation linkage should be easy to generate in aqueous solutions at neutral pH—a challenging proposition. Consequently, efforts for developing facile conjugation chemistry are being pursued vigorously.

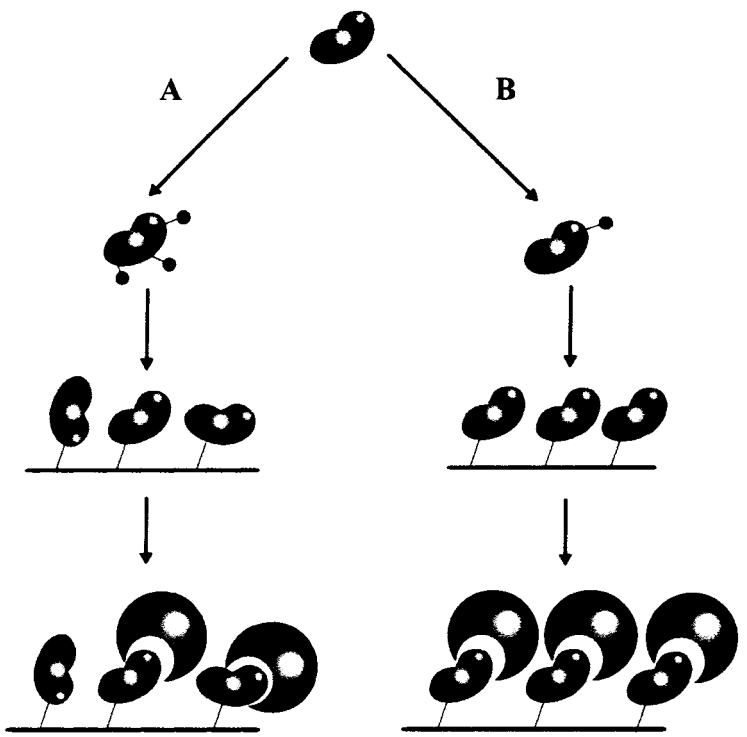


Figure 1.1 Covalent bioconjugation. Target biomolecules (blue) are treated with reactive small molecules (red). Modification of biomolecules at multiple sites results in random bioconjugation (A), and modification at exclusively one site leads to site-specific bioconjugation (B). Derivatized biomolecules can be immobilized onto surfaces, and assayed for binding to ligands (brown).

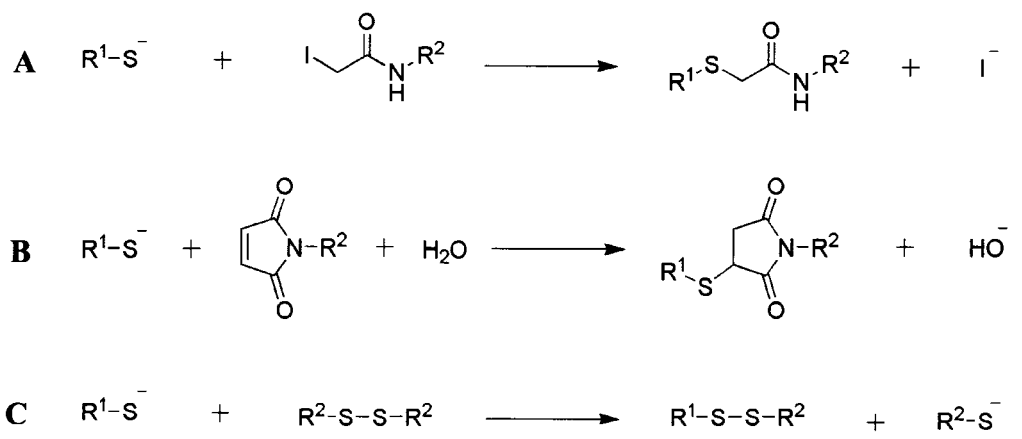


Figure 1.2 Linkages generated upon reaction with thiols. Thiolates react readily with iodoacetamides (A), maleimides (B), and disulfides (C) to form bioconjugates in aqueous solution.

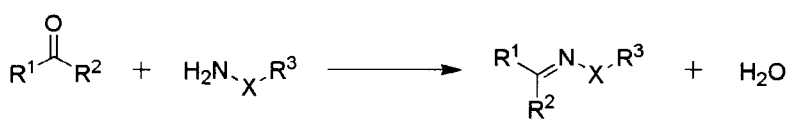


Figure 1.3 Linkages containing carbon–nitrogen double bonds. X = O in oximes, NH in alkylhydrazones, and NHC(O) in acylhydrazones.

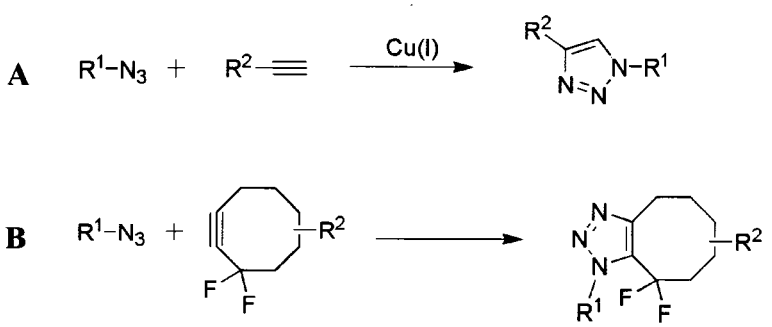


Figure 1.4 Huisgen 1,3-dipolar azide–alkyne cycloaddition. A Cu(I)-catalyzed version (A), and a Cu-free version (B).

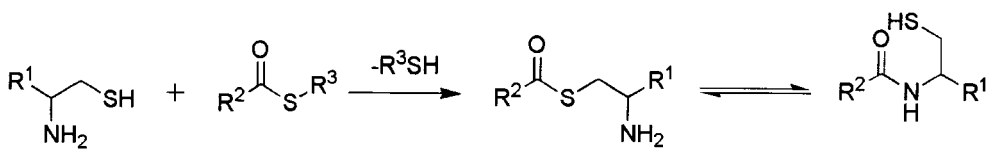


Figure 1.5 Amide bond formation via native chemical ligation. An N-terminal cysteine reacts with a thioester to undergo transthioesterification, followed by a *S*→*N* shift to form an amide bond.

.

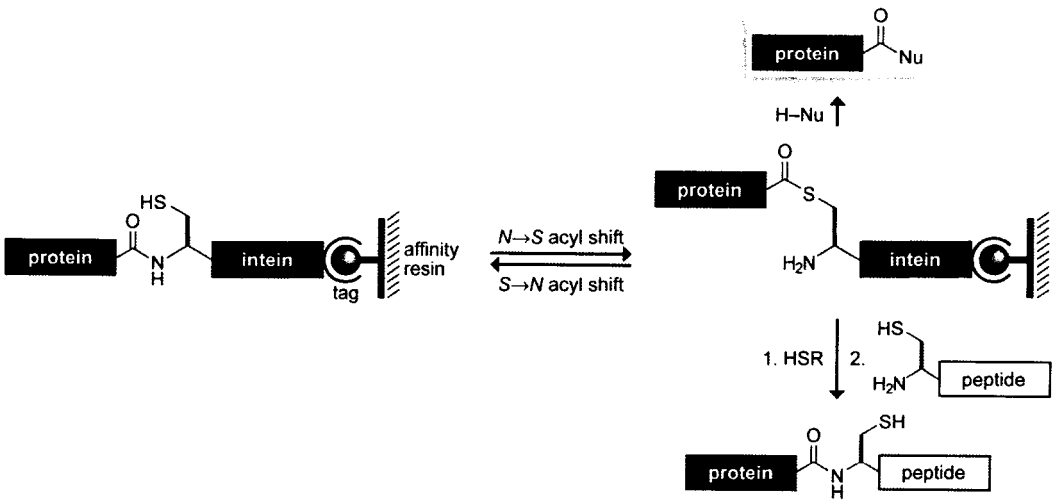


Figure 1.6 Amide bond formation via expressed protein ligation. An intein installs a thioester at the C-terminus of the target protein, which upon treatment with an N-terminal cysteine-containing peptide undergoes native chemical ligation. Protein–intein thioesters can also be treated with other nucleophiles (H–Nu) to modify the C-termini of proteins.

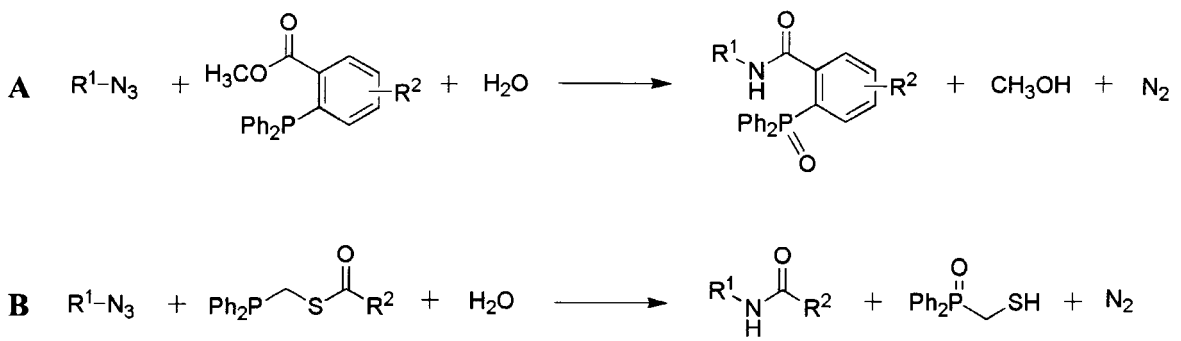


Figure 1.7 The Staudinger ligation. In version (A), the phosphine oxide remains attached to the amide product. Version (B) represents a traceless Staudinger ligation wherein the phosphine oxide is excised to yield a native peptide bond.

Chapter Two

Hydrolytic Stability of Hydrazones and Oximes

Contribution: Synthesis of conjugates, design and execution of kinetics experiments, kinetic analyses and calculations, data interpretation, composition of manuscript and figure drafts.

This chapter was accepted for publication in *Angew. Chem. Int. Ed.* as:
Kalia, J. and Raines, R. T. (2008) Hydrolytic Stability of Hydrazones and Oximes.

2.1 Abstract

Hydrazones and oximes are common conjugates, but are labile to hydrolysis. The hydrolytic stability of isostructural hydrazones and an oxime have been determined at pD 5.0–9.0. The hydrolysis of each adduct was catalyzed by acid. Rate constants for oxime hydrolysis were nearly 10^3-fold lower than those for simple hydrazones; a trialkylhydrazonium ion (formed after condensation) was even more stable than the oxime. The data suggest a general mechanism for conjugate hydrolysis.

2.2 Introduction

Molecules containing carbon–nitrogen double bonds are prevalent in both chemical and biological contexts. The foundations for our current understanding of carbon–nitrogen double-bond formation and hydrolysis were laid by classic early work on hydrazone hydrolysis and formation (Ardagh and Rutherford, 1935; Conant and Bartlett, 1932; Westheimer, 1934), and by contributions from mechanistic studies on enzymes that utilize pyridoxal phosphate (Metzler, 1957; Metzler *et al.*, 1954). In particular, the meticulous kinetic analyses of Jencks resulted in the delineation of a carbinolamine intermediate in carbon–nitrogen double-bond formation and hydrolysis, and elucidation of the general mechanism of carbonyl-group addition reactions (Anderson and Jencks, 1960; Cordes and Jencks, 1962a; Cordes and Jencks, 1962b; Cordes and Jencks, 1962c; Cordes and Jencks, 1963; Jencks, 1959; Sander and Jencks, 1968; Sayer *et al.*, 1973; Wolfenden and Jencks, 1961). These principles were summarized in a landmark review (Jencks, 1964).

Hydrazones and oximes (C^1=N^1–X^2) possess higher intrinsic hydrolytic stability than do imines. The textbook explanation for the higher stability of hydrazones (X^2 = N) and oximes (X^2 = O) involves the participation of X^2 in electron delocalization (Figure 2.1) (Carey and Sundberg, 2008). The contribution of resonance form **2.10** in alkylhydrazones and oximes, and resonance form **2.12** in acylhydrazones increases the negative-charge density on C^1 and hence reduces its electrophilicity, thereby imparting higher hydrolytic stability to hydrazones and oximes. An alternative explanation is based on the repulsion of the lone pairs of N^1 and X^2 being relieved in the conjugates (Wiberg and Glaser, 1992).

Although the higher stability of oximes and hydrazones as compared to imines is well-appreciated, a consensus on the comparative stability of hydrazones and oximes is lacking. To the best of our knowledge, the only report of a direct comparison of the rates of hydrolysis of hydrazones and oximes was from Stieglitz and coworkers in 1934 (Johnson and Stieglitz, 1934). These workers assayed the hydrolysis of benzophenonehydrazone and benzophenoneoxime in extremely acidic solutions by titrating the respective hydrazine and hydroxylamine products. This rudimentary study provided little insight. More recently, other workers have discussed the stability of the oximes and hydrazones used in particular applications (Canne *et al.*, 1995; Dirksen *et al.*, 2006b; Kochendoerfer *et al.*, 2003; Kolonko and Kiessling, 2008; Rose, 1994; Rose *et al.*, 1996), but without direct comparisons.

Here, we report the first detailed investigation of the hydrolysis of isostructural alkylhydrazones, acylhydrazones, and an oxime. Half-lives for the hydrolysis of these

conjugates were measured with ^{1}H NMR spectroscopy in deuterated buffers (pD 5.0–9.0) to obtain pD–rate profiles. In addition, pD-titrations of the conjugates were performed with ^{1}H NMR spectroscopy to determine relevant pK_a values and thereby provide mechanistic insight. Our findings establish oximes as the linkage of choice for the stable conjugation of molecules via a carbon–nitrogen double bond.

2.3 Experimental Procedures

Materials. Anhydrous DMF and CH_2Cl_2 were withdrawn from a CYCLE-TAINER® solvent delivery system (J.T. Baker, Phillipsburg, NJ). Other solvents and chemicals were from Sigma–Aldrich (St. Louis, MO). Synthetic reactions were monitored by thin-layer chromatography with visualization by UV-light, or staining with phosphomolybdic acid. Flash chromatography was performed with columns of silica gel 60, 230–400 mesh (Silicycle, Québec City, Québec, Canada).

Instrumentation. NMR spectra for compound characterization were acquired with a Bruker DMX-400 Avance spectrometer (^{1}H: 400 MHz, ^{13}C: 100 MHz) at the National Magnetic Resonance Facility at Madison (NMRFAM). Samples for compound characterization were prepared in DMSO-d_6 unless stated otherwise. NMR spectra for kinetic analysis were acquired with a Bruker AC+ 300 spectrometer (^{1}H: 300 MHz) at the Magnetic Resonance Facility in the Department of Chemistry. Mass spectra were obtained with a Micromass LCT (electrospray ionization, ESI) in the

Mass Spectrometry Facility in the Department of Chemistry. Elemental analyses were performed by Midwest Microlab LLC (Indianapolis, IN).

General procedure for the synthesis of tBuCH=NNHCH$_3$ **2.1**, *tBuCH=NN(CH$_3$)$_2$* **2.2**, *and tBuCH=NNHCOCH$_3$* **2.4**. *t*BuCHO (13.61 mL, 123.07 mmol) was stirred with the alkylhydrazine or acetylhydrazine (123.07 mmol) for 25 min at 0 °C. The mixture was allowed to warm to room temperature, and stirred for 1.5 h. Anhydrous $MgSO_4$(s) was added, and the mixture was stirred for 15 min. The solid was removed by filtration to yield the hydrazone in >90% yield. Compounds **1** and **2** were obtained as light-yellow liquids, and compound **2.4** was a white solid.

*t*BuCH=NNHCH$_3$ **2.1**: ^{1}H NMR (400 MHz, $CDCl_3$) δ = 6.83 (s, 1H), 4.93 (bs, 1H), 2.78 (s, 3H), 1.07 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 148.8, 35.3, 34.3, 28.2; anal. calcd. for $C_6H_{14}N_2$: C 63.11, H 12.36, N 24.53; found: C 62.05, H 12.00, N 23.20.

*t*BuCH=NN(CH$_3$)$_2$ **2.2**: HRMS (ESI) $[M+H]^+$ calcd. for $C_7H_{17}N_2$, 129.1392, found 129.1398; ^{1}H NMR (400 MHz, $CDCl_3$) δ = 6.58 (s, 1H), 2.69 (s, 6H), 1.07 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 147.4, 43.5, 34.4, 28.4.

*t*BuCH=NNHCOCH$_3$ **2.4**: HRMS (ESI) $[M+Na]^+$ calcd. for $C_7H_{14}N_2ONa$, 165.1004; found 165.0999; ^{1}H NMR (400 MHz, 2 rotamers) δ = 10.83 and 10.74 (s, 1H), 7.36 and 7.21 (s, 1H), 2.04 and 1.83 (s, 3H), 1.03 (s, 9H); ^{13}C NMR (100 MHz, 2 rotamers) δ = 171.5 and 165.1, 156.7 and 153.4, 34.4 and 34.2, 27.2, 21.5 and 20.1.

Synthesis of tBuCH=NOCH$_3$ **2.3**. $CH_3ONH_2 \cdot HCl$ (4.85 g, 58.08 mmol) was dissolved in DMF (15 mL), and *N,N*-diisopropylethylamine (10.11 mL, 58.08 mmol) and *t*BuCHO (6.42 mL, 58.08 mmol) were added to the resulting solution. The mixture was cooled to 0 °C, stirred for 25 min, and allowed to warm to room temperature. After stirring for 1.5 h, anhydrous $MgSO_4$(s) was added, and the mixture was stirred for 15 min. The solid was removed by filtration, and the filtrate was distilled to yield *t*BuCH=NOCH$_3$ (**2.3**) as a colorless liquid (1.62 g, 24%, b.p. = 65 °C). ^{1}H NMR (400 MHz, $CDCl_3$) δ = 7.29 (s, 1H), 3.80 (s, 3H), 1.09 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 158.3, 61.3, 33.6, 27.7; anal. calcd. for $C_6H_{13}NO$: C 62.57, H 11.38, N 12.16, O 13.89; found: C 62.71, H 11.70, N 11.83, O 13.77.

Synthesis of tBuCH=NNHCONH$_2$ **2.5**. $NH_2CONHNH_2 \cdot HCl$ (2.00 g, 17.93 mmol) was dissolved in DMF (20 mL), and Et_3N (2.75 mL, 19.73 mmol) and *t*BuCHO (2.38 mL, 21.52 mmol) were added to the resulting solution. The mixture was cooled to 0 °C, stirred for 25 min, and allowed to warm to room temperature. After stirring for 1.5 h, anhydrous $MgSO_4$(s) was added, and the mixture was stirred for 15 min. After filtration, the organic layer was concentrated under reduced pressure, and the residue was purified by flash chromatography (silica gel, 10% (*v/v*) methanol in methylene chloride) to give *t*BuCH=NNHCONH$_2$ (**2.5**) as a white solid (1.49 g, 58%). HRMS (ESI) $[M+Na]^+$ calcd. for $C_6H_{13}N_3ONa$: 166.0956, found 166.0964; ^{1}H NMR (400 MHz) δ = 9.76 (s, 1H), 7.08 (s, 1H), 6.11 (bs, 2H), 1.02 (s, 9H); ^{13}C NMR (100 MHz) δ = 157.0, 150.5, 34.0, 27.4.

Synthesis of BocNHNHCOCF$_3$ **2.8**. BocNHNH$_2$ (5.00 g, 37.83 mmol) was dissolved in CH_3CN (100 mL). The mixture was cooled to 0 °C, and Et_3N (5.8 mL, 41.61 mmol) and $(CF_3CO)_2O$ (5.25 mL, 37.77 mmol) were added. The reaction mixture was stirred for 1 h. Solvent was removed under reduced pressure, and the residue was purified by flash chromatography (silica gel, ethyl acetate). BocNHNHCOCF$_3$ was obtained as a white solid (7.15 g, 83%). HRMS (ESI) $[M+Na]^+$ calcd. for $C_7H_{11}F_3N_2O_3Na$: 251.0619, found 251.0623; ^{1}H NMR (400 MHz) δ = 11.27 (bs, 1H), 9.30 (s, 1H), 1.42 (s, 9H); ^{13}C NMR (100 MHz) δ = 156.3 (q, J(C,F) = 36.1 Hz), 154.5, 115.9 (q, J(C,F) = 288.0 Hz), 80.2, 28.0.

Synthesis of tBuCH=NHNHCOCF$_3$ **2.6**. HCl.H$_2$NNHCOCF$_3$ was synthesized by dissolving BocNHNHCOCF$_3$ (**2.8**) (5.00 g, 21.92 mmol) in HCl (4N) in dioxane (140 mL). The mixture was then stirred for 1h. The solvent was removed under reduced pressure to give an off-white powder. This powder (3.0 g) was transferred to another flask, and dissolved in DMF (20 mL). The resulting solution was cooled to 0 °C, and *t*BuCHO (2.42 mL, 21.88 mmol) and Et_3N (2.78 mL, 20.00 mmol) were added. After stirring for 30 min, anhydrous $MgSO_4$(s) was added, and the reaction mixture was allowed to warm to room temperature. After stirring for 1.5 h, the solid was removed by filtration, and the filtrate was concentrated under reduced pressure. The residue was purified by flash chromatography (silica gel, methylene chloride). *t*BuCH=NHNHCOCF$_3$ was obtained as a white solid (2.58 g, 95%). HRMS (ESI)

$[M+Na]^+$ calcd. for $C_7H_{11}F_3N_2ONa$: 219.0721, found 219.0715; ^1H NMR (400 MHz) δ = 7.75 (s, 1H), 1.08 (s, 9H), 1.04 (s, 1H); ^{13}C NMR (100 MHz) δ = 164.7, 152.5 (q, J(C,F) = 36.7 Hz), 115.9 (q, J(C,F) = 289.4 Hz), 35.0, 26.8.

Synthesis of tBuCH=NN(CH$_3$)$_3$I **2.7**. CH_3I (0.73 mL, 11.69 mmol) was added to compound **2.2** (0.50 g, 3.89 mmol), and the mixture was stirred for 15 min at rt. Unreacted CH_3I was removed under reduced pressure to yield *t*BuCH=NN$(CH_3)_3$I as a yellow solid (1.00 g, 95%). HRMS (ESI) $[M]^+$ calcd. for $C_8H_{19}N_2$: 143.1548, found 143.1543; ^1H NMR (400 MHz) δ = 8.43 (s, 1H), 3.37 (s, 9H), 1.13 (s, 9H); ^{13}C NMR (100 MHz) δ = 174.3, 54.4, 36.0, 26.1.

Kinetics of conjugate hydrolysis. Deuterated sodium phosphate buffers were prepared by dissolving Na_3PO_4 in D_2O to a concentration of 0.15 M. Acidity was adjusted by adding D_3PO_4, to pDs 5.0, 6.0, 7.0, 8.0, and 9.0 (pD = pH meter reading + 0.41) (Covington *et al.*, 1968). The conjugates were dissolved to a concentration of 25.00 mM in buffer solutions containing D_2CO (0.25 M, added from a 20% (v/v) D_2CO solution in D_2O). ^1H NMR spectra were obtained at the desired time-points, and the extent of hydrolysis was quantitated by peak integration. Hydrolysis resulted in the appearance of the aldehyde, characterized by the formyl proton at 9.4 ppm, a *t*Bu singlet at 1.0 ppm, and a *t*Bu peak of the hydrated aldehyde at 0.8 ppm. As expected, there was a concurrent decrease in the intensities of peaks due to the conjugate, namely, the *t*Bu group at ~1 ppm, and the proton attached to the double-bonded

carbon atom at ~7–8 ppm. The area under the three peaks corresponding to the *t*Bu groups were assigned a cumulative value of 9, serving as an internal standard for integration. Hydrolysis was quantitated according to the equation:

$$\%\ \text{hydrolysis} = 100\frac{A+\frac{B}{9}}{A+\frac{B}{9}+C} \tag{2.1}$$

where A is the area under the peak at 9.4 ppm, B is the area under peak at 0.8 ppm, and C is the area under the peak at ~7–8 ppm. Hydrolysis was allowed to proceed to >95% completion. % Hydrolysis was plotted versus time, and the data were fitted to equation:

$$Y = Y_{\max}(1-e^{-kt}) \tag{2.2}$$

where Y is the % hydrolysis, t is time, k is the first-order rate constant, and $Y_{\max}$ is the % hydrolysis at $t = \infty$. Kinetic traces were obtained in duplicate, and half-lives were calculated with equation:

$$t_{1/2} = \frac{0.693}{k} \tag{2.3}$$

NMR titration of conjugates. Deuterated buffers were prepared in the pD range of 0.73–13.36. Trichloroacetic acid (0.40 M)–NaOD was used as a buffer in the pD range of 0.73–2.01, chloroacetic acid (0.40 M)–NaOD was used in the pD range of 2.67–3.39, acetic acid (0.40 M)–NaOD was used in the pD range of 4.50–6.02, and sodium phosphate (0.17 M) was used in the pD range of 6.34–13.36. The ionic strength of the buffers was maintained at $I = 0.45$ M by the addition of KCl. Sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was used as an internal standard for referencing the chemical shift. The conjugates were dissolved in the buffers to a concentration of 25.00 mM, and the chemical shift of the proton attached to the double-bonded carbon was obtained at different pDs. The chemical shifts were plotted against the pDs to generate the data-points in Figure 2.4. Titration curves for methylhydrazone **2.1**, dimethylhydrazone **2.2**, and trifluoroacetylhydrazone **2.6** were obtained by fitting the data-points to equation:

$$\delta = \delta_{\text{bottom}} + \frac{\delta_{\text{top}} - \delta_{\text{bottom}}}{1 + \frac{K_a}{[D^+]}} \tag{2.4}$$

where δ is the chemical shift, δ_{bottom} is the chemical shift at high pDs where the conjugate exists as a free base, δ_{top} is the chemical shift at low pDs where the conjugate is completely protonated, and pK_a is the point of inflection of the curve.

2.4 Results and Discussion

Conjugates **2.1–2.6** were synthesized by condensation of the respective nitrogen bases with pivalaldehyde (*t*BuCHO), and removing the water byproduct with anhydrous $MgSO_4$(s) (Scheme 2.1). Pivalaldehyde was chosen because it lacks enolizable protons, thus preventing obfuscating side reactions such as aldol condensations. Methoxyamine and all the alkylhydrazines and acylhydrazines were available commercially except for trifluoroacetylhydrazine, which was generated in situ by the deprotection of Boc-trifluoroacetylhydrazine (compound **2.8**). Trimethylhydrazonium ion **2.7** was synthesized by reacting dimethylhydrazone **2.2** with methyl iodide (Scheme 2.1). The synthesis of **2.7** by the condensation of a trimethylhydrazinium ion and pivalaldehyde was unsuccessful, as reported by others (Moore and Stupp, 1990; Smith and Most, 1957; Smith and Walker, 1962); this condensation reaction was not facilitated to a detectable extent by aniline (Dirksen *et al.*, 2006a) at pD 5.0–9.0. (The likely problem is that nucleophilic attack by trimethylhydrazinium ion, $H_2N^1–N^2(CH_3)_3^+$, generates a positive charge on N^1 when N^2 already bears a positive charge.) ^{1}H NMR spectroscopy in deuterated phosphate buffers (pD 5.0–9.0) was used to monitor the appearance of the aldehydic proton of pivalaldehyde (δ = 9.4 ppm), a signal for conjugate hydrolysis.

The hydrolytic cleavage of carbon–nitrogen double bonds is reversible. An excess of a deuterated aldehyde or ketone can be used to trap the liberated nitrogen base and thereby push the hydrolysis reaction to completion, allowing the forward (hydrolysis) reaction to be monitored without interference from the reverse (condensation) reaction. Various aldehydes and ketones were tested as potential chemical traps.

Deuterated acetone was an inefficient trap—a 100-fold excess drove the hydrolysis of a methylhydrazone to only 62% completion at pD 7.0 (data not shown). Hexachloroacetone, tribromoacetaldehyde, and calcium mesoxylate could not be used due to their low aqueous solubility. Alloxan, an electrophilic ketone,was unstable in water. Finally, a 10-fold excess of deuterated formaldehyde (CD_2O) was identified as an effective trap, driving the hydrolysis reactions of all the conjugates (except that of trimethylhydrazonium ion **2.7**) to completion at pD 5.0–9.0. A typical kinetic trace obtained is shown in Figure 2.2.

At pD 5.0–9.0, the half-life of oxime **2.3** was much larger that those of each hydrazone, except for trimethylhydrazonium ion **2.7** (Table 2.1). At pD 7.0, the first-order rate constant for the hydrolysis of oxime **2.3** was approximately 600-fold lower than methylhydrazone **2.1**, 300-fold lower than acetylhydrazone **2.4**, and 160-fold lower than semicarbazone **2.5**. Although the linkage in a trialkylhydrazonium ion (such as conjugate **2.7**) is highly stable, it is not suitable for bioconjugation because its synthesis involves treatment with methyl iodide—a reagent that is not chemoselective in a biological system—subsequent to condensation. Thus, oximes are the most preferable linkages for carbon–nitrogen doubled bond-mediated bioconjugation.

The hydrolysis of the conjugates is catalyzed by acid (Figure 2.3). These data are consistent with hydrolysis being accelerated by conjugate protonation. The hydrolysis of oxime **2.3** at pD >7.0 and that of trimethylhydrazonium ion **2.7** at pD >5.0 were too slow to yield a complete kinetic trace within a reasonable time-frame.

pD-Titration experiments monitored with ^{1}H NMR spectroscopy revealed that some (but not all) of the conjugates experience a substantial change in protonation state between pD 0.7 and 13.4 (Figure 2.4). The δ value of $C^1\underline{H}$ for methylhydrazone **2.1** ($pK_a = 5.5$), dimethylhydrazone **2.2** (5.8), and trifluoroacetylhydrazone **2.6** (7.9) exhibited a sigmoidal dependence on pD. The δ value of $C^1\underline{H}$ in conjugates **2.3–2.5** and **2.7** was not a function of pD, indicating that an insignificant fraction of these conjugates is protonated at pD 0.7–13.4.

What is the site of conjugate protonation? The titration curves for methylhydrazone **2.1** and dimethylhydrazone **2.2** are presumably due to the protonation of either N^1 or N^2. The similarity of δ values for the protonated forms of **2.1** and **2.2** to the δ value for the trimethylhydrazonium ion **2.7** (Figure 2.4), in which N^2 bears a positive charge, suggests that the site of protonation of methylhydrazone **2.1** and dimethylhydrazone **2.2** is N^2 (**2.14** in Figure 2.5). This interpretation is also supported by N^2 of dimethylhydrazone **2.2** being more nucleophilic than N^1 toward methyl iodide (Scheme 2.1). The only other literature report of attempts to determine the site of hydrazone protonation reached the same conclusion (Zverev *et al.*, 1977). The observed titration of trifluoroacetylhydrazone **2.6** is due to the loss of its N^2 proton, which is made acidic by the proximal trifluoromethyl group.

The value of δ does not correlate with conjugate stability. A high δ value of $C^1\underline{H}$ is indicative of low electron density on C^1, which portends a high susceptibility to attack by nucleophiles. Surprisingly, despite having the largest δ value (Figure 2.4), trimethylhydrazonium ion **2.7** is the most stable conjugate (Figure 2.3 and Table 2.1).

Moreover, oxime **2.3** and acetylhydrazone **2.4** have similar δ values, but at pD 7.0 the half-life of oxime **2.3** is 25 d whereas that of acetylhydrazone **2.4** is 2 h (Table 2.1).

The data are consistent with a mechanism that entails protonation of N^1 (Figure 2.5). The resultant protonated species (**2.15**) would be highly susceptible to hydrolysis due to the enhanced electrophilicity of C^1. None of the conjugates is protonated to a significant extent on N^1 at pD 0.7–13.4 (Figure 2.4), indicating that the pK_a value of species **2.15** is <0.7 in each conjugate. The protonation of N^1 of trimethylhydrazonium ion **2.7** is discouraged by the adjacent quaternary ammonium group. Consequently, trimethylhydrazonium ion **2.7** is highly stable (Figure 2.3), even without the ability to access resonance form **2.10**. This finding belies the textbook (Carey and Sundberg, 2008) and alternative (Wiberg and Glaser, 1992) explanation for the stability of hydrazones and oximes being greater than that of imines. Rather, these conjugates are more stable than imines because of the inductive effect of X^2 = N or O. This explanation is analogous to one for the origin of the α-effect (Buncel and Um, 2004).

The protonation of the N^1 of oxime **2.3** is more favorable than that of trimethylhydrazonium ion **2.7**, accounting for the lower stability of oxime **2.3**. Still, the protonation of the oxime is less favorable than the protonation of alkylhydrazones **2.1**–**2.2** and acylhydrazones **2.4**–**2.6** due to the higher electronegativity of X^2 in the oxime (χ_O = 3.5 (Pauling, 1960)) *versus* the hydrazones (χ_N = 3.0). Hence, oxime **2.3** is more resistant to hydrolysis than are alkylhydrazones **2.1**–**2.2** and acylhydrazones **2.4**–**2.6**.

Finally, we note that the NMR spectra revealed no evidence of a carbinolamine intermediate (**2.16**). Accordingly, the rate-determining step in the hydrolysis reaction is the attack of water on the conjugates, and not the decomposition of the carbinolamine. Jencks and coworkers suggested that the decomposition of the carbinolamine intermediate becomes rate-determining only in extremely acidic conditions (Cordes and Jencks, 1963).

2.5 Conclusions

We have evaluated the hydrolytic stability of a series of isostructural hydrazones and an oxime. We found the oxime to be *much* more stable than the simple hydrazones. pD-Rate profiles and pD-titrations suggest that the anomalous stabilities of the oxime (as well as a trialkylhydrazonium ion) is due to its resistance to protonation. These data can inform the proper use of compounds containing carbon–nitrogen double bonds.

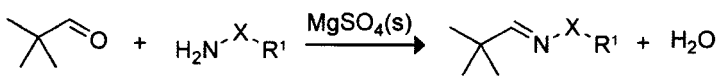
2.6 Acknowledgements

We are grateful to Prof. W. W. Cleland, Prof. S. F. Nelsen, M. D. Shoulders, and T. J. Rutkoski for contributive discussions, and Dr. C. G. Fry for providing laboratory space and unconstrained access to NMR spectrometers in the Department of Chemistry. This work was supported by Grant GM044783 (NIH), and the Materials Research Science and Engineering Center at the University of Wisconsin–Madison (NSF DMR-0520527). The Magnetic Resonance Facility was supported by Grants

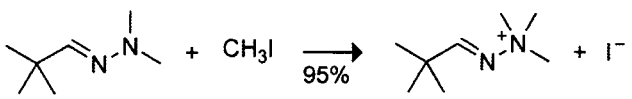
CHE-9208463 (NSF) and RR08389 (NIH). NMRFAM was supported by Grant P41RR02301 (NIH).

Conjugate	pD 5.0	pD 6.0	pD 7.0	pD 8.0	pD 9.0
2.1	9 ± 1 min	24.5 ± 0.6 min	1.0 ± 0.1 h	4.2 ± 0.6 h	19.5 ± 0.5 h
2.2	7.4 ± 0.5 min	11.3 ± 0.2 min	32 ± 3 min	2.0 ± 0.1 h	11.7 ± 0.1 h
2.3	15.7 ± 0.4 h	4.4 ± 0.3 d	~25 d	not determined	not determined
2.4	2.4 ± 0.4 min	21.4 ± 0.8 min	2.0 ± 0.2 h	10.17 ± 0.02 h	4.2 ± 0.7 d
2.5	8.5 ± 0.4 min	36 ± 2 min	3.8 ± 0.5 h	12.3 ± 0.8 h	2.9 ± 0.1 d
2.6	7.5 ± 0.9 min	12.4 ± 0.8 min	14 ± 1 min	23 ± 1 min	1.0 ± 0.1 h
2.7	10.3% hydrolysis in 17 d	not determined	no hydrolysis detected in 22 d	not determined	not determined

Table 2.1 Values of $t_{1/2}$ for the hydrolysis of conjugates **2.1**–**2.7** at pD 5.0–9.0.



2.1	X = NH	$R^1 = CH_3$
2.2	NCH_3	CH_3
2.3	O	CH_3
2.4	NH	$C(O)CH_3$
2.5	NH	$C(O)NH_2$
2.6	NH	$C(O)CF_3$



2.2 **2.7**

Scheme 2.1 Synthesis of hydrazone and oxime conjugates (**2.1–2.7**).

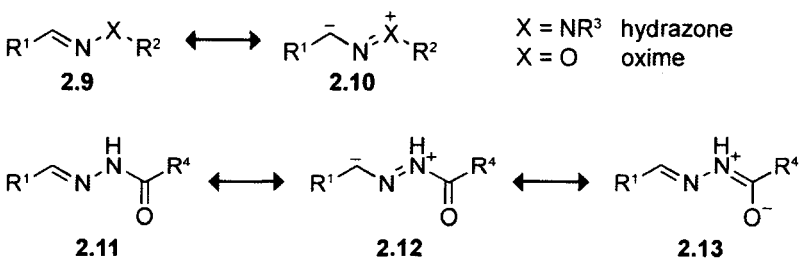


Figure 2.1 Major resonance forms of hydrazones and oximes.

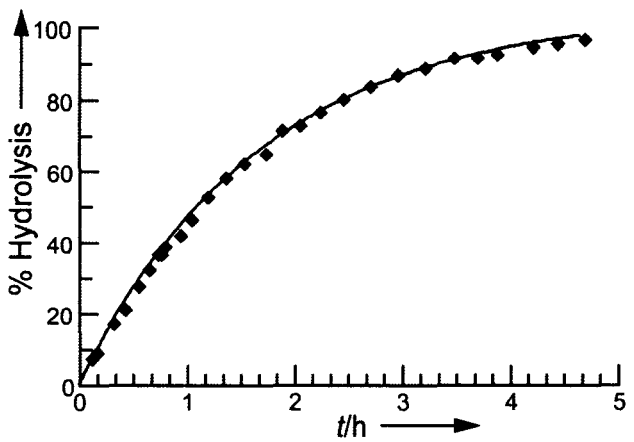


Figure 2.2 Kinetic trace for the hydrolysis of methylhydrazone **2.1** at pD 7.0 in the presence of a 10-fold molar excess of D_2CO. Each data point was obtained by integration of a ^1H NMR spectrum. Similar kinetic traces were obtained for other hydrolysis reactions.

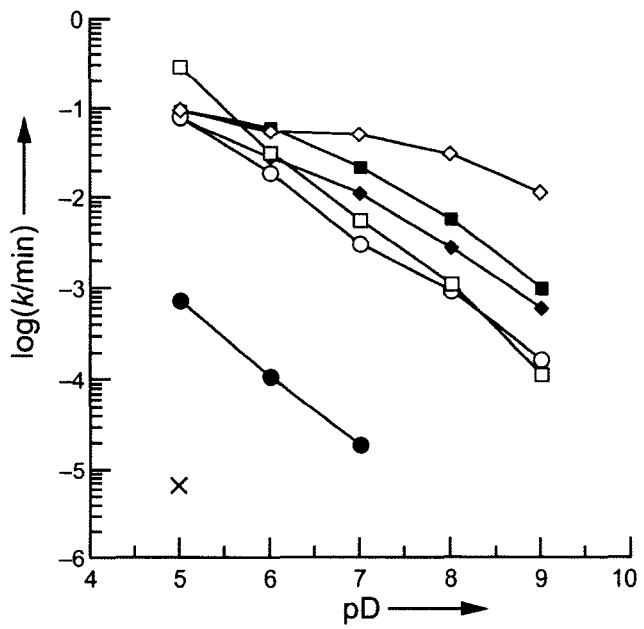


Figure 2.3 pD–rate profiles for the hydrolysis of conjugates **2.1** (◆), **2.2** (■), **2.3** (●), **2.4** (□), **2.5** (○), **2.6** (◇), and **2.7** (×). First-order rate constants (*k*) were calculated from kinetic traces (*e.g.*, Figure 2.2).

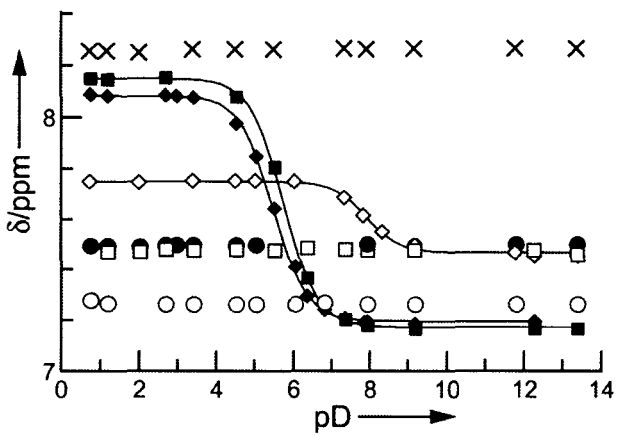


Figure 2.4 pD–Titration of the chemical shifts of $C^1\underline{H}$ of conjugates **2.1** (◆), **2.2** (■), **2.3**, (●), **2.4** (□), **2.5** (○), **2.6** (◇), and **2.7** (×).

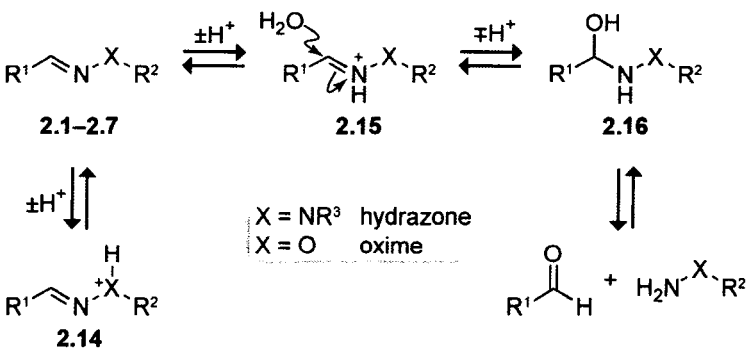
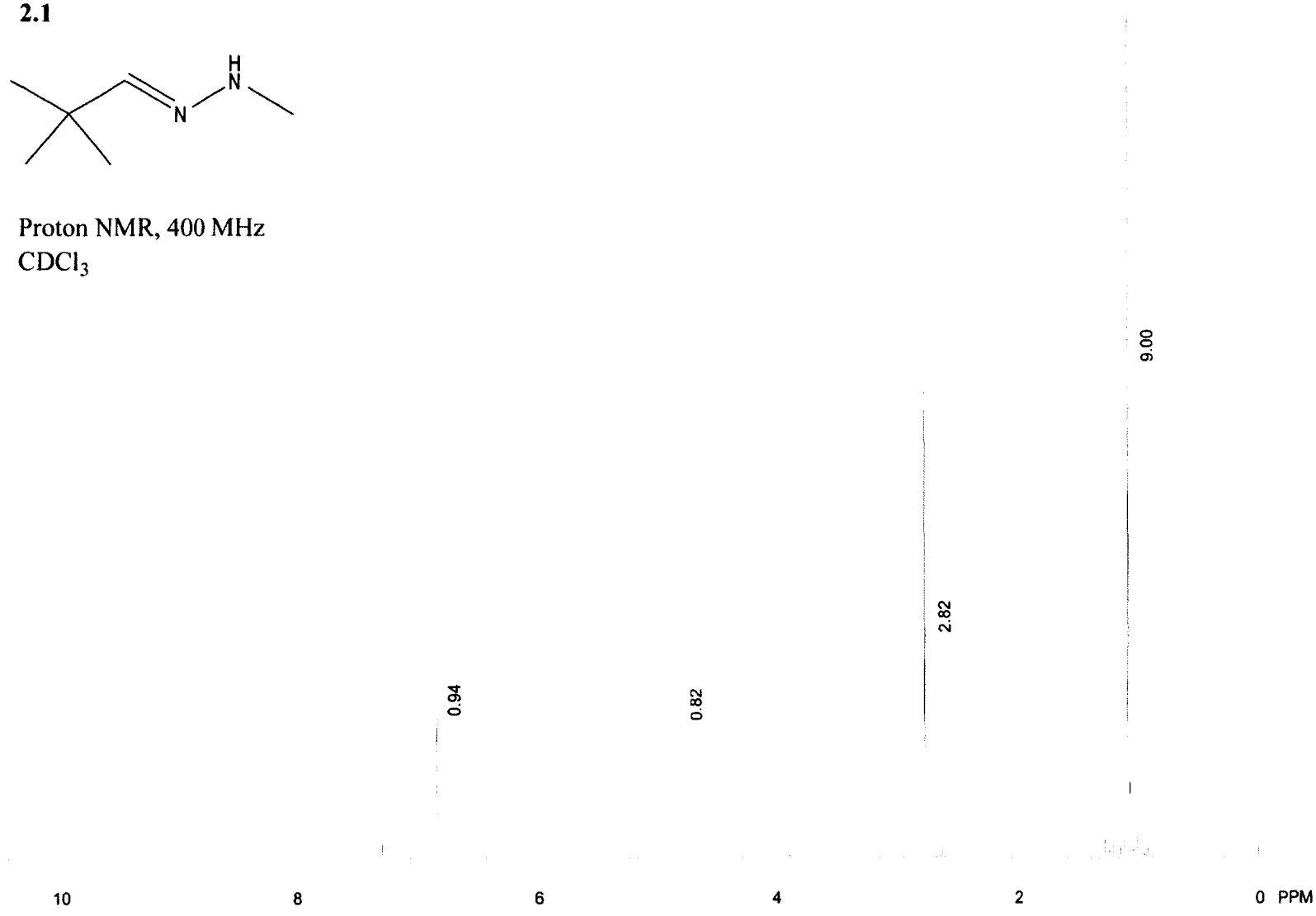


Figure 2.5 Putative mechanism for the hydrolysis of hydrazones and oximes.

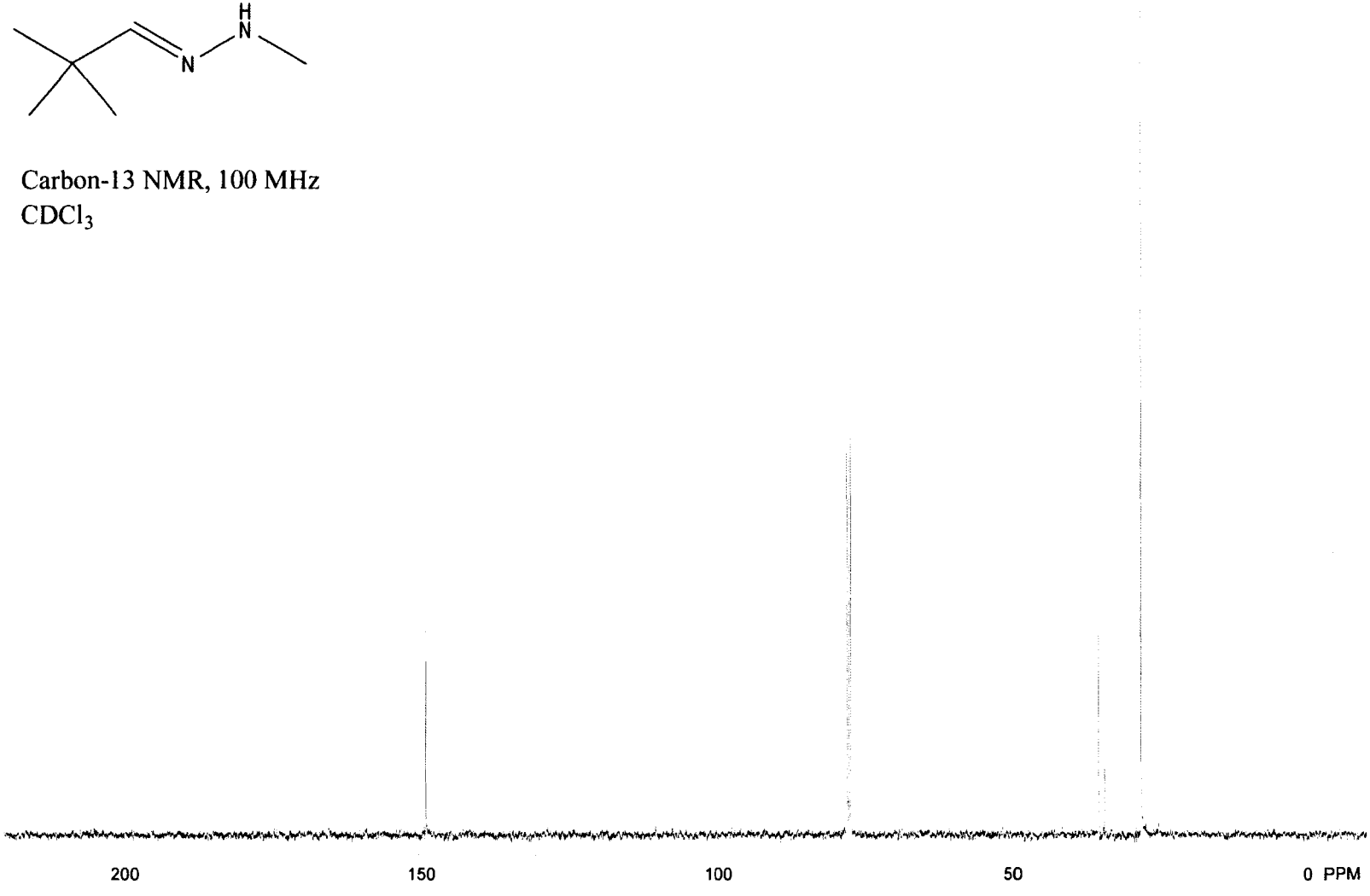
2.1

Proton NMR, 400 MHz
$CDCl_3$

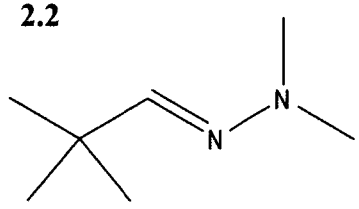


2.1

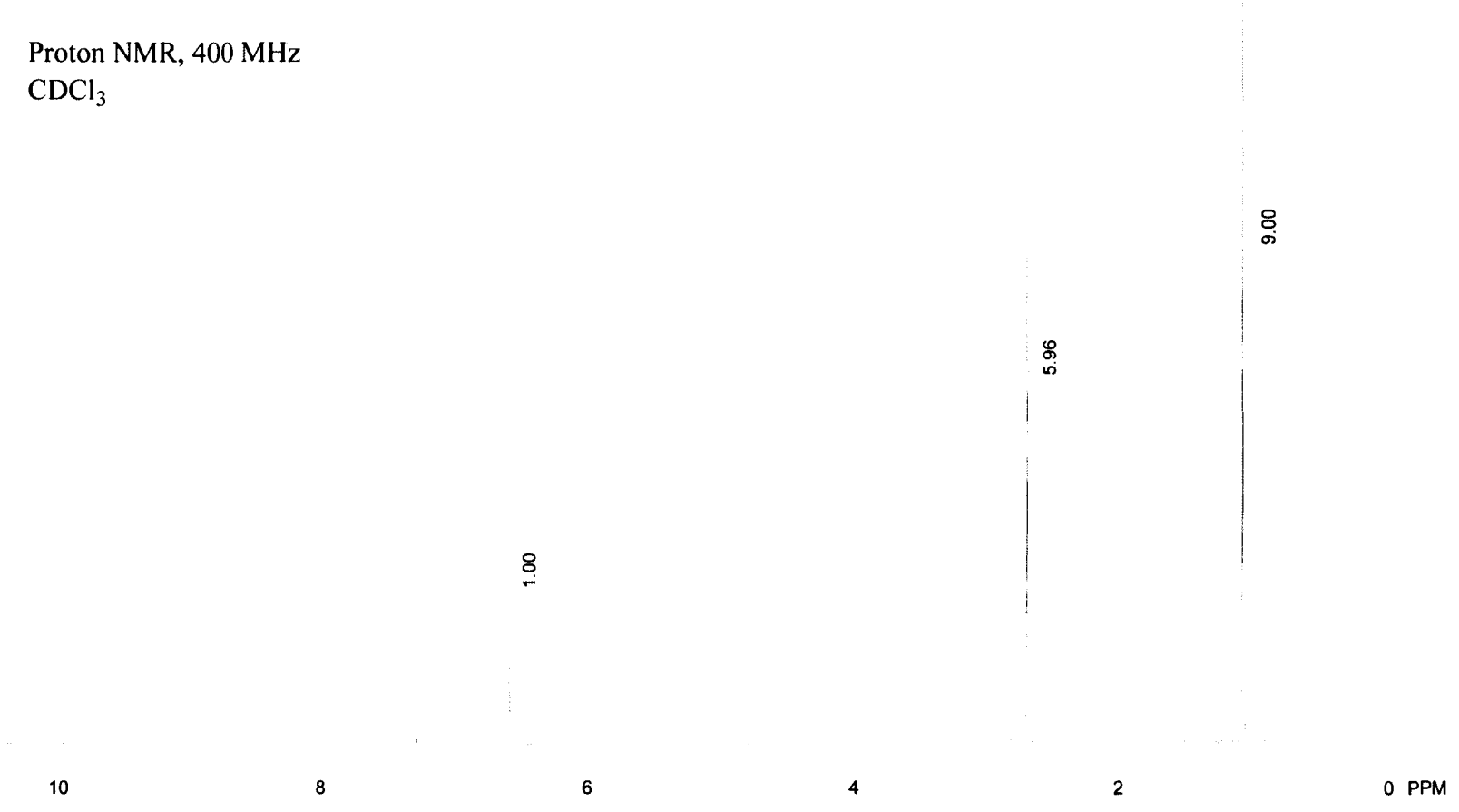
Carbon-13 NMR, 100 MHz
$CDCl_3$



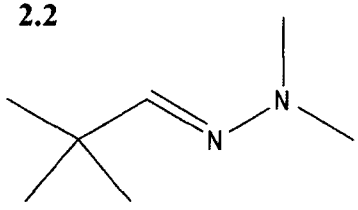
2.2



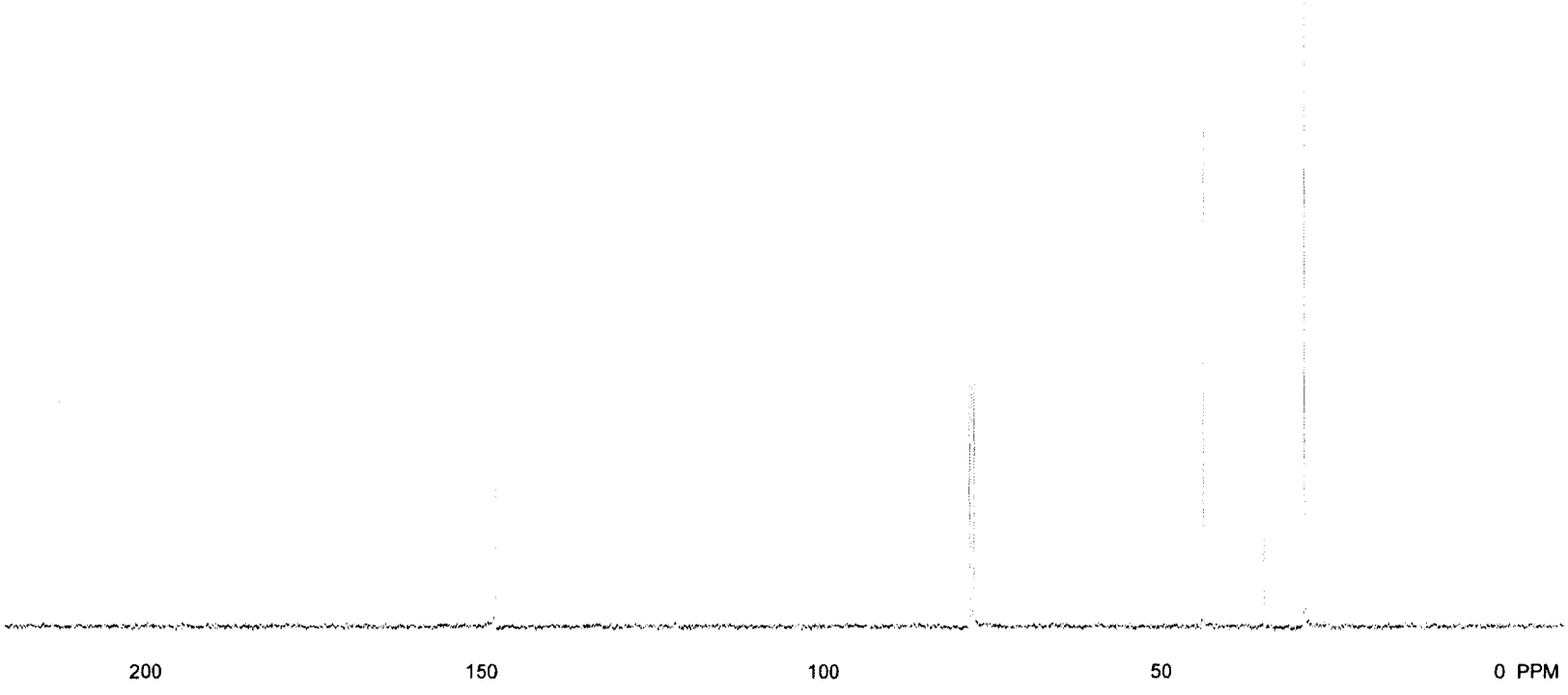
Proton NMR, 400 MHz
$CDCl_3$



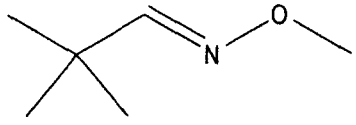
2.2



Carbon-13 NMR, 100 MHz
$CDCl_3$

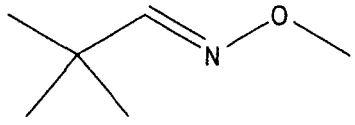


2.3

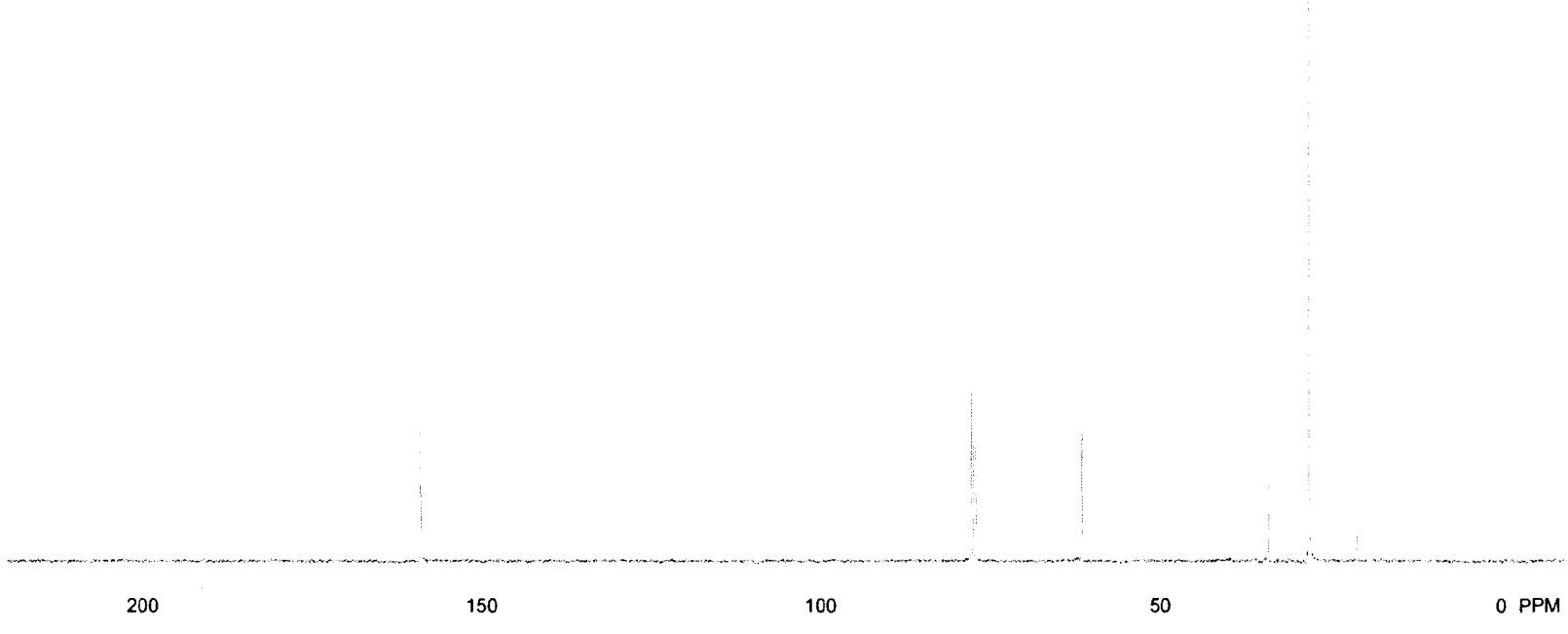


Proton NMR, 400 MHz
$CDCl_3$

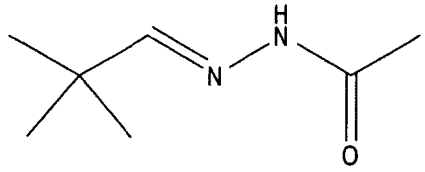
2.3



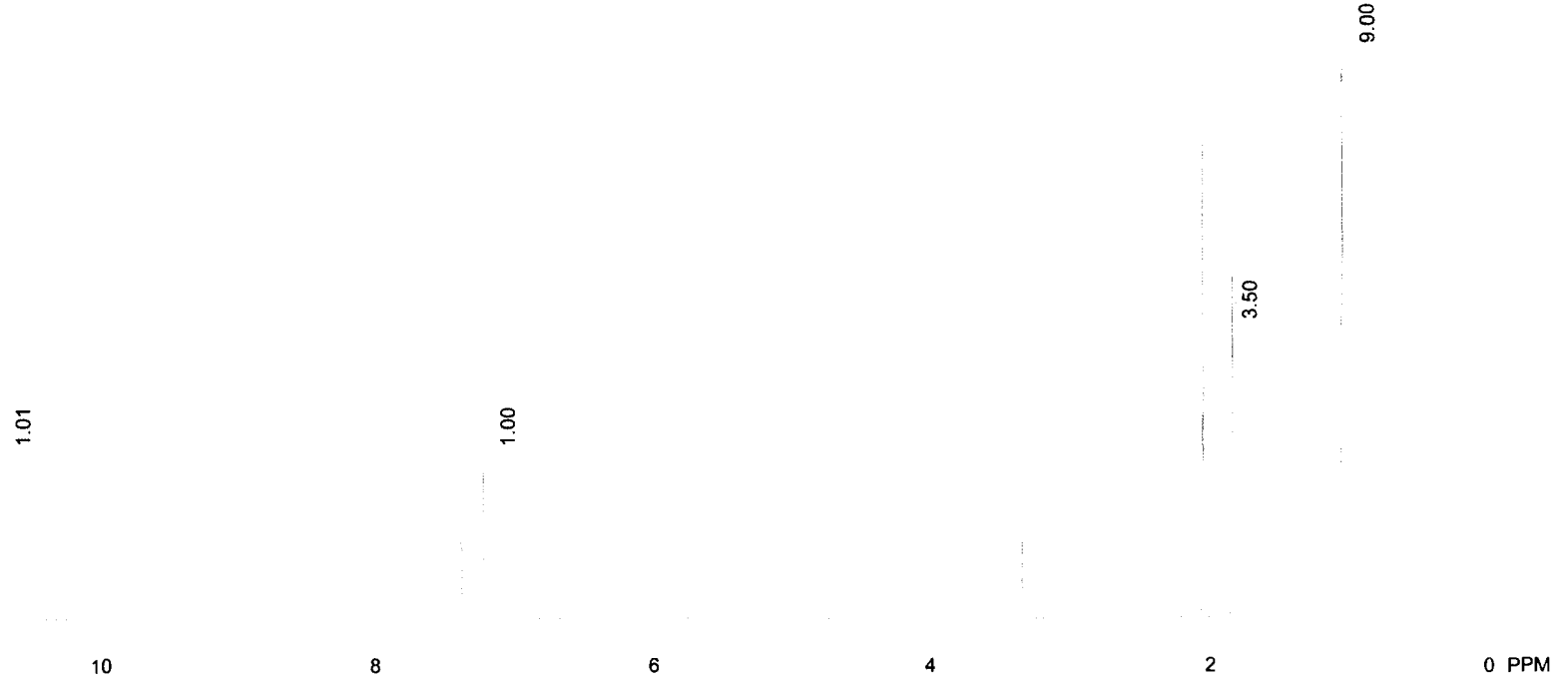
Carbon-13 NMR, 100 MHz
$CDCl_3$



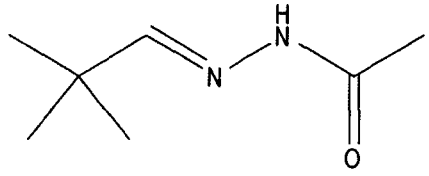
2.4



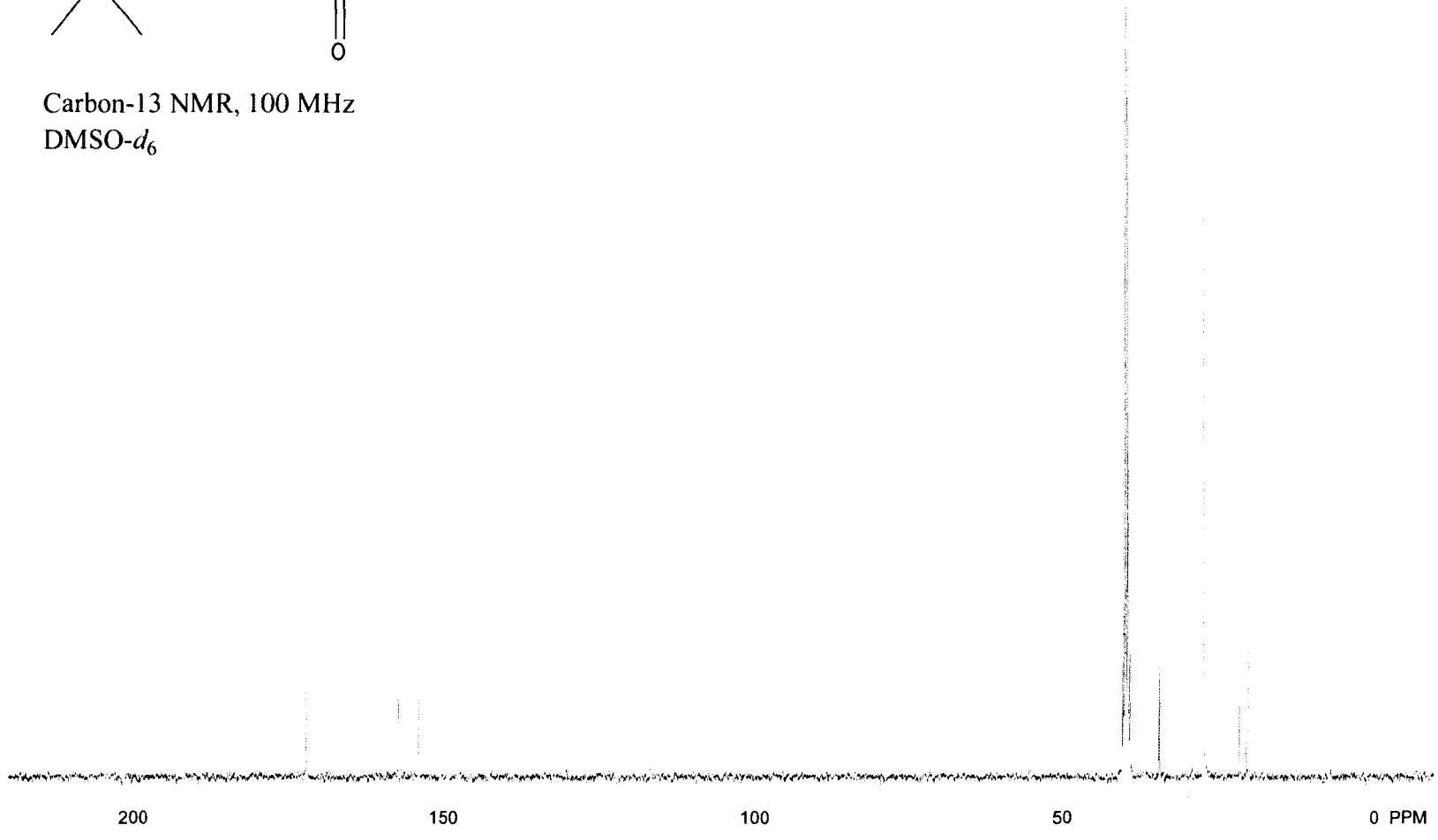
Proton NMR, 400 MHz
DMSO-d_6



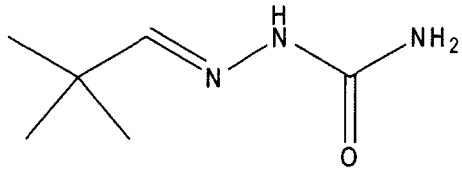
2.4



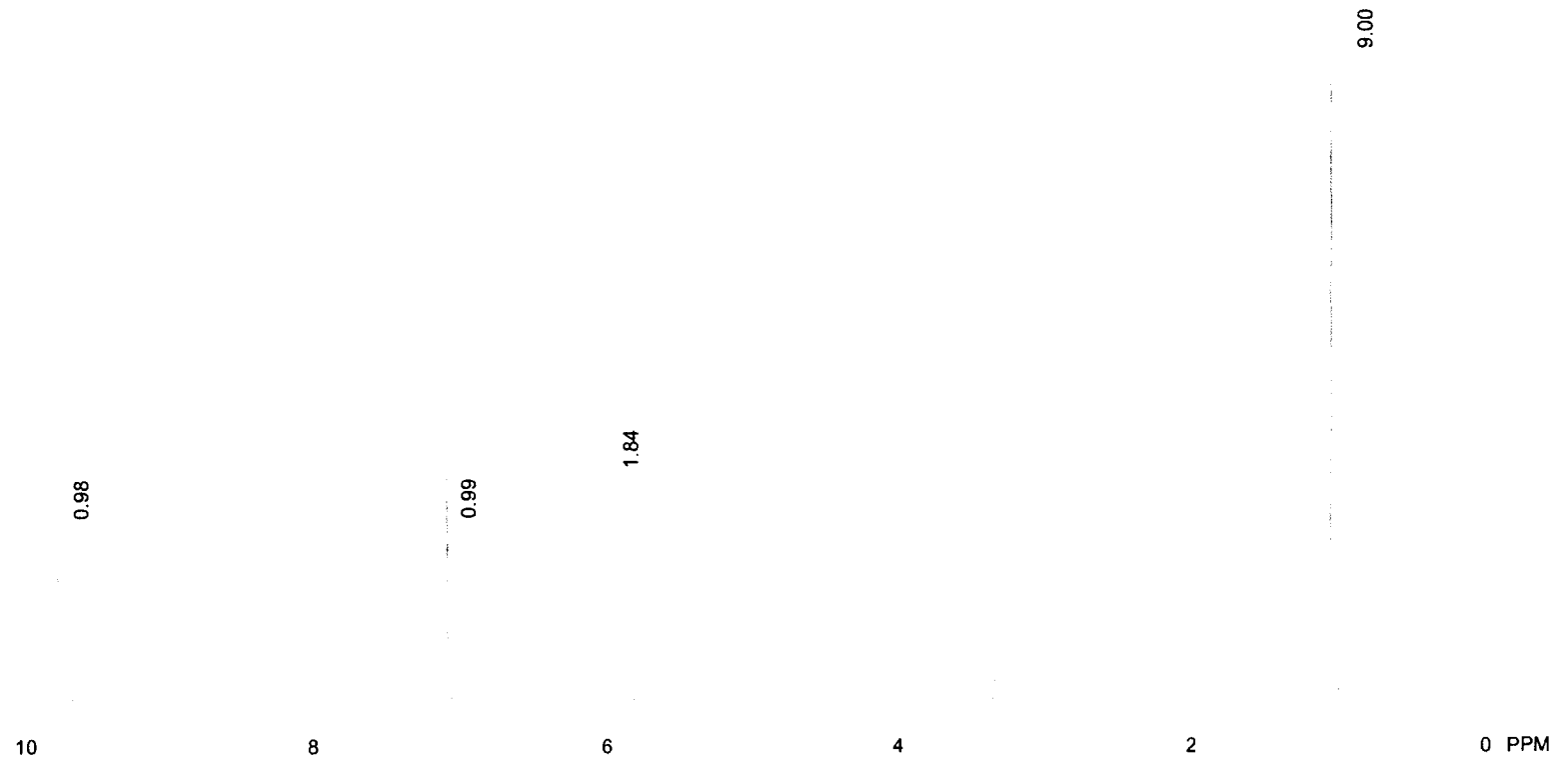
Carbon-13 NMR, 100 MHz
DMSO-d_6



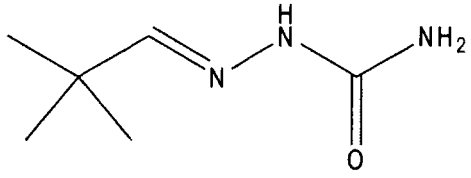
2.5



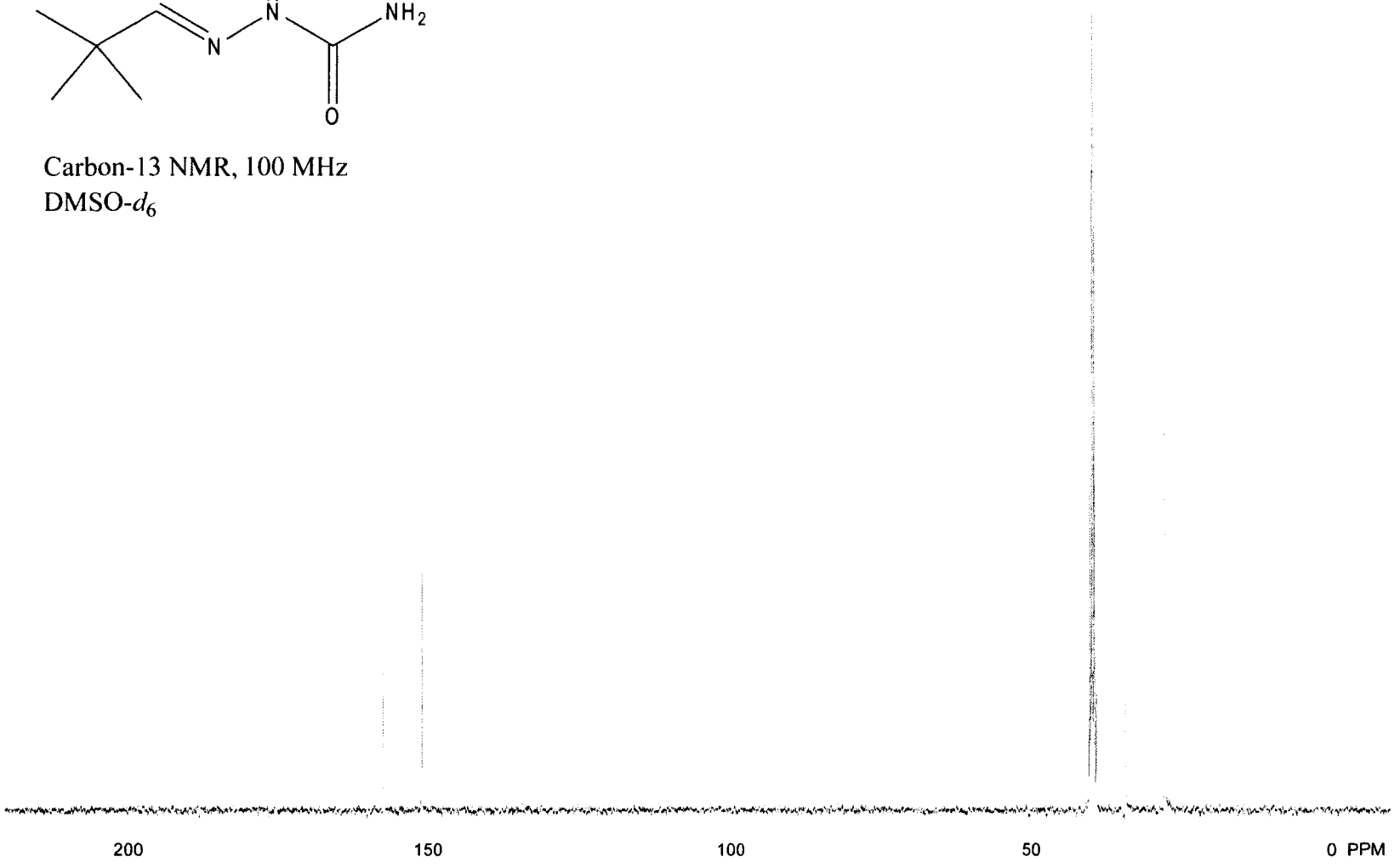
Proton NMR, 400 MHz
DMSO-d_6



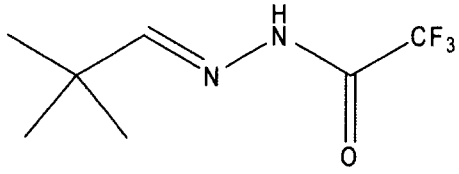
2.5



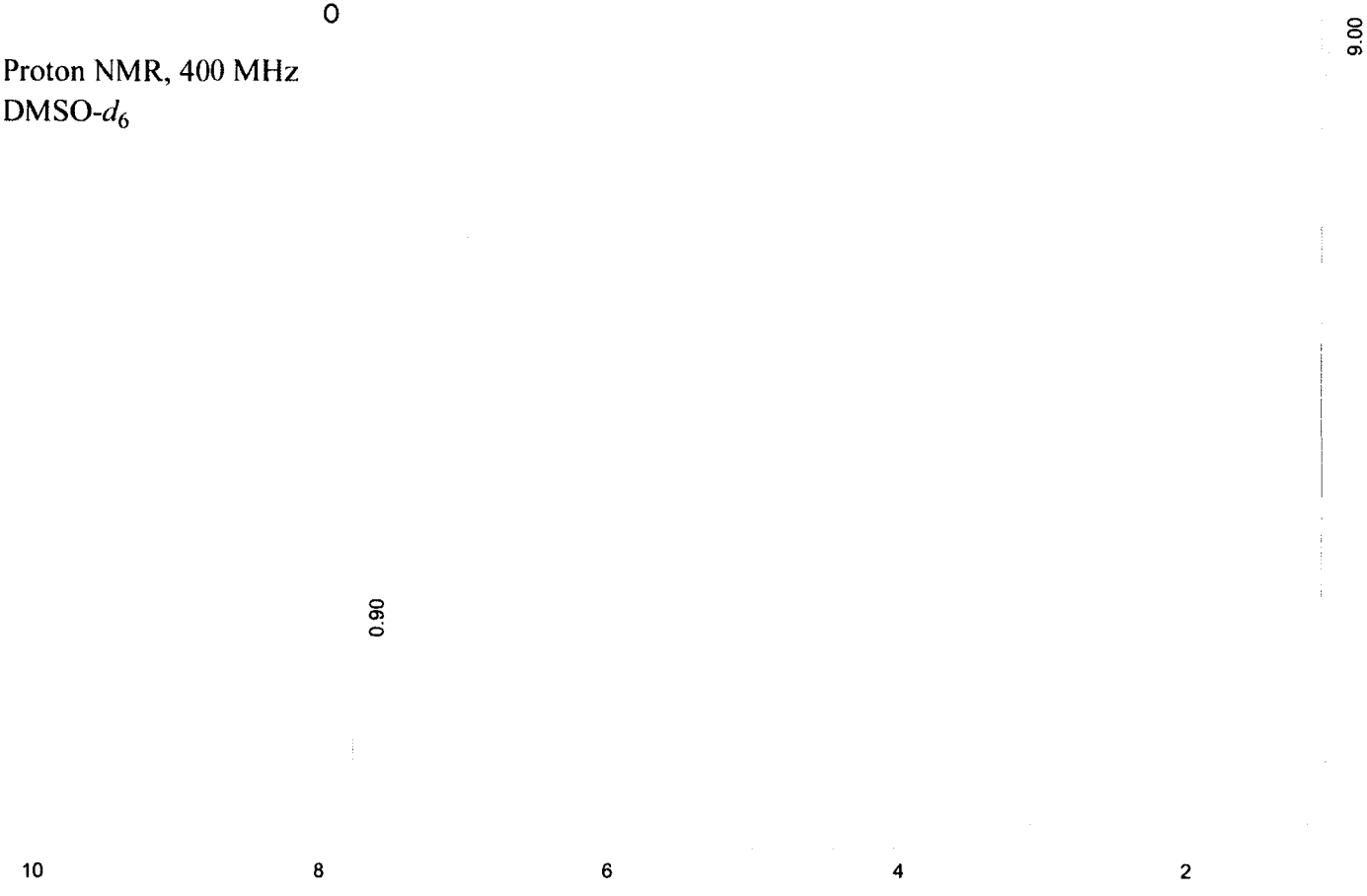
Carbon-13 NMR, 100 MHz
DMSO-d_6



2.6

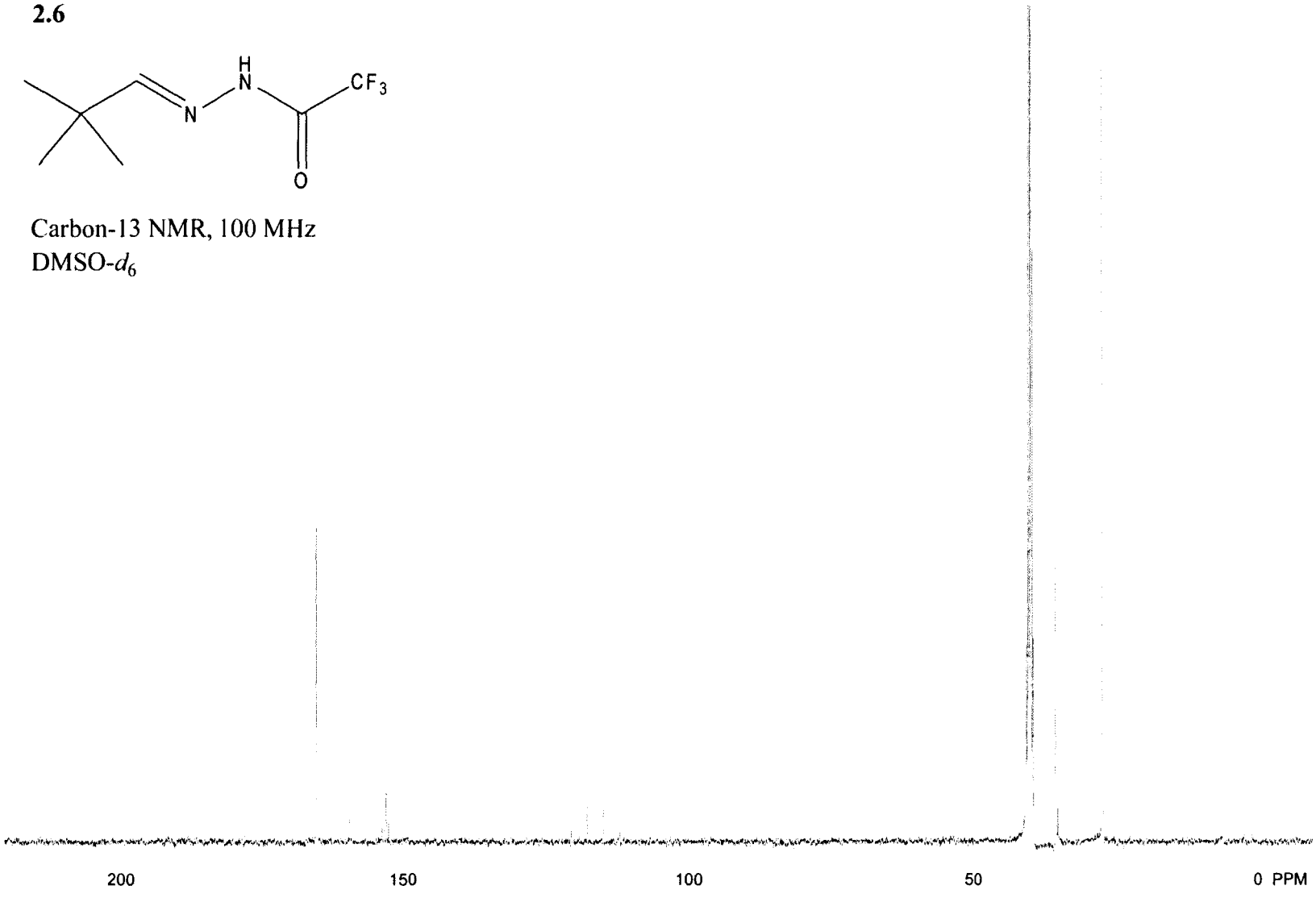


Proton NMR, 400 MHz
DMSO-d_6

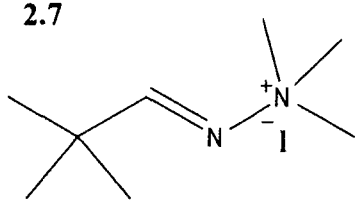


2.6

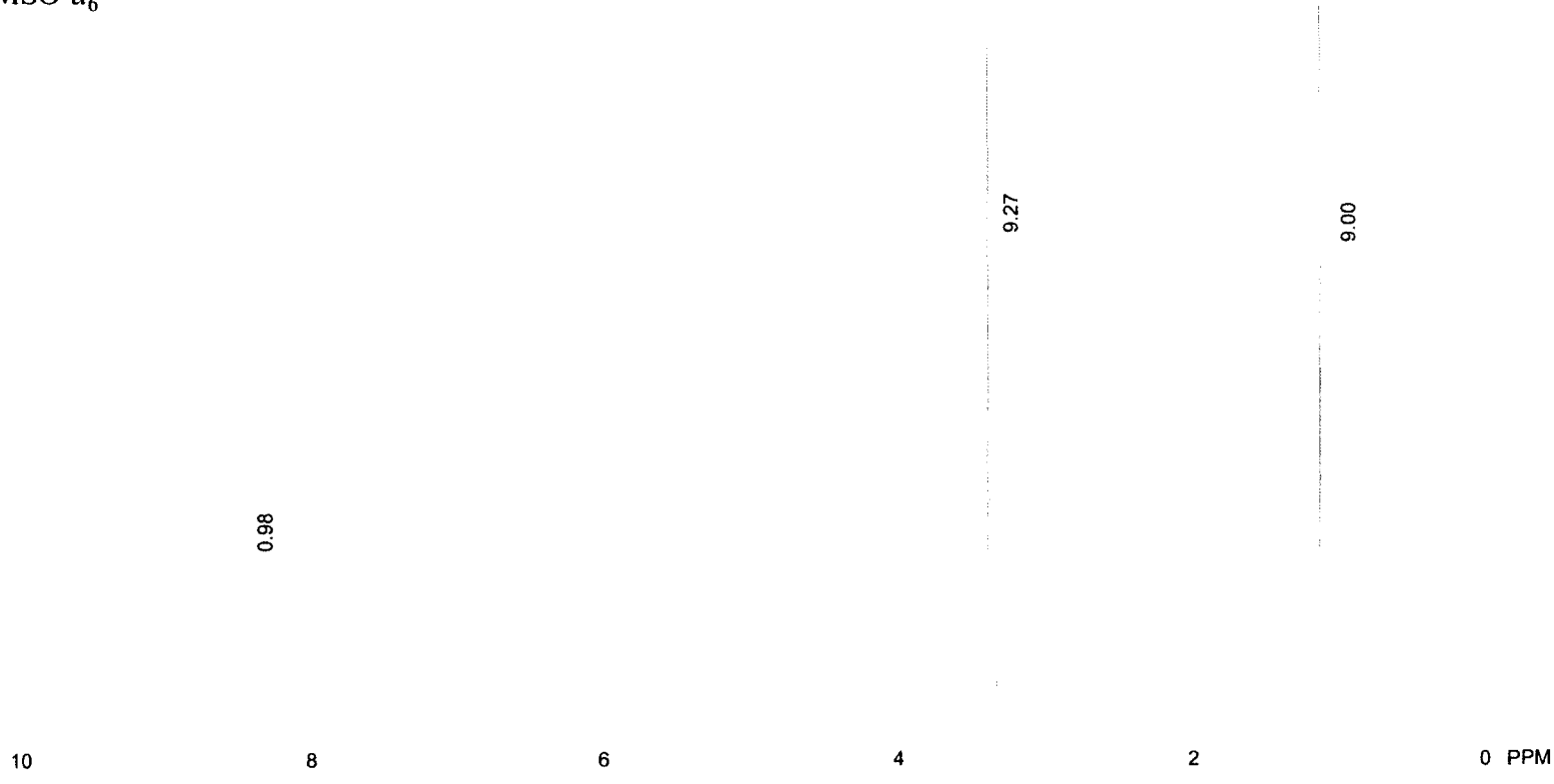
Carbon-13 NMR, 100 MHz
DMSO-d_6



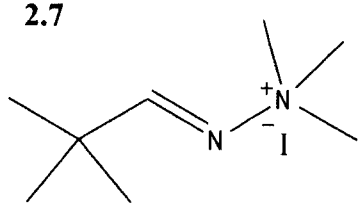
2.7



Proton NMR, 400 MHz
DMSO-d_6



2.7



Carbon-13 NMR, 100 MHz
DMSO-d_6

Chapter Three

Catalysis of Imido-group Hydrolysis in a Maleimide Conjugate

Contribution: Synthesis of chromogenic succinimide, design and execution of spectrophotometric and HPLC-based assays for monitoring hydrolysis, kinetic analyses and calculations, composition of manuscript and figure drafts.

This chapter was published as:
Kalia, J. and Raines, R. T. (2007) Catalysis of Imido-group Hydrolysis in a Maleimide Conjugate. *Bioorg. Med. Chem. Lett.* **18:** 6286–6289.

3.1 Abstract

Maleimides are often used for biomolecular conjugation with thiols. An underappreciated aspect of the imido group in a maleimide conjugate is its susceptibility to spontaneous hydrolysis, resulting in undesirable heterogeneity. Here, a chromophoric maleimide is used to demonstrate that both molybdate and chromate catalyze the hydrolysis of an imido group near neutral pH. Tungstate and 4-(dimethylamino)pyridine are less effective as catalysts. This work reveals a new mode of chemical reactivity for molybdate and chromate, and provides a strategy for decreasing the heterogeneity of bioconjugates derived from maleimides

3.2 Introduction

Chemoselective bioconjugation has revolutionized functional genomics, proteomics, and glycomics (Aslam and Dent, 1998; Lundblad, 2005; Prescher and Bertozzi, 2005). Reactions, such as the Staudinger ligation (Kalia *et al.*, 2007; Köhn and Breinbauer, 2004; Nilsson *et al.*, 2005), the Huisgen 1,3-dipolar cycloaddition (Huisgen, 1963; Kalia and Raines, 2006; Kolb and Sharpless, 2003), carbonyl group condensation reactions (Grandjean *et al.*, 2000; Shao and Tam, 1995; Tully *et al.*, 2006; Zatsepin *et al.*, 2005), and maleimide conjugation (Houseman *et al.*, 2002a; Lavis *et al.*, 2006; Slavica *et al.*, 2007; Watzke *et al.*, 2006a) enable the site-specific labeling of biomolecules, and their uniform immobilization to generate microarrays.

Bioconjugation via maleimides is especially common. This reaction involves the Michael addition of a biomolecular thiolate, often from the side chain of a cysteine

residue, to a maleimide to form a succinimidyl thioether (Bednar, 1990; Friedmann, 1952; Gregory, 1955). The reaction proceeds rapidly and in high yield in neutral aqueous solutions at room temperature, making it ideal for biological applications.

An underappreciated aspect of maleimides and succinimides is the tendency of their imido groups to undergo spontaneous hydrolysis (Clark, 2003; Nektar Therapeutics AL, 2004). As water can attack either of the two carbons of the imido group, hydrolysis of a succinimidyl thioether produces isomeric succinamic acid thioethers. Accordingly, a thiol-conjugated maleimide actually consists of a mixture of many species, some having an additional carboxylic acid that can ionize near neutral pH. As hydrolysis proceeds over time, the concentrations of these species change, introducing yet another complexity.

The heterogeneity of succinimidyl thioethers can alter the activity of bioconjugates and convolute the interpretation of experimental data. For example, an *N*-(1-pyrene)maleimide-conjugate of α,α-tropomyosin was observed to hydrolyze in aqueous solution, causing a time-dependent increase in fluorescence (Ishii and Lehrer, 1986). The cellular internalization of some peptides and proteins relies on their cationic charge (Fuchs and Raines, 2006; Johnson *et al.*, 2007), and the presence of a hydrolyzed succinimidyl thioether could complicate data analysis by reducing that charge. Surface-based assays can be especially sensitive to heterogeneity within the immobilized species. For example, the orientation of liquid crystals near a surface is likely to highly sensitive to non-uniformity (Luk *et al.*, 2004). Heterogeneity can also affect data analysis in experiments that are independent of molecular charge. For

example, assays based on mass spectrometry can be confounded by the additional 18 amu introduced upon hydrolysis of a succinimidyl thioether (Rutkoski and Raines; Underbakke and Kiessling).

We reasoned that the heterogeneity of succinimidyl thioethers could be reduced by intentionally *increasing* their rate of hydrolysis. The hydrolysis of maleimides and succinimides and its catalysis has been studied previously. In 1955, Gregory reported that the rate constant for maleimide hydrolysis is considerably less than that for its reaction with a thiol (Gregory, 1955). Subsequently, the hydrolysis of both maleimides and succinimides was found to be catalyzed by base (Matsui and Aida, 1978; Niyaz Khan and Aziz Khan, 1975). In general, however, bioconjugates are not stable at elevated pH. Here, we report on catalysts of imido group hydrolysis that are effective near neutral pH.

3.3 Experimental Procedures

Materials. *N*-(*p*-nitrophenyl) maleimide (**3.1**) was from TCI America (Portland, OR). Ethanethiol was purchased from Sigma-Aldrich (St. Louis, MO). Synthetic reactions were monitored by thin-layer chromatography with visualization by UV-light or staining with phosphomolybdic acid. Flash chromatography was performed with columns of silica gel 60, 230–400 mesh (Silicycle, Quebec City, Quebec, Canada).

Instrumentation. A Cary 400 Bio UV/VIS spectrophotometer (Varian, Palo Alto, CA) was used to perform kinetic assays. NMR spectra were acquired with a Bruker DMX-

400 Avance spectrometer (^{1}H: 400 MHz, ^{13}C: 100 MHz) at the NMR Facility at Madison (NMRFAM). Carbon-13 spectra were proton-decoupled. Mass spectra on small organic molecules were obtained with electrospray ionization (ESI) techniques. Analytical HPLC was performed with a Varian C_{18} analytical column. Semi-preparative HPLC was performed with a Varian C_{18} semi-prep column. Elemental analysis was performed by Midwest Microlab LLC (Indianapolis, IN).

Synthesis of chromogenic succinimide **3.2**. *N*-(*p*-Nitrophenyl)maleimide (2.0 g, 9.17 mmol) was dissolved in acetonitrile (100 mL) and water (20 mL). Ethanethiol (5.5 mL, 74.25 mmol) was added, and the resulting solution was stirred for 2.5 h. Solvent and unreacted ethanethiol were removed under reduced pressure, and the residue was purified by flash chromatography (silica gel, 60% v/v hexanes in ethyl acetate) to give the desired product as an off-white solid (1.29 g, 50%). ^{1}H NMR (400 MHz, CD_3CN) δ = 8.35 (d, J = 9.2 Hz, 2H), 7.60 (d, J = 9.2 Hz, 2H), 4.05 (dd, J = 9.3, 4.3 Hz, 1H), 3.37 (dd, J = 18.7, 9.3 Hz, 1H), 2.87 (m, 2H), 2.72 (dd, J = 18.7, 4.6 Hz, 1H), 1.32 (t, J = 7.4 Hz, 3H); ^{13}C NMR (100 MHz, $CDCl_3$) δ = 174.8, 173.0, 147.2, 137.3, 127.0, 124.5, 38.9, 36.2, 26.4, 14.2; anal. calcd. for $C_{12}H_{12}N_2SO_4$: C 51.41, H 4.32, N 10.00, S 11.44, O 22.83; found: C 51.76, H 4.26, N 9.96, S 11.81, O 22.21.

Spectrophotometric assay for monitoring hydrolysis. Putative catalysts were dissolved in 0.50 M HEPES–NaOH buffer at pH 7.5 containing NaCl (0.10 M). Na_2SO_4 was added to maintain a constant ionic strength of I = 1.6. The catalyst

solution (157 μL) was added to a cuvette, and the absorbance at 319 nm was adjusted to zero. Hydrolysis was initiated by adding maleimide **3.1** or succinimide **3.2** (3 μL of a 1.65 mM solution in acetonitrile) to the cuvette, and the absorbance at 319 nm was monitored with time. Kinetic traces were fitted to the equation:

$$(A_{319\,\mathrm{nm}})_t = (A_{319\,\mathrm{nm}})_{t=\infty}(1 - e^{-kt}) \tag{3.1}$$

to give the rate constant, k. The uncatalyzed rate constant for hydrolysis was subtracted from the catalyzed rate constant to yield the first-order rate constant, k_1. The k_1 values were plotted versus catalyst concentration, and the data were fitted to the equation: $k_1 = k_2$[catalyst], where k_2 is the second-order rate constant.

HPLC assay for monitoring hydrolysis. Succinimide **3.2** (382 μL of a 1.65 mM solution in acetonitrile) was added to a solution (20 mL) of 50 mM HEPES–NaOH buffer at pH 7.5 containing catalyst (0.10 M) and NaCl (0.10 M). The reaction mixture was stirred for 2 h, and extracted with ethyl acetate (2 × 20 mL). The organic layer was concentrated under reduced pressure, and the residue was analyzed by C18 HPLC using a linear gradient of water (80–20% v/v) in acetonitrile containing TFA (0.1% v/v) over 50 min. Hydrolyzed succinimide **3.2** eluted in two peaks at 40 min; intact succinimide **3.2** eluted at 45 min. To compare molybdate and chromate as catalysts, reactions were performed on 90-mL scale, and 20-mL aliquots were removed at four time points and subjected to HPLC analysis.

3.4 Results and Discussion

As a model system that enables facile quantitation of the rate of imido group hydrolysis, we chose a maleimide that is activated for hydrolysis and that undergoes a spectral change upon hydrolysis: *N*-(*p*-nitrophenyl)maleimide (**3.1**). We synthesized a model succinimidyl thioether by reacting maleimide **3.1** with ethanethiol to give succinimide **3.2**, as shown in Scheme 3.1. This conjugate has an absorbance maximum at 271 nm before hydrolysis that shifts to 319 nm after hydrolysis, enabling a continuous spectrophotometric assay of hydrolysis.

The molybdate anion is known to act as a nucleophilic catalyst for the hydrolysis of oxoesters (Wikjord and Byers, 1992a; Wikjord and Byers, 1992b), thioesters (Wikjord and Byers, 1992b), phosphoesters (Ahn *et al.*, 2000), 2,4-dinitrofluorobenzene (Wikjord and Byers, 1992a), and methyl iodide (Wikjord and Byers, 1992a). The nucleophilicity of molybdate is remarkable, given its low basicity (pK_a 4.1) (Sasaki *et al.*, 1959). To determine if molybdate could also act as a catalyst of imido-group hydrolysis, we incubated succinimide **3.2** with varying concentrations of Na_2MoO_4 in 0.50 M HEPES–NaOH buffer at pH 7.5, containing NaCl (0.10 M). The resulting kinetic traces were indicative of substantial catalysis (Figure 3.1A). Similar kinetic traces were obtained with maleimide **3.1**. First-order rate constants were derived from exponential fits of the kinetic traces obtained with different concentrations of molybdate, and were plotted versus the molybdate concentration (Figure 3.1B). The ensuing second-order rate constant for the hydrolysis of

maleimide **3.1** (k_2 = 2.5 $M^{-1}min^{-1}$) was 9-fold greater than that for succinimide **3.2** (k_2 = 0.28 $M^{-1}min^{-1}$).

Kinetic analyses were also performed with other group VI oxometallates, namely, those of chromium and tungsten. Chromate has a high absorbance at 271 and 319 nm, obviating the use of the spectrophotometric assay. To circumvent this obstacle, an HPLC-based assay was developed to monitor catalysis by chromate. An aqueous solution of succinimide **3.2** was incubated with Na_2CrO_4 (0.10 M) for 2 h. The reaction was quenched by extraction, and analyzed by using analytical HPLC and monitoring at 271 nm. Unhydrolyzed succinimide **3.2** elutes at ~45 min (Figure 3.2); hydrolyzed succinimide **3.2** elutes in two closely spaced peaks at ~40 min and was characterized by performing hydrolysis on a larger scale and isolating the hydrolysis products by semi-preparative HPLC. A solution of Na_2SO_4 (0.10 M) was used to report on the uncatalyzed rate of hydrolysis. The data indicated that chromate is a comparable catalyst to molybdate, whereas tungstate is a poor catalyst (Figure 3.2). 4-(Dimethylamino)pyridine (DMAP), a common catalyst for acylation reactions (Fu, 2004), accelerated the hydrolysis to a significantly lesser extent than did molybdate or chromate (Figure 3.2). To compare catalysis by molybdate and chromate directly, succinimide **2** was incubated with molybdate or chromate (0.10 M) for different time periods, and its hydrolysis was quantitated by peak integration of HPLC traces. The results indicate that the apparent first-order rate constant in the presence of molybdate was 1.3-fold higher than that in the presence of chromate (Figure 3.3).

The mechanism of molybdate-catalyzed ester hydrolysis has been proposed to proceed via the nucleophilic attack of a molybdate oxygen on the ester carbon, followed by the spontaneous hydrolysis of the resulting acid molybdate intermediate (Wikjord and Byers, 1992a; Wikjord and Byers, 1992b). An analogous mechanism for the molybdate- and chromate-catalyzed hydrolysis of an imido group is shown in Scheme 3.2. Why are these two group VI oxometallates effective catalysts? To search for a correlation between reactivity and basicity, Um and coworkers subjected the reaction of *p*-nitrophenylacetate with group VI oxometallates to a Brϕnsted analysis (Ahn *et al.*, 2000). Although tungstate conformed to the expected trend, molybdate and chromate showed significant positive deviations. The anomalous nucleophilicity of molybdate has been attributed to its poor solvation (Wikjord and Byers, 1992b). In addition, the high reactivity of molybdate with oxoesters and thioesters has been attributed to Lewis acid catalysis via formation of a transient Mo–O bond with the nonbridging ester oxygen in a tetrahedral intermediate (Wikjord and Byers, 1992a; Wikjord and Byers, 1992b).

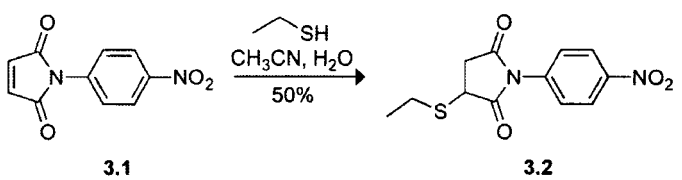
3.5 Conclusions

We have discovered that molybdate and chromate catalyze the hydrolysis of the imido group of maleimide conjugates. This discovery adds to the unexpected nucleophilic reactivity of molybdate and chromate, and provides a potential strategy for hydrolyzing the unactivated succinimidyl thioethers found in bioconjugation reactions (Aslam and Dent, 1998; Lundblad, 2005; Prescher and Bertozzi, 2005).

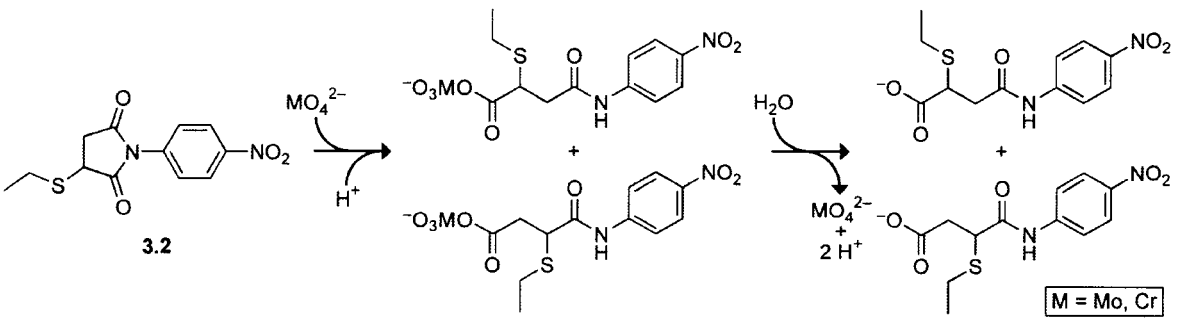
These catalysts contain but five atoms, and are readily separable from bioconjugates by size-exclusion chromatography, gel-filtration chromatography, or dialysis. Although chromate reacts with thiols to form chromium(VI) thioesters (Connett and Wetterhahn, 1985; Levina and Lay, 2004), maleimide conjugates would presumably be devoid of thiol groups. Finally, we anticipate that our work will increase awareness of the heterogeneity of maleimide conjugates and inspire means to diminish that heterogeneity.

3.6 Acknowledgements

We are grateful to L. E. Strong for contributive discussions. This work was supported by grant GM44783 (NIH) and the Materials Research Science and Engineering Center at the University of Wisconsin–Madison (NSF DMR-0520527).



Scheme 3.1 Route for the synthesis of succinimide **3.2**.



Scheme 3.2 Putative mechanism for the molybdate- and chromate-catalyzed hydrolysis of succinimide **3.2**.

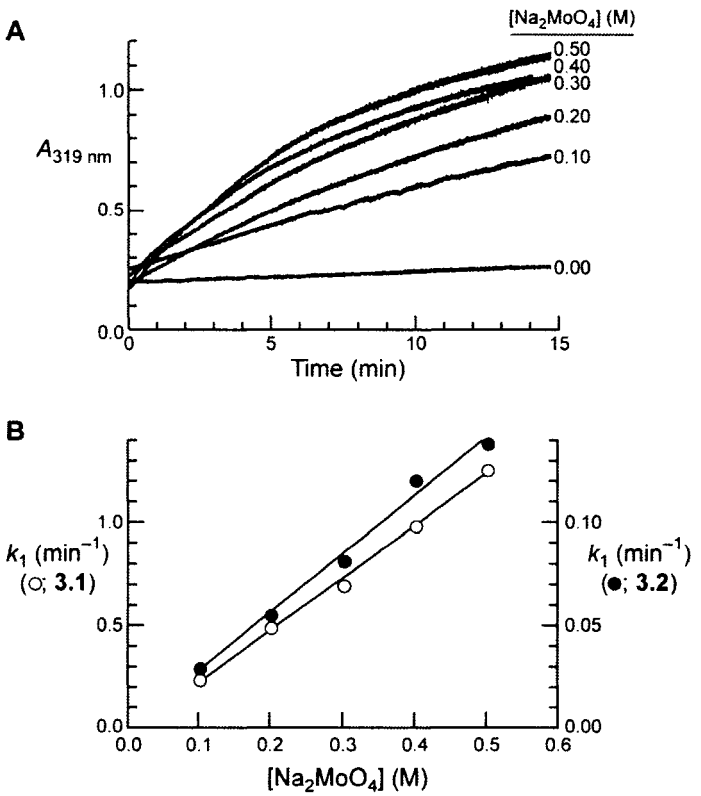


Figure 3.1 Molybdate-catalyzed hydrolysis of imido groups. A) Raw data for catalysis of the hydrolysis of succinimide **3.2**. B) Transformed data for the hydrolysis of maleimide **3.1** ($k_2 = 2.5\ M^{-1}\ min^{-1}$) and succinimide **3.2** ($k_2 = 0.28\ M^{-1}\ min^{-1}$).

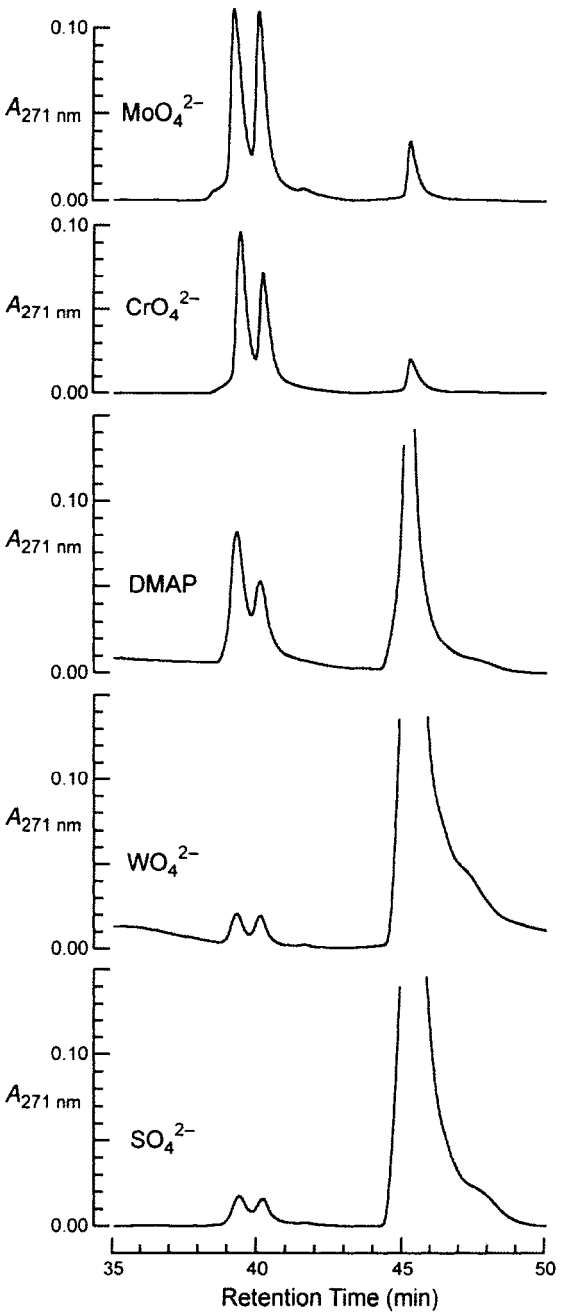


Figure 3.2 HPLC analysis of succinimide **3.2** after incubation for 2h at room temperature with a putative catalyst (0.10 M) of hydrolysis. The twin peaks at ~40 min correspond to the hydrolysis products; the peak at ~45 min corresponds to intact succinimide **3.2**.

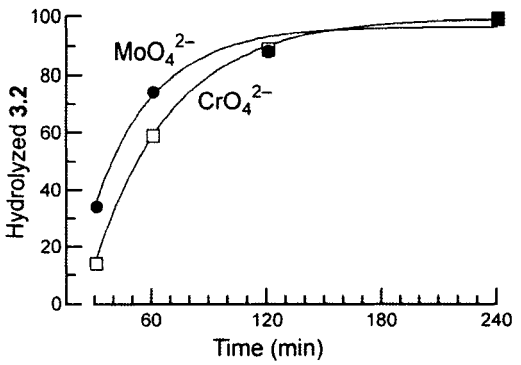
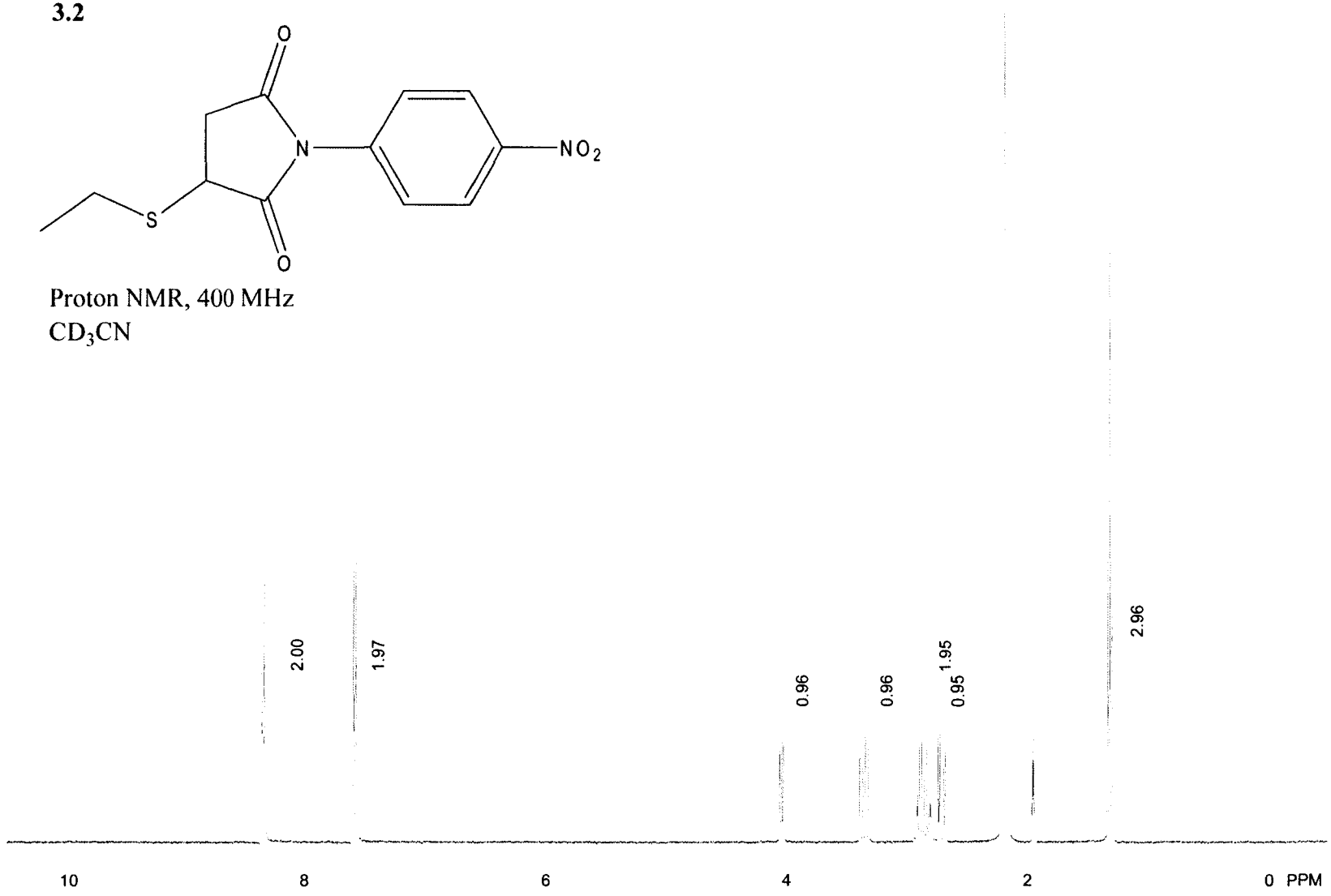


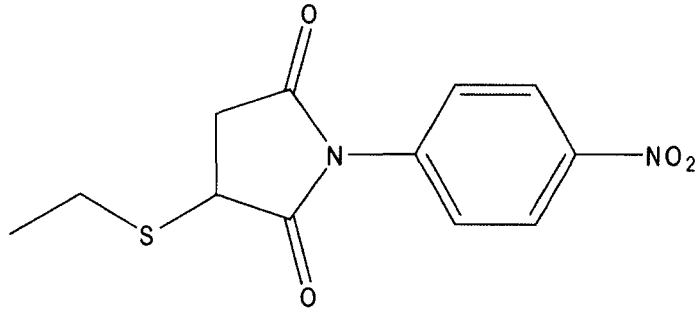
Figure 3.3 Comparison of molybdate and chromate as catalysts for the hydrolysis of succinimide **3.2**. Assays were performed by HPLC, as shown in Figure 3.2. Data were fitted to the equation: [hydrolyzed **3.2**]$_t$ = [hydrolyzed 3.2]$_{t=\infty}(1 - e^{-kt})$, where k is the apparent first-order rate constant.

3.2

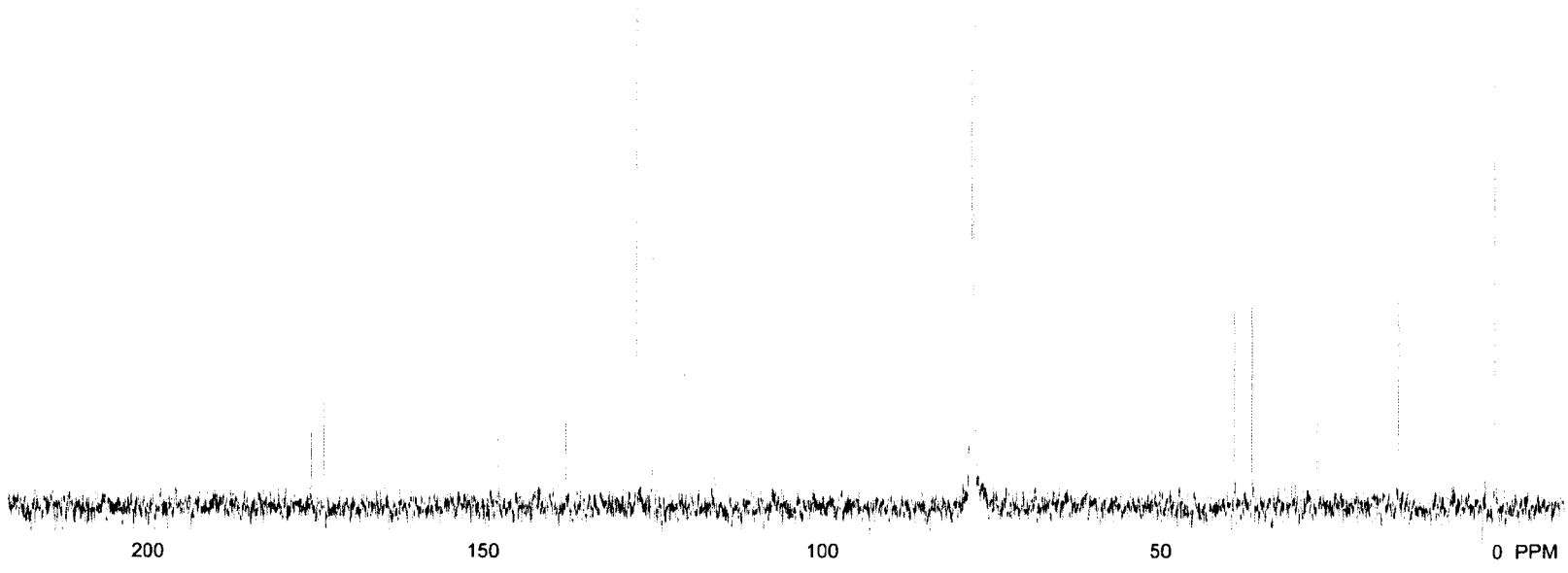
Proton NMR, 400 MHz
CD_3CN



3.2



Carbon-13 NMR, 100 MHz
$CDCl_3$



Chapter Four

Azido–Proteins: Semisynthesis and Applications

Contribution: Kinetic assays and calculations, synthesis of bifunctional molecules, alkynyl fluorescein, and triazole ligand, site-directed mutagenesis, protein expression in *E. coli*, protein purification and characterization, composition of manuscript and figure drafts.

This chapter was published as:
Kalia, J. and Raines, R. T. (2006) Reactivity of Intein Thioesters: Appending a Functional Group to a Protein. *ChemBioChem* **7:** 1375–1383.

4.1 Abstract

The success of genome sequencing has heightened the demand for new means to manipulate proteins. An especially desirable goal is the ability to modify a target protein at a specific site with a functional group of orthogonal reactivity. Here, we achieve that goal by exploiting the intrinsic electrophilicity of the thioester intermediate formed during intein-mediated protein splicing. Detailed kinetic analyses of the reaction of nitrogen nucleophiles with a chromogenic small-molecule thioester revealed that the α-hydrazino acetyl group was the optimal nucleophile for attacking a thioester at neutral pH to form a stable linkage. A bifunctional reagent bearing an α-hydrazino acetamido and azido group was synthesized in high overall yield. This reagent was used to attack the thioester linkage between a target protein and intein, and thereby append an azido group to the target protein in a single step. The azido-protein retained full biological activity, and its azido group was available for chemical modification and its azido group was available for chemical modification by Huisgen 1,3-dipolar azide–alkyne cycloaddition. Thus, the mechanism of intein-mediated protein splicing provides the means to install a useful functional group at a specific site—the C-terminus—of virtually any protein.

4.2 Introduction

Proteins are nucleophilic. Their side chains contain no electrophiles, other than the disulfide bonds of cystines or functional groups installed by post-translational modification. Accordingly, the chemical reactivity of proteins necessarily entails

nucleophilic side chains, such as those of lysine (MacBeath and Schreiber, 2000) and cysteine (Falsey *et al.*, 2001; Luk *et al.*, 2004). The prevalence of these residues obviates control over the regiochemistry of reactions (Lundblad, 2005; Wallace, 1999), producing heterogeneous reaction products often at the expense of biological function (Luk *et al.*, 2004; Soellner *et al.*, 2003).

An intermediate that forms during the intein-mediated splicing of proteins contains an electrophile—a thioester (Scheme 4.1) (Paulus, 2000). The orthogonal reactivity of this functional group can be exploited for the site-specific modification of a protein by reaction with cysteine derivatives (Tan and Yao, 2005; Tolbert and Wong, 2000) or tandem reaction with a small-molecule thiol and amine (Borodovsky *et al.*, 2002; Cottingham *et al.*, 2001). Although thiols are potent nucleophiles for thioesters, the resultant thioesters are inherently unstable to hydrolysis (Connors and Bender, 1961; Noda *et al.*, 1953), making the simple transthioesterification of an intein-derived thioester unsuitable for the chemical modification of proteins. The powerful methods of native chemical ligation (Dawson and Kent, 2000) and expressed protein ligation (Evans and Xu, 2002; Muir, 2003; Muralidharan and Muir, 2006) offer an ingenious solution to this problem. After transthioesterification with a cysteine residue, $S \rightarrow N$ acyl transfer regenerates the thiol and forms a stable amide linkage. This approach, which has been used for protein modification and immobilization (Lue *et al.*, 2004; Watzke *et al.*, 2006b; Wood *et al.*, 2004), introduces a residual thiol that can be the focal point for undesirable side reactions. For example, cysteine is by far the most reactive residue toward disulfide bonds,

O_2(g), and other common electrophiles (Bednar, 1990; Raines, 1997; Schneider and de Weck, 1965). In addition, the sulfhydryl group of cysteine can suffer β-elimination to generate dehydroalanine (Friedman, 1999), or disrupt self-assembled monolayers on gold or silver surfaces (Terrettax *et al.*, 2002). These detrimental attributes caused us to search for an alternative means to exploit an intein-derived thioester for the installation of an orthogonal functional group.

In contrast to sulfur nucleophiles, nitrogen nucleophiles can, in theory, react directly with the thioesters formed during intein-mediated protein splicing to form inert linkages. This reaction has been neither explored nor exploited previously. Moreover, we reasoned that an appropriate bifunctional nitrogen nucleophile could both attack an intein-derived thioester to form a stable linkage and install an orthogonal (and thus useful) functional group.

The azido group can serve as an orthogonal functional group, being absent from natural proteins, nucleic acids, and carbohydrates (Bräse *et al.*, 2005; Hang and Bertozzi, 2001). Moreover, chemical reactions of the azido group, such as the Cu(I)-catalyzed Huisgen 1,3-dipolar azide–alkyne cycloaddition (Huisgen, 1963; Kolb and Sharpless, 2003; van Steenis *et al.*, 2005) and Staudinger ligation (Köhn and Breinbauer, 2004; Nilsson *et al.*, 2005), can lead to site-specific modification or immobilization. Azido–proteins have been produced previously. For example, Schultz and coworkers have developed a method for incorporating azidolysine into proteins (Noren *et al.*, 1989). Their approach involves producing a suppressor tRNA charged with azidolysine that inserts that residue into a protein as directed by an

engineered gene. This method, although site-specific, is labor intensive and low yielding. Tirrell, Bertozzi, and coworkers have incorporated an azido group into a protein by using azidohomoalanine, which replaces methionine in proteins produced in methionine-depleted bacterial cultures (Kiick *et al.*, 2002). This method is not site-specific for proteins containing more than one methionine residue.

Here, we report a general strategy for producing proteins labeled site-specifically with an azido group. We produce these azido–proteins by semisynthesis using a variation of expressed protein ligation (Scheme 4.1) (Evans and Xu, 2002; Muir, 2003; Muralidharan and Muir, 2006). Our strategy involves producing the protein of interest as a fusion protein with an intein and a tag for affinity chromatography. On-resin cleavage of the intein-derived thioester is induced with a bifunctional reagent bearing a nucleophile for thioesters and an azido group. Kinetic analyses with a chromogenic thioester are used to identify the optimal nucleophile for this purpose, and the desired bifunctional reagent is synthesized on a multi-gram scale. This strategy is used to produce an azido-protein that maintains full biological activity and displays at its C-terminus an azido group that is available for chemoselective modification. Thus, exploiting the intrinsic and orthogonal reactivity of the thioester produced during intein-mediated protein splicing enables the site-specific chemical modification of a protein.

4.3 Experimental Procedures

Materials. The chromogenic thioester $AcGlySC_6H_4$-*p*-NO_2 was a generous gift from Dr. B. L. Nilsson (Hondal *et al.*, 2001), and was purified by re-crystallization from methylene chloride and stored in a tightly sealed vial in a desiccator to prevent hydrolysis due to moisture present in air. Fluorescein–NHS ester was a generous gift from L. D. Lavis. All other chemicals were commercial reagent grade or better, and were used as received except for benzyl azide, which was purified by flash chromatography before use. Anhydrous THF, DMF, and CH_2Cl_2 were obtained from a CYCLE-TAINER® solvent delivery system (J. T. Baker, Phillipsburg, NJ). Other anhydrous solvents were obtained in septum-sealed bottles. Synthetic reactions were monitored by thin-layer chromatography (TLC) with visualization by UV-light or staining with vanillin, ninhydrin, or I_2. In all reactions involving anhydrous solvents, glassware was flame-dried. Flash chromatography was performed with columns of silica gel 60, 230–400 mesh (Silicycle, Quebec City, Quebec, Canada).

Instrumentation. A Cary Model 3 UV/VIS spectrophotometer (Varian, Palo Alto, CA) was used to perform kinetic assays and measure ultraviolet absorbance. NMR spectra were acquired with a Bruker AC+ 300 spectrometer (^{1}H: 300 MHz, ^{13}C: 75 MHz) at the Magnetic Resonance Facility in the Department of Chemistry or (as indicated) Bruker DMX-400 Avance spectrometer (^{1}H: 400 MHz, ^{13}C: 100 MHz) at the National Magnetic Resonance Facility at Madison (NMRFAM). Carbon-13 spectra were proton-decoupled. Mass spectra on small organic molecules were

obtained with electrospray ionization (ESI) techniques. Mass spectra of proteins were obtained with matrix-assisted laser desorption ionization-time-of-flight (MALDI–TOF) mass spectrometry using a Voyager-DE-PRO Biospectrometry workstation (Applied Biosystems, Foster City, CA) and a 3,5-dimethoxy-4-hydroxycinnamic acid (sinapinic acid) matrix (Aldrich). Fluorescence measurements were made with a QuantaMaster 1 photon counting fluorometer equipped with sample stirring (Photon Technology International, South Brunswick, NJ). A Typhoon 9410 variable mode fluorimager (Amersham Biosciences) was used to visualize fluorescein-labeled protein after SDS–PAGE.

Kinetics of thioester cleavage. $AcGlySC_6H_4$-*p*-NO_2 was dissolved in anhydrous acetonitrile to a concentration of 0.45 mM and used immediately. Amine solutions (except $CF_3CH_2NH_2 \cdot HCl$) were prepared by dissolving the amine hydrochloride salts in 0.10 M sodium phosphate buffer at pH 7.0 to a concentration of 0.15 M. Reaction mixtures were prepared in a 1.1-mL volume and equilibrated at 25 °C. All reactions contained ≤0.3% v/v acetonitrile. The ionic strength of each reaction mixture was I = 0.22–0.25. All reactions were carried out with a large excess of amine and followed pseudo-first-order kinetics. In a cuvette containing 0.10 M sodium phosphate buffer at pH 7.0 (967 μL), amine solution (0.15 M, 100 μL) was added and the absorbance at 410 nm was adjusted to zero. The thioester (33 μL of a 0.45 mM solution in acetonitrile) was then added to the cuvette, and the absorbance at 410 nm was monitored with time (ε = 11,230 $M^{-1}cm^{-1}$ for *p*-nitrothiophenolate anion (Hondal *et*

al., 2001)). The pK_a of HSC_6H_4-*p*-NO_2 is 4.77 in 40% v/v ethanol in water (Danehy and Parameswaran, 1968). This value is likely to be lower in an aqueous buffer, such that the ionization of the *p*-nitrothiophenol product is virtually complete in all our assays. The final concentration of the thioester in the reaction was 13.64 μM and the final concentration of the amines (except $CF_3CH_2NH_2 \cdot HCl$) was 13.64 mM. Each reaction was performed in triplicate. The reactions were allowed to go to completion, and the $t_{1/2}$ was determined from the kinetic trace. The $t_{1/2}$ was corrected by subtracting the $t_{1/2}$ for hydrolysis from the observed value. Pseudo-first-order rate constants were calculated by using the equation $k_1 = 0.693/t_{1/2}$. Second-order rate constants were obtained by dividing each observed first-order rate constant by the concentration of free amine. Logarithmic values of second-order rate constants were plotted against the pK_a values of the conjugate acids of the respective amines to yield a Brønsted plot.

Kinetics of Thioester Cleavage by $CF_3CH_2NH_2$. The procedure for the kinetics of thioester cleavage (*vide supra*) was found to be problematic for $CF_3CH_2NH_2$ because this amine is a poor nucleophile at pH 7.0. Indeed, the rate of hydrolysis was found to be greater than the rate of aminolysis by $CF_3CH_2NH_2$ under standard reaction conditions. Using a higher concentration of the hydrochloride salt (40.92 mM) resulted in the changing of the final reaction pH to 5.8. Using a higher buffer concentration was not ideal, as that would increase the ionic strength to a much value much larger than that in other reaction mixtures. To overcome these problems, the

reaction with $CF_3CH_2NH_2$ was carried out using 40.92 mM $CF_3CH_2NH_2 \cdot HCl$ and allowing the pH of the reaction mixture to decrease to 5.8. The observed $t_{1/2}$ was corrected by subtracting the $t_{1/2}$ for hydrolysis at pH 5.8. The second-order rate constant was calculated by accounting for the concentration of free amine at pH 5.8, and the resulting value was used in the Brønsted plot. The first-order rate constant thus obtained was not compared to those for the other nucleophiles.

Synthesis of $BocNHCH_2CH_2N_3$ **4.3**. $BocNHCH_2CH_2Br$ (10.00 g, 44.62 mmol) was dissolved in DMF (200 mL). NaN_3 (14.48 g, 223.1 mmol) was added, and the mixture was stirred at 110 °C for 12 h. The solvent was removed under reduced pressure, and the residue was dissolved in water (200 mL). The resulting aqueous solution was extracted with ethyl acetate (2 × 200 mL). The organic layers were combined and dried over anhydrous $MgSO_4$(s). After filtering, the organic layer was concentrated under reduced pressure and the residue was dissolved in methylene chloride (10–20 mL) and purified by flash chromatography (silica gel, methylene chloride). $BocNHCH_2CH_2N_3$ (6.60 g, 80%) was isolated as a colorless oil. HRMS (ESI) $[M+Na]^+$ calcd for $C_7H_{14}N_4O_2Na$, 209.1014; found, 209.1010; ^1H NMR (400 MHz, $CDCl_3$) δ 4.88 (bs, 1H), 3.42 (t, $J = 5.3$ Hz, 2H), 3.34–3.26 (m, 2H), 1.45 (s, 9H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 155.8, 79.8, 51.4, 40.2, 28.5.

Synthesis of $(Boc)_2NN(Boc)CH_2CONHCH_2CH_2N_3$ **4.5**. $HCl \cdot H_2NCH_2CH_2N_3$ (**4.4**) was synthesized by dissolving azide **4.3** (2.11 g, 11.33 mmol) in 4 N HCl in dioxane (100

mL) and stirring at room temperature for 1 h. The solvent was removed under reduced pressure to give a dirty white powder. $(Boc)_2NN(Boc)CH_2CO_2H$ (4.42 g, 11.33 mmol) was then added, and the mixture was dissolved in methylene chloride/DMF (70:45 mL). The mixture was cooled to 0 °C, and PyBOP (5.9 g, 11.33 mmol) and Et_3N (3.2 mL, 22.66 mmol) were added. The reaction mixture was allowed to warm to room temperature and stirred under Ar(g) for 21 h. The solvent was removed under reduced pressure, and the residue was purified by flash chromatography (silica gel, 50% v/v ethyl acetate in hexanes) to give $(Boc)_2NN(Boc)CH_2CONHCH_2CH_2N_3$ as a colorless viscous oil (4.9 g, 95%). HRMS (ESI) $[M+Na]^+$ calcd for $C_{19}H_{34}N_6O_7Na$, 481.2387; found, 481.2389; ^{1}H NMR (400 MHz, $CDCl_3$, 2 rotamers) δ 8.51 and 8.27 (bs, 1H), 4.06 and 3.99 (s, 2H), 3.49–3.40 (m, 4H), 1.56–1.44 (m, 27H); ^{13}C NMR (100 MHz, $CDCl_3$, 2 rotamers) δ 169.5, 169.2, 154.0, 153.5, 151.4, 151.2, 85.2, 85.1, 83.5, 82.8, 56.5, 54.8, 50.6, 38.9, 34.1, 29.9, 28.2, 28.1.

Synthesis of $H_2NNHCH_2CONHCH_2CH_2N_3$ **4.1**. Azide **4.5** (4.66 g, 10.17 mmol) was dissolved in 4 N HCl in dioxane (200 mL), and the solution was stirred at room temperature for 5 h. Solvent was removed under reduced pressure, and the residue was dissolved in water (15 mL) and purified by cation-exchange chromatography (Dowex 50WX8-200 ion-exchange resin, 1 M NH_4OH) to give $H_2NNHCH_2CONHCH_2CH_2N_3$ as a yellow oil (1.52 g, 95%). HRMS (ESI) $[M+Na]^+$ calcd for $C_4H_{10}N_6ONa$, 181.0814; found, 181.0805; ^{1}H NMR (300 MHz, d_6-DMSO) δ

8.03 (app bs, 1H), 3.80–3.40 (bs, 3H), 3.40–3.34 (m, 2H), 3.32–3.24 (m, 2H), 3.16 (s, 2H); ^{13}C NMR (75 MHz, d_6-DMSO) δ 171.2, 57.1, 50.0, 37.8.

Synthesis of BocNHN=CHCH$_2$CH$_2$CO$_2$H **4.7**. 4-Pentenoic acid (10.00 g, 99.88 mmol) was dissolved in methylene chloride (150 mL), and the mixture was cooled to –78 °C under N_2(g). Ozone was bubbled through the reaction mixture, and the course of the reaction was monitored by TLC. After TLC showed disappearance of the starting material (*ca.* 2.5 h), methyl sulfide (15.0 mL, 205.23 mmol) was added and the reaction mixture was allowed to warm to room temperature. Subsequently, solvent was removed under reduced pressure, and the residue was dissolved in THF (200 mL). BocNHNH$_2$ (13.2 g, 99.88 mmol) was then added, and the reaction mixture was refluxed overnight under Ar(g). Solvent was removed under reduced pressure, and the residue was purified by flash chromatography (silica gel, 6% v/v methanol in methylene chloride) to give BocNHN=CHCH$_2$CH$_2$CO$_2$H as a white solid (13.60 g, 63%). HRMS (ESI) [M–H]$^-$ calcd for $C_9H_{15}N_2O_4$, 215.1032; found, 215.1035; ^{1}H NMR (300 MHz, d_6-DMSO) δ 12.16 (s, 1H), 10.45 (s, 1H), 7.30 (app bs, 1H), 2.45–2.30 (m, 4H), 1.41 (s, 9H); ^{13}C NMR (100 MHz, d_6-DMSO) δ 173.5, 152.4, 145.6, 78.9, 30.4, 28.1, 27.2.

Synthesis of BocNHN=CHCH$_2$CH$_2$CONHCH$_2$CH$_2$N$_3$ **4.8**. HCl·H$_2$NCH$_2$CH$_2$N$_3$ (**4.4**) was synthesized as described earlier. Compound **4.7** (4 g, 18.5 mmol) and HCl·H$_2$NCH$_2$CH$_2$N$_3$ (2.27 g, 18.5 mmol) were dissolved in methylene chloride/DMF

(180:70 mL). The mixture was cooled to 0 °C, and PyBOP (9.63 g, 18.5 mmol) and Et_3N (20 mL, 142.8 mmol) were then added. The reaction mixture was allowed to warm to room temperature and stirred for 21 h. Solvent was removed under reduced pressure, and the residue was purified by flash chromatography (silica gel, ethyl acetate). BocNHN=$CHCH_2CH_2CONHCH_2CH_2N_3$ was obtained as a white solid (3.94 g, 75%). HRMS (ESI) $[M+Na]^+$ calcd for $C_{11}H_{20}N_6O_3Na$, 307.1495; found, 307.1494; ^{1}H NMR (300 MHz, d_6-DMSO) δ 10.43 (bs, 1H), 8.13 (t, J = 5.3 Hz, 1H), 7.28 (app bs, 1H), 3.33 (t, J = 5.7 Hz, 2H), 3.28–3.19 (m, 2H), 2.4–2.3 (m, 2H), 2.29–2.22 (m, 2H), 1.41 (s, 9H); ^{13}C NMR (100 MHz, d_6-DMSO) δ 171.6, 152.5, 146.1, 79.0, 50.0, 38.3, 32.1, 28.2, 27.8.

Synthesis of BocNHNH$CH_2CH_2CH_2$CONH$CH_2CH_2N_3$ **4.9**. Compound **4.8** (1.12 g, 3.9 mmol) was dissolved in acetonitrile (27 mL), and acetic acid (3.7 mL). $NaCNBH_3$ (1.11 g, 16.75 mmol) was added to the resulting solution. The reaction mixture was stirred at room temperature for 3 h. Solvent was removed under low pressure, and the residue was dissolved in water (45 mL). The pH of the solution was increased to 13.0 by adding 10 N NaOH, and the aqueous solution was extracted with ether (3 × 45 mL) and then methylene chloride (50 mL). The organic layers were combined and dried over anhydrous $MgSO_4$(s). After filtering, the organic layer was concentrated under reduced pressure, and the residue was purified by flash chromatography (silica gel, 2% v/v methanol in methylene chloride) to give BocNHNH$CH_2CH_2CH_2$CONH$CH_2CH_2N_3$ as a colorless oil (0.78 g, 70%). HRMS

(ESI) $[M+Na]^+$ calcd for $C_{11}H_{22}N_6O_3Na$, 309.1651; found, 309.1641; ^{1}H NMR (300 MHz, d_6-DMSO, 2 rotamers) δ 8.15 (bs, 1H), 8.09–7.90 (m, 1 H), 4.34 (bs, 1H), 3.32 (t, J = 5.7 Hz, 2H), 3.26–3.18 (m, 2H), 2.66–2.53 (m, 2H), 2.22–2.07 (m, 2H), 1.64–1.48(m, 2H), 1.46–1.22 (m, 9 H); ^{13}C NMR (100 MHz, d_6-DMSO, 2 rotamers) δ 172.5, 156.4, 78.2, 56.1, 54.9, 50.4, 50.0, 38.2, 33.0, 28.2, 23.5, 22.9.

Synthesis of $CF_3CO_2H{\cdot}H_2NNHCH_2CH_2CH_2CONHCH_2CH_2N_3$ **4.2**. Compound **4.9** (197.8 mg, 0.69 mmol) was dissolved in methylene chloride (6.4 mL), and trifluoroacetic acid (6.4 mL) was added to the resulting solution. The reaction mixture was stirred at room temperature for 15 min. Solvent was removed under reduced pressure to afford $CF_3CO_2H{\cdot}H_2NNHCH_2CH_2CH_2CONHCH_2CH_2N_3$ as a yellow oil (130 mg, 100%). HRMS (ESI) $[M+H]^+$ calcd for $C_6H_{15}N_6O$, 187.1307; found, 187.1299; ^{1}H NMR (400 MHz, d_6-DMSO) δ 8.19 (t, J = 4.8, 1H), 7.9–5.5 (bs, 3H), 3.34 (t, J = 5.8, 2H), 3.28–3.20 (m, 2H), 2.89 (t, J = 7.5 Hz, 2H), 2.21–2.11(m, 2H), 1.85–1.68 (m, 2H); ^{13}C NMR (100 MHz, d_6-DMSO) δ 171.9, 50.1, 50.0, 38.2, 32.2, 21.0.

Synthesis of alkynyl fluorescein **4.10**. Fluorescein–NHS ester (100 mg, 0.21 mmol) was dissolved in THF (10 mL) and propargylamine (23.27 mg, 0.42 mmol) was added. The reaction mixture was stirred at room temperature for 4 h. Solvent was removed under reduced pressure, and the residue was purified by flash chromatography (silica gel, 35% v/v hexanes in ethyl acetate containing 1% v/v

AcOH) to give the desired product as a red solid (72.80 mg, 84%). HRMS (ESI) $[M+Na]^+$ calcd for $C_{24}H_{15}NO_6Na$, 436.0797; found, 436.0777; ^{1}H NMR (400 MHz, d_6-DMSO) δ 10.20 (bs, 2H), 9.31 (t, J = 5.1 Hz, 1H), 8.47 (s, 1H), 8.25 (dd, J = 7.9, 1.4 Hz, 1H), 7.39 (d, 8.3 Hz, 1H), 6.68 (d, J = 2.3 Hz, 2H), 6.59 (d, J = 8.6 Hz, 2H), 6.54 (dd, J = 8.7, 2.3 Hz, 2H), 4.11 (dd, J = 5.3, 2.2 Hz, 2H), 3.17 (t, J = 2.5 Hz, 1H); ^{13}C NMR (100 MHz, d_6-DMSO) δ 168.1, 164.4, 159.7, 154.8, 151.9, 135.5, 134.7, 129.2, 126.7, 124.4, 123.5, 112.8, 109.1, 102.3, 83.8, 81.0, 73.1, 28.8.

Synthesis of polytriazole ligand. A polytriazole ligand for the Cu(I) catalyst was synthesized essentially as described (Wang *et al.*, 2003b). Tripropargyl amine (0.56 g, 4.28 mmol) was dissolved in acetonitrile (5.7 mL), and benzyl azide (2 g, 15.02 mmol) and 2,6-lutidine (0.46 g, 4.28 mmol) were added to the resulting solution. The reaction mixture was cooled to 0 °C, and $Cu(CH_3CN)_4PF_6$ (81.25 mg, 0.22 mmol) was added. The reaction mixture was allowed to warm to room temperature and stirred under Ar(g) for $2^1/_2$ days. The reaction mixture was filtered, and the white precipitate obtained was dried under high vacuum to yield the polytriazole ligand (0.34 g, 15%). HRMS (ESI) $[M+Na]^+$ calcd for $C_{30}H_{30}N_{10}Na$, 553.2553; found, 553.2570; ^{1}H NMR (400 MHz, $CDCl_3$) δ 7.66 (s, 3H), 7.40–7.20 (m, 15H), 5.50 (s, 6H), 3.70 (s, 6H); ^{13}C NMR (100 MHz, $CDCl_3$) δ 134.8, 129.2, 128.8, 128.1, 123.9, 54.2, 47.2.

Production of RNase A–intein–chitin-binding domain fusion protein. A plasmid that directs the expression of an RNase A–*mxe* intein–chitin-binding domain fusion protein was a generous gift from U. Arnold. A glycine codon was introduced between the RNase A and the *mxe* intein genes using the QuickChange site-directed mutagenesis kit from Stratagene (LaJolla, CA). The resulting plasmid, pJK01, was transformed into *E. coli* BL21(DE3) cells, and the production of Met(–1)RNase A–Gly–intein-chitin binding domain fusion protein was induced as described previously (Arnold *et al.*, 2003; Arnold *et al.*, 2002).

Production of azido–ribonuclease A. Cells were resuspended in 20 mM 3-(N-morpholino)propane sulfonic acid (MOPS)–NaOH buffer at pH 6.8 containing NaCl (0.50 M), ethylenediaminetetraacetic acid (EDTA; 0.10 mM), and Triton X-100 (0.1% v/v). Cells were lysed with a French pressure cell, and the cell lysate was subjected to centrifugation at 15,000*g* for 30 min. The supernatant was diluted to a final volume of 25 mL (per g of cells) and applied (flow rate: 0.75 mL/min) to a chitin column that had been equilibrated with the same buffer. The loaded resin was washed with two column-volumes of buffer and then with four column-volumes of 0.5 M MOPS–NaOH buffer at pH 7.0 containing NaCl (0.50 M) and EDTA (0.10 mM). Azide **4.1** was dissolved in the latter buffer to a concentration of 450 mM, and three column-volumes of this solution were loaded onto the resin, out of which, two column-volumes were allowed to flow through and one column-volume was allowed to sit on top of the resin. This incubation was carried out for three days at

room temperature to enable the reaction to proceed to completion. The hydrazide product was eluted with three column-volumes of an aqueous solution of NaCl (2 M). Azido-RNase A was precipitated out of the eluate by adding an aqueous solution of sodium deoxycholate (NaDOC; to 0.72 mM) and trichloroacetic acid (TCA; to 260 mM). This precipitate was washed with acetone and dissolved in an aqueous solution of guanidine–HCl (4 M).

The solution of protein was added with gentle stirring in 20-μL aliquots into a refolding solution (50 mL) consisting of 100 mM Tris–HCl buffer at pH 8.0 containing NaCl (100 mM), reduced glutathione (1 mM), and oxidized glutathione (0.2 mM). The final concentration of guanidine–HCl was 0.05 M. The refolding solution was incubated at room temperature for 24 h.

The refolding solution was dialyzed for 12 h against 50 mM sodium acetate buffer at pH 5.0. The azido-protein was purified by cation-exchange chromatography as described previously (Leland *et al.*, 1998).

Ribonucleolytic activity. Values of k_{cat}/K_M for the enzymatic cleavage of a fluorogenic substrate, 6-carboxyfluorescein-dArU(dA)$_2$-6-carboxytetramethylrhodamine, were determined as described previously (Smith *et al.*, 2003).

Huisgen 1,3-dipolar cycloaddition. To a solution of azido–RNase A (9.6 μM) in 0.10 M potassium phosphate buffer at pH 8.0 (41.9 μL) were added alkynyl fluorescein **4.10** (1.1 μL, 2.23 mM suspension in 20% v/v ethanol in water), tris(2-carboxyethyl)

phosphine hydrochloride (1.0 μL, 50 mM), $CuSO_4 \cdot 5H_2O$ (1.0 μL, 50 mM), and polytriazole ligand (5.0 μL, 20 mM suspension in 80% v/v *t*-butanol in water). The reaction mixture was agitated gently, and incubated at room temperature for 1 h and then at 4 °C for 16 h. The same procedure was followed for the control reaction with unmodified wild-type RNase A (Sigma Chemical, St. Louis, MO). Protein precipitation was observed in both reaction mixtures, as is common (but not well appreciated) during Cu(I)-catalyzed Huisgen 1,3-dipolar azide–alkyne cycloaddition to a protein (Speers *et al.*, 2003), obviating the calculation of a yield for this reaction. The reaction mixtures were subjected to centrifugation at 5900*g* for 4 min, and the supernatant was discarded. The pellet was resuspended in 2× denaturing buffer (20 μL) and subjected to SDS–PAGE. The resulting gel was visualized with a fluorescence imager and was then stained with Coomassie blue.

4.4 Results and Discussion

Identification of the optimal nucleophile for thioesters. To identify the optimal nitrogen nucleophile for a thioester, kinetic studies were performed on a model chromogenic thioester: $AcGlySC_6H_4$-*p*-NO_2. The rate of release of the thiophenolate anion was monitored by measuring the change in absorbance at 410 nm (Scheme 4.2). Nitrogen nucleophiles with conjugate-acid pK_a values ranging from 4.6 to 10.6 (Table 4.1) were used in the experiments. The logarithmic values of the second-order rate constants (k_2) of the unprotonated primary amines were plotted against the pK_a values

of their conjugate acids to yield the Brønsted plot shown in Figure 4.1A. The data were fitted to the equation (Hupe and Jencks, 1977):

$$\log k_2 = \log (\mathrm{AB}) + (\beta + \beta')\mathrm{p}K_\mathrm{a} - \log(A10^{\beta \mathrm{p}K_\mathrm{a}} + B10^{\beta' \mathrm{p}K_\mathrm{a}}) \qquad (4.1)$$

In eq 4.1, A and B are constants, β' is the slope of the former part of the Brønsted plot, and β is the slope of the latter part.

The Brønsted plot in Figure 4.1A is biphasic. The slope changes from 0.81 with nucleophiles of low pK_a to 0.42 with nucleophiles of high pK_a. This change is due to the known change in rate-determining step from the formation of a tetrahedral zwitterionic intermediate to the decomposition of the intermediate into products (Castro and Ureta, 1989; Gresser and Jencks, 1977; Hupe and Jencks, 1977). The value of the slope obtained ($\beta' = 0.81$) is in agreement with that for the aminolysis of oxygen esters (Jencks and Gilchrist, 1968). As expected (Bruice *et al.*, 1963; Bruice *et al.*, 1967; Gregory and Bruice, 1967), "α-effect" nucleophiles (alkoxy amines, α-hydrazino acetyl, and alkyl hydrazine) exhibit much greater nucleophilicity than that predicted from their pK_a values. In water, the α-effect could arise from the nucleophile being less solvated and hence more reactive because of the inductive withdrawal of electrons by the adjacent heteroatom (Buncel and Um, 2004).

The Brønsted plot reports on the nucleophilicity of unprotonated (that is, neutral) amines. According to Figure 4.1A, the best nucleophile for thioesters in an environment in which all the amines are deprotonated is ethyl amine. Indeed, other

simple amines have been used in tandem with a small-molecule thiol to modify an intein-derived thioester (Borodovsky *et al.*, 2002; Cottingham *et al.*, 2001). At the high pH necessary to deprotonate an amine, however, thioesters are prone to undergo hydrolysis (Connors and Bender, 1961; Noda *et al.*, 1953). Moreover, proteins are subject to unfolding and subsequent aggregation at high pH values (Antosiewicz *et al.*, 1994). Performing the reaction at pH 7.0 provides an acceptable trade-off between deprotonation of the nitrogen nucleophile and hydrolysis of the thioester. Data on the first-order rate constant (k_1) for the reaction of various nitrogen nucleophiles at pH 7.0 under pseudo-unimolecular reaction conditions are shown in Figure 4.1B. At pH 7.0, the α-hydrazino acetyl and alkyl hydrazine functionalities are much more nucleophilic than are the simple amines and alkoxy amine (*e.g.*, 30- and 100-fold greater k_1 value than ethyl amine, respectively), and are therefore the optimal nucleophiles for reacting with a thioester to form a stable linkage.

Jencks (Jencks and Carriuolo, 1960; Jencks and Gilchrist, 1968), Bruice (Bruice *et al.*, 1963; Bruice *et al.*, 1967; Gregory and Bruice, 1967), Castro (Castro, 1999; Castro and Ureta, 1989), and others have reported in detail on the mechanism and kinetics of the nucleophilic attack on esters and, to a lesser extent, thioesters. This previous work was not, however, performed with the intent of making a stable linkage with a bifunctional reagent. For example, hydroxylamine was identified in 1950 as an exceptional nucleophile for a thioester (Chou and Lipmann, 1952; Lynen *et al.*, 1951). Indeed, this attribute of hydroxylamine has led to its use in revealing transient thioesters formed during intein-mediated protein splicing (Bruice and Fedor,

1964a; Bruice and Fedor, 1964b; Bruice and Fedor, 1964c). The cleavage of thioesters by hydroxylamine relies, however, on the formation of an *O*-acylated hydroxylamine intermediate (Jencks, 1958a; Jencks, 1958b) that is inaccessible during the attack of an alkoxy amine on a thioester. Accordingly, an alkoxy amine is not an especially potent nucleophile for a thioester (Figure 4.1)

Synthesis of bifunctional azides. After identification of two optimal nucleophiles, we proceeded to synthesize two bifunctional reagents bearing those nucleophiles on one end and an azido group on the other. Azides **4.1** and **4.2** are both amides of 1-azido-2-aminoethane. Azide **4.1** has an α-hydrazino acetamido group, which is a more stable analog of the α-hydrazino acetyl group of $C_2H_5O(O)CCH_2NHNH_2$ (Table 4.1; Figure 4.1); azide **4.2** has a γ-hydrazino acetamido group and is effectively an alkyl hydrazine.

Azide **4.1** was synthesized by the route in Scheme 4.3. Briefly, Boc-protected 1-azido-2-aminoethane was synthesized from Boc-protected 1-bromo-2-aminoethane. After Boc-deprotection, the amine was coupled to tri–Boc-protected α-hydrazino acetic acid. The Boc groups were removed, and azide **4.1** was isolated as a free base after cation-exchange chromatography with an overall yield of 72%.

Azide **4.2** was synthesized by the route in Scheme 4.4. Briefly, 4-pentenoic acid was subjected to ozonolysis, and the resulting aldehyde was reacted *in-situ* with Boc-protected hydrazine. The azido group was installed by coupling 1-azido-2-aminoethane (**4.4**) to the carboxylic acid. The hydrazone was then reduced selectively

with $NaCNBH_3$. The Boc group was removed to produce azide **4.2** as its trifluoroacetic acid (TFA) salt with an overall yield of 26% (which includes the 80% yield for the synthesis of azide **3**). Attempts to produce the free hydrazine base by cation-exchange chromatography resulted in decomposition of the molecule by an (as yet) unknown mechanism. Likewise, the TFA salt was unstable even upon storage under vacuum, and hence was used immediately after its synthesis.

Kinetics of thioester cleavage. Kinetic studies were performed by reacting azides **4.1** and **4.2** with a model chromogenic thioester (Scheme 4.2). The rate constants (k_2 and k_1) for azide **4.1** were found to be indistinguishable from those of the α-hydrazino acetyl group. The rate constants for azide **4.2** were, surprisingly, much lower than those of methylhydrazine. This result is contrary to our finding that methyl hydrazine is a somewhat better nucleophile than the α-hydrazino acetyl functional group (Figure 4.1). The intrinsic instability of azide **4.2** is likely to be responsible for this apparent decrease in reactivity.

Production of an azido–protein. Next, we sought to use our bifunctional reagents to install an azido group at the C-terminus of a model protein. As our protein, we chose bovine pancreatic ribonuclease (RNase A), which has been the object of much seminal work in protein chemistry (Raines, 1998) and has been manipulated previously with expressed protein ligation (Arnold *et al.*, 2003; Arnold *et al.*, 2002; Evans *et al.*, 1998; Hondal *et al.*, 2001; Nilsson *et al.*, 2003). RNase A has valine as

its C-terminal residue. A valine residue at the C-terminus of a target protein is known to have a debilitating effect on the cleavage efficiency of protein–intein thioesters (Lue *et al.*, 2004; Yee *et al.*, 2003). To avert this problem, we inserted a glycine residue between the C-terminus of RNase A and the intein. The resulting Met(–1)RNase A–Gly–*mxe* intein–chitin-binding domain fusion protein (MW ~36 kDa) was produced in *E. coli*, and the cell lysate was loaded onto chitin resin. Azides **4.1** and **4.2** were used to induce the on-resin cleavage of the fusion protein. As expected from the kinetic studies, azide **4.1** was found to be much more effective than azide **4.2** in cleaving the Met(–1)RNase A–Gly–*mxe* intein thioester. Its shorter and higher yielding synthesis, superior stability, and higher cleavage efficiency makes azide **4.1** the optimal bifunctional reagent for the semisynthesis of proteins labeled with the azido group. The purity of the azido–RNase A (even upon elution from the chitin column) was apparent from SDS–PAGE analysis (Figure 4.2A); the integrity of the azido–RNase A was verified by using MALDI–TOF mass spectrometry (Figure 4.2B). This procedure produced an overall yield of ~1 mg of purified azido–RNase A per liter of *E. coli* culture.

Incubating a protein with a potent nucleophile, such as the α-hydrazino acetamido group of azide **4.1**, could compromise the structure of the protein. For example, the target protein in this study has 142 amide bonds in its main chain and side chains that could be attacked by the α-hydrazino acetamido group, but only one thioester bond. In addition, the target protein has eleven amino groups that could serve as intramolecular nucleophiles for that thioester bond. Enzymatic catalysis provides an

extremely sensitive measure of native protein structure (Knowles, 1987). This measure is especially useful for detecting the inadvertent modification of RNase A, as one of its eleven amino groups is both especially reactive and critical to enzymatic activity (Messmore *et al.*, 1995; Murdock *et al.*, 1966; Raines, 1998). Purified azido-RNase A had $k_{cat}/K_M = (3.2 \pm 1.0) \times 10^7\ M^{-1}s^{-1}$ for the cleavage of RNA. This value was in gratifying agreement with that of the wild-type enzyme, which had $k_{cat}/K_M = (5.2 \pm 0.4) \times 10^7\ M^{-1}s^{-1}$ (Rutkoski *et al.*, 2005).

The α-hydrazino acetamido group was found to be the optimal nitrogen nucleophile for producing azido-proteins. Still, we had to use a 450 mM solution of azide **4.1** to produce the desired hydrazide product. Using lower concentrations lead to hydrolyzed protein (that is, the protein with a C-terminal carboxyl group) being a dominant product. In contrast, a 50 mM solution of thiol typically suffices for transthioesterification during expressed protein ligation (Evans and Xu, 2002; Muir, 2003; Muralidharan and Muir, 2006). The resulting thioester must then, however, react with a peptide (present in vast excess) that contains an N-terminal cysteine residue (Lue *et al.*, 2004; Watzke *et al.*, 2006b; Wood *et al.*, 2004). The ability to obtain an azido–protein in a single step by on-resin cleavage and the absence of the residual sulfhydryl group installed during expressed protein ligation are noteworthy advantages of our strategy (Scheme 4.1). These attributes are of particular importance for high-throughput procedures, such as the fabrication of protein microarrays (Cretich *et al.*, 2006).

Huisgen 1,3-dipolar cycloaddition to an azido–protein. For our strategy to be useful, the azido group in azido–RNase A must be available for further reaction. We used a chemoselective reaction, Cu(I)-catalyzed Huisgen 1,3-dipolar azide–alkyne cycloaddition (Huisgen, 1963; Kolb and Sharpless, 2003; Speers *et al.*, 2003; van Steenis *et al.*, 2005; Wang *et al.*, 2003b), to probe for the availability of the azide functionality. To affect this "functional group test", alkynyl fluorescein **4.10** was synthesized by the route in Scheme 4.5. Azido–RNase A was reacted with **4.10** in the presence of the Cu(I) catalyst and its polytriazole ligand (Wang *et al.*, 2003b). The resulting protein had a molecular mass of m/z = 14449, which agreed well with that expected for the conjugate ($[C_{610}H_{936}N_{180}O_{201}S_{13}]$ = 14424). The protein was also subjected to SDS–PAGE and visualized by staining with Coomassie blue and fluorescence imaging. The azido–RNase A was found to be fluorescent as a result of the cycloaddition, whereas wild-type RNase A treated in the same manner was not fluorescent (Figure 4.3). Neither the mass spectrum nor the SDS–PAGE gel showed evidence of cleavage products, as have been observed in an azido-protein exposed to reducing agents (Back *et al.*, 2005). Thus, an azido group was not only installed into a specific site on RNase A, but also was available for reaction. In on-going work, the Staudinger ligation (Köhn and Breinbauer, 2004; Nilsson *et al.*, 2005) is being used for the site-specific immobilization of proteins produced by the novel route shown in Scheme 4.1.

4.5 Conclusions

We have exploited the mechanism of intein-mediated protein splicing to develop a general strategy for installing a functional group at the C-terminus of a protein. The strategy is based on the capture of an intein-derived thioester with a nucleophile that was selected by using Brønsted analysis. We used this strategy to append an azido group to a model protein. The azido group did not affect the function of the protein and was available for Huisgen 1,3-dipolar azide–alkyne cycloaddition of a fluorophore. We anticipate that such a C-terminal azido group could also be used for site-specific protein immobilization, which can be preferable to random immobilization (Luk *et al.*, 2004; Soellner *et al.*, 2003), and to modification by Staudinger ligation (Köhn and Breinbauer, 2004; Nilsson *et al.*, 2005). Finally, we note that our strategy can be used to append other functional groups with orthogonal reactivity (such as an alkene, alkyne, or nitrile) to a target protein.

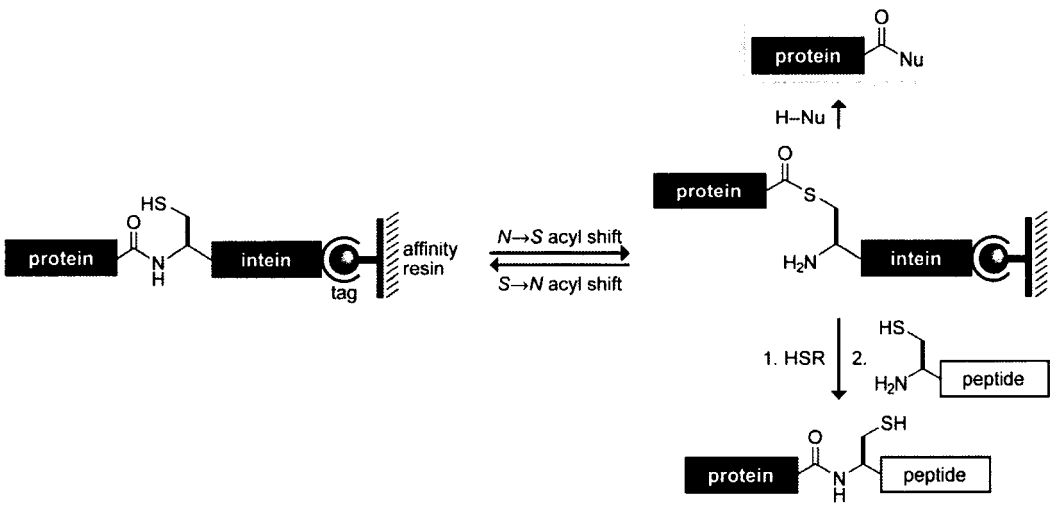
4.6 Acknowledgements

We are grateful to V. I. Timokhin and D. Lee for assistance with ozonolysis, B. G. Miller and J. E. Lee for help with site-directed mutagenesis, B. L. Nilsson for $AcGlySC_6H_4$-p-NO_2, L. D. Lavis for fluorescein–NHS ester, and U. Arnold for a plasmid that directs the expression of the RNase A–*mxe* intein–chitin-binding domain fusion protein. This work was supported by Grant GM44783 (NIH) and the Materials Research Science and Engineering Center at the University of Wisconsin–Madison (NSF-DMR 0079983). The Magnetic Resonance Facility was supported by Grants

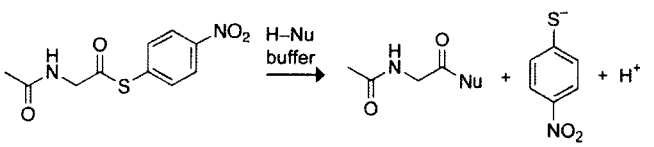
CHE-9208463 (NSF) and RR08389 (NIH). NMRFAM was supported by Grant P41RR02301 (NIH).

Nucleophile	pK_a
CH_3ONH_2	4.60 (Hall, 1957)
$F_3CCH_2NH_2$	5.40 (Kluger and Hunt, 1984)
$C_2H_5O(O)CCH_2NHNH_2$	6.45 (Bonnet *et al.*, 2001)
$CH_3O(O)CCH_2NH_2$	7.75 (Hall, 1957)
CH_3NHNH_2	7.87 (Hinman, 1958)
$FCH_2CH_2NH_2$	9.19 (Kluger and Hunt, 1984)
$HOCH_2CH_2NH_2$	9.50 (Hall, 1957)
$CH_3CH_2NH_2$	10.63 (Hall, 1957)

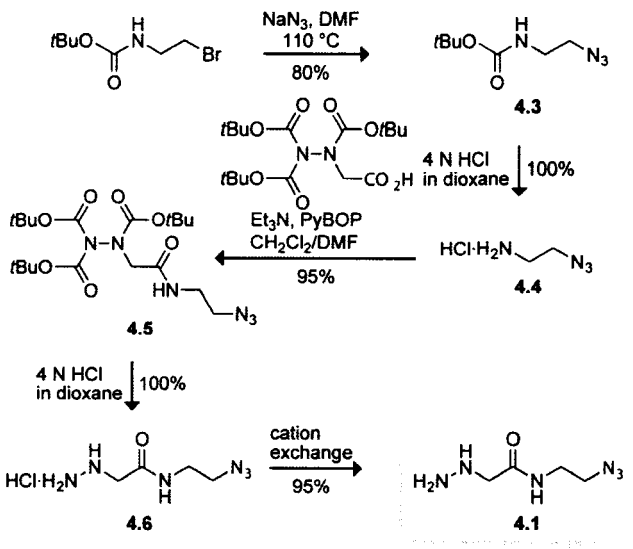
Table 4.1 Nucleophiles used in this study and the pK_a values of their conjugate acids.



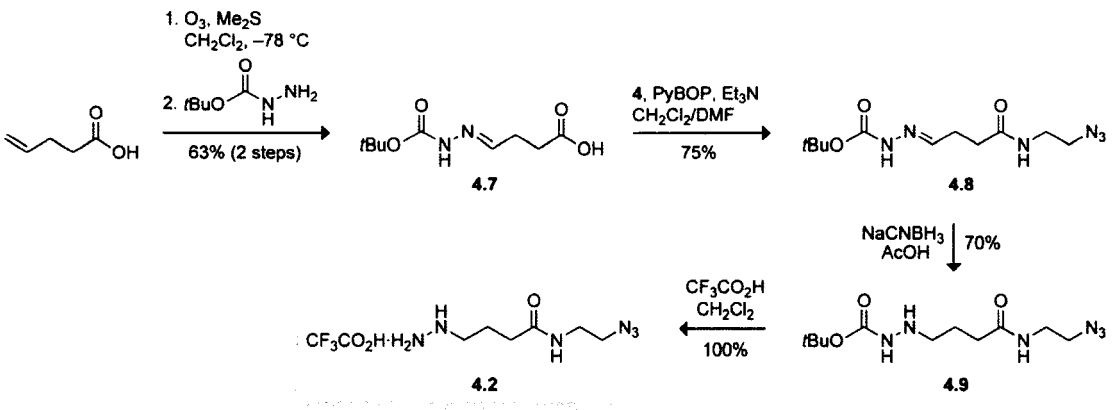
Scheme 4.1 Mechanism of expressed protein ligation and on-resin capture of the thioester intermediate with a small-molecule nucleophile (H–Nu).



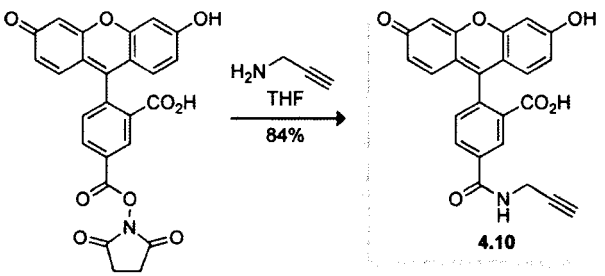
Scheme 4.2 Cleavage reaction of a model chromogenic thioester.



Scheme 4.3 Synthetic route to azide **4.1**.



Scheme 4.4 Synthetic route to azide **4.2**.



Scheme 4.5 Synthetic route to alkynyl fluorescein **4.10**.

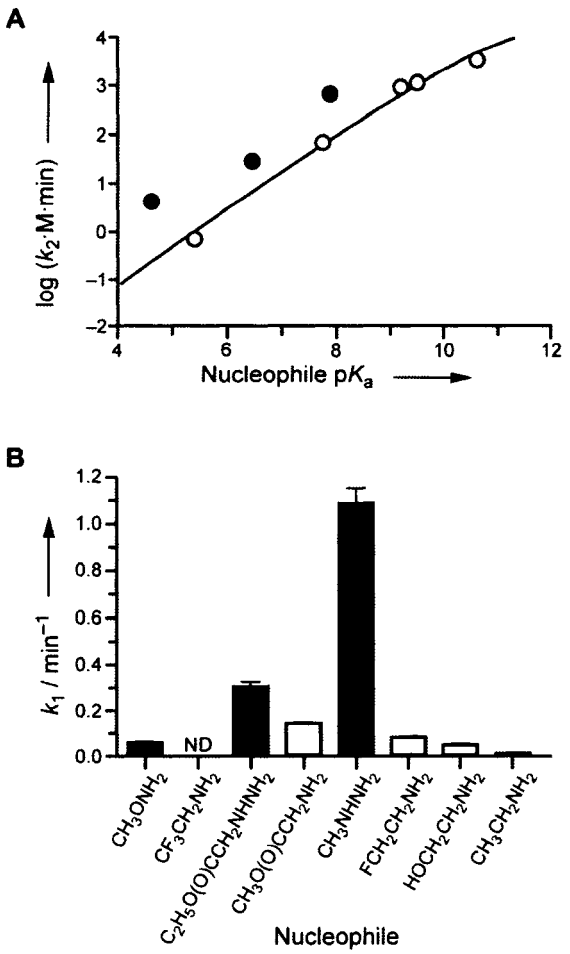


Figure 4.1 Rate constants for the attack of nitrogen nucleophiles on a thioester (Scheme 4.1). (A) Brønsted plot for the reaction of simple amines (open circles) and α-effect nucleophiles (filled circles) with $AcGlySC_6H_4$-*p*-NO_2 at 25 °C. Relevant pK_a values are listed in Table 1. Data were fitted to eq 1 with $A = 0.21$, $B = 4.0 \times 10^{-5}$, $\beta = 0.42$, and $\beta' = 0.81$. (B) Values of the first-order rate constant (k_1) at pH 7.0 and 25 °C. ND: not determined.

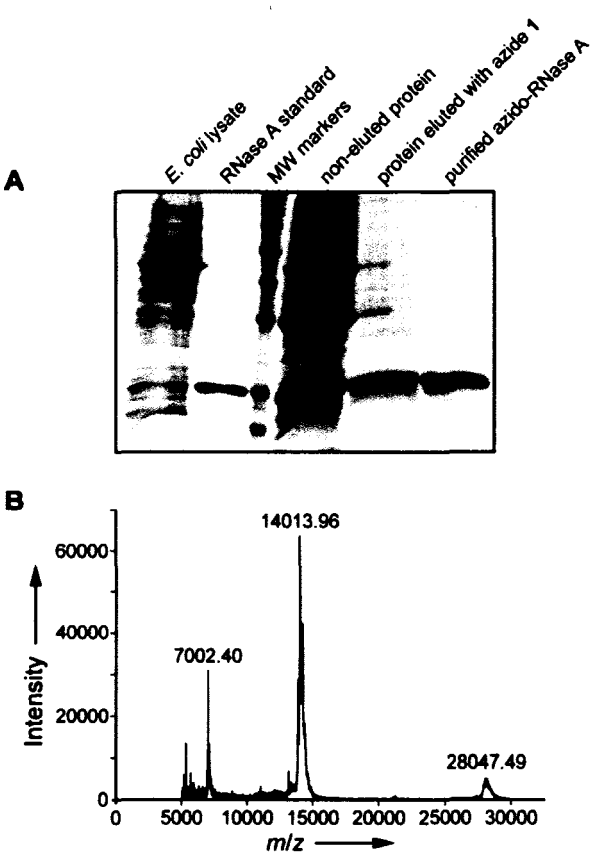


Figure 4.2 Characterization of azido–RNase A. (A) SDS–PAGE analysis of the preparation of azido–RNase A. (B) MALDI–TOF mass spectrum of azido-RNase A (expected for Met(–1)RNase A–Gly–$NHNHCH_2C(O)NHCH_2CH_2N_3$ [$C_{586}H_{92a}N_{179}O_{195}S_{13}$] = 14011).

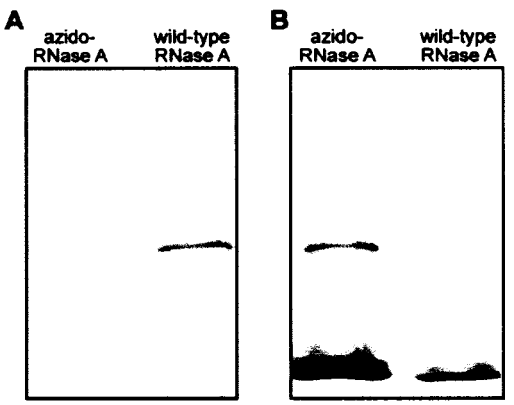
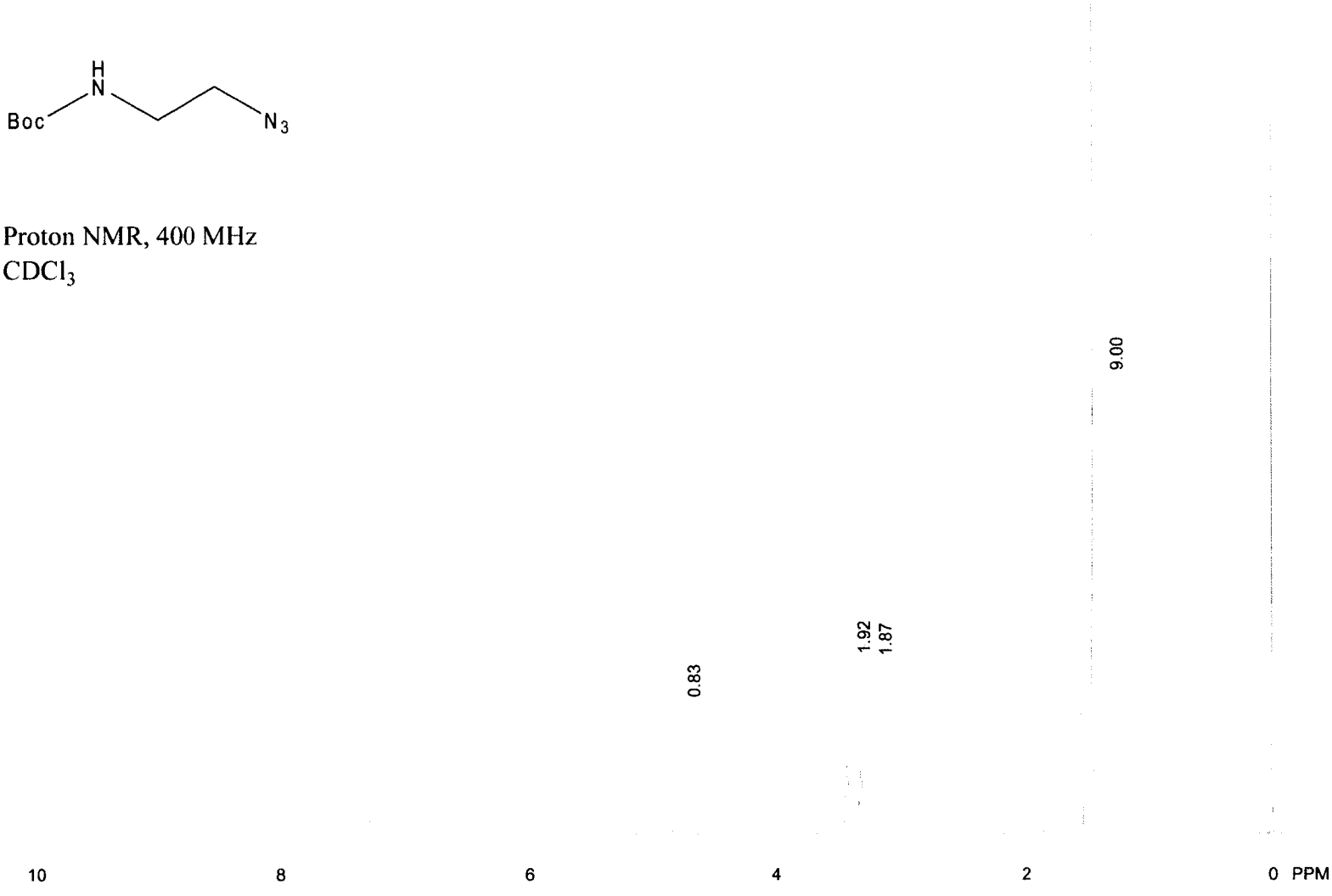


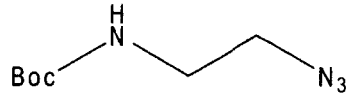
Figure 4.3 SDS–PAGE analysis of the reaction of azido-RNase A and wild-type RNase A with alkynyl fluorescein **4.10**. (A) Visualization with Coomassie-blue staining. (B) Visualization with fluorescence imaging.

4.3

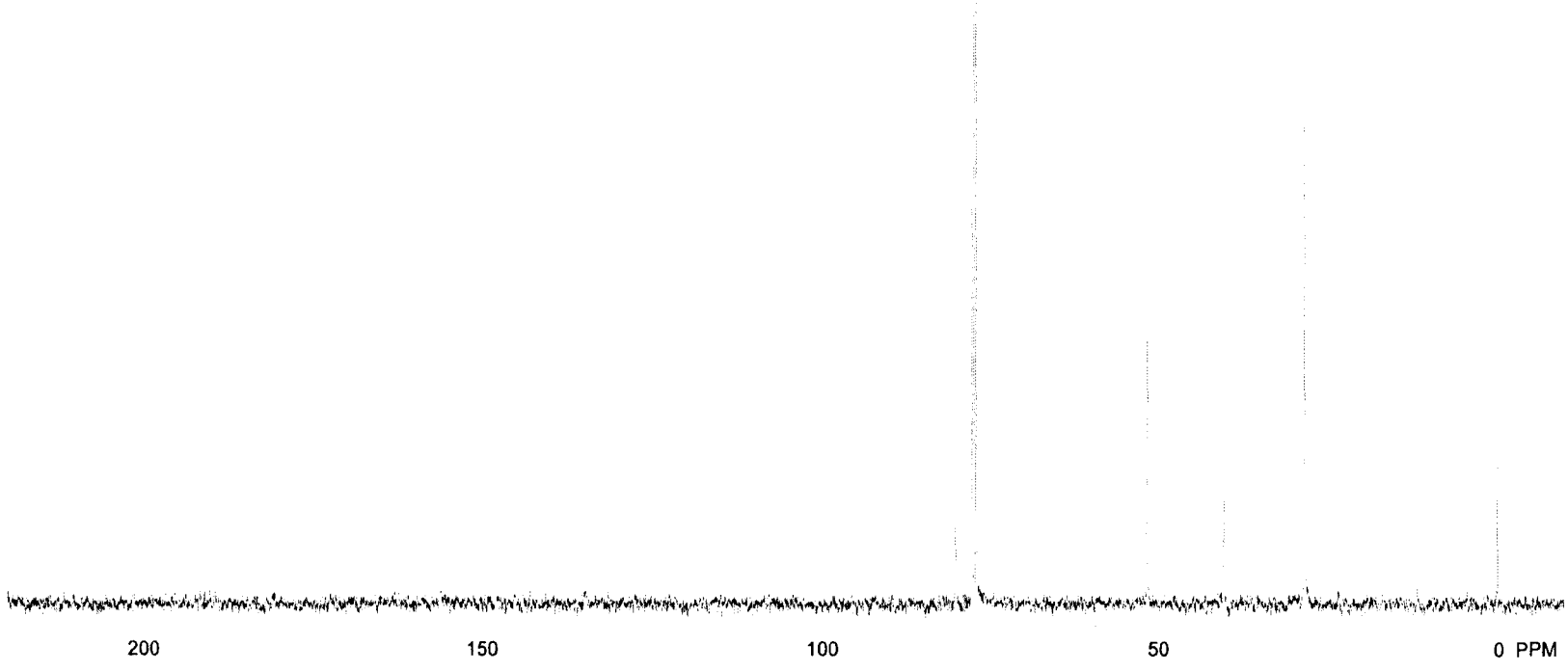
Proton NMR, 400 MHz
$CDCl_3$



4.3



Carbon-13 NMR, 100 MHz
$CDCl_3$

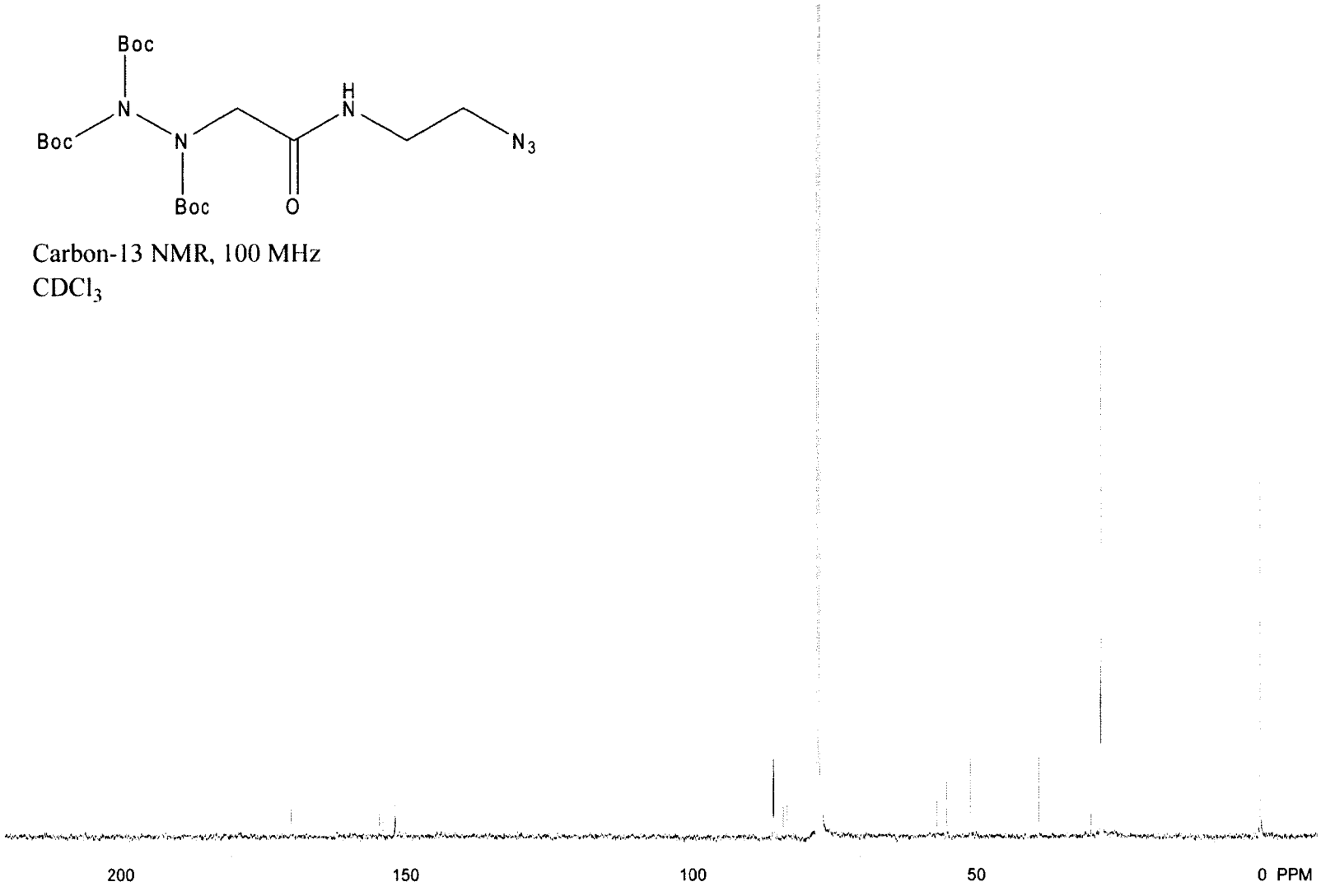


4.5

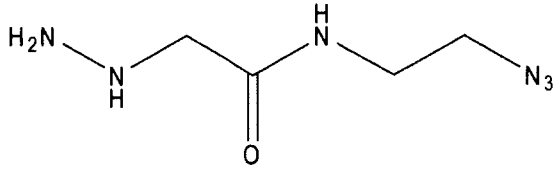
Proton NMR, 400 MHz
$CDCl_3$

4.5

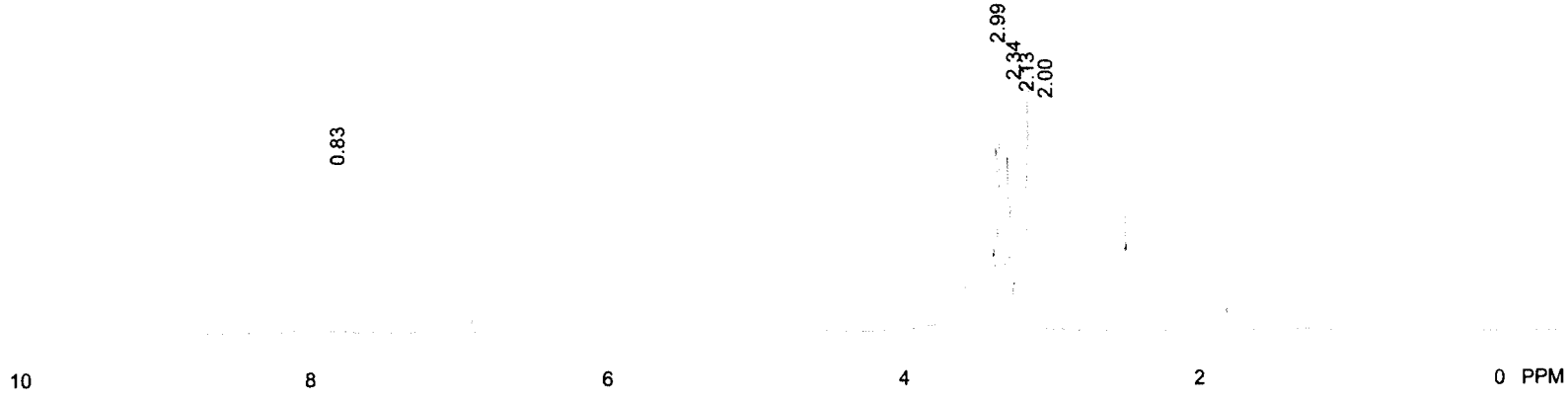
Carbon-13 NMR, 100 MHz
$CDCl_3$



4.1

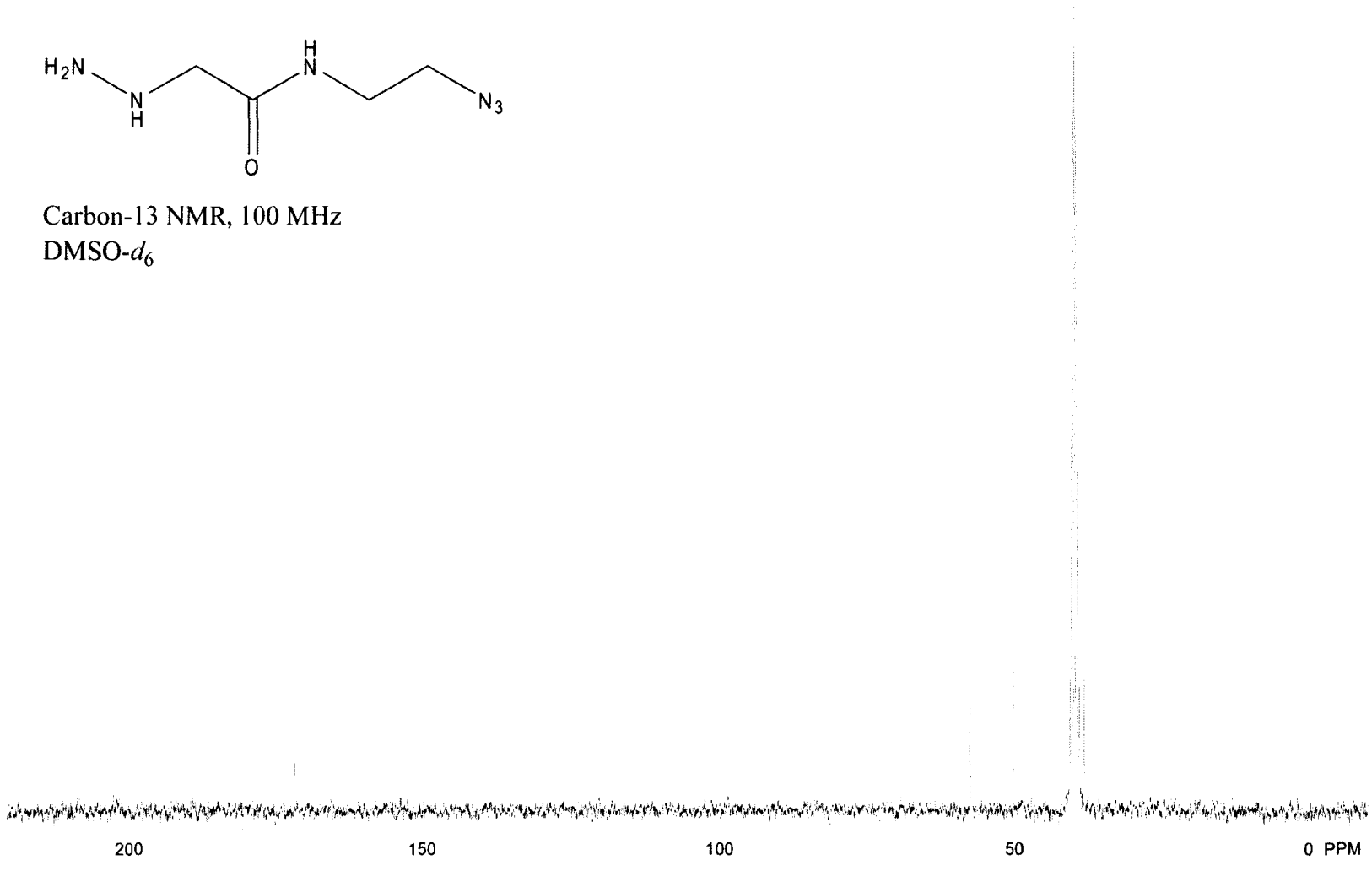


Proton NMR, 400 MHz
DMSO-d_6

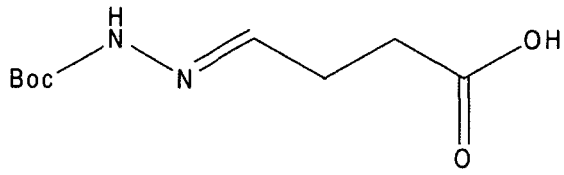


4.1

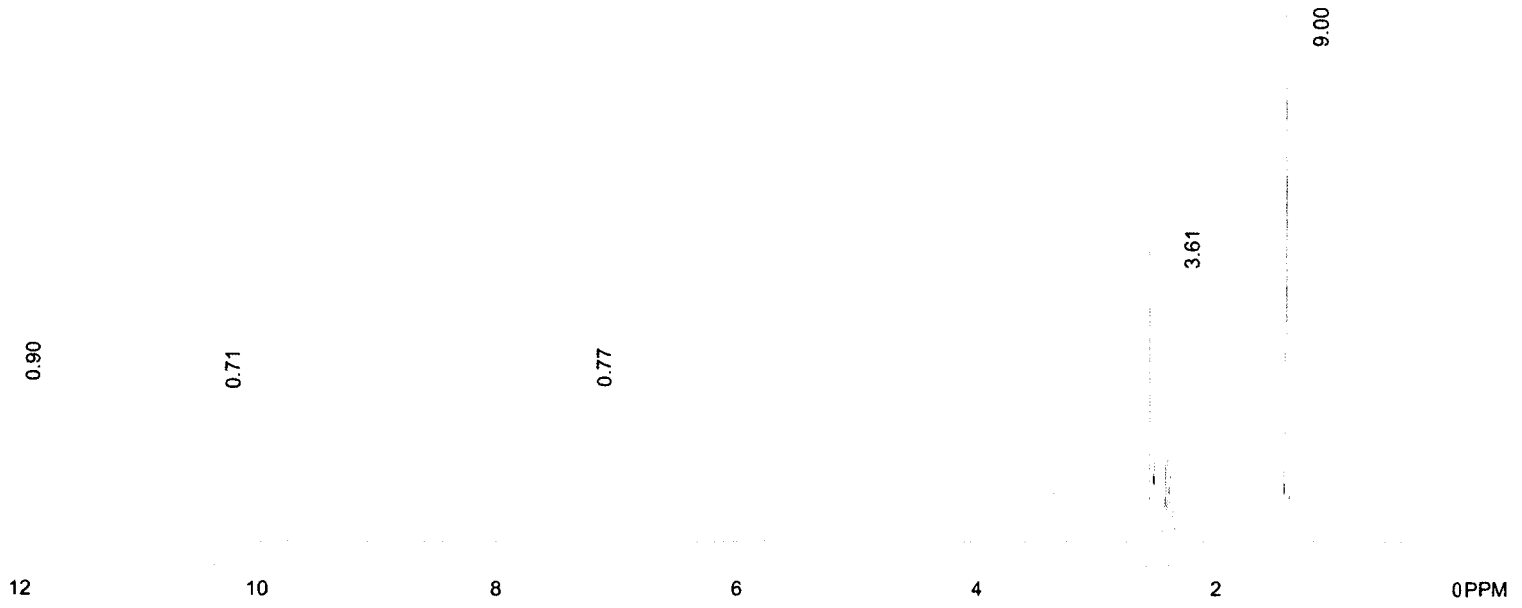
Carbon-13 NMR, 100 MHz
DMSO-d_6



4.7

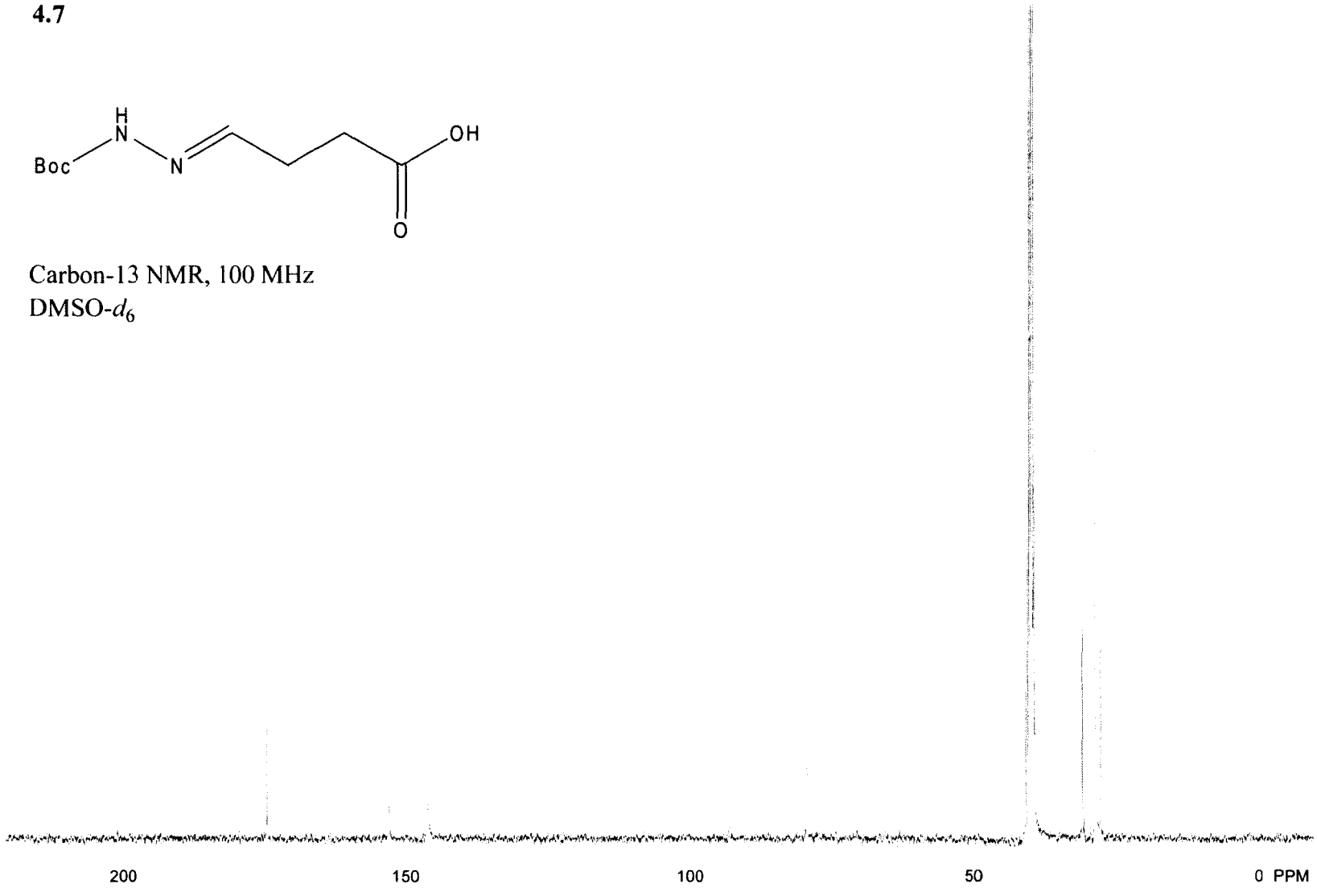


Proton NMR, 400 MHz
DMSO-d_6



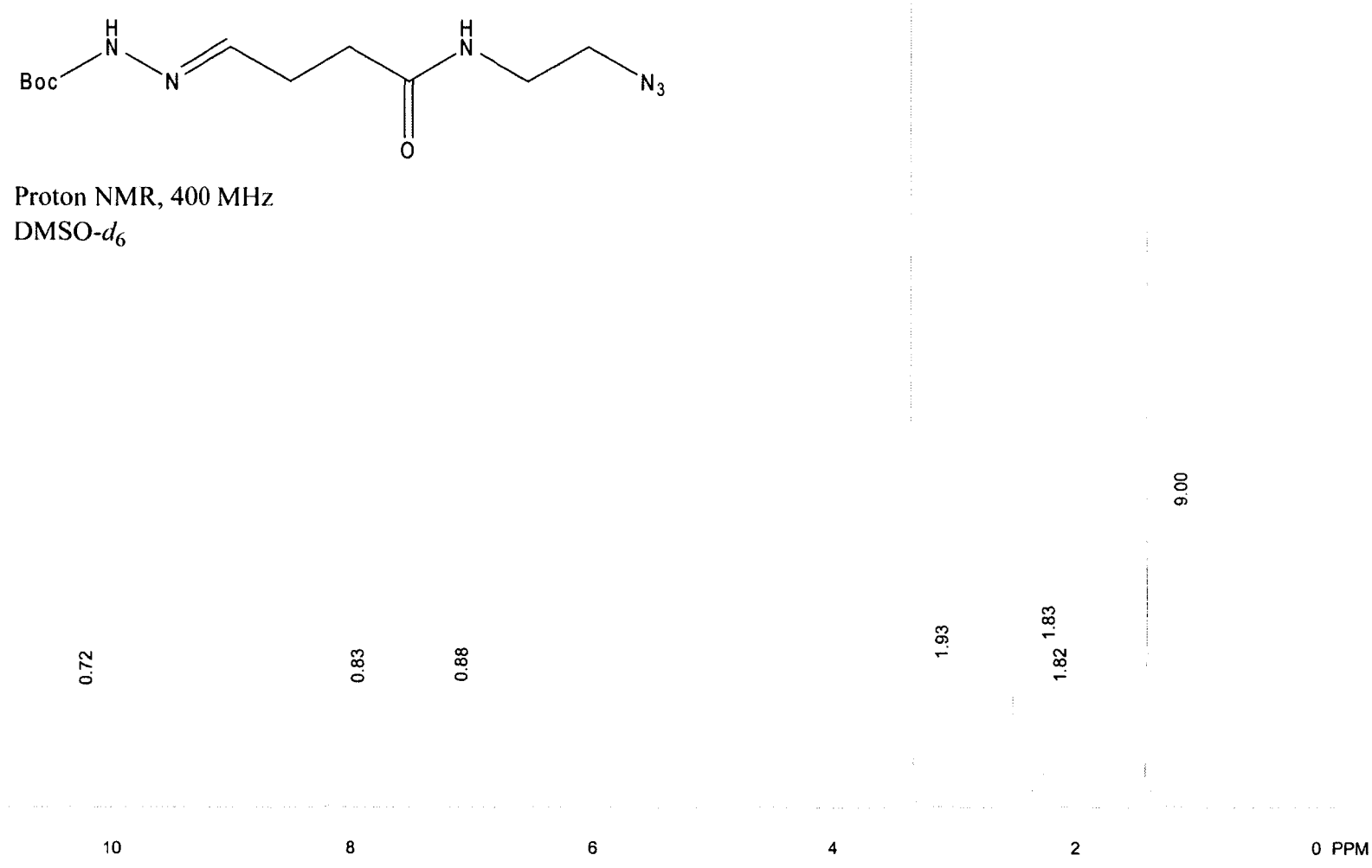
4.7

Carbon-13 NMR, 100 MHz
DMSO-d_6



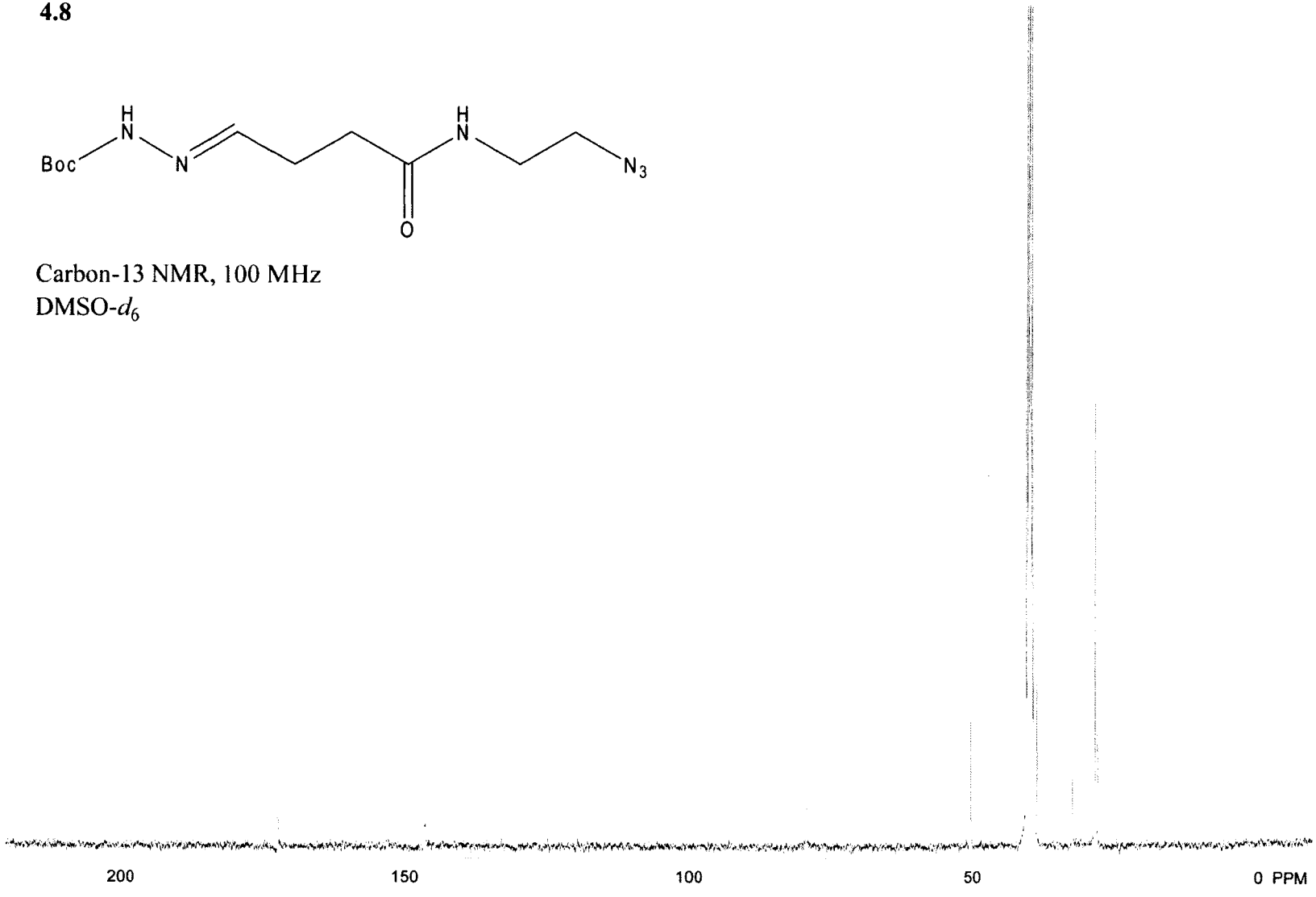
4.8

Proton NMR, 400 MHz
DMSO-d_6



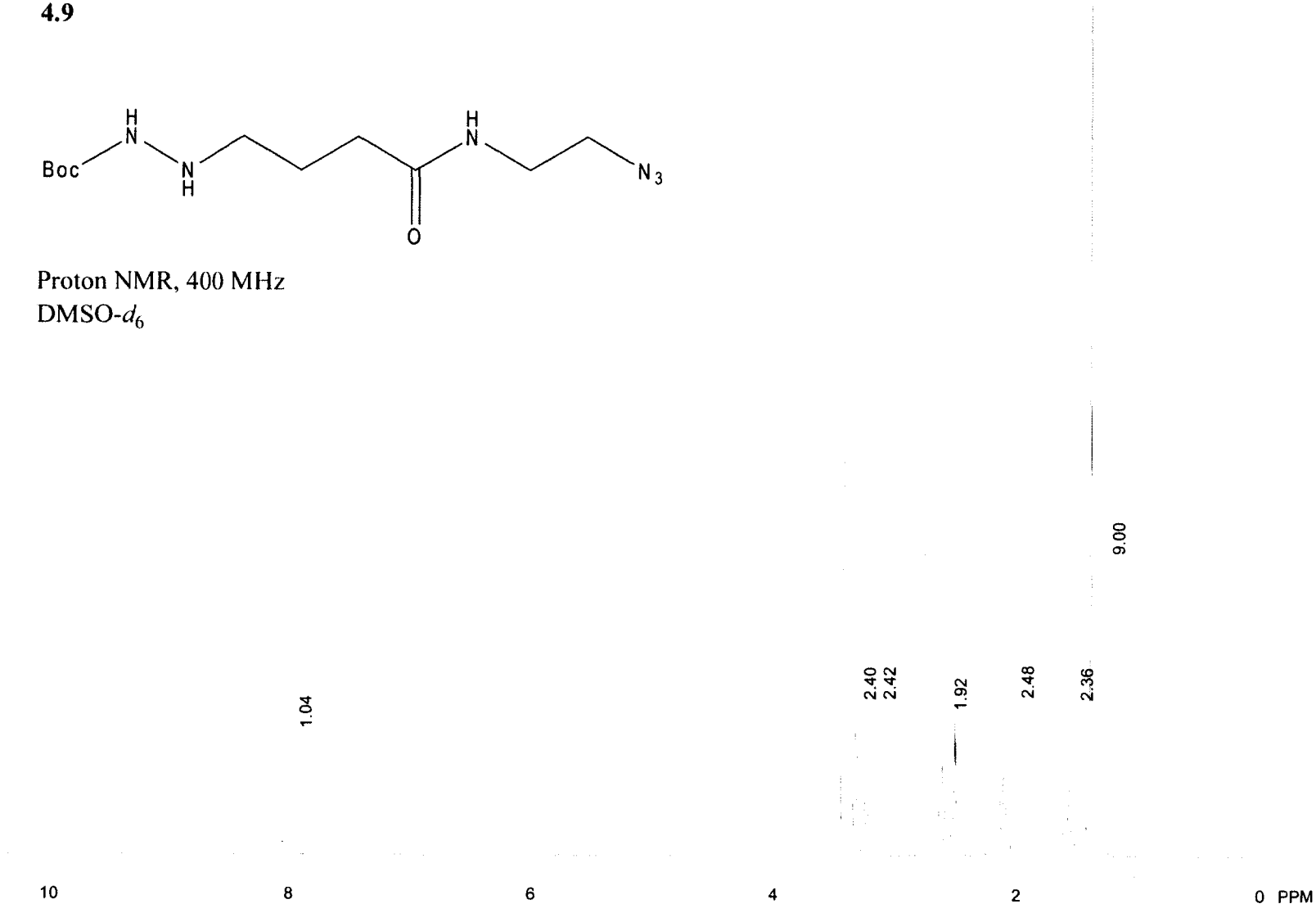
4.8

Carbon-13 NMR, 100 MHz
DMSO-d_6



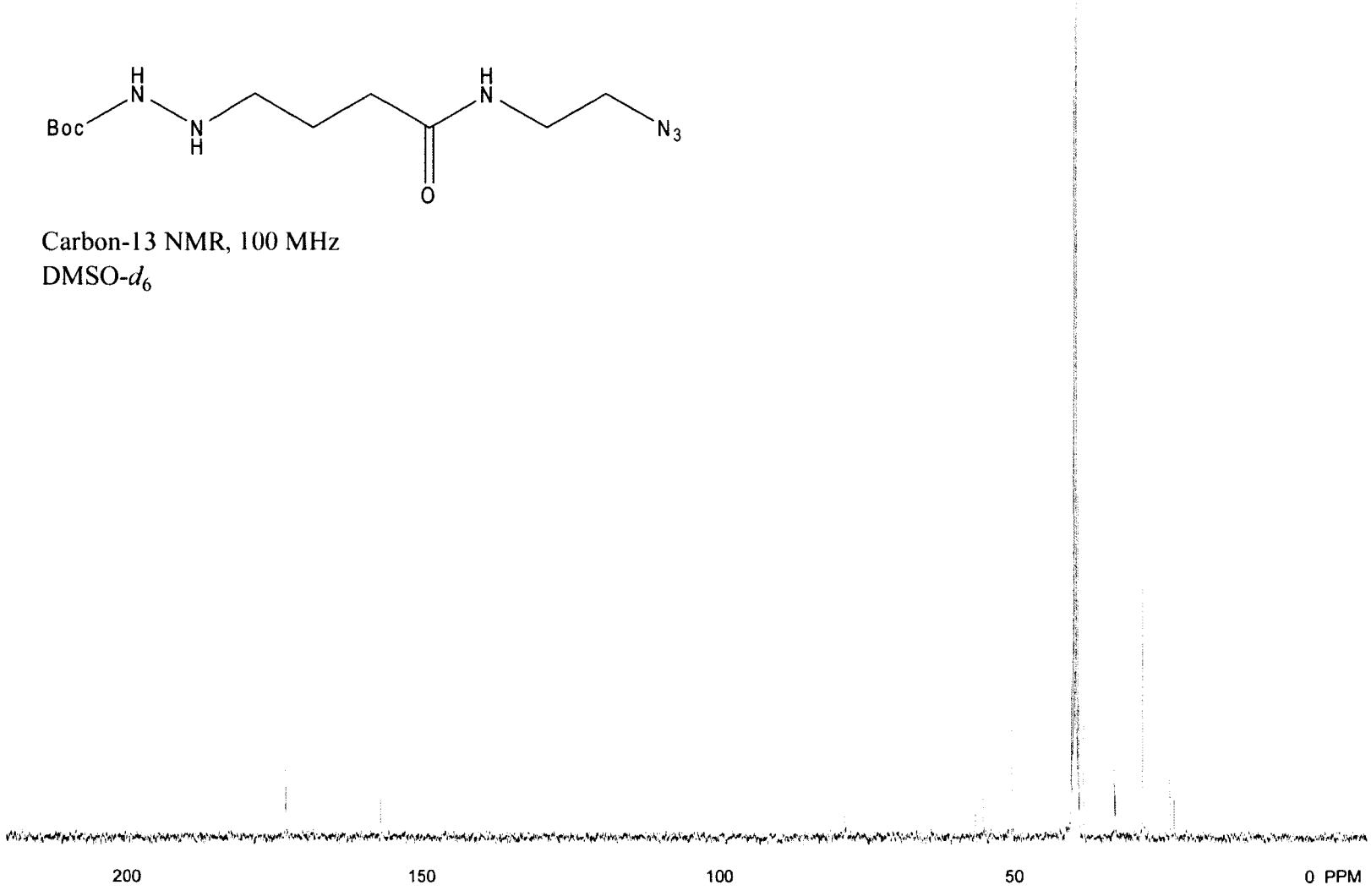
4.9

Proton NMR, 400 MHz
DMSO-d_6

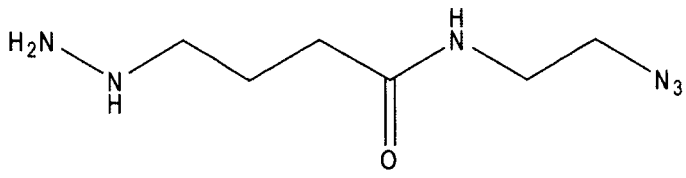


4.9

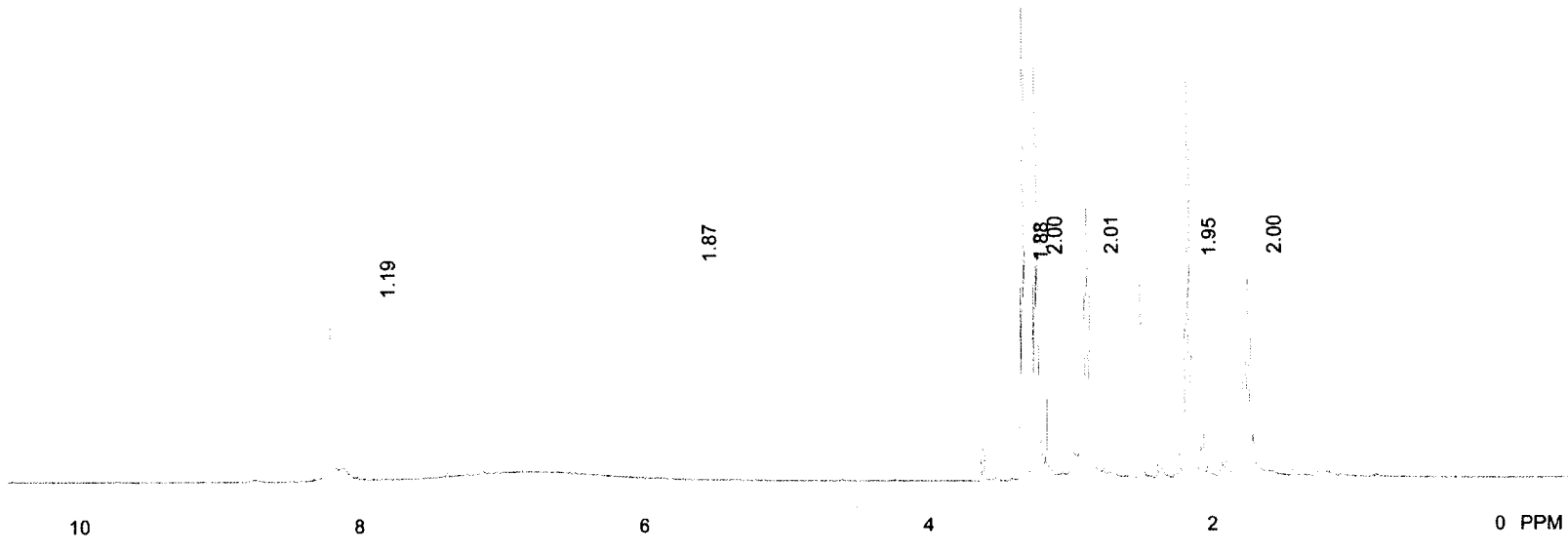
Carbon-13 NMR, 100 MHz
DMSO-d_6



4.2

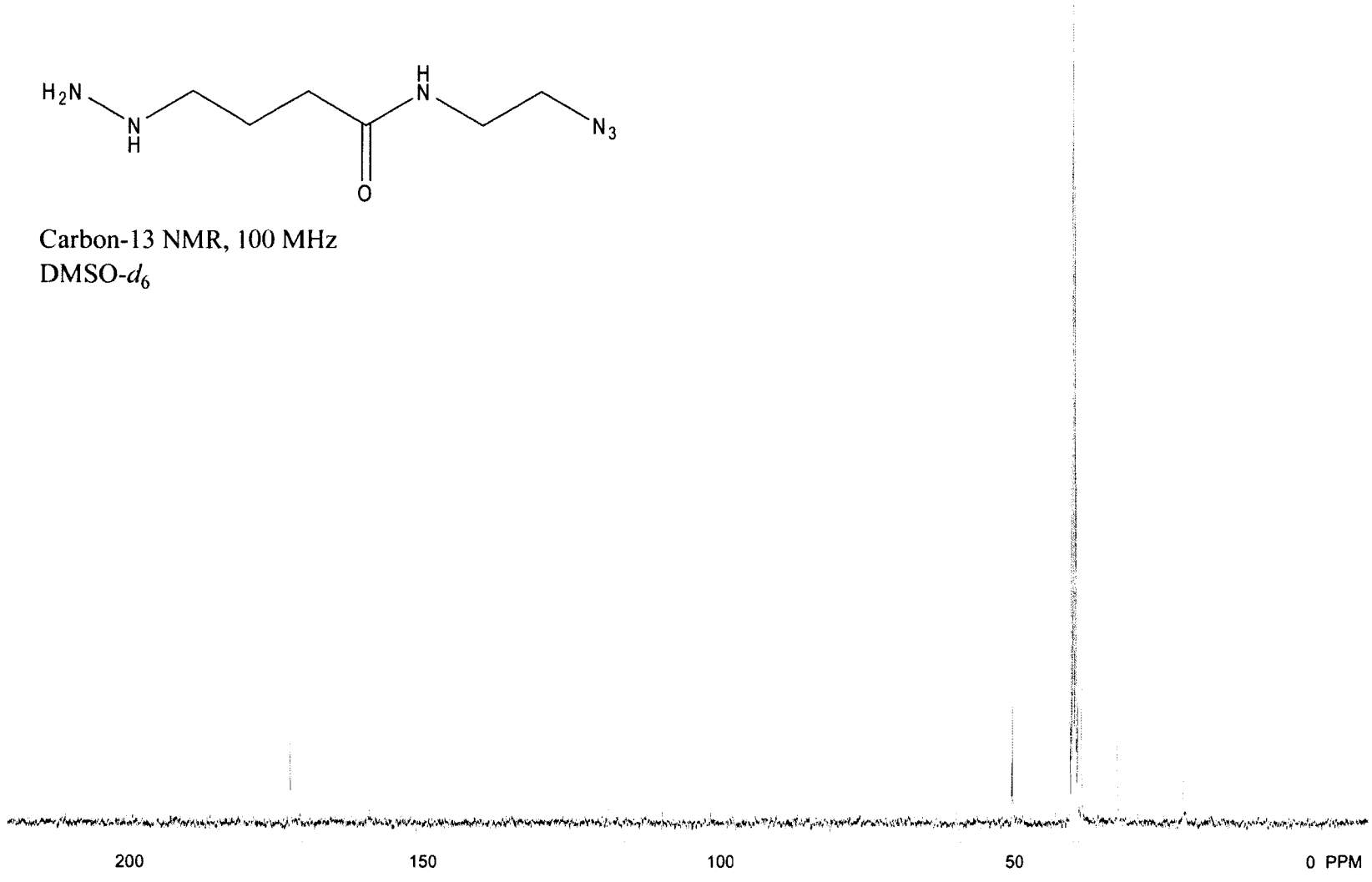


Proton NMR, 400 MHz
DMSO-d_6

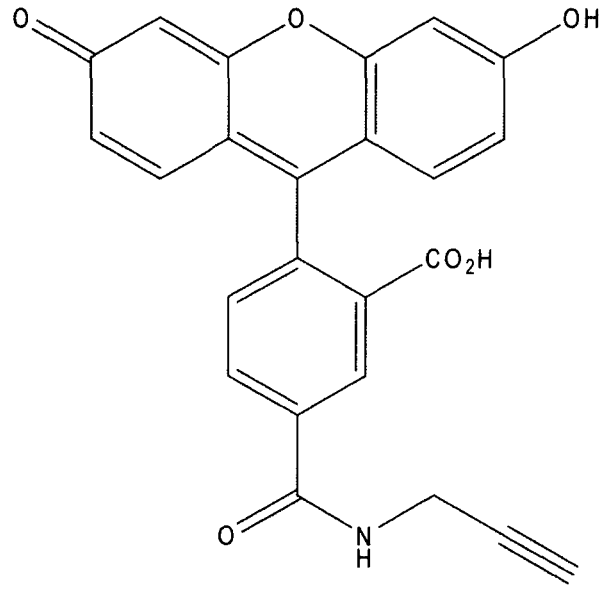


4.2

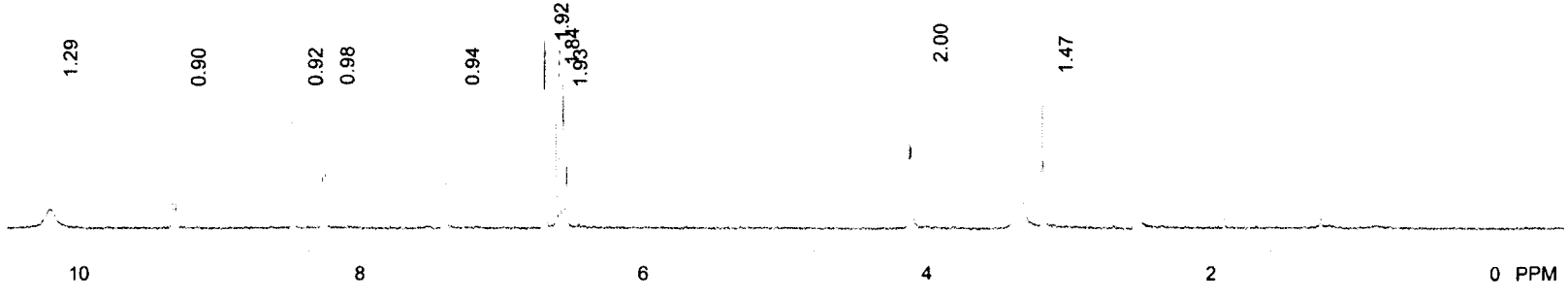
Carbon-13 NMR, 100 MHz
DMSO-d_6



4.10

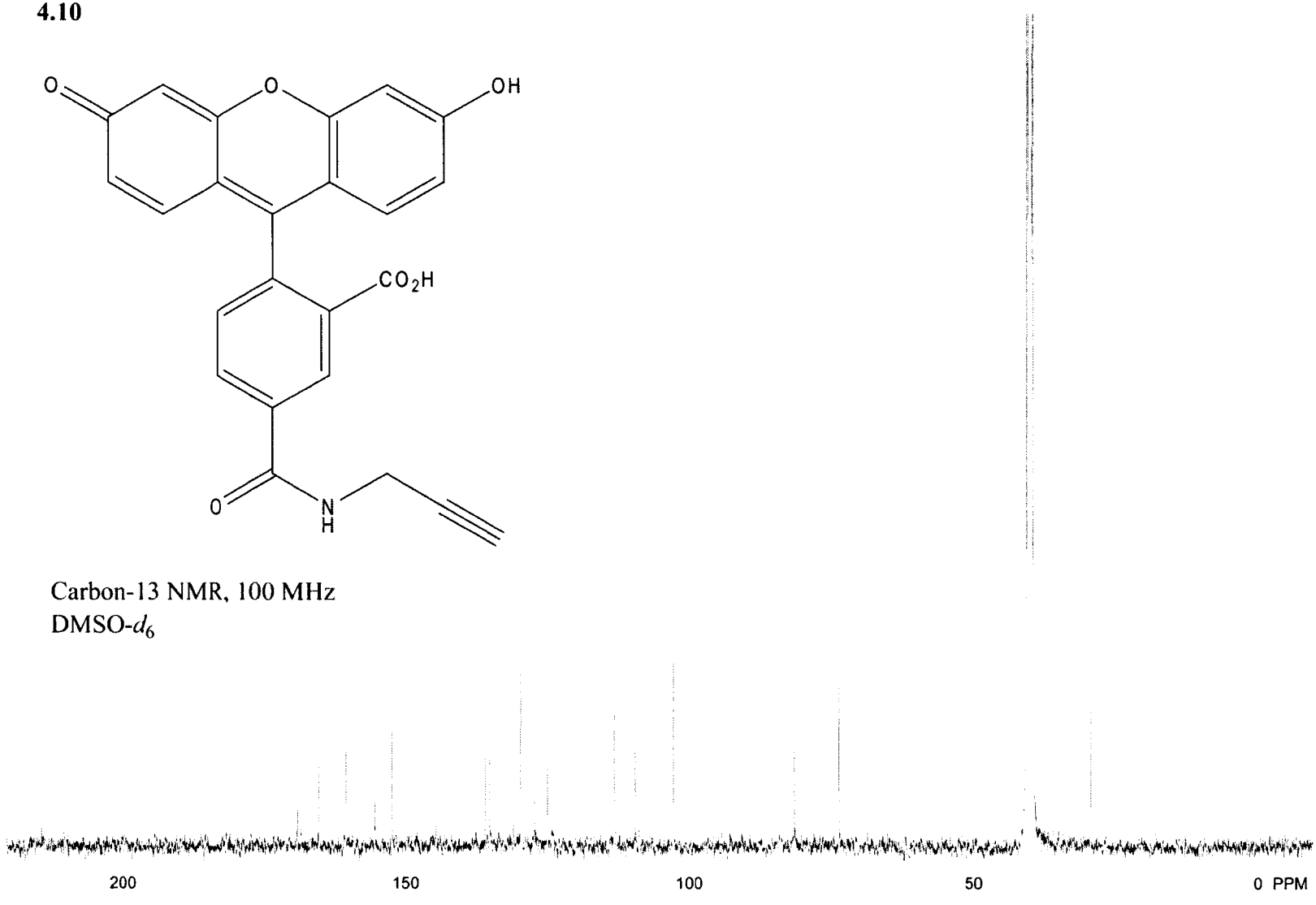


Proton NMR, 400 MHz
DMSO-d_6



4.10

Carbon-13 NMR, 100 MHz
DMSO-d_6



Chapter Five

General Method for Site-Specific Protein Immobilization by Staudinger Ligation

Contribution: Synthesis of phosphinothiol, surface chemistry, protein immobilization experiments, protein characterization on surface using fluorimager and ellipsometry, enzyme assays and Ribonuclease Inhibitor-binding experiments on immobilized RNase A, composition of manuscript and figure drafts.

This chapter was published as:
Kalia, J. and Raines, R. T. (2007) General Method for Site-Specific Protein Immobilization by Staudinger Ligation. *Bioconjugate Chem.* **18:** 1064–1069.

5.1 Abstract

Protein microarrays are playing an increasingly important role in the discovery and characterization of protein–ligand interactions. The uniform orientation conferred by site-specific immobilization is a demonstrable advantage in using such microarrays. Here, we report on a general strategy for fabricating gold surfaces displaying a protein in a uniform orientation. An azido group was installed at the C-terminus of a model protein, bovine pancreatic ribonuclease, by using the method of expressed protein ligation and a synthetic bifunctional reagent. This azido protein was immobilized by Staudinger ligation to a phosphinothioester-displaying self-assembled monolayer on a gold surface. Immobilization proceeded rapidly and selectively via the azido group. The immobilized enzyme retained its catalytic activity and was able to bind to its natural ligand, the ribonuclease inhibitor protein. This strategy provides a general means to fabricate microarrays displaying proteins in a uniform orientation.

5.2 Introduction

Protein–protein, protein–DNA, and protein–small-molecule interactions underlie biological chemistry. High-throughput approaches for the discovery and characterization of these various protein–ligand interactions can elucidate complicated biological pathways. In addition, the screening of small-molecule libraries for agonists and antagonists can provide leads for drug development. Recyclable devices that detect biomarkers in body fluids can be used to diagnosis human diseases. All of these challenges could be met with the use of protein

microarrays (Hu *et al.*, 2006; Kambhampati, 2004; Merkel *et al.*, 2005; Schena, 2004; Wilson and Nock, 2003).

Microarrays that present proteins in a uniform orientation exhibit both a markedly greater capacity to bind ligands (Peluso *et al.*, 2003; Sweeney *et al.*, 2000) (Du and Saavedra, 2003; Luk *et al.*, 2004) and more reproducible biological function (Cha *et al.*, 2005) than do protein arrays generated by random immobilization. Microarrays of uniformly oriented DNA have long been produced on a large scale and used with notable success in studies of gene function (Brown and Botstein, 1999; Hoheisel, 2006; Lipshutz *et al.*, 1999; Singh-Gasson *et al.*, 1999). The production of DNA microarrays is made facile by the ease with which oligonucleotides can be prepared by chemical synthesis (Caruthers, 1991) and incorporated into large fragments by DNA polymerase (Mullis, 1994). By comparison, the fabrication of protein microarrays is much more arduous. Notable advances are being made in the chemical synthesis of proteins (Nilsson *et al.*, 2005), but this technology is still nascent. The immobilization of natural proteins is complicated by their chemical complexity, as each of the twenty proteinogenic amino acids bears a side chain with distinct physicochemical properties, and the reactivity of the same amino acid can vary according to its position in the three-dimensional structure. Post-translational modifications confer additional diversity (Walsh, 2005). Other difficulties result from the intrinsic fragility of proteins, which have modest conformational stability and can be damaged upon chemical modification. Accordingly, new strategies are needed to immobilize proteins while maintaining their integrity.

Proteins have been immobilized on surfaces, both noncovalently and covalently (Yeo *et al.*, 2004). Noncovalent immobilization has been achieved by physical adsorption (Butler *et al.*, 1992; Duburcq *et al.*, 2004) and affinity tag-mediated complex formation (Paborsky *et al.*, 1996; Zhu *et al.*, 2001). Covalent immobilization, however, results in more robust arrays. A common strategy, immobilization by attack of the nucleophilic side chain of lysine or cysteine residues, typically proceeds at multiple sites and thus yields an array in which the proteins have a random orientation (Lahiri *et al.*, 1999; MacBeath and Schreiber, 2000). In contrast, site-specific covalent immobilization, via a unique natural (Liu *et al.*, 2006; Luk *et al.*, 2004; Sweeney *et al.*, 2000) or nonnatural residue, affords a uniformly oriented protein array.

The azido group is chemically inert to the functional groups found in nature. Accordingly, the azido group is being exploited widely by chemical biologists (Bräse *et al.*, 2005), including in the site-specific immobilization of proteins and peptides by the Huisgen 1,3-dipolar azide–alkyne cycloaddition (Duckworth *et al.*, 2006; Gauchet *et al.*, 2006; Sun *et al.*, 2006) and Staudinger ligation (Gauchet *et al.*, 2006; Soellner *et al.*, 2003; Watzke *et al.*, 2006b). A limitation, however, has been the difficulty of installing an azido group in a particular position in a protein. Recently, we developed a general method for appending the C-terminus of proteins with an azido group (Kalia and Raines, 2006). Specifically, we used a bifunctional reagent bearing an α-hydrazino acetamido and azido group to cleave the transient thioester generated on the C-terminus of the target protein fused to an intein, thereby labeling the protein

with an azido group on the C-terminus (Scheme 5.1). This azido group was available for site-specific installation of a fluorophore by the Huisgen 1,3-dipolar azide–alkyne cycloaddition.

Here, we report a general method for the site-specific immobilization of a protein. As a model protein, we choose one of the most thoroughly studied enzymes, bovine pancreatic ribonuclease (RNase A (Raines, 1998); EC 3.1.27.5). The immobilized protein was detected with a standard immunoassay, which does not report on the integrity of the protein. The choice of RNase A provides two rigorous means to assess its structure and function. First, RNase A is an efficient catalyst of RNA cleavage, and its catalytic activity provides a measure of its conformational integrity. Second, the affinity of the ribonuclease inhibitor protein (RI) for RNase A is extremely high (Dickson *et al.*, 2005), and this interaction provides an independent report on the structure and function of the immobilized protein. Our strategy was to react an azido–RNase A (Scheme 5.1) (Kalia and Raines, 2006) with a phosphinothioester group displayed on self-assembled monolayers (SAMs) of alkane thiols on a gold surface (Scheme 5.2). The resulting Staudinger ligation (Köhn and Breinbauer, 2004; Nilsson *et al.*, 2001; Nilsson *et al.*, 2005; Soellner *et al.*, 2006) immobilized the protein in a form that retained its biological function.

5.3 Experimental Procedures

Materials. Silicon chips coated with gold prepared by physical vapor deposition as described previously (Jang *et al.*, 2006) and were a generous gift from C.-H. Jang.

Alkane thiols $HS(CH_2)_{11}(OCH_2CH_2)_6OCH_2CO_2H$ and $HS(CH_2)_{11}(OCH_2CH_2)_3OH$ were from Prochimia (Gdansk, Poland). A fluorogenic ribonuclease substrate, 6-carboxyfluorescein–dArU$(dA)_2$–6-carboxytetramethylrhodamine (6-FAM–dArU$(dA)_2$–6-TAMRA), was from Integrated DNA Technologies (Coralville, IA). Two-hand AtmosBags™ were from Sigma–Aldrich (St. Louis, MO).

Azido–RNase A was produced as described previously (Scheme 5.1) (Kalia and Raines, 2006). Wild-type RNase A was from Sigma–Aldrich. Human RI was prepared as described previously and was a generous gift from R. J. Johnson and G. A. Ellis. Anti-RNase A rabbit primary antibody was from Biodesign International (Kennebunk, ME). Alexa Fluor® 488-conjugated anti-rabbit secondary antibody was from Molecular Probes (Eugene, OR). Anti-RI chicken primary antibody was from Genetel (Madison, WI). Fluorescein-conjugated anti-chicken secondary antibody was from Abcam (Cambridge, MA).

Instrumentation. Fluorescence measurements for assaying ribonucleolytic activity were made with a QuantaMaster 1 photon-counting spectrofluorimeter equipped with sample stirring (Photon Technology International, South Brunswick, NJ). Immunoassays of immobilized RNase A and RI bound to the immobilized RNase A were visualized with a GeneTac UC4×4 fluorescence scanner (Genomic Solutions, Ann Arbor, MI). The optical thickness of SAMs and proteins on the gold chips was determined with a AutoEL ellipsometer (Rudolf Research, Flanders, NJ).

Preparation of phosphinothioester-displaying SAMs of alkane thiols on gold chips. Alkane thiol solutions were prepared by dissolving $HS(CH_2)_{11}(OCH_2CH_2)_6OCH_2CO_2H$ and $HS(CH_2)_{11}(OCH_2CH_2)_3OH$ in ethanol to a final concentration of 0.25 mM each. The oligo(ethylene oxide) $HS(CH_2)_{11}(OCH_2CH_2)_3OH$ is included to discourage the nonspecific adsorption of protein (Prime and Whitesides, 1993), which can be problematic (Wood *et al.*, 1997). Gold chips were cleaned under a stream of Ar(g) and immersed in the alkane thiol solution for at least 18 h. After rinsing thoroughly with ethanol and drying under a stream of Ar(g), the chips were overlaid with an aqueous solution containing *N*-hydroxysuccinimide (NHS, 50 mM) and 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC, 200 mM) for 7 min to generate succinimidyl esters in situ. Diphenylphosphinomethanethiol (PPh_2CH_2SH) was synthesized as described earlier previously (Soellner *et al.*, 2002), and was dissolved in anhydrous DMF to a concentration of 0.10 M. *N,N*-Diisopropylethylamine (DIEA) was added to a concentration of 0.12 M. The resulting solution (in a 25 mL flask), succinimidyl ester displaying chips (in a Petri dish containing a water-soaked absorbent paper to serve as a humid chamber), and azido–RNase A (10 μM in 5% v/v DMF(aq)) were placed in a two-hand AtmosBag™ along with vials containing anhydrous DMF (20 mL/vial), water (20 mL/vial), and 25 mM sodium phosphate buffer at pH 7.5 (20 mL/vial). The AtmosBag™ was sealed and connected to the house vacuum on one end and an Ar(g) supply on the other end. The air inside the bag was removed by using the house vacuum and flushing twice with Ar(g) before finally filling up the bag with Ar(g).

The diphenylphosphinomethanethiol–DIEA solution was transferred into a small empty vial and succinimidyl ester displaying chips were incubated in that vial for 2 h. The phosphinothioester-displaying chips thus produced were rinsed with DMF (2 × 20 mL) and then with water (2 × 20 mL).

Immobilization of azido–RNase A by Staudinger ligation. Azido–RNase A (1 μl, 10 μM in 5% v/v DMF(aq)) was incubated on phosphinothioester-displaying chips for 1, 5, 10, and 15 min inside the AtmosBag™ (Figure 5.1). The chips were subsequently rinsed with 25 mM sodium phosphate buffer at pH 7.5 (20 mL) and removed from the AtmosBag™. They were then incubated in 25 mM sodium phosphate buffer at pH 7.5 (1.0 mL) for 30 min to remove protein bound nonspecifically.

Detection of immobilized RNase A. A chip displaying immobilized RNase A was overlaid with anti-RNase A rabbit primary antibody (100 μg/mL in 25 mM sodium phosphate buffer at pH 7.5) for 30 min. The chip was then rinsed with 25 mM sodium phosphate buffer at pH 7.5 (2 × 20 mL), and incubated in the same buffer for 15 min. The chip was subsequently overlaid with Alexa Fluor® 488-conjugated anti-rabbit secondary antibody (2.0 μg/mL in 25 mM sodium phosphate buffer at pH 7.5) for 30 min. Finally, the chip was rinsed with 25 mM sodium phosphate buffer at pH 7.5 (2 × 20 mL), incubated in the same buffer for 15 min, and scanned with a Genomic Solutions® GeneTac UC4×4 Fluorescence Scanner using a gain of 50.

Ribonucleolytic activity assay on immobilized RNase A. A phosphinothioester-displaying chip was spotted with azido–RNase A (10 μM in 5% v/v DMF (aq)) in an Ar(g) atmosphere to immobilize RNase A. The chip was rinsed with 25 mM sodium phosphate buffer at pH 7.5 (2 × 20 mL) and then introduced into a vial containing a solution (22 mL) of 0.10 M 2-(N-morpholino)ethanesulfonic acid (MES)–NaOH buffer at pH 6.0 containing 0.10 M NaCl. A fluorogenic substrate of RNase A, 6-FAM–dArU$(dA)_2$–6-TAMRA (10 μL of a 40 μM solution) was introduced into the vial, and the vial was placed on a shaker. Aliquots (2.0 mL) were withdrawn at known time intervals, and their fluorescence was measured (λ_{ex} = 493 nm, λ_{em} = 515 nm). After 1 h, excess wild-type RNase A (20 μL of a 2.8 mg/mL solution) was added to obtain a value of fluorescence after complete cleavage of the substrate. A phosphinothioester-displaying chip was overlaid with wild-type RNase A (10 μM in 5% v/v DMF(aq)), and the assay was performed again. As a negative control, a similar assay was performed on a phosphinothioester-displaying chip which was not overlaid with RNase A. The percentage of the fluorescent substrate cleaved at various time intervals was subtracted from the corresponding values for wild-type RNase A, and plotted against time to generate the graph in Figure 5.2.

Detection of RI bound to immobilized RNase A. Azido–RNase A (1 μL, 10 μM in 5% v/v DMF(aq)) was spotted on phosphinothioester-displaying SAMs on a gold chip for 15 min in an AtmosBag™ filled with Ar(g). The chip was removed from the bag, rinsed with 25 mM sodium phosphate buffer at pH 7.5 (2 × 20 mL), and incubated in

the same buffer for 30 min to remove non-specifically bound RNase A. RI (10 μM in 25 mM sodium phosphate buffer at pH 7.5 containing 10 mM DTT) was overlaid on the chip for 10 min at 4 °C. The chip was then rinsed with 25 mM sodium phosphate buffer at pH 7.5 containing 10 mM DTT (2 × 20 mL) and incubated in the same buffer for 15 min.

Immunoassay. After RI binding, the chip was overlaid with bovine serum albumin (2.0 mg/mL in 25 mM sodium phosphate buffer at pH 7.5) for 15 min. The chip was rinsed with 25 mM sodium phosphate buffer at pH 7.5 (2 × 20 mL), and subsequently incubated with anti-RI chicken primary antibody (12.5 μg/mL in 25 mM sodium phosphate buffer at pH 7.5) for 15 min. The chip was rinsed with 25 mM sodium phosphate buffer at pH 7.5 (2 × 20 mL), incubated in the same buffer for 15 min, and subsequently overlaid with fluorescein-conjugated anti-chicken secondary antibody (1.5 μg/mL in 25 mM sodium phosphate buffer at pH 7.5) for 15 min. The chip was rinsed with 25 mM sodium phosphate buffer at pH 7.5 (2 × 20 mL), incubated in the same buffer for 15 min, and then scanned with a Genomic Solutions® GeneTac UC4×4 fluorescence scanner using a gain of 50.

Ellipsometry. Ellipsometric measurements were performed to determine the optical thicknesses of the SAMs, RNase A, and RI on the gold surface. The measurements were performed with a wavelength of 632 nm and an angle of incidence of 70°. The optical thickness reported is the average from seven different surface locations. A

slab model was used to interpret these constants in which the slab (SAM and protein) was assumed to have an index of refraction of 1.46.

5.4 Results and Discussion

Generation of phosphinothioester-displaying SAMs on gold. SAMs of alkane thiols on gold were prepared by incubating gold-coated silicon chips in an ethanolic solution containing equimolar quantities of $HS(CH_2)_{11}(OCH_2CH_2)_6OCH_2CO_2H$ and $HS(CH_2)_{11}(OCH_2CH_2)_3OH$. Subsequently, the chips were overlaid with an aqueous solution of EDC/NHS to produce succinimidyl ester-displaying chips. Diphenylphosphinomethanethiol was synthesized in multigram quantities by a procedure described previously (Soellner *et al.*, 2002), and was reacted with succinimidyl ester-displaying chips to generate phosphinothioester-displaying chips. The resultant phosphinothioester-displaying SAMs of alkane thiols on gold surfaces provide a highly homogeneous surface for performing the Staudinger ligation. These surfaces are significantly better defined than that of a glass slide, which had been used previously for Staudinger ligation (Köhn *et al.*, 2003; Soellner *et al.*, 2003; Watzke *et al.*, 2006b) thereby enabling the use of high-precision techniques such as ellipsometry, surface plasmon resonance (Green *et al.*, 2000), and liquid crystal methods (Gupta *et al.*, 1998; Luk *et al.*, 2004) for the detection of ligand binding.

Azido–RNase A immobilization by Staudinger ligation. Azido-RNase A (in 5% v/v DMF(aq)) was incubated on phosphinothioester-displaying chips for immobilization

by Staudinger ligation (Scheme 5.2). A time-course of immobilization of azido-RNase A is depicted in Figure 5.1. The immobilized RNase A was detected by an immunoassay utilizing a primary anti-RNase A antibody and a fluorescent secondary antibody. Protein immobilization proceeded rapidly, with a significant fluorescence signal being clearly observable after only 1 min. The rapidity of this reaction is reminiscent of that for the immobilization of azido-peptides to glass slides with a phosphinothioester-mediated Staudinger ligation (Soellner *et al.*, 2003).

Exposure of a protein to a phosphinothioester-displaying chip could lead to undesirable protein immobilization via a side chain of the protein. For example, RNase A has 10 lysine residues, any one of which could attack the phosphinothioester via its nucleophilic ε-amino group, thereby resulting in random immobilization. Additionally, the protein could be immobilized noncovalently onto the surface by hydrophobic interactions with the phenyl groups of the phosphinothioester. To probe for these possible immobilization routes, wild-type RNase A was incubated adjacent to azido–RNase A on phosphinothioester-displaying SAMs on gold. No significant immobilization of wild-type RNase A was observed during the timescale required for substantial azido–RNase A immobilization (Figure 5.1), thereby establishing that the immobilization of azido–RNase A proceeded exclusively via the azido group. Longer incubation times (>1 h) did result in detectable immobilization of wild-type RNase A (data not shown), highlighting the benefit of the rapidity of the Staudinger ligation.

Attempts to immobilize azido–RNase A in purely aqueous solutions of various pH were unsuccessful (data not shown). The presence of a small quantity of polar

organic solvent (here, 5% v/v DMF) is required, perhaps for adequate solubilization of the immobilized phosphinothioester. A version of the Staudinger ligation has been used to immobilize proteins in purely aqueous solutions (Watzke *et al.*, 2006b), though this method requires a long incubation period (4 h) and installs a racemic nonnatural moiety—a phosphine oxide—into the immobilized protein.

Activity of immobilized RNase A. Enzymatic activity assays were performed with phosphinothioester-displaying chips incubated with azido–RNase A for 10 min. These chips were incubated in solutions containing 6-FAM–dArU$(dA)_2$–6-TAMRA, which is a fluorogenic substrate for RNase A (Abel *et al.*, 2002). After 1 h, the fluorescence intensity from a chip incubated with azido-RNase A had increased substantially relative to one incubated with wild-type RNase A (Figure 5.2). Thus, the azido-RNase A was immobilized preferentially, and immobilization did not destroy its enzymatic activity.

Binding to immobilized RNase A. A phosphinothioester-displaying chip was overlaid with azido–RNase A, and subsequently overlaid with RI. Another phosphinothioester-displaying chip was incubated with wild-type RNase A, and then overlaid with RI. Both chips were assayed for the presence of RI by an immunoassay that involved a primary antibody to RI and a fluorescent secondary antibody. The chip that was overlaid with azido–RNase A had a high fluorescence signal, whereas the chip incubated with wild-type RNase A produced negligible fluorescence (Figure

5.3A). This result demonstrated that the RI detected on the surface was bound to the immobilized RNase A, and was not binding to the surface non-specifically.

Ellipsometric measurements corroborated these findings (Figure 5.3B). Upon incubation of azido–RNase A on phosphinothioester chips for 10 min, the ellipsometric thickness of the surface increased by 16.9 Å. Exposure to RI resulted in a further increase of 14.3 Å. That the increase in thickness upon RI binding was less than the increase in thickness upon azido–RNase A immobilization is surprising, as RI (molecular mass: ~51 kDa) is a larger protein than RNase A (molecular mass: ~14 kDa). One explanation is that the immobilization of RNase A via its C-terminus forces RI to bind obliquely to the surface, causing a smaller increase in the thickness than expected from lineal binding. Another explanation is that the binding of RI molecules to a few immobilized RNase A molecules creates steric hinderance to other incoming RI molecules. Consequently, not all of the immobilized RNase A molecules would form a complex with RI, producing a smaller than expected increase in ellipsometric thickness. To obtain additional information, we measured the ribonucleolytic activity of the immobilized RNase A after incubation with excess RI. Only half of the activity was lost, consistent with steric hindrence preventing half of the RNase A molecules from forming a complex with RI.

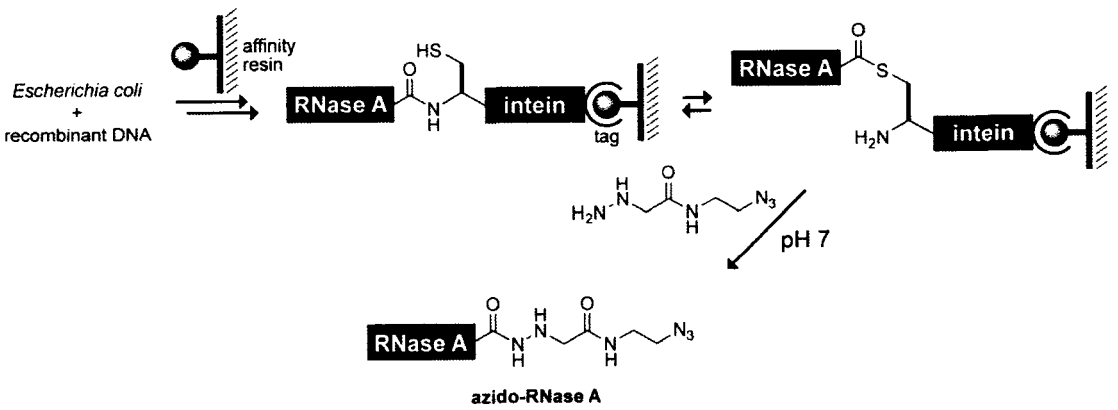
5.5 Conclusions

We have developed a general strategy for the site-specific immobilization of a protein on SAMs on a gold surface. Immunoassays, assays of enzymatic activity, and

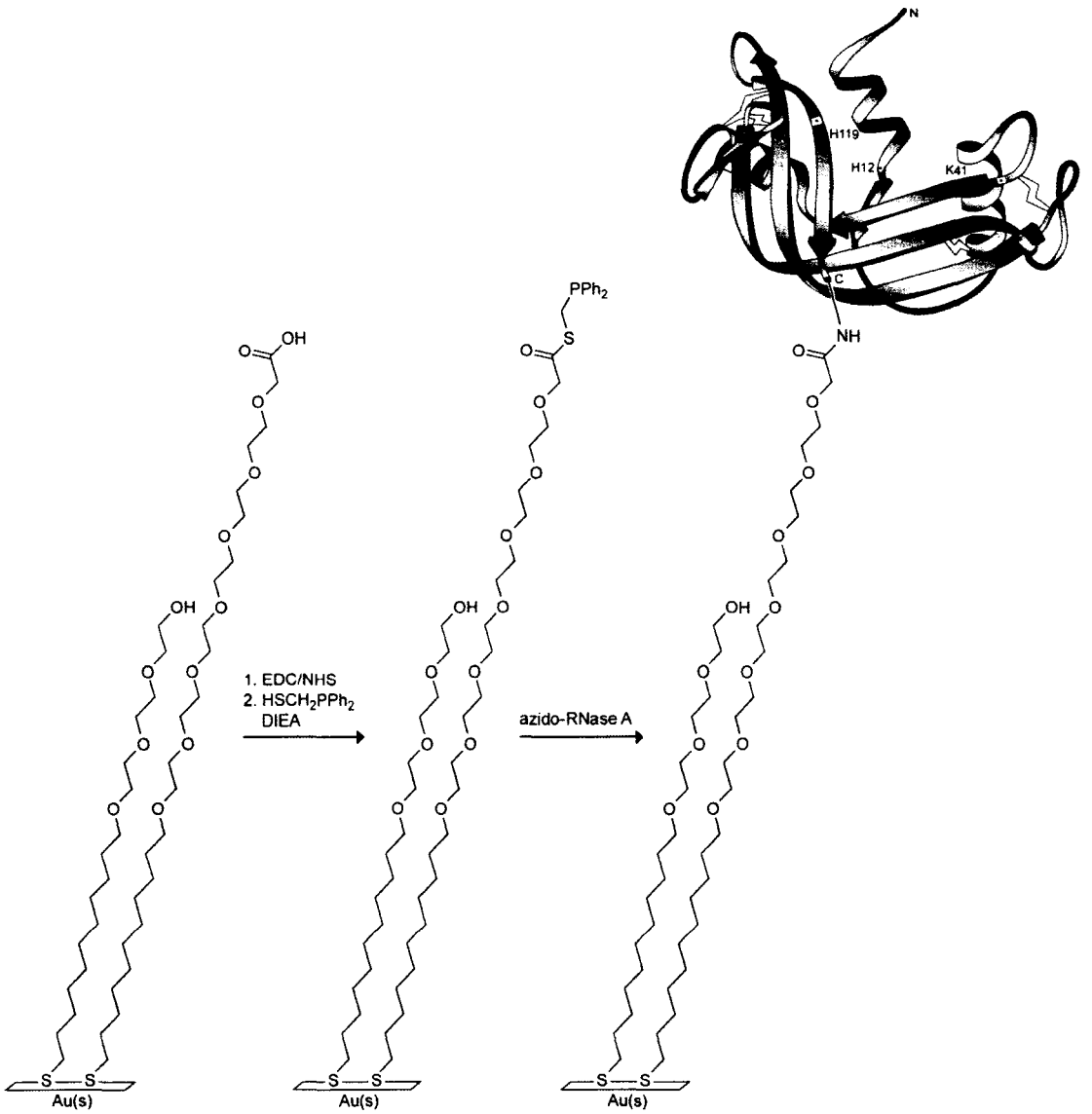
ellipsometric measurements revealed that the immobilized protein retained its structure and function. We envisage numerous applications of our strategy, which could be applicable to any protein. For example, we expect that our strategy could be useful for the high-throughput screening of ligands, in diagnostic applications that detect disease markers in clinical samples, and in industrial processes that utilize immobilized enzymes as recyclable catalysts (Chibata *et al.*, 1979).

5.6 Acknowledgments

We are grateful to C.-H. Jang, R. J. Johnson, and G. A. Ellis for supplying reagents, and T. J. Rutkoski for contributive discussions. This work was supported by Grants GM44783 and CA108467 (NIH), and the Materials Research Sciences and Engineering Center at the University of Wisconsin–Madison (NSF DMR-0520527).



Scheme 5.1 Semisynthesis of azido–RNase A.



Scheme 5.2. Site-specific protein immobilization on a SAM by Staudinger ligation.

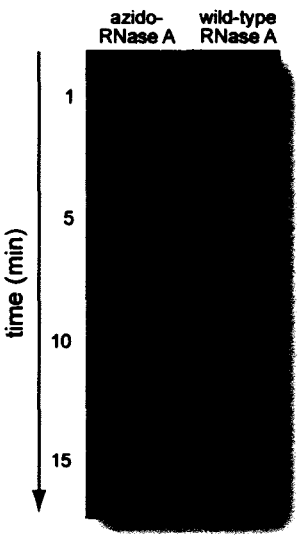


Figure 5.1. Time course and specificity of immobilization of azido–RNase A. Azido–RNase A and wild-type RNase A (each 10 μM) were spotted on a phosphinothioester-displaying chip for various times. The chip was overlaid with an anti-RNase A primary antibody (10 μg/mL) and a secondary antibody conjugated to AlexaFluor 488 (2 μg/mL).

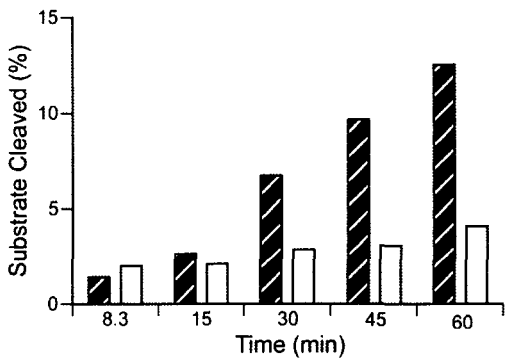


Figure 5.2. Catalytic activity of immobilized RNase A. Phosphinothioester-displaying chips were incubated with azido–RNase A (hatched bars) or wild-type RNase A (open bars), and assayed for ribonucleolytic activity by using a fluorogenic substrate.

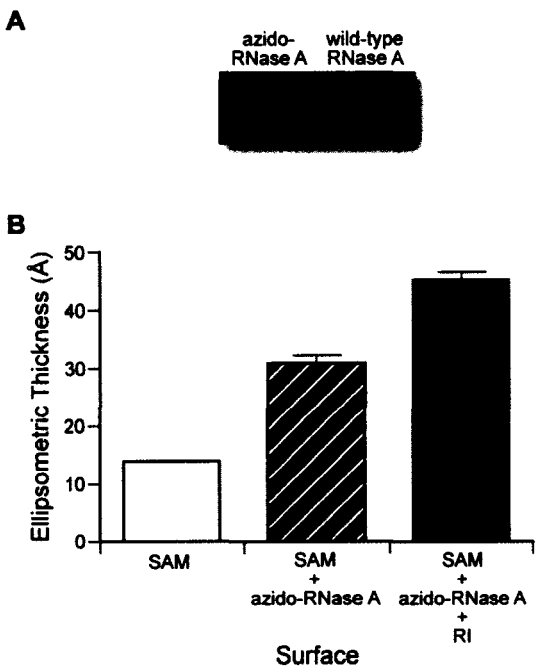


Figure 5.3. Binding of RI to immobilized RNase A. (A) Immunoassay with a primary anti-RI antibody and a secondary antibody conjugated to fluorescein. (B) Ellipsometric thickness before and after incubation with azido–RNase A, and after the addition of RI.

Chapter Six

Conclusions and Future Directions

Bioconjugation is a burgeoning field of research. Novel methods for specific derivatization of proteins, DNA, RNA, and carbohydrates are being developed for various applications. These powerful methods owe their existence to rigorous physical organic chemistry performed over the past 50 years. In fact, almost all modern ligation and bioconjugation methods are built on the solid foundations laid by earlier work. For example, Staudinger ligation is a creative modification of the Staudinger reduction, click chemistry is a Cu(I)-catalyzed version of Huisgen's 1,3-dipolar cycloaddition, and native chemical ligation is an extension of earlier studies on the reaction of N-terminal cysteines with thioesters.

It is imperative for the modern biochemist to not forget the fundamental principles that underlie the design of bioconjugation linkages they use almost on a daily basis. In recognition of this fact, Chapters 2 and 3 of this thesis attempted to re-learn the basics of chemical principles underlying three ubiquitous bioconjugation linkages, namely, maleimides, hydrazones and oximes. During this process, we stumbled upon some startling discoveries—maleimide conjugates hydrolyze to generate heterogeneity, and oximes are more stable than hydrazones by several orders of magnitude. Our observations profoundly influence data interpretation and experimental planning of research that utilize these linkages. We proceeded to understand the mechanistic basis of these results, propose means to minimize undesirable features of the linkages, and delineate the optimal conditions under which each linkage should be employed. We believe that such detailed investigations need to be performed on other commonly utilized bioconjugation linkages—there are

treasures of knowledge on physical organic chemistry to be unearthed, and erroneously performed and interpreted experiments corrected. For example, iodoacetamides and maleimides are common linkages for bioconjugation via thiol groups. Which one of these linkages has the optimal combination of chemoselectivity, hydrolytic stability, and rate of formation? Answering this question is a potential future project.

Chapters 4 and 5 describe a novel method for site-specific bioconjugation that entails generating proteins labeled at their C-termini with the azido group, which is utilized to perform conjugation chemistry such as Staudinger ligation and Huisgen 1,3-dipolar cycloaddition. This method exploits the ability of inteins to install an electrophilic center—the thioester linkage—at the C-terminus of proteins. It would be interesting to study the chemistry of hydrolysis of thioesters in more detail. The discovery of a catalyst for the hydrolysis of thioesters at neutral pH would be useful in purifying proteins by intein chemistry. Existing methods utilize reagents such as thiols and hydroxylamine to cleave protein–intein thioesters. The resultant purified protein is not the wild-type protein, since the reagent remains attached to the C-terminus of the protein, which is undesirable, especially if the C-terminal region of the protein is important for the activity of the protein. A potent but selective catalyst for the hydrolysis of protein–intein thioesters would solve this problem. Molybdate, which was utilized as a catalyst for the hydrolysis of maleimides in Chapter 2, is a potential catalyst for the hydrolysis of protein–intein thioesters.

Finally, we have integrated intein chemistry with yeast surface display to generate azido–proteins efficiently (Appendix I). In future work, protein–intein fusion proteins displayed on yeast can be treated with fluorescent molecules appended with an α-hydrazino acetamido group—an optimal nucleophile for thioesters (Chapter 4)—to label proteins on their C-termini. Furthermore, semisynthesis of proteins by expressed protein ligation can be performed by treating peptide–intein thioesters displayed on the yeast surface with peptides bearing N-terminal cysteines. This approach would circumvent the cell lysis, and chitin-column chromatography steps of traditional expressed protein ligation protocols, thus enabling highly efficient semisynthesis of proteins. Moreover, the eukaryotic protein production machinery of yeast would allow semisynthesis of proteins that are not efficiently expressed in *E. coli*, but can be readily expressed in yeast, thereby increasing the scope of expressed protein ligation.

Appendix I

Production of Azido–Single-Chain Antibodies

This work was done in collaboration with N. Agarwal, Department of Chemical and Biological Engineering at the University of Wisconsin–Madison.

Contribution: Synthesis of hydrazino azide, alkynyl fluorescein, alkynyl biotin, hydrazino biotin, and phosphino thiol reagents, cloning and production of sm3E–*mxe* intein–chitin-binding domain fusion protein in *E. coli*, purification and characterization of azido–sm3E, and experimental design and troubleshooting of immobilization experiments.

AI.1 Abstract

We have developed two methods for generating azido–scFvs based on intein chemistry—one that utilizes an *E. coli* protein expression system and the other that employs yeast surface display. ScFv–intein thioesters generated by both these methods were treated with a bifunctional reagent bearing an α-hydrazino acetamido and an azido group to generate azido–scFvs. Azido–scFvs retained antigen-binding activity. Furthermore, their azido groups were utilized for site-specific labeling with fluorescein and biotin, and immobilization via Huisgen 1,3-dipolar azide–alkyne cycloaddition. These two methods of producing azido–proteins provide the means to perform highly selective bioconjugation on both prokaryotic and eukaryotic proteins.

AI.2 Introduction

Antibodies have tremendous potential for applications in clinical diagnoses and therapy (Waldman, 1991). However, the production of full-length antibodies is confounded due to their inefficient expression in *E. coli.* Moreover, regions of the antibody that are not involved in antigen-binding often bind to other cellular components, causing undesirable background signals. For example, when used as tumor markers, considerable binding of heavy constant regions of antibodies to non-tumor cells was observed (Wahl *et al.*, 1983). To counter these problems, Whitlow and coworkers introduced single-chain antibodies (scFvs) which are much smaller, "minimal" antibody molecules (Bird *et al.*, 1988). ScFvs are recombinant antibody fragments, consisting of only the variable heavy chain (V_H) and the variable light

chain (V_L) linked to each other via an oligopeptide linker. Due to their small size, scFvs exhibit rapid pharmacokinetics and tumor penetration *in vivo*, and can be readily expressed in *E. coli*. ScFvs have been prepared against more than 30 different antigens (Huston *et al.*, 1991; Raag and Whitlow, 1995).

In vitro affinity maturation methods such as phage display (Burton, 1993; Winter *et al.*, 1994), *E. coli* display (Francisco *et al.*, 1993; Georgiou *et al.*, 1997), and yeast surface display (Boder and Wittrup, 1997; Gai and Wittrup, 2007) have revolutionized the field of antibody engineering. These methods enable generation of high-affinity antibodies (Hawkins *et al.*, 1992; Rajpal *et al.*, 2005; Schier *et al.*, 1996) which are extremely useful for therapeutic purposes (Schlom *et al.*, 1992). Yeast surface display is especially useful because it employs an eukaryotic host possessing the biosynthetic apparatus for allowing posttranslational modifications and oxidative protein folding.

Utilizing yeast surface display, Wittrup and coworkers generated sm3E—a scFv that binds the carcinoembryonic antigen (CEA) with picomolar-affinity (Graff *et al.*, 2004). CEA is a tumor-associated antigen, and is present in colonic tumors and in tumors of epithelial origin, such as breast, lung, and pancreas in concentrations greater than 60-fold higher than in normal body fluids (Boucher *et al.*, 1989). CEA has been used for tumor-monitoring and also for antibody-targeting (Thompson *et al.*, 1991). In fact, the use of radiolabeled scFvs against CEA as tumor-targeting agents is the first example of the use of scFvs in patients (Begent *et al.*, 1996). We reasoned that site-specifically immobilized sm3E could be useful as a diagnostic tool for

cancer. A method analogous to the one described in Chapters 4 and 5 for the production and immobilization of azido–RNase A was attempted. We generated azido–sm3E by utilizing a method based on intein chemistry discussed in detail in Chapter 4 (Figure AI.1). *E. coli* cells expressing the sm3E–*mxe* intein–chitin-binding domain fusion protein were lysed and the cell lysate was loaded on a chitin column. Subsequently, the desired azido–scFv was obtained by performing on-column cleavage of scFv–intein thioesters with a bifunctional reagent bearing a potent nucleophile for thioesters—the α-hydrazino acetamido group—and the azido group (compound **4.1** in Chapter 4).

A drawback of the above method for producing azido–proteins is its inefficiency due to the cell lysis and affinity chromatography steps. Moreover, the *E. coli* expression system is not suitable for expressing numerous eukaryotic proteins, and proteins possessing multiple cystines. To remove these limitations, we developed a novel method that combines intein chemistry with yeast surface display (Figure AI.2). Yeast produces the scFv–*mxe* intein fusion protein, and displays the fusion protein on its surface. Upon treatment with azide **4.1**, the scFv–intein thioesters are cleaved to generate the desired azido–scFv, circumventing the cell lysis and chitin column chromatography steps of the former method. Furthermore, since yeast surface display utilizes a eukaryotic expression system, this method is more conducive for producing eukaryotic azido–proteins. Yeast surface display was utilized for the production of two azido–scFvs—azido–anti-epidermal growth factor receptor (EGFR), and azido–anti-fluorescein. Azido–scFvs were labeled site-specifically by fluorescein and biotin,

and immobilized in a uniform orientation by utilizing Cu(I)-catalyzed Huisgen 1,3-dipolar azide–alkyne cycloaddition.

AI.3 Experimental Procedures

Materials. The cDNA of sm3E scFv was obtained from Professor K. D. Wittrup at the Massachusetts Institute of Technology. Plasmid pTXB1 was from New England Biolabs (Beverly, MA). Anti-FLAG M1 resin was from Sigma–Aldrich (St. Louis, MO). Silicon chips coated with gold were a generous gift from Professor N. L. Abbott, Department of Chemical and Biological Engineering at the University of Wisconsin–Madison. Alkane thiols $HS(CH_2)_{11}(OCH_2CH_2)_6OCH_2CO_2H$ and $HS(CH_2)_{11}(OCH_2CH_2)_3OH$ were from Prochimia (Gdansk, Poland). Carcinoembryonic antigen (CEA) was purchased from Fitzgerald (Concord, MA). Alexa Fluor 488 succinimidyl ester was from Molecular Probes (Eugene, OR). *N*-Hydroxysuccinimidobiotin was from Pierce (Rockford, IL). Phosphate-buffered saline (PBS) contained (in 1.0 L) NaCl (8.0 g), KCl (2.0 g), $Na_2HPO_4.7H_2O$, KH_2PO_4 (2.0 g), and NaN_3 (0.10 g), and had pH 7.4. Azide **4.1** and alkynyl fluorescein **4.10** were synthesized as described in Chapter 4. Alkynyl biotin **AI.1** (Figure AI.3) was synthesized by following a procedure reported earlier (Lin *et al.*, 2007).

Instrumentation. Immunoassays on immobilized scFvs were visualized with a GeneTac UC4×4 fluorescence scanner (Genomic Solutions, Ann Arbor, MI).

Production of sm3E–mxe intein–chitin-binding domain in E. coli. DNA primers were designed to amplify the cDNA encoding sm3E, and to incorporate *NdeI* and *SapI* restriction sites at the 5′ and 3′ ends respectively. The PCR products were purified with Promega's Wizard® SVGel and PCR Clean-Up system, digested with the restriction enzymes, and inserted into plasmid pTXB1 which directs the expression of the *mxe*-intein fused to a chitin-binding domain. The resultant plasmid was transformed into *E. coli* BL21(DE3) cells, and the production of Met(–1)sm3E–Ala–Ala–Gly–*mxe* intein–chitin-binding domain fusion protein was induced by adding 0.5 mM IPTG (see Figure AI.4A for test inductions).

Production of azido–sm3E by using an E. coli expression system. The procedure for producing sm3E–intein–chitin-binding domain fusion protein was based on an earlier reported method for expressing scFv fusion proteins by Nock and coworkers (Sydor *et al.*, 2002). Cell paste from a 1 L culture was suspended in 20 mM 3-(N-morpholino)propanesulfonic acid (MOPS)–NaOH buffer at pH 6.8 containing NaCl (0.50 M), ethylenediaminetetraacetic acid (EDTA; 0.10 mM), and Triton X-100 (0.1% v/v). Cells were lysed with a French pressure cell, and the cell lysate was subjected to centrifugation at 15,000*g* for 30 min. SDS–PAGE analysis demonstrated that the sm3E–intein–chitin-binding domain fusion protein was expressed as an insoluble protein, and was present in the pellet (Figure AI.4B). The pellet was suspended in 100 mM Tris–HCl buffer at pH 7.0 containing 1 mM EDTA (10 mL). To the resultant solution was added an aqueous solution (5 mL) containing EDTA (60

mM), Triton X-100 (6.0% v/v), and NaCl (1.5 M). After stirring at 4 °C for 30 min, the suspension was subjected to centrifugation at 12,000*g* for 20 min, and the pellet was resolubilized in 50 mM sodium phosphate buffer at pH 7.5 containing 1 mM EDTA, and centrifuged again at 12,000*g* for 20 min. The resultant pellet was solubilized in 50 mM sodium phosphate buffer at pH 7.5 containing 1 mM EDTA, 6 M guanidinium–HCl , and 100 mM DTT (5 mL). The solution was stirred for 2 h at room temperature, the pH adjusted to ~4.0 by adding HCl, and the solution was centrifuged at 10,000*g* for 30 min to remove insoluble cellular debris. The supernatant was filtered and the filtrate was added drop-wise into a refolding buffer (50 mL)—50 mM sodium phosphate buffer at pH 7.5 containing EDTA (1 mM), and urea (4 M). The refolding solution was dialyzed against the refolding buffer (1 L) for 12 h. The dialyzed protein was subjected to centrifugation at 10,000*g* for 30 min, and the supernatant was loaded on a chitin column (8 mL) equilibrated with the refolding buffer. Azide **4.1** was dissolved in the refolding buffer to a concentration of 500 mM, the pH was adjusted to 7.5, and three column-volumes of this solution were loaded onto the resin, out of which, two column-volumes were allowed to flow through and one column-volume was allowed to sit on top of the resin. This incubation was carried out for three days at room temperature to enable the reaction to proceed to completion. Azido–sm3E (band II in Figure AI.4C) was eluted with four column-volumes of the refolding buffer. These fractions were contaminated with a high-molecular weight impurity (band I in Figure AI.4C). This contaminant was sm3E–intein–chitin-binding domain protein that leached off the chitin column. This impurity

was removed by applying the protein solution to a small chitin column, and collecting the flow-through (Figure AI.4D).

Testing activity of azido–sm3E. Alkane thiol solutions were prepared by dissolving $HS(CH_2)_{11}(OCH_2CH_2)_6OCH_2CO_2H$ and $HS(CH_2)_{11}(OCH_2CH_2)_3OH$ in ethanol to a final concentration of 0.25 mM each. Gold chips were cleaned under a stream of Ar(g) and immersed in the alkane thiol solution for at least 18 h. After rinsing thoroughly with ethanol and drying under a stream of Ar(g), the chips were overlaid with an aqueous solution containing *N*-hydroxysuccinimide (NHS, 50 mM) and 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC, 200 mM) for 7 min to generate succinimidyl esters-displaying chips. Subsequently, azido–sm3E (1 μl, 20 mM in PBS) was randomly immobilized by incubating it on succinimidyl ester-displaying chips for 10 min. The chip was then rinsed with water (2 × 20 mL), and overlaid with AlexaFluor 488-labeled CEA (0.8 μM in PBS) for 30 min. Finally, the chip was rinsed and visualized under a fluorimager. AlexaFluor 488-labeled CEA was generated by treating CEA with Alexa Fluor 488 succinimidyl ester in $NaHCO_3$ (0.1 M) buffer at pH 8.3, and passing the reaction mixture through a desalt spin-column to remove unreacted dye.

Production of azido–anti-EGFR and azido–anti-fluorescein by yeast surface display. Yeast cells displaying scFv–intein thioesters were suspended in MOPS–NaOH (0.5 M) buffer at pH 7.0 containing NaCl (0.50 M), EDTA (0.10 mM) and azide **4.1** (500

mM) for 3 days at room temperature. Subsequently, the suspension was centrifuged at 14000*g* for 5 min, and the supernatant was incubated with anti-FLAG M1 resin in the presence of $CaCl_2$ (10 mM) for 2 h. Subsequently, the resin was washed 3 times with Tris–HCl buffer (50 mM) at pH 7.4, containing NaCl (150 mM), and purified azido–scFv was eluted by incubating the resin in a Tris–HCl buffer (50 mM) at pH 7.4 containing NaCl (150 mM) and EDTA (2 mM) for 30 min, followed by 6 more incubations in the same buffer for 10 min each.

Huisgen 1,3-dipolar cycloaddition to label azido–sm3E with fluorescein. To a solution of azido–sm3E (20 μL, 4.6 μM) in potassium phosphate (0.10 M) buffer at pH 8.0 were added alkynyl fluorescein **4.10** (1.1 μL, 2.23 mM suspension in 20% v/v ethanol in water), tris(2-carboxyethyl) phosphine hydrochloride (1.0 μL, 50 mM), $CuSO_4 \cdot 5H_2O$ (1.0 μL, 50 mM), and polytriazole ligand (5.0 μL, 20 mM suspension in 80% v/v *t*-butanol in water). The reaction mixture was agitated gently, and incubated at room temperature for 1 h and then at 4 °C for 16 h. The reaction mixtures was suspended in 2× denaturing buffer (20 μL) and subjected to SDS–PAGE. The resulting gel was visualized with a fluorescence imager and was then stained with Coomassie blue.

Huisgen 1,3-dipolar cycloaddition to label azido–anti-EGFR and azido–anti-fluorescein with biotin. To a solution of azido–scFv (100 μL, ~1 μM) in 0.10 M potassium phosphate buffer at pH 8.0 was added alkynyl biotin **AI.1** (5.6 μL,

3.0 mM), tris(2-carboxyethyl) phosphine hydrochloride (5.0 μL, 50 mM), $CuSO_4 \cdot 5H_2O$ (5.0 μL, 50 mM), and polytriazole ligand (25.0 μL, 20 mM suspension in 80% v/v *t*-butanol in water). The reaction mixture was agitated gently, and incubated at 4 °C for 16 h. The same procedure was followed for the control reaction with wild-type scFv. An aliquot of the reaction mixture (10 μL) was suspended in 2× denaturing buffer (10 μL) and subjected to SDS–PAGE. Western blot-analysis was then performed by utilizing a primary antibody against the biotin, and a fluorescent secondary antibody. The presence of protein in both the samples was demonstrated by a western blot-analysis that utilized a primary antibody against a FLAG epitope present at the N-terminus of the scFv, and a fluorescent secondary antibody.

Preparation of alkyne-displaying SAMs of alkane thiols on gold chips. Alkane thiol solutions were prepared by dissolving $HS(CH_2)_{11}(OCH_2CH_2)_6OCH_2CO_2H$ and $HS(CH_2)_{11}(OCH_2CH_2)_3OH$ in ethanol to a final concentration of 0.25 mM each. Gold chips were cleaned under a stream of Ar(g) and immersed in the alkane thiol solution for at least 18 h. After rinsing thoroughly with ethanol and drying under a stream of Ar(g), the chips were overlaid with an aqueous solution containing *N*-hydroxysuccinimide (NHS, 50 mM) and 1-ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC, 200 mM) for 7 min to generate succinimidyl esters in situ. Propargylamine was dissolved in anhydrous DMF to a concentration of 0.20 M. Triethylamine (TEA) was added to the resultant solution to a concentration of 0.20 M, and succinimidyl ester-displaying chips were incubated in this solution for 4 h.

The alkyne-displaying chips thus produced were rinsed with water (2 × 6 mL), and dried under Ar(g).

Immobilization of azido–scFvs by Huisgen 1,3-dipolar cycloaddition. Azido–scFvs (10 μl, ~10 μM in PBS), tris(2-carboxyethyl) phosphine hydrochloride (0.5 μL, 50 mM), $CuSO_4 \cdot 5H_2O$ (0.5 μL, 50 mM), and polytriazole ligand (2.5 μL, 20 mM suspension in 80% v/v *t*-butanol in water) were mixed in a vial, and 1 μl of this mixture was incubated on an alkyne-displaying chip for 24 h at 4 °C. The chip was subsequently rinsed with 25 mM sodium phosphate buffer at pH 7.5 containing Tween-20 (0.1% v/v) (2 × 6 mL), and incubated in the same buffer (1.0 mL) for 48 h to remove protein bound nonspecifically. Non-azido scFv was spotted as a negative control. Finally, the chip was then incubated in a solution (1 mL) of ethanolamine (1.0 M in 25 mM sodium phosphate buffer at pH 8.6), rinsed with 25 mM sodium phosphate buffer at pH 7.5 (2 × 6 mL), and incubated in 25 mM sodium phosphate buffer containing Tween-20 (0.1% v/v) at pH 7.5 for 10 min.

Detection of immobilized scFvs. A chip displaying immobilized scFv was overlaid with anti-FLAG primary antibody (10 μg/mL in 25 mM sodium phosphate buffer at pH 7.5 containing 0.1% (v/v) Tween-20) for 30 min. The chip was then rinsed with 25 mM sodium phosphate buffer (2 × 6 mL) at pH 7.5 containing 0.1% (v/v) Tween-20, and incubated in the same buffer for 15 min. The chip was subsequently overlaid with Alexa Fluor 488-conjugated anti-rabbit secondary antibody (2 μg/mL in 25 mM

sodium phosphate buffer at pH 7.5 containing 0.1% (v/v) Tween-20) for 30 min. Finally, the chip was rinsed with 25 mM sodium phosphate buffer (2 × 6 mL) at pH 7.5 containing 0.1% (v/v) Tween-20, incubated in the same buffer for 15 min, and scanned with a fluorimager.

AI.4 Results and Discussion

Generation of azido–scFvs using an E. coli expression system. Azido–sm3E was produced in *E. coli* by generating sm3E–*mxe* intein–chitin-binding domain fusion protein, and treating sm3E–intein thioesters with azide **4.1**. The fusion protein was expressed as an insoluble protein presumably due to the reducing environment of the *E. coli* cytosol that prevents formation of cystines in the scFv. Therefore, the fusion protein was refolded using the procedure described by Nock and coworkers (Sydor *et al.*, 2002) before loading the cell lysate on a chitin column. The thioester-cleavage reaction was performed by incubating azide **4.1** dissolved in a refolding buffer containing 4 M urea on the chitin column(Figure AI.4C). This result demonstrates the robustness of the azide **4.1** reagent—the α–hydrazino acetamido group of the reagent retained its nucleophilicity in a solution containing a high concentration of salts. The eluted azido–sm3E (band II in Figure AI.4C) was contaminated with uncleaved sm3E–intein–chitin-binding domain fusion protein (band I). Such an impurity was not observed when a similar method was utilized to generate azido–RNase A (Chapter 4). We believe that the presence of 4M urea in the cleavage buffer reduced the strength of the interaction between the chitin-binding domain of the fusion protein and the

chitin resin, resulting in dissociation of the complex. Nevertheless, this contaminant was easily removed by passing the the protein solution over another chitin column (Figure AI.4D).

The integrity of azido–sm3E was verified by mass spectrometry analysis (Figure AI.5A). Upon random immobilization via lysine residues on a succinimidyl ester-displaying chip, the azido-sm3E was found to be able to bind its antigen–the CEA protein (Figure AI.5B). This result demonstrated that derivatization of sm3E by azide **4.1** did not compromise the activity of the scFv, and the refolding process was successful.

Generation of azido–scFvs by yeast surface display. Initial attempts to produce azido–scFvs by treating scFv–intein thioesters displayed on the yeast surface with azide **4.1**, and subsequent analysis by SDS–PAGE resulted in numerous bands on the gel (data not shown). For chemoselective labeling and immobilization of azido–scFvs, it is crucial for the scFv to be the only azide-labeled protein in the protein solution—if several azide-labeled proteins are present, the azido group of all these proteins would react with labeling reagents and surfaces, resulting in heterogeneity. Therefore, to determine whether the α–hydrazino acetamido group of azide **4.1** reacts with yeast proteins and labels them with the azido group, an α–hydrazino acetamido biotin reagent was synthesized. Utilization of this reagent to treat scFv–intein thioester-displaying yeast cells, followed by a western blot analysis for detecting biotin revealed that numerous proteins were labeled by biotin (data not shown). These

results suggested that the α–hydrazino acetamido group of azide **4.1** was not reacting specifically with scFv–intein thioesters, but was also modifying other proteins expressed in yeast. To purify azido–scFvs from contaminating azide-labeled proteins, we introduced a FLAG-tag at the N-terminus of the scFv in the scFv–intein construct, and applied the protein solution to an anti-FLAG resin. A western blot analysis performed after this purification step for azido–anti-EGFR scFv is shown in Figure AI.6C (lane 2). A control experiment in which azide **4.1** was treated with yeast displaying anti-EGFR scFv *without* the intein fused to it yielded no azido–scFv (lane 1).

Production of azido–proteins by yeast surface display extends the scope of our technology to proteins that cannot be expressed in *E. coli*, for example eukaryotic proteins, and proteins that prefer an oxidative cytosolic environment for functional expression. Moreover, yeast displaying different protein–intein fusion proteins can be grown in a 96-well format, and treated with azide **4.1** to yield multiple azido–proteins in a high-throughput fahion. This approach would be more efficient than expressing proteins and purifying them individually (Zhu *et al.*, 2001).

Site-specific labeling of azido–scFvs. We used a chemoselective reaction, Cu(I)-catalyzed Huisgen 1,3-dipolar azide–alkyne cycloaddition (Huisgen, 1963; Kolb and Sharpless, 2003; Speers *et al.*, 2003; van Steenis *et al.*, 2005; Wang *et al.*, 2003b), to probe for the availability of the azide functionality. Azido–sm3E was reacted with alkynyl fluorescein **4.10** in the presence of the Cu(I) catalyst and its polytriazole

ligand (Wang *et al.*, 2003b). The protein was also subjected to SDS–PAGE and visualized by staining with Coomassie blue and fluorescence imaging. The azido-sm3E was found to be fluorescent as a result of the cycloaddition (Figure AI.6A). Azido–anti-fluorescein and azido–anti-EGFR scFvs were treated with alkynyl biotin (**AI.1**) under conditions similar to ones for labeling of azido–sm3E, and western blot analyses were performed to detect biotin. Azido–anti-fluorescein (not shown) and azido–anti-EGFR scFvs were site-specifically labeled (Figure AI.6E). Both these experiments demonstrated that the azido groups installed in scFvs were available for further chemical derivatization.

Site-specific immobilization of azido–scFvs. SAMs of alkane thiols on gold were prepared by incubating gold-coated silicon chips in an ethanolic solution containing equimolar quantities of $HS(CH_2)_{11}(OCH_2CH_2)_6OCH_2CO_2H$ and $HS(CH_2)_{11}(OCH_2CH_2)_3OH$. Subsequently, the chips were overlaid with an aqueous solution of EDC/NHS to produce succinimidyl ester-displaying chips. Succinimidyl ester-displaying chips were treated with propargylamine to generate alkyne-displaying chips. Immobilization of azido–scFvs was performed in presence of tris(2-carboxyethyl) phosphine hydrochloride , $CuSO_4 \cdot 5H_2O$, and polytriazole ligand—standard reagents for Cu(I)-catalyzed Huisgen 1,3-dipolar azide–alkyne cycloaddition. The reaction-time was 24 h at 4 °C. The immobilization proceeded exclusively via the azido group—no significant immobilization of non-azido scFv was observed (Figure AI.7).

Attempts to immobilize azido–scFvs site-specifically by Staudinger ligation on phosphinothioester chips (as reported for immobilization of azido–RNase A in Chapter 5) resulted in nonspecific attachment of the scFvs (data not shown). Alterations in various parameters, such as pH of the protein solution, %DMF in protein solutions, %OH-displaying SAMs on the surface, time of incubation of protein on the surface, and stringent washing of chips did not reduce the nonspecific attachment of scFvs. Presumably, the hydrophobicity of the phosphinothioester restricts its availability in aqueous solutions, consequently preventing Staudinger ligation. A water soluble phosphinothiol reagent can be employed to solve this problem (Tam *et al.*, 2007).

AI.5 Conclusions

The methods described here can be utilized to generate prokaryotic and eukaryotic azido–proteins, and any azido–scFv. The use of yeast surface display extends our method of intein-mediated generation of azido–proteins to the realm of proteins that undergo posttranslational modifications or require oxidative folding for functional expression. Yeast surface display is compatible with a high-throughput format, which augurs well for applications in proteomics, screening, and diagnostics. We demonstrate that azido–proteins are amenable to extremely selective derivatization via reactions of the azido group, such as Staudinger ligation and Huisgen 1,3-dipolar cycloaddition, enabling site-specific labeling with small molecules and oriented immobilization on surfaces.

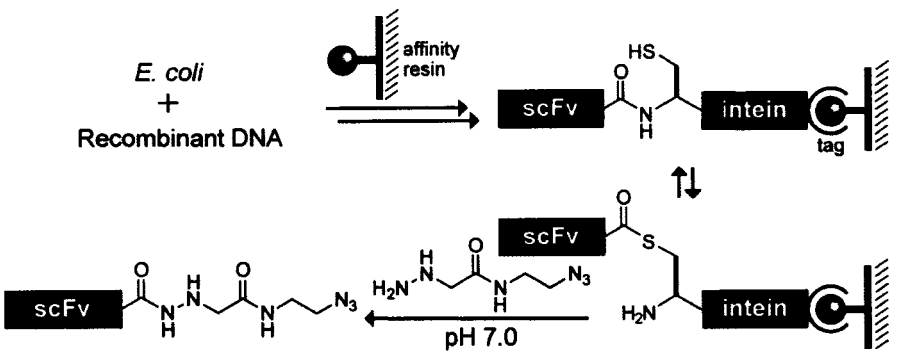


Figure AI.1 Production of azido–scFvs by using an *E. coli* expression system. ScFv–intein–chitin-binding domain fusion proteins are expressed in *E. coli*, the cells are lysed, and the lysate is loaded on a chitin resin. ScFv–intein thioesters displayed on the resin are treated with a bifunctional molecule to generate scFvs appended with the azido group on their C-termini.

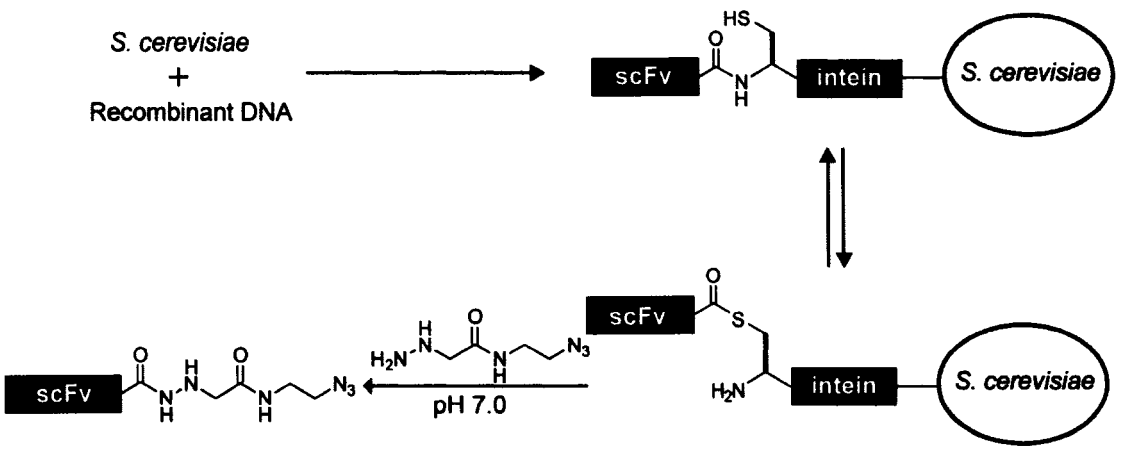


Figure AI.2 Production of azido–scFvs by using yeast surface display. ScFv–intein fusion proteins are displayed on yeast cells, and the cells are treated with a bifunctional molecule that cleaves the thioester to generate scFvs appended with the azido group on their C-termini.

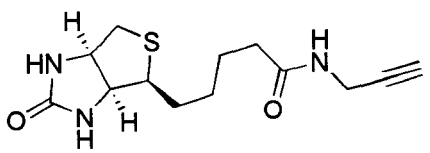
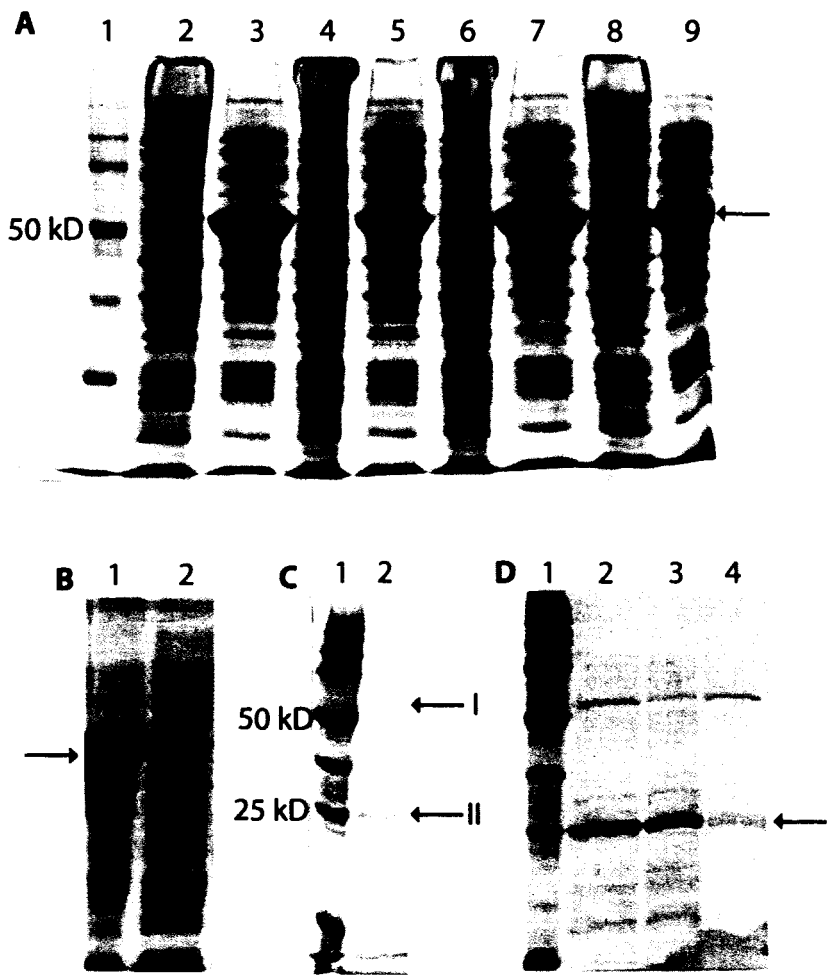


Figure AI.3 Alkynyl biotin **AI.1**.

Figure AI.4 Production of azido–sm3E by using an *E. coli* expression system. Gels were stained with Coomassie blue. Molecular weight markers were loaded in lane 1 of gels shown in panels A, C and D. (A) Test induction of sm3E–*mxe* intein–chitin-binding domain protein in BL21(DE3) *E. coli* cells. Uninduced cell lysates were loaded in lanes 2, 4, 6, and 8, cell lysates of cultures induced with 0.1 mM IPTG at 25 °C and 37 °C were loaded in lanes 3 and 5 respectively, and cell lysates of cultures induced with 0.5 mM IPTG at 25 °C and 37 °C were loaded in lanes 7 and 9 respectively. (B) Expression of sm3E–*mxe* intein–chitin-binding domain protein as an insoluble protein in *E. coli* (lane 1). The supernatant was loaded in lane 2. (C) Elution of azido–sm3E with azide **4.1**. Uncleaved fusion protein (band I) coelutes with azido–sm3E (band II) (D) Removal of fusion protein contaminant from azido–sm3E by application through a chitin column. Protein solution before purification was loaded in lane 2, the purified protein collected as flow-through was loaded in lane 3, and chitin resin after elution was loaded in lane 4.



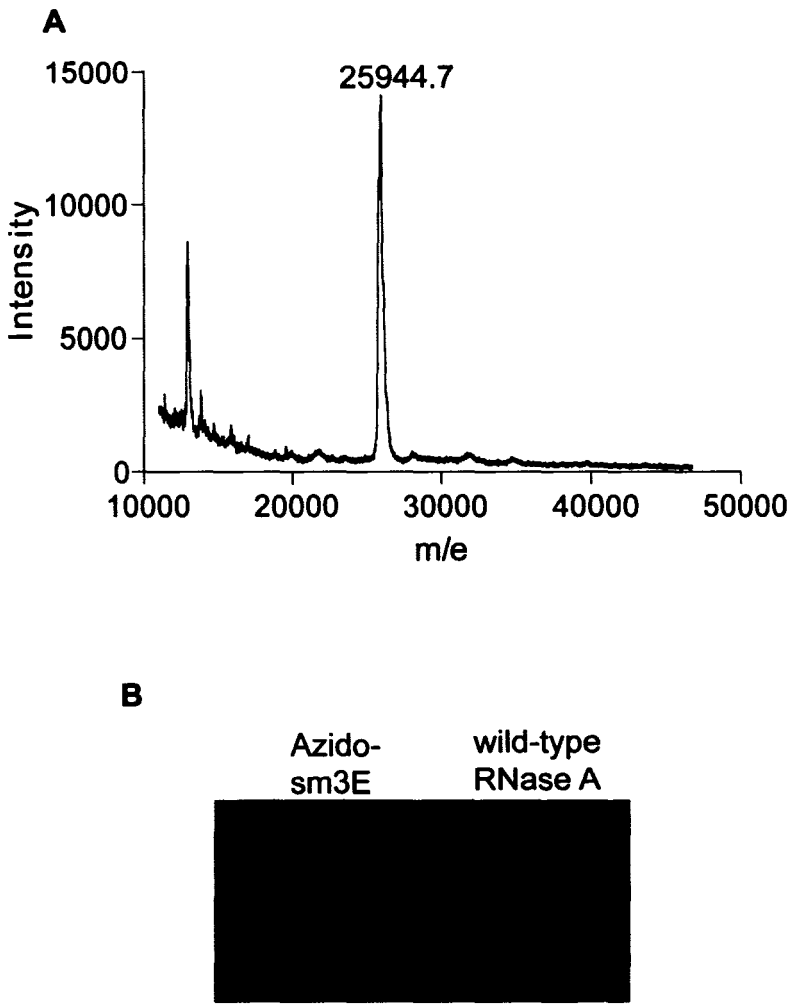
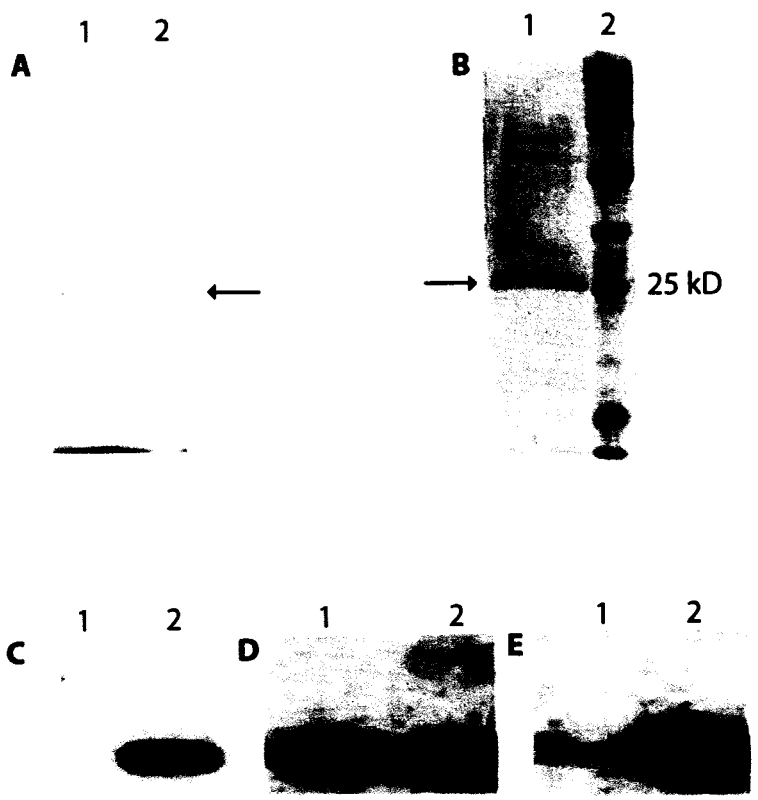


Figure AI.5 Characterization of azido–sm3E. (A) MALDI-TOF mass spectrum of azido–sm3E (expected for Met(–1)sm3E–Ala–Ala–Gly–NHNHCH$_2$C(O)NH CH$_2$CH$_2$N$_3$ = 25955.6). (B) Activity of azido–sm3E. Azido–sm3E was randomly immobilized on succinimidyl ester-displaying chips, and the chip was overlaid with CEA labeled with AlexaFluor 488. Wild-type RNase A was spotted as a negative control.

Figure AI.6 ScFv-labeling by Huisgen 1,3-dipolar cycloaddition. (A) SDS–PAGE analysis of reaction between azido–sm3E and alkynyl fluorescein **4.10**. The reaction mixture was loaded in lane 1, and molecular weight markers were loaded in lane 2. The gel was visualized under a fluorimager. (B) The gel shown in (A) stained with Coomassie blue. (C) Western blot analysis utilizing an anti-FLAG primary antibody demonstrating production of azido–anti-EGFR scFv by yeast surface display. Supernatant of a suspension of yeast cells displaying scFv *without* intein on its surface in azide **4.1** was loaded in lane 1, and supernatant of a suspension of yeast cells displaying scFv–intein fusion protein on its surface in azide **4.1** was loaded in lane 2. The band in lane 2 corresponds to azido–scFv. (D) Western blot analysis utilizing an anti-FLAG primary antibody on a reaction mixture for Huisgen 1,3-dipolar azide–alkyne cycloaddition between azido–anti-fluorescein scFv and alkynyl biotin **AI.1** (lane 2). Non-azido scFv was treated with alkynyl biotin **AI.1** under identical conditions (lane 1) (E) Lane 1 and lane 2 are the same as in (D). Western blot analysis was performed by utilizing an anti-biotin primary antibody.



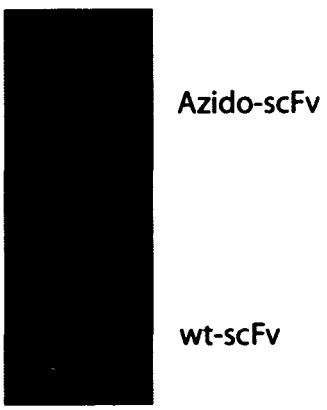


Figure AI.7 Immobilization of azido–scFvs by Huisgen 1,3-dipolar azide–alkyne cycloaddition. Azido–scFv and wild-type scFv were incubated on alkyne-displaying chips for 24 h. The chip was overlaid with primary anti-FLAG, and fluorescent secondary antibody.

Appendix II

Synthesis of Novel Disulfide Coupling Reagents

This work was done in collaboration with Professor Ivan Rayment, Department of Biochemistry at the University of Wisconsin–Madison.

Contribution: Synthesis and characterization of disulfide molecules.

AII.1 Abstract

We have synthesized two disulfide molecules for crosslinking proteins via thiol groups of cysteine side chains. Each molecule consists of two labile disulfides attached via an oligothyleneglycol backbone. These molecules can be utilized to generate dimers of proteins bearing cysteine residues.

AII.2 Experimental Procedures

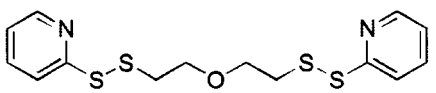
Materials. 2,2′-Dithiopyridine disulfide was purchased from Sigma-Aldrich (St. Louis, MO). Synthetic reactions were monitored by thin-layer chromatography with visualization by UV-light or staining with phosphomolybdic acid. Flash chromatography was performed with columns of silica gel 60, 230–400 mesh (Silicycle, Quebec City, Quebec, Canada).

Instrumentation. NMR spectra were acquired with a Bruker AC+ 300 spectrometer (^{1}H: 300 MHz, ^{13}C: 75 MHz) at the Magnetic Resonance Facility in the Department of Chemistry at Madison. ^{13}C spectra were proton-decoupled. Mass spectra on small organic molecules were obtained with electrospray ionization (ESI) techniques.

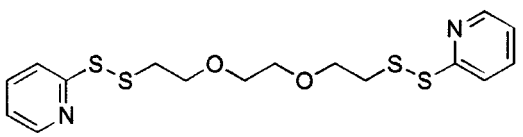
Synthesis of 1,9-bis(2-pyridyl)-1,2,8,9-tetrathia-5-oxanonane **AII.1**. 2,2′-Dithiopyridine disulfide (1.00 g, 4.44 mmol) was dissolved in 20 mL of dichloromethane, and to this solution was added 3-oxa-1,5-disulfanylpentane (0.32 g, 2.22 mmol). The color of the reaction mixture changed to yellow within 2 min after

addition of the dithiol. The reaction mixture was stirred at room temperature for 6 h. Solvent was then removed under reduced pressure, and the residue was purified by flash chromatography (80% v/v hexane in ethyl acetate). Bisdisulfide **AII.1** was obtained as a translucent, yellow liquid in 53% yield. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.48–8.44 (m, 2 H), 7.78–7.73 (m, 2 H), 7.68–7.61 (m, 2 H), 7.11–7.05 (m, 2 H), 3.67 (t, J = 6.3 Hz, 4 H), 2.97 (t, J = 6.3 Hz, 4 H) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 160.5, 149.5, 137.2, 120.8, 119.9, 68.9, 38.7 ppm. MS (ESI) *m/z* 357.2 (MH^+ $[C_{14}H_{17}N_2OS_4]$ = 357.55).

Synthesis of 1,12-bis(2-pyridyl)-1,2,11,12-tetrathia-5,8-dioxadodecane **AII.2**. 2,2′-Dithiopyridine disulfide (0.94 g, 4.18 mmol) was dissolved in 20 mL of dichloromethane, and to this solution was added 3,6-dioxa-1,8-disulfanyloctane (0.40 g, 2.09 mmol). The color of the reaction mixture changed to yellow instantaneously upon addition of the dithiol. The reaction mixture was stirred at room temperature for 6 h. Solvent was then removed under reduced pressure, and the residue was purified by flash chromatography (70% v/v hexane in ethyl acetate). Bisdisulfide **AII.2** was obtained as a translucent, yellow liquid in 33% yield. ^{1}H NMR (300 MHz, $CDCl_3$) δ 8.47–8.43 (m, 2 H), 7.79–7.75 (m, 2 H), 7.68–7.62 (m, 2 H), 7.11–7.05 (m, 2 H), 3.73 (t, 4 H, 6.5 Hz), 3.56 (m, 4 H), 2.99 (t, 4 H, 6.4 Hz) ppm; ^{13}C NMR (75 MHz, $CDCl_3$) δ 160.6, 149.7, 137.2, 120.8, 119.8, 70.5, 69.3, 38.6 ppm; MS (ESI) *m/z* 401.2 (MH^+ $[C_{16}H_{21}N_2O_2S_4]$ = 401.6).



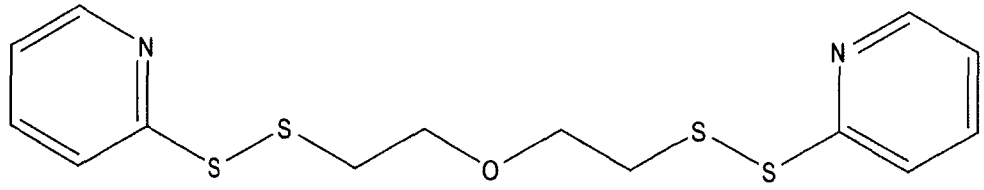
AII.1



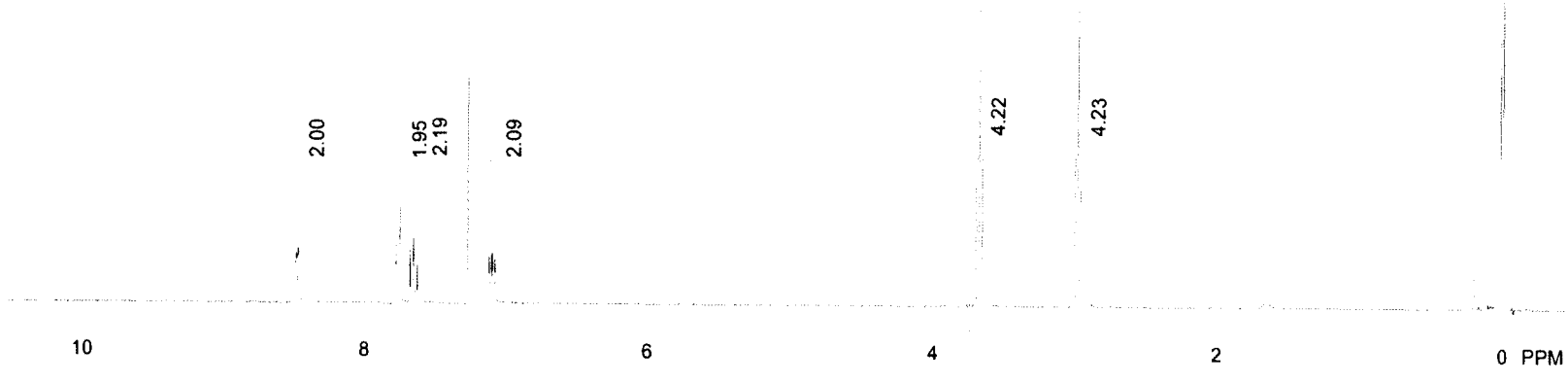
AII.2

Figure AII.1 Disulfide coupling reagents.

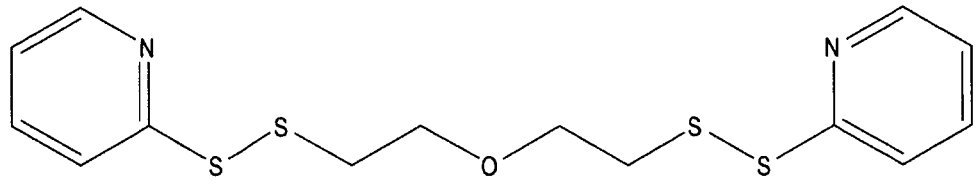
AII.1



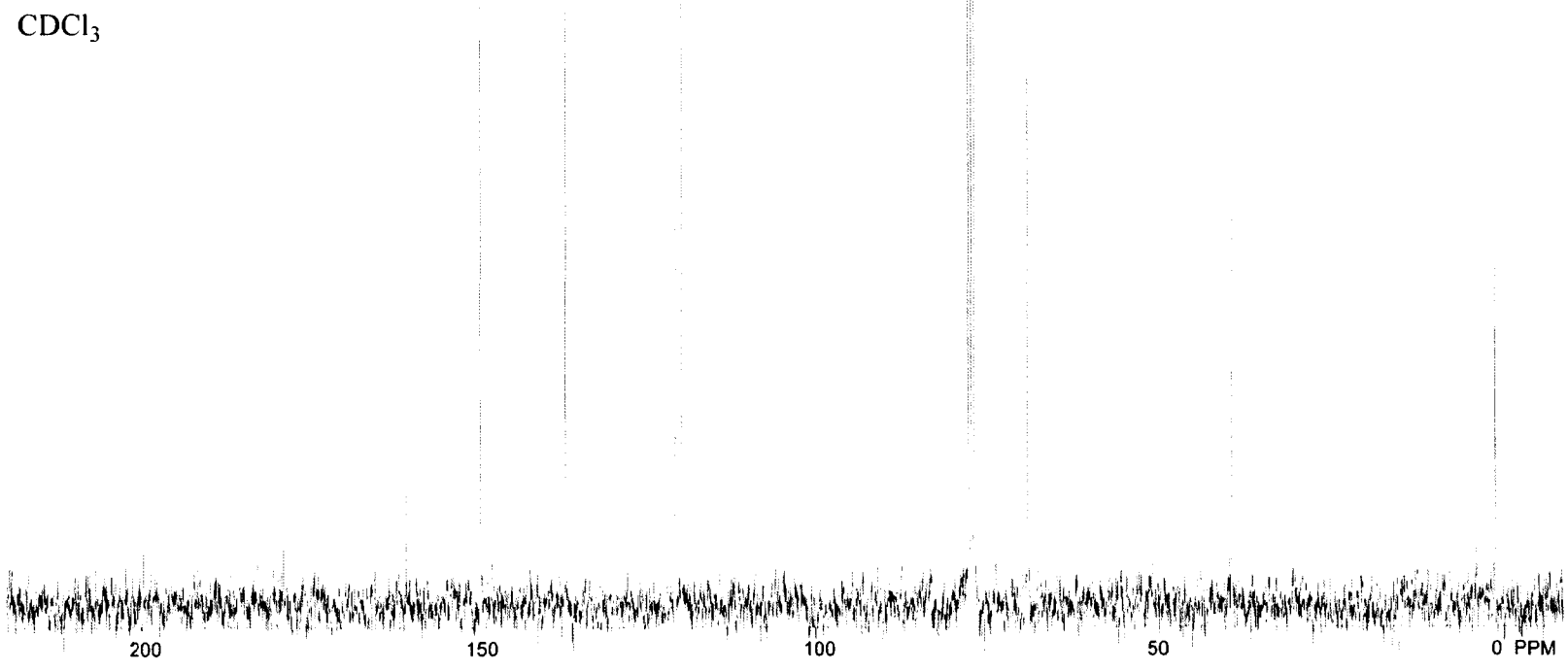
Proton NMR, 300 MHz
$CDCl_3$



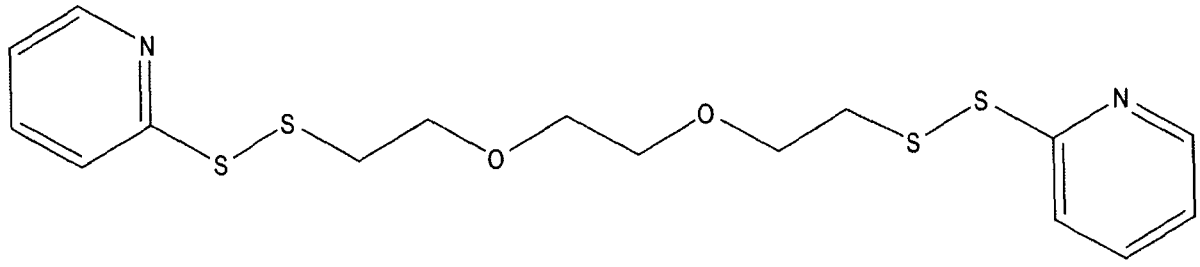
AII.1



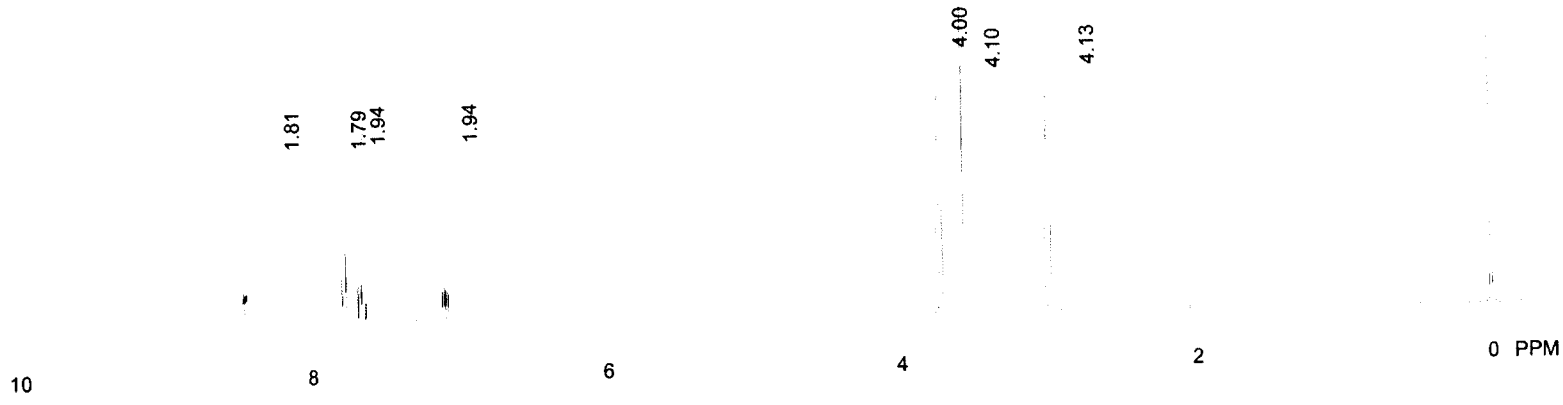
Carbon-13 NMR, 75 MHz
$CDCl_3$



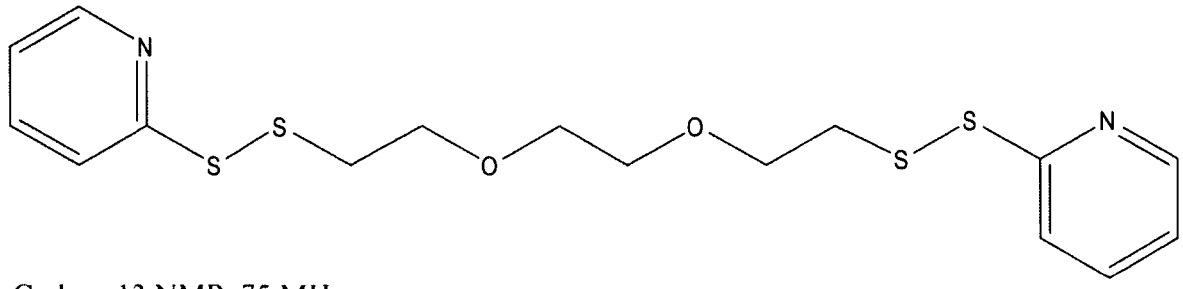
AII.2



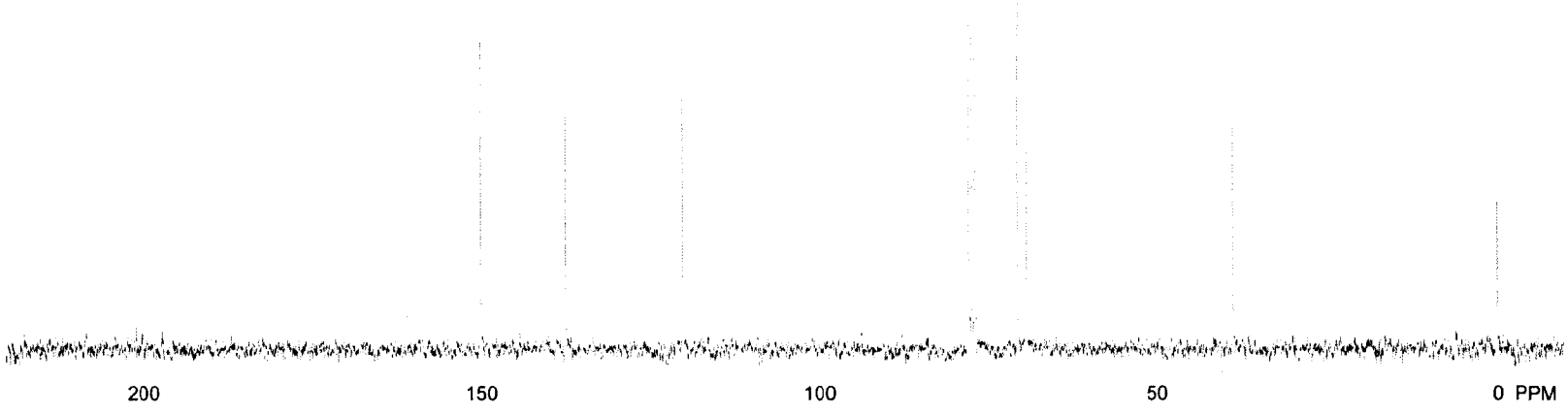
Carbon-13 NMR, 75 MHz
$CDCl_3$



AII.2



Carbon-13 NMR, 75 MHz
$CDCl_3$



References

Abel, R. L., Haigis, M. C., Park, C., and Raines, R. T. (2002). Fluorescence assay for binding of ribonuclease A to the ribonuclease inhibitor protein. *Anal. Biochem.* **306,** 100-107.

Agard, N. J., Baskin, J. M., Prescher, J. A., Lo, A., and Bertozzi, C. R. (2006). A comparative study of bioorthogonal reactions with azides. *ACS Chem. Biol.* **1,** 644-648.

Agard, N. J., Prescher, J. A., and Bertozzi, C. R. (2004). A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J. Am. Chem. Soc.* **126,** 15046-15047.

Ahn, B.-T., Park, H.-S., Lee, E.-J., and Um, I.-H. (2000). Catalytic effect of MO_4^{2-} (M=Cr, Mo and W) on hydrolyses of carbon and phosphorus esters. *Bull. Korean Chem. Soc.* **21,** 905-908.

Allen, M. J., Raines, R. T., and Kiessling, L. L. (2006). Contrast agents for magnetic resonance imaging synthesized with ring-opening metathesis polymerization. *J. Am. Chem. Soc.* **128,** 6534-5.

Anderson, B. M., and Jencks, W. P. (1960). The effect of structure on reactivity in semicarbazone formation. *J. Am. Chem. Soc.* **82,** 1773-1777.

Antosiewicz, J., McCammon, J. A., and Gilson, M. K. (1994). Prediction of pH-dependent properties of proteins. *J. Mol. Biol.* **238,** 415-436.

Ardagh, E. G. R., and Rutherford, F. C. (1935). Studies from some hydrazone and osazone reactions. *J. Am. Chem. Soc.* **57,** 1085-1088.

Armstrong, N., Jasti, J., Beich-Frandsen, M., and Gouaux, E. (2006). Measurement of conformational changes accompanying desensitization in an ionotropic glutamate receptor. *Cell* **127,** 85-97.

Arnold, U., Hinderaker, M. P., Köditz, J., Golbik, R., Ulbrich-Hoffmann, R., and Raines, R. T. (2003). Protein prosthesis: A nonnatural residue accelerates folding and increases stability. *J. Am. Chem. Soc.* **125,** 7500-7501.

Arnold, U., Hinderaker, M. P., Nilsson, B. L., Huck, B. R., Gellman, S. H., and Raines, R. T. (2002). Protein prosthesis: A semisynthetic enzyme with a β-peptide reverse turn. *J. Am. Chem. Soc.* **124,** 8522-8523.

Aslam, M., and Dent, A. (1998). "Bioconjugation: Protein Coupling Techniques for the Biomedical Sciences." Macmillan Reference Ltd., London.

Aslan, K., Lakowicz, J. R., and Geddes, C. D. (2005). Plasmon light scattering in biology and medicine: New sensing approaches, visions and perspectives. *Curr. Opin. Chem. Biol.* **9,** 538-544.

Back, J. W., David, O., Kramer, G., Masson, G., Kasper, P. T., de Koning, L. J., de Jong, L., van Maarseveen, J. H., and de Koster, C. G. (2005). Mild and chemoselective peptide-bond cleavage of peptides and proteins at azido homoalanine. *Angew. Chem. Int. Ed. Engl.* **44,** 7946-7950.

Baskin, J. M., Prescher, J. A., Laughlin, S. T., Agard, N. J., Chang, P. V., Miller, I. A., Lo, A., Codelli, J. A., and Bertozzi, C. R. (2007). Copper-free click chemistry for dynamic *in vivo* imaging. *Proc. Natl. Acad. Sci. U.S.A.* **104,** 16793-16797.

Bednar, R. A. (1990). Reactivity and pH dependence of thiol conjugation to *N*-ethylmaleimide: Detection of a conformational change in chalcone isomerase. *Biochemistry* **29,** 3684-3690.

Begent, R. H., Verhaar, M. J., Chester, K. A., Casey, J. L., Green, A. J., Napier, M. P., Hope-Stone, L. D., Cushen, N., Keep, P. A., Johnson, C. J., Hawkins, R. E., Hilson, A. J., and Robson, L. (1996). Clinical evidence of efficient tumor targeting based on single-chain Fv antibody selected from a combinatorial library. *Nat. Med.* **2,** 979-984.

Bertics, P. J., Ozer, B. H., Wiepz, G. J., Guadarrama, A. G., Abbott, N. L., Lowe, A. M., Bai, Y. B., Shusta, E. V., Agarwal, N., Raines, R. T., and Kalia, J. (2007).

Liquid crystal-based analytic technology: Enabling a molecular view of cancer. *BIOforum Europe* **11,** 42-45.

Bird, R. E., Hardman, K. D., Jacobson, J. W., Johnson, S., Kaufman, B. M., Lee, S., Lee, T., Pope, S. H., Riordan, G. S., and Whitlow, M. (1988). Single-chain antigen binding proteins. *Science* **242,** 423-426.

Boder, E. T., and Wittrup, K. D. (1997). Yeast surface display for screening combinatorial polypeptide libraries. *Nat. Biotech.* **15,** 553-555.

Bokinsky, G., and Zhuang, X. (2005). Single-molecule RNA folding. *Acc. Chem. Res.* **38,** 566-73.

Bonnet, D., Ollivier, N., Gras-Masse, H., and Melnyk, O. (2001). Chemoselective acylation of fully deprotected hydrazino acetyl peptides. Application to the synthesis of lipopeptides. *J. Org. Chem.* **66,** 443-449.

Borodovsky, A., Ovaa, H., Kolli, N., Gan-Erdene, T., Wilkinson, K. D., Ploegh, H. L., and Kessler, B. M. (2002). Chemistry-based functional proteomics reveals novel members of the deubiquitinating enzyme family. *Chem. Biol.* **9,** 1149-1159.

Boucher, D., Cournoyer, D., Stanners, C. P., and Fuks, A. (1989). Studies on the control of gene expression of the carcinoembryonic antigen family in human tissue. *Cancer Res.* **49,** 847-852.

Bräse, S., Gil, C., Knepper, K., and Zimmermann, V. (2005). Organic azides: An exploding diversity of a unique class of compounds. *Angew. Chem. Int. Ed. Engl.* **44,** 5188-5240.

Brown, P. O., and Botstein, D. (1999). Exploring the new world of the genome with DNA microarrays. *Nat. Genetics* **21.**

Bruice, T. C., Bruno, J. J., and Chou, W.-S. (1963). Nucleophilic displacement reactions at the thiol-ester dond of δ-thiolvalerolactone. *J. Am. Chem. Soc.* **85,** 1659-1669.

Bruice, T. C., Donzel, A., Huffman, R. W., and Butler, A. R. (1967). Aminolysis of phenyl acetates in aqueous solutions. VII.[1] Observations on the influence of salts, amine structure, and base strength. *J. Am. Chem. Soc.* **89,** 2106-2121.

Bruice, T. C., and Fedor, L. R. (1964a). The kinetic demonstration of a metastable intermediate in a nucleophilic displacement at a thiol-ester bond. *J. Am. Chem. Soc.* **86,** 738-739.

Bruice, T. C., and Fedor, L. R. (1964b). Nucleophilic displacement reactions at the thiolester bond. III. Kinetic demonstration of metastable intermediates in the hydroxylaminolysis and methoxylaminolysis of thiolesters and thiolactones in aqueous solutions. *J. Am. Chem. Soc.* **86,** 4886-4897.

Bruice, T. C., and Fedor, L. R. (1964c). O- *vs*. N-Attack of hydroxylamine on *n*-butyl thiolacetate and the establishment that N-attack passes through a metastable intermediate. *J. Am. Chem. Soc.* **86,** 739-740.

Buncel, E., and Um, I.-H. (2004). The α-effect and its modulation by solvent. *Tetrahedron* **60,** 7801-7825.

Burtis, C. A., and Ashwood, E. R. (1994). "Tietz Textbook of Clinical Chemistry." W. B. Saunders Co., Philadelphia, PA.

Burton, D. R. (1993). Monoclonal antibodies from combinatorial libraries. *Acc. Chem. Res.* **26,** 405-411.

Butler, J. E., Ni, L., Nessler, R., Joshi, K. S., Suter, M., Rosenberg, B., Chang, J., Brown, W. R., and Cantarero, L. A. (1992). The physical and functional behavior of capture antibodies adsorbed on polystyrene. *J. Immuol. Methods* **150,** 77-90.

Camarero, J. A., Kwon, Y., and Coleman, M. A. (2004). Chemoselective attachment of biologically active proteins to surfaces by expressed protein ligation and its application for "protein chip" fabrication. *J. Am. Chem. Soc.* **126,** 14730-14731.

Canne, L. E., Ferre-D'Amare, A. R., Burley, S. K., and Kent, S. B. H. (1995). Total chemical synthesis of a unique transcription factor-related protein: cMyc-Max. *J. Am. Chem. Soc.* **117,** 2998-3007.

Caravan, P. (2006). Strategies for increasing the sensitivity of Gadolinium based MRI contrast agents. *Chem. Soc. Rev.* **35,** 512-523.

Carey, F. A., and Sundberg, R. J. (2008). "Advanced Organic Chemistry." pp. 650-651, Springer, New York.

Caruthers, M. H. (1991). Chemical synthesis of DNA and DNA analogues. *Acc. Chem. Res.* **24,** 278-284.

Castro, E. A. (1999). Kinetics and mechanisms of reactions of thiol, thiono, and dithio analogues of carboxylic esters with nucleophiles. *Chem. Rev.* **99,** 3505-3524.

Castro, E. A., and Ureta, C. (1989). Structure–reactivity correlations in the aminolysis of phenyl and *p*-nitrophenyl thiolacetates. *J. Org. Chem.* **54,** 2153-2159.

Cha, T., Guo, A., and Zhu, X. Y. (2005). Enzymatic activity on a chip: The critical role of protein orientation. *Proteomics* **5,** 416-419.

Chapman, A. P. (2002). PEGylated antibodies and antibody fragments for improved therapy: a review. *Adv. Drug. Deliv. Rev.* **54,** 531-545.

Chibata, I., Tosa, T., Sato, T., and Mori, T. (1979). "Immobilized Enzymes: Research and Development." John Wiley & Sons, New York.

Chou, T. C., and Lipmann, F. (1952). Separation of acetyl transfer enzymes in pigeon liver extract. *J. Biol. Chem.* **196,** 89-103.

Clark, S. (2003). Ageing as war between chemical and biochemical processes: Protein methylation and the recognition of age-damaged proteins for repair. *Ageing Res. Rev.* **2,** 263-285.

Conant, J. B., and Bartlett, P. D. (1932). A quantitative study of semicarbazone formation. *J. Am. Chem. Soc.* **54,** 2881-2899.

Connett, P. H., and Wetterhahn, K. E. (1985). *In vitro* reaction of the carcinogen chromate with cellular thiols and carboxylic acids. *J. Am. Chem. Soc.* **107,** 4282-4288.

Connors, K. A., and Bender, M. L. (1961). The kinetics of alkaline hydrolysis and *n*-butylaminolysis of ethyl *p*-nitrobenzoate and ethyl *p*-nitrothiolbenzoate. *J. Org. Chem.* **26,** 2498-2504.

Cordes, E. H., and Jencks, W. P. (1962a). General acid catalysis of semicarbazone formation. *J. Am. Chem. Soc.* **84,** 4319-4328.

Cordes, E. H., and Jencks, W. P. (1962b). Nucleophilic catalysis of semicarbazone formation by anilines. *J. Am. Chem. Soc.* **84,** 826-831.

Cordes, E. H., and Jencks, W. P. (1962c). On the mechanism of Schiff base formation and hydrolysis. *J. Am. Chem. Soc.* **84,** 832-837.

Cordes, E. H., and Jencks, W. P. (1963). The mechanism of hydrolysis of Schiff bases derived from aliphatic amines. *J. Am. Chem. Soc.* **85,** 2843-2848.

Cornish, P. V., and Ha, T. (2007). A survey of single-molecule techniques in chemical biology. *ACS Chem. Biol.* **2,** 53-61.

Cottingham, I. R., Millar, A., Emslie, E., Colman, A., Schnieke, A. E., and McKee, C. (2001). A method for the amidation of recombinant peptides expressd as intein fusion proteins in *Escherichia coli*. *Nat. Biotechnol.* **19,** 974-977.

Covington, A. K., Pabbo, M., Robinson, R. A., and Bates, R. G. (1968). Use of glass electrode in deuterium oxide and the relation betweeen the standardized pD (paD) scale and operational pH in heavy water. *Anal. Chem.* **40,** 700-706.

Cretich, M., Damin, F., Pirri, G., and Chiari, M. (2006). Protein and peptide arrays: Recent trends and new directions. *Biomol. Eng.* **23,** 77-88.

Crick, F. H. (1958). On protein synthesis. *Symp. Soc. Exp. Biol.* **12,** 139-163.

Crick, F. H. (1970). Central dogma of molecular biology. *Nature* **227,** 561-563.

Danehy, J. P., and Parameswaran, K. N. (1968). Acidic dissociation constants of thiols. *J. Chem. Eng. Data* **13,** 386-389.

Dawson, P. E., and Kent, S. B. H. (2000). Synthesis of native proteins by chemical ligation. *Annu. Rev. Biochem.* **69,** 923-960.

Dawson, P. E., Muir, T. W., Clarklewis, I., and Kent, S. B. H. (1994). Synthesis of proteins by native chemical ligation. *Science* **266,** 776-779.

Dickson, K. A., Haigis, M. C., and Raines, R. T. (2005). Ribonuclease Inhibitor: Structure and function. *Prog. Nucleic Acid Res. Mol. Biol.* **80,** 349-374.

Dirksen, A., Dirksen, S., Hackeng, T. M., and Dawson, P. E. (2006a). Nucleophilic catalysis of hydrazone formation and transimination: implications for dynamic covalent chemistry. *J. Am. Chem. Soc.* **128,** 15602-15603.

Dirksen, A., Hackeng, T. M., and Dawson, P. E. (2006b). Nucleophilic catalysis of oxime ligation. *Angew. Chem. Int. Ed. Engl.* **45,** 7581-7584.

Du, Y.-Z., and Saavedra, S. S. (2003). Molecular orientation distributions in protein films. V. Cytochrome c adsorbed to a sulfonate-terminated self-assembled monolayer. *Langmuir* **19,** 6443-6448.

Duburcq, X., Olivier, C., Malingue, F., Desmet, R., Bouzidi, A., Zhou, F., Auriault, C., Gras-Masse, H., and Melnyk, O. (2004). Peptide-pathogen microarrays for the simultaneous detection of pathogen infections. *Bioconjugate Chem.* **15,** 307-316.

Duckworth, B. P., Xu, J., Taton, T. A., Guo, A., and Distefano, M. D. (2006). Site-specific, covalent attachment of proteins to a solid surface. *Bioconjugate Chem.* **17,** 967-974.

Ducoux, M., Urbach, S., Baldacci, G., Hubscher, U., Koundrioukoff, S., Christensen, J., and Hughes, P. (2001). Mediation of proliferating cell nuclear antigen (PCNA)-dependent DNA replication through a conserved p21(Cip1)-like PCNA-binding motif present in the third subunit of human DNA polymerase delta. *J. Biol. Chem.* **276,** 49258-66.

Evans, T. C., Benner, J., and Xu, M. Q. (1998). Semisynthesis of cytotoxic proteins using a modified protein splicing element. *Protein Sci.* **7,** 2256-2264.

Evans, T. C., Jr., and Xu, M.-Q. (2002). Mechanistic and kinetic considerations of protein splicing. *Chem. Rev.* **102,** 4869-4883.

Falsey, J. R., Renil, M., Park, S., Li, S., and Lam, K. S. (2001). Peptide and small molecule microarray for high throughput cell adhesion and functional assays. *Bioconjugate Chem.* **12,** 346-353.

Francisco, J. A., Campbell, R., Iverson, B. L., and Georgiou, G. (1993). Production and fluorescence-activated cell sorting of *Escherichia coli* expressing a functional antibody fragment on the external surface. *Proc. Natl. Acad. Sci. U.S.A.* **90,** 10444-10448.

Friedman, M. (1999). Lysinoalanine in food and in antimicrobial proteins. *Adv. Exp. Med. Biol.* **459,** 145-159.

Friedmann, E. (1952). Spectrophotometric investigation of the interaction of glutathione with maleimide and N-ethylmaleimide. *Biochim. et Biophys. Acta.* **9,** 65-75.

Fu, G. C. (2004). Asymmetric catalysis with "planar-chiral" derivatives of 4-(dimethylamino)pyridine. *Acc. Chem. Res.* **37,** 542-547.

Fuchs, S. M., and Raines, R. T. (2006). Internalization of cationic peptides: The road less (or more?) traveled. *Cell. Mol. Life Sci.* **63,** 1819-1822.

Gai, S. A., and Wittrup, K. D. (2007). Yeast surface display for protein engineering and characterization. *Curr. Opin. Struct. Biol.* **17,** 467-473.

Gama, C. I., Tully, S. E., Sotogaku, N., Clark, P. M., Rawat, M., Vaidehi, N., Goddard, W. A., 3rd, Nishi, A., and Hsieh-Wilson, L. C. (2006). Sulfation patterns of glycosaminoglycans encode molecular recognition and activity. *Nat. Chem. Biol.* **2,** 467-473.

Gauchet, C., Labadie, G. R., and Poulter, C. D. (2006). Regio- and chemoselective covalent immobilization of proteins through unnatural amino acids. *J. Am. Chem. Soc.* **128,** 9274-9275.

Georgiou, G., Stathopoulos, C., Daugherty, P. S., Nayak, A. R., Iverson, B. L., and Curtiss, R., 3rd. (1997). Display of heterologous proteins on the surface of microorganisms: From the screening of combinatorial libraries to live recombinant vaccines. *Nat. Biotech.* **15,** 29-34.

Gierlich, J., Burley, G. A., Gramlich, P. M., Hammond, D. M., and Carell, T. (2006). Click chemistry as a reliable method for the high-density postsynthetic functionalization of alkyne-modified DNA. *Org. Lett.* **8,** 3639-3642.

Gololobov, Y. G., and Kasukhin, L. F. (1992). Recent advances in the Staudinger reaction. *Tetrahedron* **48,** 1353-1406.

Graff, C. P., Chester, K., Begent, R., and Wittrup, K. D. (2004). Directed evolution of an anti-carcinoembryonic antigen scFv with a 4-day monovalent dissociation half-time at 37 degrees C. *Protein Eng. Des. Select.* **17,** 293-304.

Gramlich, P. M., Warncke, S., Gierlich, J., and Carell, T. (2008). Click-click-click: Single to triple modification of DNA. *Angew. Chem. Int. Ed. Engl.* **47,** 3442-3444.

Grandjean, C., Rommens, C., Gras-Masse, H., and Melnyk, O. (2000). One-pot synthesis of antigen-bearing lysine-based cluster mannosides using two orthogonal chemoselective ligation reactions. *Angew. Chem. Int. Ed.* **39,** 1068-1072.

Grandjean, C., Santraine, V., Fruchart, J.-S., Melnyk, O., and Gras-Masse, H. (2002). Combined thioether/hydrazone chemoselective ligation reactions for the synthesis of glycocluster-antigen peptide conjugates. *Bioconjugate Chem.* **13,** 887-892.

Green, R. J., Frazier, R. A., Shakesheff, K. M., Davies, M. C., Roberts, C. J., and Tendler, S. J. (2000). Surface plasmon resonance analysis of dynamic biological interactions with biomaterials. *Biomaterials* **21,** 1823-1835.

Gregory, J. D. (1955). The stability of N-ethylmaleimide and its reaction with sulfhydryl groups. *J. Am. Chem. Soc.* **77,** 3922-3923.

Gregory, M. J., and Bruice, T. C. (1967). Nucleophilic displacement reactions at the thiol ester bond. V. Reactions of 2,2,2-trifluoroethyl thiolacetate. *J. Am. Chem. Soc.* **89,** 2121-2127.

Gresser, M. J., and Jencks, W. P. (1977). Ester aminolysis. Structure–reactivity relationships and the rate-determining step in the aminolysis of substituted diphenyl carbonates. *J. Am. Chem. Soc.* **77,** 6963-6970.

Gupta, V. K., Skaife, J. J., Dubrovsky, T. B., and Abbott, N. L. (1998). Optical amplification of ligand-receptor binding using liquid crystals. *Science* **279,** 2077-2080.

Gurd, F. R. N. (1972). Carboxymethylation. *Methods Enzymol.* **25,** 424-438.

Ha, T., Enderle, T., F., O. D., Chemla, D. S., Selvin, P. R., and Weiss, S. (1996). Probing the interaction between two single molecules: Fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl. Acad. Sci. U.S.A* **93,** 6264-6268.

Hall, H. K. (1957). Correlation of the base strengths of amines. *J. Am. Chem. Soc.* **79,** 5441-5444.

Hang, H. C., and Bertozzi, C. R. (2001). Chemoselective approaches to protein glycoassembly. *Acc. Chem. Res.* **34,** 727-736.

Hartmann-Petersen, R., and Gordon, C. (2005). Quantifying protein-protein interactions in the ubiquitin pathway by surface plasmon resonance. *Methods. Enzymol.* **399,** 164-177.

Hawkins, R. E., Russell, S. J., and Winter, G. (1992). Selection of phage antibodies by binding affinity. Mimicking affinity maturation. *J. Mol. Biol.* **226,** 889-896.

Heller, A., and Feldman, B. (in press). Electrochemical glucose sensors and their applications in diabetes management. *Chem. Rev.*

Hinman, R. L. (1958). Notes—Base strengths of some alkylhydrazines. *J. Org. Chem.* **23,** 1587-1588.

Hoheisel, J. D. (2006). Microarray technology: Beyond transcript profiling and genotype analysis. *Nat. Rev. Genet.*

Holohan, K. N., Murphy, R. F., Flanagan, R. W., Buchanan, K. D., and Elmore, D. T. (1973). Enzymic iodination of the histidyl residue of secretin: A radioimmunoassay of the hormone. *Biochim. Biophys. Acta.* **322,** 178-180.

Hondal, R. J., Nilsson, B. L., and Raines, R. T. (2001). Selenocysteine in native chemical ligation and expressed protein ligation. *J. Am. Chem. Soc.* **123,** 5140-5141.

Houseman, B. T., Gawalt, E. S., and Mrksich, M. (2002a). Maleimide-functionalized self-assembled monolayers for the preparation of peptide and carbohydrate biochips. *Langmuir* **19,** 1522-1531.

Houseman, B. T., Huh, J. H., Kron, S. J., and Mrksich, M. (2002b). Peptide chips for the quantitative evalution of protein kinase activity. *Nat. Biotechnol.* **20,** 270-274.

Houseman, B. T., and Mrksich, M. (2002). Carbohydrate arrays for the evaluation of protein binding and enzymatic modification. *Chem. Biol.* **9,** 443-454.

Hu, Y., Uttamchandani, M., and Yao, S. Q. (2006). Microarray: A versatile platform for high-throughput functional proteomics. *Comb. Chem. High Throughput Screen.* **9,** 203-212.

Huisgen, R. (1963). 1,3-Dipolar cycloadditions. Past and future. *Angew. Chem. Int. Ed. Engl.* **2,** 565-598.

Hupe, D. J., and Jencks, W. P. (1977). Nonlinear structure-reactivity correlations. Acyl transfer between sulfur and oxygen nucleophiles. *J. Am. Chem. Soc.* **99,** 451-464.

Huston, J. S., Mudhett-Hunter, M., Mei-Sheng, T., McCartney, J., Warren, F., Haber, E., and Oppermann, H. (1991). Protein engineering of single-chain Fv analogs and fusion proteins. *Methods Enzymol.* **203,** 46-88.

Ishii, Y., and Lehrer, S. S. (1986). Effects of the state of the succinimido-ring on the fluorescence and structural properties of pyrene maleimide-labeled aa-Tropomyosin. *Biophys. J.* **50,** 75-80.

Jang, C.-H., Cheng, L.-L., Olsen, C. W., and Abbott, N. L. (2006). Anchoring of nematic liquid crystals on viruses with different envelope structures. *Nano Lett.* **6,** 1053-1058.

Jannatipour, M., Dion, P., Khan, S., Jindal, H., Fan, X., Laganiere, J., Chishti, A. H., and Rouleau, G. A. (2001). Schwannomin isoform-1 interacts with syntenin via PDZ domains. *J. Biol. Chem.* **276,** 33093-100.

Jares-Erijman, E. A., and Jovin, T. M. (2003). FRET imaging. *Nat. Biotech.* **21,** 1387-1395.

Jencks, W. P. (1958a). The reaction of hydroxylamine with activated acyl groups. I. Formation of O-acylhydroxylamine. *J. Am. Chem. Soc.* **80,** 4581-4584.

Jencks, W. P. (1958b). The reaction of hydroxylamine with activated acyl groups. II. Mechanism of the reaction. *J. Am. Chem. Soc.* **80,** 4585-4588.

Jencks, W. P. (1959). Studies on the mechanism of oxime and semicarbazone formation. *J. Am. Chem. Soc.* **81,** 475-481.

Jencks, W. P. (1964). Mechanism and catalysis of simple carbonyl group reactions. *Prog. Phys. Org. Chem.* **2,** 63-128.

Jencks, W. P., and Carriuolo, J. (1960). Reactivity of nucleophilic reagents towards esters. *J. Am. Chem. Soc.* **82,** 1778-1786.

Jencks, W. P., and Gilchrist, M. (1968). Nonlinear structure–reactivity correlations. The reactivity of nucleophilic reagents toward esters. *J. Am. Chem. Soc.* **90,** 2622-2637.

Jiang, Y., Ruta, V., Chen, J., Lee, A., and MacKinnon, R. (2003). The principle of gating charge movement in a voltage-dependent K+ channel. *Nature* **423,** 42-48.

Johnson, R. J., Chao, T.-Y., Lavis, L. D., and Raines, R. T. (2007). Cytotoxic ribonucleases: Dichotomy of Coulombic forces. *Biochemistry* **46,** 10308-10316.

Johnson, R. J., and Stieglitz, J. (1934). The velocity of hydrolysis of stereoisomeric hydrazones and oximes. *J. Am. Chem. Soc.* **56,** 1904-1908.

Kalia, J., Abbott, N. L., and Raines, R. T. (2007). General method for site-specific protein immobilization by Staudinger ligation. *Bioconjugate Chem.* **18,** 1064-1069.

Kalia, J., and Raines, R. T. (2006). Reactivity of intein thioesters: Appending a functional group to a protein. *Chembiochem* **7,** 1375-1383.

Kambhampati, D. (2004). "Protein microarray technology." Wiley-VCH, Weinheim, Germany.

Kent, S. (2003). Total chemical synthesis of enzymes. *J. Pept. Sci.* **9,** 574-593.

Kersteen, E. A., Barrows, S. R., and Raines, R. T. (2005). Catalysis of protein disulfide bond isomerization in a homogeneous substrate. *Biochemistry* **44,** 12168-78.

Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002). Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 19-24.

Kingsmore, S. F. (2006). Multiplexed protein measurement: Technologies and applications of protein and antibody arrays. *Nat. Rev. Drug Discov.* **5,** 310-320.

Kluger, R., and Hunt, J. C. (1984). Aminolysis of maleic anhydride. Kinetics and thermodynamics of amine formation. *J. Am. Chem. Soc.* **106,** 5667-5670.

Knowles, J. R. (1987). Tinkering with enzymes: What are we learning? *Science* **236,** 1252-1258.

Kobayashi, M., Nagasawa, T., and Yamada, H. (1992). Enzymatic synthesis of acrylamide: A success story not yet over. *Trends Biotechnol.* **10,** 402-408.

Kochendoerfer, G. G., Chen, S.-Y., Mao, F., Cressman, S., Traviglia, S., Shao, H., Hunter, C. L., Low, D. W., Cagle, E. N., Carnevali, M., Gueriguian, V., Keogh, P. J., Porter, H., Stratton, S. M., Wiedeke, M. C., Wilken, J., Tang, J., Levy, J. J., Miranda, L. P., Crnogorac, M. M., Kalbag, S., Botti, P., Schindler-Horvat, J., Savatski, L., Adamson, J. W., Kung, A., Kent, S. B. H., and Bradburne, J. A. (2003). Design and chemical synthesis of a homogeneous polymer-modified erythropoiesis protein. *Science* **299,** 884-887.

Köhn, M., and Breinbauer, R. (2004). The Staudinger ligation—A gift to chemical biology. *Angew. Chem. Int. Ed. Engl.* **43,** 3106-3116.

Köhn, M., Wacker, R., Peters, C., Schröder, H., Soulère, L., Breinbauer, R., Niemeyer, C. M., and Waldmann, H. (2003). Staudinger ligation: A new immobilization strategy for the preparation of small-molecule arrays. *Angew. Chem. Int. Ed. Engl.* **42,** 5830-5834.

Kolb, H. C., and Sharpless, K. B. (2003). The growing impact of click chemistry on drug discovery. *Drug Discov. Today* **15,** 1128-1137.

Kolonko, E. M., and Kiessling, L. L. (2008). A polymeric domain that promotes cellular internalization. *J. Am. Chem. Soc.* **130,** 5626-5627.

Kremsky, J. N., Wooters, J. L., Dougherty, J. P., Meyers, R. E., Collins, M., and Brown, E. L. (1987). Immobilization of DNA via oligonucleotides containing an aldehyde or carboxylic acid group at the 5' terminus. *Nucleic Acids Res.* **15,** 2891-2909.

Lagerlof, E., Nathorst-Westfelt, L., Ekstrom, B., and Sjoberg, B. (1976). Production of 6-aminopenicillanic acid with immobilized *Escherichia coli* acylase. *Methods Enzymol.* **44,** 759-768.

Lahiri, J., Isaacs, L., Tien, J., and Whitesides, G. M. (1999). A strategy for the generation of surfaces presenting ligands for studies of binding based on an active ester as a common reactive intermediate: a surface plasmon resonance study. *Anal. Chem.* **71,** 777-790.

Laughlin, S. T., Baskin, J. M., Amacher, S. L., and Bertozzi, C. R. (2008). *In vivo* imaging of membrane-associated glycans in developing zebrafish. *Science* **320,** 664-667.

Lavis, L. D., Chao, T. Y., and Raines, R. T. (2006). Fluorogenic label for biomolecular imaging. *ACS Chem. Biol.* **1,** 252-60.

Leland, P. A., Schultz, L. W., Kim, B.-M., and Raines, R. T. (1998). Ribonuclease A variants with potent cytotoxic activity. *Proc. Natl. Acad. Sci. U.S.A.* **98,** 10407-10412.

Levina, A., and Lay, P. (2004). Solution structures of Chromium(VI) complexes with glutathione and model thiols. *Inorg. Chem.* **43,** 324-335.

Li, M., Yamato, K., Ferguson, J. S., Singarapu, K. K., Szyperski, T., and Gong, B. (2008). Sequence-specific, dynamic covalent crosslinking in aqueous media. *J. Am. Chem. Soc.* **130,** 491-500.

Lin, F. L., Hoyt, H. M., van Halbeek, H., Bergman, R. G., and Bertozzi, C. R. (2005). Mechanistic investigation of the Staudinger ligation. *J. Am. Chem. Soc.* **127,** 2686-2695.

Lin, P. C., Ueng, S. H., Yu, S. C., Jan, M. D., Adak, A. K., Yu, C. C., and Lin, C. C. (2007). Surface modification of magnetic nanoparticle via Cu(I)-catalyzed alkyne-azide [2 + 3] cycloaddition. *Org. Lett.* **9,** 2131-2134.

Lipshutz, R. J., Fodor, S. P. A., Gingeras, T. R., and Lockhard, D. J. (1999). High density synthetic oligonucleotide arrays. *Nat. Genetics* **21,** 20-24.

Liu, X., C.-H., J., Zheng, F., Jürgensen, A., Denslinger, J. D., Dickson, K. A., Raines, R. T., Abbott, N. L., and Himpsel, F. J. (2006). Characterization of protein immobilization at sliver surfaces by near edge X-ray absorption fine structure spectroscopy. *Langmuir* **22,** 7719-7725.

Lowe, A. M., Bertics, P. J., and Abbott, N. L. (2008). Quantitative methods based on twisted nematic liquid crystals for mapping surfaces patterned with bio/chemical functionality relevant to bioanalytical assays. *Anal. Chem.* **80,** 2637-45.

Luchansky, S. J., Hang, H. C., Saxon, E., Grunwell, J. R., Yu, C., Dube, D. H., and Bertozzi, C. R. (2003). Constructing azide-labeled cell surfaces using polysaccharide biosynthetic pathways. *Methods Enzymol.* **362,** 249-272.

Lue, R. Y., Chen, G. Y., Hu, Y., Zhu, Q., and Yao, S. Q. (2004). Versatile protein biotinylation strategies for potential high-throughput proteomics. *J. Am. Chem. Soc.* **126,** 1055-1062.

Luk, Y.-Y., Tingey, M. L., Dickson, K. A., Raines, R. T., and Abbott, N. L. (2004). Comparison of the binding activity of randomly oriented and uniformly oriented proteins immobilized by chemoselective coupling to a self-assembled monolayer. *J. Am. Chem. Soc.* **126,** 9024-9032.

Lundblad, R. L. (2005). "Chemical Reagents for Protein Modification, 3rd ed." CRC Press, Boca Raton, FL.

Lynen, F., Reichert, E., and Rueff, L. (1951). Biological degradation of acetic acid. VI. Isolation and chemical nature of activated acetic acid. *Ann. Chem.* **574,** 1-32.

MacBeath, G., and Schreiber, S. L. (2000). Printing proteins as microarrays for high-throughput function determination. *Science* **289,** 1760-1763.

Macpherson, L. J., Dubin, A. E., Evans, M. J., Marr, F., Schultz, P. G., Cravatt, B. F., and Patapoutian, A. (2007). Noxious compounds activate TRPA1 ion channels through covalent modification of cysteines. *Nature* **445,** 541-545.

Matsui, S., and Aida, H. (1978). Hydrolysis of some N-alkylmaleimides. *J. Chem. Soc. Perkin II,* 1277-1280.

McCaldon, P., and Argos, P. (1988). Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences. *Proteins* **4,** 99-122.

Melnyk, O., Duburcq, X., Olivier, C., Urbes, F., Auriault, C., and Gras-Masse, H. (2002). Peptide arrays for highly sensitive and specific antibody-binding fluorescence assays. *Bioconjugate Chem.* **13,** 713-720.

Merkel, J. S., Michaud, G. A., Salcius, M., Schweitzer, B., and Predki, P. F. (2005). Functional protein microarrays: Just how functional are they? *Curr. Opin. Biotechnol.* **16,** 447-452.

Messmore, J. M., Fuchs, D. N., and Raines, R. T. (1995). Ribonuclease A: Revealing structure–function relationships with semisynthesis. *J. Am. Chem. Soc.* **117,** 8057-8060.

Metzler, D. E. (1957). Equilibria between pyridoxal and amino acids and their imines. *J. Am. Chem. Soc.* **79,** 485-490.

Metzler, D. E., Ikawa, M., and Snell, E. (1954). A general mechanism of vitamin B_6-catalyzed reactions. *J. Am. Chem. Soc.* **76,** 648-652.

Moore, J. S., and Stupp, S. I. (1990). Cleavage of aldehyde hydrazonium iodides under mild conditions. A convenient route to chiral nitriles of high enantiomeric purity. *J. Org. Chem.* **55,** 3374-3377.

Moses, J. E., and Moorhouse, A. D. (2007). The growing applications of click chemistry. *Chem. Soc. Rev.* **36,** 1249-1262.

Muir, T. W. (2003). Semisynthesis of proteins by expressed protein ligation. *Annu. Rev. Biochem.* **72,** 249-289.

Mullis, K. B. (1994). The polymerase chain reaction (Nobel lecture). *Angew. Chem. Int. Ed. Engl.* **33,** 1209-1213.

Muralidharan, V., and Muir, T. W. (2006). Protein ligation: An enabling technology for the biophysical analysis of proteins. *Nat. Methods* **3,** 429-438.

Murdock, A. L., Grist, K. L., and Hirs, C. H. W. (1966). On the dinitrophenylation of bovine pancreatic ribonuclease A. Kinetics of the reaction in water and 8 M urea. *Arch. Biochem. Biophys.* **114,** 375-390.

Nektar Therapeutics AL, C. (2004). World Patent WO 2004/06095.

Neuner, P., Gallo, P., Orsatti, L., Fontana, L., and Monaci, P. (2003). An efficient and versatile synthesis of bisPNA-peptide conjugates based on chemoselective oxime formation. *Bioconjugate Chem.* **14,** 276-281.

Newman, R. H., and Zhang, J. (2008). Visualization of phosphatase activity in living cells with a FRET-based calcineurin activity sensor. *Mol. Biosyst.* **4,** 496-501.

Nilsson, B. L., Hondal, R. J., Soellner, M. B., and Raines, R. T. (2003). Protein assembly by orthogonal chemical ligation methods. *J. Am. Chem. Soc.* **125,** 5268-5269.

Nilsson, B. L., Kiessling, L. L., and Raines, R. T. (2000). Staudinger ligation: A peptide from a thioester and azide. *Org. Lett.* **2,** 1939-1941.

Nilsson, B. L., Kiessling, L. L., and Raines, R. T. (2001). High-yielding Staudinger ligation of a phosphinothioester and azide to form a peptide. *Org. Lett.* **3,** 9-12.

Nilsson, B. L., Soellner, M. B., and Raines, R. T. (2005). Chemical synthesis of proteins. *Annu. Rev. Biophys. Biomolec. Struct.* **34,** 91-118.

Niyaz Khan, N., and Aziz Khan, A. (1975). Kinetics and mechanism of hydrolysis of succinimide in highly alkaline medium. *J. Org. Chem.* **40,** 1793-1794.

Noda, L. H., Kuby, S. A., and Lardy, H. A. (1953). Properties of thiolesters: Kinetics of hydrolysis in dilute aqueous media. *J. Am. Chem. Soc.* **75,** 913-917.

Noren, C. J., Anthony-Cahill, S. J., Griffith, M. C., and Schultz, P. G. (1989). A general method for site-specific incorporation of unnatural amino acids into proteins. *Science* **244,** 182-188.

Paborsky, L. R., Dunn, K. E., Gibbs, C. S., and Dougherty, J. P. (1996). A nickel chelate microtiter plate assay for six histidine-containing proteins. *Anal. Biochem.* **234,** 60-65.

Parker, D. (1990). Tumor targeting with radio-labeled macrocycle-antibody conjugates. *Chem. Soc. Rev.* **19,** 271-291.

Pauling, L. (1960). "The Nature of the Chemical Bond." Cornell University Press, Ithaca, NY.

Paulus, H. (2000). Protein splicing and related forms of protein autoprocessing. *Annu. Rev. Biochem.* **69,** 447-496.

Pellecchia, M., Sem, D. S., and Wuthrich, K. (2002). NMR in drug discovery. *Nat. Rev. Drug Discov.* **1,** 211-9.

Peluso, P., Wilson, D. S., Do, D., Tran, H., Venkatasubbaiah, M., Quincy, D., Heidecker, B., Poindexter, K., Tolani, H., Phelan, M., Witte, K., Jung, L. S., Wagner, P., and Nock, S. (2003). Optimizing antibody immobilization

strategies for the construction of protein microarrays. *Anal. Biochem.* **312,** 113-124.

Perez-Howard, G. M., Weil, P. A., and Beechem, J. M. (1995). Yeast TATA binding protein interaction with DNA: fluorescence determination of oligomeric state, equilibrium binding, on-rate, and dissociation kinetics. *Biochemistry* **34,** 8005-17.

Peschke, B., Zundel, M., Bak, S., Clausen, T. R., Blume, N., Pedersen, A., Zaragoza, F., and Madsen, K. (2007). C-Terminally PEGylated hGH-derivatives. *Bioorg. Med. Chem.* **15,** 4382-4395.

Prescher, J. A., and Bertozzi, C. R. (2005). Chemistry in living systems. *Nat. Chem. Biol.* **1,** 13-21.

Price, C. P., and Newman, D. J. (1997). "Principles and Practice of Immunoassay." Macmillan Reference Ltd., London.

Prime, K. L., and Whitesides, G. M. (1993). Adsorption of proteins onto surfaces containing end-attached oligo(ethylene oxide): A model system using self-assembled monolayers. *J. Am. Chem. Soc.* **115,** 10714-10721.

Raag, R., and Whitlow, M. (1995). Single-chain Fvs. *FASEB J.* **9,** 73-80.

Radzicka, A., and Wolfenden, R. (1996). Rates of uncatalyzed peptide bond hydrolysis in neutral solution and the transition state affinities of proteases. *J. Am. Chem. Soc.* **118,** 6105-6109.

Raines, R. T. (1997). Nature's transitory covalent bond. *Nat. Struct. Biol.* **4,** 424-427.

Raines, R. T. (1998). Ribonuclease A. *Chem. Rev.* **98,** 1045-1066.

Rajpal, A., Beyaz, N., Haber, L., Cappuccilli, G., Yee, H., Bhatt, R. R., Takeuchi, T., Lerner, R. A., and Crea, R. (2005). A general method for greatly improving the affinity of antibodies by using combinatorial libraries. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 8466-8471.

Roberts, M. J., Bentley, M. D., and Harris, J. M. (2002). Chemistry for peptide and protein PEGylation. *Adv. Drug. Deliv. Rev.* **54,** 459-479.

Rose, K. (1994). Facile synthesis of homogeneous artificial proteins. *J. Am. Chem. Soc.* **116,** 30-33.

Rose, K., Zeng, W., Ragamey, P.-O., Chernushevich, I. V., Standing, K. G., and Gaertner, H. F. (1996). Natural peptides as building blocks for the synthesis of large protein like molecules with hydrazone and oxime linkages. *Bioconjugate Chem.* **7,** 552-556.

Rup, B., and O'Hara, D. (2007). Critical ligand binding reagent preparation/selection: When specificity depends on reagents. *AAPS J.* **9,** E148-55.

Rutkoski, T. J., Kurten, E. L., Mitchell, J. C., and Raines, R. T. (2005). Disruption of shape-complementarity markers to create cytotoxic variants of ribonuclease A. *J. Mol. Biol.* **354,** 41-54.

Rutkoski, T. J., and Raines, R. T. *unpublished results.*

Sander, E. G., and Jencks, W. P. (1968). Equilibria for additions to the carbonyl group. *J. Am. Chem. Soc.* **90,** 6154-6162.

Sasaki, Y., Lindqvist, I., and Sillen, L. G. (1959). On the equilibrium steps in the acidification of the molybdate anion. *J. Inorg. Nuc. Chem.* **9,** 93-94.

Saxon, E., and Bertozzi, C. R. (2000). Cell surface engineering by a modified Staudinger reaction. *Science* **287,** 2007-2010.

Saxon, E., Luchansky, S. J., Hang, H. C., Yu, C., Lee, S. C., and Bertozzi, C. R. (2002). Investigating cellular metabolism of synthetic azidosugars with the Staudinger ligation. *J. Am. Chem. Soc.* **124,** 14893-14902.

Sayer, J. M., Peskin, M., and Jencks, W. P. (1973). Imine forming elimination reactions. I. General base and acid catalysis and influence of the nitrogen

substituent on rates and equilibria for carbinolamine dehydration. *J. Am. Chem. Soc.* **95,** 4277-4287.

Schena, M. (2004). "Protein Microarrays." Jones and Bartlett, Boston, MA.

Schier, R., McCall, A., Adams, G. P., Marshall, K. W., Merritt, H., Yim, M., Crawford, R. S., Weiner, L. M., Marks, C., and Marks, J. D. (1996). Isolation of picomolar affinity anti-c-erbB-2 single-chain Fv by molecular evolution of the complementarity determining regions in the center of the antibody binding site. *J. Mol. Biol.* **263,** 551-567.

Schlom, J., Eggensperger, D., Colcher, D., Molinolo, A., Houchens, D., Miller, L. S., Hinkle, G., and Siler, K. (1992). Therapeutic advantage of high-affinity anticarcinoma radioimmunoconjugates. *Cancer Res.* **52,** 1067-1072.

Schneider, C. H., and de Weck, A. L. (1965). Studies on the direct neutral penicilloylation of functional groups occurring on proteins. *Biochim. Biophys. Acta* **168,** 27-35.

Schuler, B., and Eaton, W. A. (2008). Protein folding studied by single-molecule FRET. *Curr. Opin. Struct. Biol.* **18,** 16-26.

Shao, J., and Tam, J. P. (1995). Unprotected peptides as building blocks for the synthesis of peptide dendrimers with oxime, hydrazone, and thiazolidine linkages. *J. Am. Chem. Soc.* **117,** 3893-3899.

Shipp, E. L., and Hsieh-Wilson, L. C. (2007). Profiling the sulfation specificities of glycosaminoglycan interactions with growth factors and chemotactic proteins using microarrays. *Chem. Biol.* **14,** 195-208.

Shreve, P., and Aisen, A. M. (1986). Monoclonal antibodies labeled with polymeric paramagnetic ion chelates. *Magn. Reson. Med.* **3,** 336-340.

Sigala, P. A., Fafarman, A. T., Bogard, P. E., Boxer, S. G., and Herschlag, D. (2007). Do ligand binding and solvent exclusion alter the electrostatic character within the oxyanion hole of an enzymatic active site? *J. Am. Chem. Soc.* **129,** 12104-12105.

Singh-Gasson, S., Green, R. D., Yue, Y., Nelson, C., Blattner, F., Sussman, M. R., and Cerrina, F. (1999). Maskless fabrication of light-directed oligonucleotide microarrays using a digital micromirror array. *Nat. Biotech.* **17,** 974-978.

Slavica, A., Dib, I., and Nidetzky, B. (2007). Selective modification of surface-exposed thiol groups in Trigonopsis variabilis D-amino acid oxidase using poly(ethylene glycol) maleimide and its effect on activity and stability of the enzyme. *Biotechnol. Bioeng.* **96,** 9-17.

Smith, B. D., Soellner, M. B., and Raines, R. T. (2003). Potent inhibition of ribonuclease A by oligo(vinylsulfonic acid). *J. Biol. Chem.* **278,** 20934-20938.

Smith, G. J., Lee, K. T., Qu, X., Xie, Z., Pesic, J., Sosnick, T. R., Pan, T., and Scherer, N. F. (2008). A large collapsed-state RNA can exhibit simple exponential single-molecule dynamics. *J. Mol. Biol.* **378,** 941-51.

Smith, P. A. S., and Most, E. E. (1957). Quaternary hydrazones and their rearrangements. *J. Org. Chem.* **22,** 358-362.

Smith, R. F., and Walker, L. E. (1962). A facile conversion of aldehydes to nitriles. *J. Org. Chem.* **27,** 4372-4375.

Soellner, M. B., Dickson, K. A., Nilsson, B. L., and Raines, R. T. (2003). Site-specific protein immobilization by Staudinger ligation. *J. Am. Chem. Soc.* **125,** 11790-11791.

Soellner, M. B., Nilsson, B. L., and Raines, R. T. (2002). Staudinger ligation of α-azido acids retains stereochemistry. *J. Org. Chem.* **67,** 4993-4996.

Soellner, M. B., Nilsson, B. L., and Raines, R. T. (2006). Reaction mechanism and kinetics of the traceless Staudinger ligation. *J. Am. Chem. Soc.* **128,** 8820-8828.

Speers, A. E., Adam, G. C., and Cravatt, B. F. (2003). Activity-based protein profiling *in vivo* using a copper(I)-catalyzed azide–alkyne [3 + 2] cycloaddition. *J. Am. Chem. Soc.* **125,** 4686-4687.

Speers, A. E., and Cravatt, B. F. (2005). A tandem orthogonal proteolysis strategy for high-content chemical proteomics. *J. Am. Chem. Soc.* **127,** 10018-10019.

Staudinger, H., and Meyer, J. (1919). New organic compounds of phosphorus. III. Phosphinemethylene derivatives and phosphinimines. *Helv. Chim. Acta* **2,** 635-46.

Sun, X. L., Stabler, C. L., Cazalis, C. S., and Chaikof, E. L. (2006). Carbohydrate and protein immobilization onto solid surfaces by sequential Diels-Alder and azide-alkyne cycloadditions. *Bioconjugate Chem.* **17,** 52-57.

Sweeney, R. Y., Kelemen, B. R., Woycechowsky, K. J., and Raines, R. T. (2000). A highly active immobilized ribonuclease. *Anal. Biochem.* **286,** 312-314.

Sydor, J. R., Mariano, M., Sideris, S., and Nock, S. (2002). Establishment of intein-mediated protein ligation under denaturing conditions: C-Terminal labeling of a single-chain antibody for biochip screening. *Bioconjugate Chem.* **13,** 707-712.

Tam, A., Soellner, M. B., and Raines, R. T. (2007). Water-soluble phosphinothiols for traceless staudinger ligation and integration with expressed protein ligation. *J. Am. Chem. Soc.* **129,** 11421-11430.

Tan, L. P., and Yao, S. Q. (2005). Intein-mediated, *in vitro* and *in vivo* protein modifications with small molecules. *Protein Pept. Lett.* **12,** 769-775.

Terrettax, W., Ulrich, W. P., Vogel, H., Hong, Q., Dover, L. G., and Lakey, J. H. (2002). Stable self-assembly of a protein engineering scaffold on gold surfaces. *Protein Sci.* **11,** 1917-1925.

Thompson, J. A., Grunert, F., and Zimmerman, W. (1991). Carcinoembryonic antigen gene family: Molecular biology and clinical perspectives. *J. Clin. Lab. Anal.* **5,** 344-366.

Tiefenauer, L. X., Kuhne, G., and Andres, R. Y. (1993). Antibody-magnetite nanoparticles: *in vitro* characterization of a potential tumor-specific contrast agent for magnetic resonance imaging. *Bioconjugate Chem.* **4,** 347-352.

Tolbert, T. J., and Wong, C.-H. (2000). Intein-mediated synthesis of proteins containing carbohydrates and other molecular probes. *J. Am. Chem. Soc.* **122,** 5421-5428.

Tosa, T., Mori, T., Fuse, N., and Chibata, I. (1967). Studies on continuous enzyme reactions. IV. Preparation of a DEAE-sephadex-aminocyclase column and continuous optical resolution of acyl-DL-amino acids. *Biotechnol. Bioeng.* **9,** 603-615.

Tully, S. E., Rawat, M., and Hsieh-Wilson, L. C. (2006). Discovery of a TNF-alpha antagonist using chondroitin sulfate microarrays. *J. Am. Chem. Soc.* **128,** 7740-7741.

Underbakke, E. S., and Kiessling, L. L. *personal communication.*

van Steenis, D. J. V. C., David, O. R. P., van Strijdonck, G. P. F., van Maarseveen, J. H., and Reek, J. N. H. (2005). Click-chemistry as an efficient synthetic tool for the preparation of novel conjugated polymers. *Chem. Commun.*, 4333-4335.

Wahl, R. L., Parker, C. W., and Philpott, G. W. (1983). Improved radioimaging and tumor localization with monoclonal F(ab')2. *Invest. Nucl. Med.* **24,** 316-325.

Waldman, T. A. (1991). Monoclonal antibodies in diagnosis and therapy. *Science* **252,** 1657-1662.

Wallace, C. J. A. (1999). "Protein Engineering by Semisynthesis." CRC Press, Boca Raton, FL.

Walsh, C. T. (2005). "Posttranslational modification of proteins: Expanding Nature's inventory." Roberts, Englewood, CO.

Wang, C. C., Seo, T. S., Li, Z., Ruparel, H., and Ju, J. (2003a). Site-specific fluorescent labeling of DNA using Staudinger ligation. *Bioconjugate Chem.* **14,** 697-701.

Wang, Q., Chan, T. R., Hilgraf, R., Fokin, V. V., Sharpless, K. B., and Finn, M. G. (2003b). Bioconjugation by copper(I)-catalyzed azide–alkyne [3 + 2] cycloaddition. *J. Am. Chem. Soc.* **125,** 3192-3193.

Watzke, A., Gutierrez-Rodriguez, M., Kohn, M., Wacker, R., Schroeder, H., Breinbauer, R., Kuhlmann, J., Alexandrov, K., Niemeyer, C. M., Goody, R. S., and Waldmann, H. (2006a). A generic building block for C- and N-terminal protein-labeling and protein-immobilization. *Bioorg. Med. Chem.* **14,** 6288-306.

Watzke, A., Kohn, M., Gutierrex-Rodriguez, M., Wacker, R., Schroder, H., Breinbauer, R., Kuhlmann, J., Alexandrov, K., Niemeyer, C. M., Goody, R. S., and Waldmann, H. (2006b). Site-selective protein immobilization by Staudinger ligation. *Angew. Chem. Int. Ed. Engl.* **45,** 1408-1412.

Westheimer, F. H. (1934). Semicarbazone formation in sixty per cent. methyl cellosolve. *J. Am. Chem. Soc.* **56,** 1962-1965.

Wiberg, K. B., and Glaser, R. (1992). Resonance interactions in acyclic systems. 4. Stereochemistry, energetics, and electron distributions in 3-center-four-.pi.-electron systems A:BC. *J. Am. Chem. Soc.* **114,** 841-850.

Wikjord, B., and Byers, L. D. (1992a). Nucleophilic reactions of molybdate. *J. Org. Chem.* **57,** 6814-6817.

Wikjord, B., and Byers, L. D. (1992b). The remarkable nucleophilic reactivity of molybdate. *J. Am. Chem. Soc.* **114,** 5553-5554.

Wilchek, M., and Bayer, E. A. (1987). Labeling glycoconjuagtes with hydrazide reagents. *Methods Enzymol.* **138,** 429-442.

Wilson, D. S., and Nock, S. (2003). Recent developments in protein microarray technology. *Angew. Chem. Int. Ed.* **42,** 494-500.

Winter, G., Griffiths, A. D., Hawkins, R. E., and Hoogenboom, H. R. (1994). Making antibodies by phage display technology. *Annu. Rev. Immunol.* **12,** 433-455.

Wolfenden, R., and Jencks, W. P. (1961). The effect of o-substituents on benzaldehyde semicarbazone formation. *J. Am. Chem. Soc.* **83,** 2763-2768.

Wood, L. L., Cheng, S.-S., Edmiston, P. L., and Saavedra, S. S. (1997). Molecular orientation in protein films. 2. Site-directed immobilization of yeast cytochrome c on thiol-capped, self-assembled monolayers. *J. Am. Chem. Soc.* **119,** 571-576.

Wood, R. J., Pascoe, D. D., Brown, Z. K., Medlicott, E. M., Kriek, M., Neylon, C., and Roach, P. L. (2004). Optimized conjugation of a fluorescent label to proteins via intein-mediated activation and ligation. *Bioconjugate Chem.* **15,** 366-372.

Yan, L. Z., and Dawson, P. E. (2001). Synthesis of peptides and proteins without cysteine residues by native chemical ligation combined with desulfurization. *J. Am. Chem. Soc.* **123,** 526-533.

Yee, C. S., Seyedsayamdost, M. R., Chang, M. C., Nocera, D. G., and Stubbe, J. (2003). Generation of the R2 subunit of ribonucleotide reductase by intein chemistry: Insertion of 3-nitrotyrosine at residue 356 as a probe of the radical initiation process. *Biochemistry* **42,** 14541-14552.

Yeo, D. S., Panicker, R. C., Tan, L. P., and Yao, S. Q. (2004). Strategies for immobilization of biomolecules in a microarray. *Comb. Chem. High. Throughput Screen.* **7,** 213-221.

Yi, H., Hinklin, R. J., Chang, J., and Kiessling, L. L. (2004). Stereoselective *N*-glycosylation by Staudinger ligation. *Org. Lett.* **6,** 4479-4482.

Zalipsky, S. (1995). Functionalized poly(ethylene glycol) for preparation of biologically relevant conjugates. *Bioconjugate Chem.* **6,** 150-165.

Zatsepin, T. S., Stetsenko, D. A., Gait, M. J., and Oretskaya, T. S. (2005). Use of carbonyl group addition-elimination reactions for synthesis of nucleic acid conjugates. *Bioconjugate Chem.* **16,** 471-489.

Zhang, S., Metelev, V., Tabatadze, D., Zamecnik, P. C., and Bogdanov, A., Jr. (2008). Fluorescence resonance energy transfer in near-infrared fluorescent oligonucleotide probes for detecting protein-DNA interactions. *Proc. Natl. Acad. Sci. U.S.A* **105,** 4156-61.

Zhou, Z., and Fahrni, C. J. (2004). A fluorogenic probe for the copper(I)-catalyzed azide-alkyne ligation reaction: modulation of the fluorescence emission via 3(n,pi)-1(pi,pi) inversion. *J. Am. Chem. Soc.* **126,** 8862-8863.

Zhu, H., Bilgin, M., Bangham, R., Hall, D., Casamayor, A., Bertone, P., Lau, N., Jansen, R., Bidlingmaier, S., Houfek, T., Mitchell, T., Miller, P., Dean, R. A., Gerstein, M., and Snyder, M. (2001). Global analysis of protein activities using proteome chips. *Science* **293,** 2101-2105.

Zverev, V. V., Pyalaeva, T. N., Ermolaeva, L. V., Filippova, N. A., and Kitaev, Y. P. (1977). Basicities and ionization potentials of the hydrazones. *Bull. Acad. Sci. USSR, Div. Chem. Sci.* **26,** 1865-1869.