Development and Application of a Mass Spectrometry-Based Assay for the High Throughput Analysis of Protein-Ligand Binding

by

Erin Davis Hopper

Department of Chemistry
Duke University

Date:_____

Approved:

_____
Michael C. Fitzgerald, Supervisor

_____
Katherine J. Franz

_____
William M. Reichert

_____
Barbara R. Shaw

_____
Chandra L. Tucker

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Chemistry in the Graduate School
of Duke University

2009

ABSTRACT

Development and Application of a Mass Spectrometry-Based Assay for the High Throughput Analysis of Protein-Ligand Binding

by

Erin Davis Hopper

Department of Chemistry
Duke University

Date:_____

Approved:

_____
Michael C. Fitzgerald, Supervisor

_____
Katherine J. Franz

_____
William M. Reichert

_____
Barbara R. Shaw

_____
Chandra L. Tucker

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Chemistry in the Graduate School of Duke University

2009

Abstract

Many of the biological roles of proteins are modulated through protein-ligand interactions, making proteins important targets for drug therapies and diagnostic imaging probes. The discovery of novel ligands for a protein of interest often relies on the use of high throughput screening (HTS) technologies designed to detect protein-ligand binding. The basis of one such technology is a recently reported mass spectrometry-based assay termed SUPREX (stability of unpurified proteins from rates of H/D exchange). SUPREX is a technique that uses H/D exchange and MALDI-mass spectrometry for the measurement of protein stabilities and protein-ligand binding affinities. The single-point SUPREX assay is an abbreviated form of SUPREX that is capable of detecting protein-ligand interactions in a high throughput manner by exploiting the change in protein stability that occurs upon ligand binding.

This work is focused on the development and application of high throughput SUPREX protocols for the detection of protein-ligand binding. The first step in this process was to explore the scope of SUPREX for the analysis of non-two-state proteins to determine whether this large subset of proteins would be amenable to SUPREX analyses. Studies conducted on two model proteins, Bcl-x_L and alanine:glyoxylate aminotransferase, indicate that SUPREX can be used to detect and quantify the strength of protein-ligand binding interactions in non-two-state proteins.

The throughput and efficiency of a high throughput SUPREX protocol (i.e., single-point SUPREX) was also evaluated in this work. As part of this evaluation, cyclophilin A, a protein target of diagnostic and therapeutic significance, was screened against the 880-member Prestwick Chemical Library to identify novel ligands that might be useful as therapeutics or imaging agents for lung cancer. This screening not only established the analytical parameters of the assay, but it revealed a limitation of the technique: the efficiency of the assay is highly dependent on the precision of each mass measurement, which generally decreases as protein size increases.

To overcome this limitation and improve the efficiency and generality of the assay, a new SUPREX protocol was developed that incorporated a protease digestion step into the single-point SUPREX protocol. This new protocol was tested on two model proteins, cyclophilin A and alanine:glyoxylate aminotransferase, and was found to result in a significant improvement in the efficiency of the SUPREX assay in HTS applications. This body of work resulted in advancements in the use of SUPREX for high throughput applications and laid the groundwork for future HTS campaigns on target proteins of medical significance.

Contents

List of Tables

List of Figures

List of Abbreviations

ΔG_{app}	Apparent folding free energy
ΔG_f	Folding free energy
ΔMass	Increase in mass due to the uptake of deuterons
ACN	Acetonitrile
AGT	Alanine:glyoxylate aminotransferase
AGT144	Minor allele alanine:glyoxylate aminotransferase with I244T mutation
AGT170	Minor allele alanine:glyoxylate aminotransferase with G170R mutation
AGTma	Major allele alanine:glyoxylate aminotransferase
AGTmi	Minor allele alanine:glyoxylate aminotransferase
AOA	Aminooxyacetic acid
$C^{1/2}_{SPROX}$	Denaturant concentration at the transition midpoint of a SPROX curve
$C^{1/2}_{SUPREX}$	Denaturant concentration at the transition midpoint of a SUPREX curve
CsA	Cyclosporin A
CypA	Cyclophilin A
DIOS	Desorption/ionization on silicon
DMSO	Dimethyl sulfoxide
ESI-MS	Electrospray ionization mass spectrometry
FAC-MS	Frontal affinity chromatography mass spectrometry

GdmCl	Guanidinium chloride
H/D	Hydrogen/deuterium
HTS	High throughput screening
K_d	Dissociation constant
LC	Liquid chromatography
MALDI	Matrix-assisted laser desorption/ionization
MS	Mass spectrometry
PH1	Primary hyperoxaluria type I
PLP	Pyridoxal 5′-phosphate
PUF-MS	Pulsed ultrafiltration mass spectrometry
SA	Sinapinic acid
SPROX	Stability of proteins from rates of oxidation
SUPREX	Stability of unpurified proteins from rates of H/D exchange
TCA	Trichloroacetic acid
TFA	Trifluoroacetic acid

Acknowledgements

Although my graduate career has been full of ups and downs, I have probably learned more about myself by pursuing a doctoral degree than I could have learned in any other undertaking. I would like to start by thanking my research advisor, Professor Michael C. Fitzgerald, who remained supportive and eternally optimistic throughout the good times and the bad times of my graduate research. Of the many lessons I learned as a graduate student, the most important lesson I learned from Michael was that research failures should be seen as a springboard for future success.

Special thanks to my committee members, Dr. William M. Reichert, Dr. Katherine J. Franz, Dr. Barbara R. Shaw, and Dr. Chandra L. Tucker, for offering advice and monitoring my progress throughout graduate school. Thanks also to my collaborators who made this dissertation research possible. I am grateful to Dr. Edward F. Patz, Jr. and Dr. Michael J. Campa for working with us on the cyclophilin A studies and for providing the purified cyclophilin A samples. In particular, I would like to thank Dr. Chandra L. Tucker, who was a close collaborator in the AGT studies and served as an important role model in my development as a scientist. Thanks also to Adrianne M. C. Pittman for expressing and purifying the AGT constructs.

I would also like to thank the Fitzgerald Lab members for their constant support and friendship. Thanks to the past group members (Susie Dai, Yan Tong, Min Wang,

Xiaoye Yang, Petra Roulhac, Liangjie Tang, and Tabitha George) for supporting me in my early years of graduate study. In particular, thanks to Susie for keeping me optimistic and to Petra for keeping me realistic. Petra and Tabitha provided much-needed friendship through the tough times of graduate school. Thanks also to the present Fitzgerald Lab members (Graham West, Patrick DeArmond, Ying Xu and Victor Anbalagan) for your support, companionship, and humor. Patrick and Ying, I will miss our lunchtime conversations more than anything else and will always be grateful for your friendship.

I have had the good fortune to work with two talented undergraduates: Vaibhav Upadhyay and Pim Dangkulwanich. I would like to especially thank Pim, who was an extremely productive researcher and an equal contributor to the gel-based SPROX experiments discussed in Chapter 7. Vaibhav and Pim, the two of you have taught me more than I could ever teach you, and I wish you the best.

My family has been incredibly supportive throughout my graduate years as well. Thanks to my parents, John and Anne Davis, who were always willing to provide a listening ear and to my sisters, Christina Catrett and Lauren Dotson. Thanks also to my "other parents," Steve and Hazel Hopper, who welcomed me into their family as if I was their own. I am extremely grateful to my grandparents, Calvin and Dorothy Hucks, for always believing in me and for giving me their unconditional love throughout my life.

Finally and most importantly, I want to thank Steven Hopper, my husband and my best friend. First, I want to thank you for coding the MATLAB programs for analyzing raw mass spectra and the high throughput screening results – these programs saved me countless hours over the course of my graduate work. On a more personal note, thanks for always seeming to know exactly what I need, even if I may not agree with you at the time. Thank you for always keeping my best interests in mind and for supporting me through the good and the bad. I love you dearly and credit you with helping me become the person that I am.

1. Introduction

1.1 Background

Proteins perform a wide variety of functions within the cell, including catalysis, transport, regulation of biological processes, and provision of structural components (1). Many of these biological roles are modulated through protein-ligand interactions, making proteins important targets for drug therapies (2). Acceptable drug targets are proteins that are capable of modifying a disease state and that are amenable to interactions with small, drug-like molecules. While research in both academia and industry has resulted in the identification of a number of suitable drug targets, the development of compounds that modulate the activity of the protein targets has been slow (2).

In addition to the identification of lead compounds for the development of novel pharmaceuticals, the ability to screen for protein-ligand interactions is critical in the area of molecular imaging (3). One common strategy in this field is direct imaging, in which a labeled probe directly binds a target of interest; this target is usually a biomarker for the disease state (4). The work in this dissertation is largely focused on two model protein systems in which the identification of novel ligands could be useful for both therapeutic and imaging purposes. The first protein system is alanine:glyoxylate aminotransferase (AGT), a liver enzyme that is involved in primary hyperoxaluria type I (PH1) and is a target for therapeutic intervention. The second protein system is cyclophilin A (CypA), a

protein that is overexpressed in lung cancer cells (5) and appears to be required for tumor growth (6). A tight binding ligand to CypA could be developed into a highly specific lung tumor imaging agent.

The discovery of novel ligands that bind a protein target with high affinity and specificity is a notoriously slow process, partially because of the vast number of compounds that must be tested for binding to the target. One strategy for increasing the speed of discovery involves the use of high throughput screening (HTS) technologies to rapidly test large libraries of compounds. Most of the commonly used HTS assays fall into three general categories, including radiometric, fluorescence-based, and cell-based techniques. Techniques in the first two categories are highly sensitive and amenable to automation. In addition, radiometric and fluorescence-based methods often do not require time consuming separation, filtration, and wash steps (7). These qualities make them well suited for HTS applications.

Unfortunately, assays in these two screening categories (i.e., radiometric and fluorescence-based assays) have significant drawbacks. Fluorescence-based assays require labeling of the target, the ligand, or both, and in some cases, labeling may affect protein-ligand interactions (8). Radiometric methods cause safety concerns and require the use of reagents with limited stability (8). Most radiometric assays also require target immobilization, which may alter ligand binding behavior. Cell-based techniques are likely the most physiologically relevant assays, but most require time-consuming

incubation steps that decrease throughput (7). In an attempt to overcome some of the weaknesses inherent in these types of assays, mass spectrometry-based screening methods are beginning to emerge as a promising alternative approach.

1.2 Mass Spectrometry-Based Screening Methods

1.2.1 Function-Based Screening

The speed and sensitivity of modern mass spectrometers make them attractive tools for HTS assays, and a number of mass spectrometry-based HTS assays have been developed in recent years (9-13). Assays involving mass spectrometry generally select ligands based on protein activity (function-based screening) or protein-ligand binding affinity (affinity-based screening). In function-based screening, the function of the protein (e.g., enzymatic activity) is exploited to report on ligand binding. These assays often rely on quantifying either the disappearance of a substrate or the appearance of a product of an enzymatic reaction (13).

The quantitative capacity of electrospray ionization mass spectrometry (ESI-MS) makes it particularly attractive for function-based assays. However, ESI-MS is very sensitive to the presence of salts, so a rapid liquid chromatography (LC) step is generally required before mass spectral analysis. Unfortunately, this LC step requires additional analysis time, making LC-MS techniques generally lower in throughput than fluorescence-based techniques (13). Matrix-assisted laser desorption/ionization mass spectrometry (MALDI-MS) is an alternative soft ionization method that is much more

tolerant of salts than ESI-MS. Although MALDI-MS is generally regarded as less quantitative than ESI-MS, Greis and coworkers recently developed a function-based assay using MALDI-MS that was effective for inhibitor screening (14). Unfortunately, this method can only be used for enzymes with peptide substrates due to an inherent limitation of MALDI-MS. Low molecular weight compounds are difficult to detect using MALDI-MS because the matrix used in the sample preparation gives a high background signal in the low m/z region of the spectra. Because of this limitation, most MALDI-MS-based HTS approaches are only suitable for screening libraries of peptides or proteins (13).

Although MALDI is generally not well-suited for the screening of small molecules, one exception to this limitation is a technique called desorption/ionization on silicon (DIOS)-MS. This technique is very similar to MALDI-MS except that instead of incorporating the sample into a matrix, the sample is deposited on a silicon surface. Similar to MALDI-MS, the sample is desorbed with a laser pulse, but since no matrix is used, the low m/z region of the DIOS-MS spectrum is free of matrix-related ion signals. This makes the strategy effective for enzyme inhibitor screening using small molecule substrates (15). Although quantitation with DIOS-MS is quite reliable, this technique is currently limited to volatile buffer solutions (12, 13), and reproducible preparation of silicon surfaces is difficult; thus, DIOS-MS is not yet applicable for general use.

Function-based screens have the advantage of selecting for compounds that

directly influence the activity of the target protein. This advantage is of critical importance to the development of pharmaceuticals. However, these techniques require that the target possess an assayable activity. Since different proteins perform different functions, a new assay must be developed for each target protein, and this step can be quite time-consuming. A more general approach involves assaying for ligands that interact with the target protein using affinity-based methods. Because affinity-based methods do not assay for modulation of the target's function, a new assay does not need to be developed for each new target protein.

1.2.2 Affinity-Based Screening

Affinity-based methods detect protein-ligand binding by either direct or indirect methods. In direct methods, a target protein is exposed to a mixture of library compounds and then subjected to ESI-MS. Experimental conditions are chosen such that the noncovalent protein-ligand interactions are preserved throughout ESI-MS analysis. Interactions are detected by measuring the mass of the protein-ligand complexes (12). The strength of this strategy is that direct evidence of the interaction is observed. However, since the complex is observed in the gas phase, its relevance to solution-phase interactions is unclear (16).

Unlike direct affinity-based methods, indirect methods involve the identification of binding compounds after the compounds have been separated from the target protein (12, 13). In some cases, indirect methods require the immobilization of the target protein

or ligand. For example, frontal affinity chromatography mass spectrometry (FAC-MS) requires the immobilization of the target protein within a column that is connected to an ESI mass spectrometer (17, 18). A mixture of compounds is passed through the column, and nonbinding compounds are eluted in the void volume. Interacting compounds elute in order from weak binders to strong binders, and the ligands are identified by mass using the mass spectrometer. This approach has the advantage of providing information about relative ligand binding affinities, but the immobilization of the target protein may interfere with ligand binding. The library size for a single analysis is also limited because the eluting compounds must be well-resolved for accurate mass spectral identification (12).

To avoid the problems associated with immobilization of the target protein, solution-based indirect methods can be used. In these techniques, the target protein is incubated with a mixture of ligands. Nonbinding ligands are removed from the solution, and binding ligands are separated from the target and identified by mass spectrometry. One example of this strategy is SpeedScreen, a technique in which unbound ligands are separated from bound ligands using size exclusion chromatography in 96-well plates (19). In size exclusion chromatography, large molecules such as protein-ligand complexes elute quickly, while small molecules are retained in the column for longer periods of time. The quickly eluting protein-ligand complexes are collected and subjected to LC-MS for separation and identification of binding compounds.

An additional example of a solution-based indirect method is pulsed ultrafiltration mass spectrometry (PUF-MS). This strategy is similar to SpeedScreen except that nonbinding compounds are separated from protein-ligand complexes using an ultrafiltration membrane (20, 21). While solution-based indirect methods have the advantage of detecting protein-ligand interactions in the solution phase rather than the gas phase, these techniques have disadvantages as well. For example, the throughput of SpeedScreen is limited by the amount of organic material that can be loaded onto the size exclusion bed, and in PUF-MS, protein-ligand binding can be affected by interactions between the target protein and the ultrafiltration membrane (12).

The majority of the current mass spectrometry-based HTS methods require the use of ESI-MS for identification of interacting compounds. As mentioned above, ESI-MS generally requires an LC sample cleanup step prior to mass spectral analysis. Although this step can be automated, it reduces the throughput of most mass spectrometry-based HTS techniques. This limitation makes MALDI-MS-based HTS strategies attractive, but these strategies generally are not ideal for screening libraries of small molecules due to a high matrix-related background in the low m/z region of MALDI mass spectra (see above). Since low molecular weight compounds are difficult to detect using MALDI-MS, most MALDI-based HTS approaches are more suitable for screening libraries of peptides or proteins than for screening small molecule libraries (13). To overcome these limitations, the work described in this dissertation has involved the use of a technique

called SUPREX (stability of unpurified proteins from rates of H/D exchange), which uses hydrogen/deuterium (H/D) exchange and MALDI-MS to detect protein-ligand binding.

1.3 Hydrogen Exchange Mass Spectrometry and SUPREX

Similar to the affinity-based techniques described above, SUPREX is capable of detecting protein-ligand binding without any knowledge of protein function. However, SUPREX is distinct from these techniques in that the interaction is not detected by separating bound from unbound compounds but rather by measuring the stability of the protein in the presence and absence of each compound. These stability measurements are based upon the difference in free energy between the folded and unfolded forms of the protein (ΔG_f), as described in equation 1-1 (22).

$$\Delta G_f = G_{folded} - G_{unfolded} \qquad \text{Equation 1-1}$$

Stability measurements are frequently performed using conventional equilibrium unfolding experiments. In these experiments, the ΔG_f value for a protein is often measured by using a chemical denaturant such as guanidinium chloride (GdmCl) or urea to perturb the folding equilibrium. A plot of ΔG_f as a function of denaturant concentration is linear and can be extrapolated to 0 M denaturant to determine the ΔG_f value for the protein in the absence of denaturant (23). In these conventional equilibrium unfolding experiments, an optical signal (e.g., fluorescence or circular dichroism) is used to monitor the unfolding reaction at each denaturant concentration (24). In contrast,

SUPREX relies on the progress of an H/D exchange reaction rather than the spectroscopic signal of an optical probe to measure protein folding free energies.

1.3.1 Hydrogen Exchange Mass Spectrometry

All proteins contain backbone amide hydrogen atoms that are capable of exchanging with hydrogen atoms in the surrounding aqueous environment (25). When the protein is in a folded state, amide hydrogen atoms that are exposed to solvent (i.e., unprotected amide hydrogen atoms) freely exchange with solvent hydrogen atoms. However, amide hydrogen atoms that are located in the interior regions of a protein structure or that are involved in hydrogen bonds (i.e., protected amide hydrogen atoms) are only free to exchange when the protein globally unfolds. This process can be summarized by equation 1-2 (26, 27).

$$\text{Closed (NH)} \underset{k_{cl}}{\overset{k_{op}}{\leftrightarrow}} \text{Open (NH)} \overset{k_{int}}{\leftrightarrow} \text{Exchanged (ND)} \qquad \text{Equation 1-2}$$

Equation 1-2 assumes a two-state folding reaction in which amide hydrogen atoms that are originally protected from exchange (closed NH) are made exchange competent (open NH) through a global protein unfolding reaction. In H/D exchange studies, the protein of interest is introduced into a deuterated buffer solution, and deuterium atoms are incorporated into the protein through hydrogen exchange, resulting in an increase in the mass of the protein. This mass change can be detected using mass spectrometry.

Hydrogen exchange coupled to mass spectrometry has been used extensively to study protein structure, stability, and ligand binding characteristics (28-34). In a typical H/D exchange experiment, the protein is first labeled with deuterium using either pulse labeling or continuous labeling (34). In pulse labeling experiments, the protein is equilibrated under conditions that perturb its structure (e.g., high temperature, chemical denaturant, extremes in pH, or interacting compounds), and the protein is exposed to a brief pulse of deuterated solvent. In continuous labeling experiments, the protein is incubated in a deuterated solvent for a variety of time periods. In both cases, the exchange reaction is quenched after the specified exchange time, and the protein samples are analyzed by either MALDI-MS or ESI-MS. In some cases, the protein is digested with a protease prior to mass spectral analysis to gain information about the locations of various types of exchange behavior.

In addition to studies of protein structure and dynamics, H/D exchange has been widely used to examine protein-ligand binding (30, 31). Protein-ligand interactions often result in the increased protection of a subset of the amide hydrogen atoms. This increase in protection can sometimes be used to map protein-ligand binding surfaces. However, ligand binding at one site can sometimes alter hydrogen exchange rates at sites not directly involved in ligand binding (34, 35). Thus, in protein-ligand binding studies, hydrogen exchange should be regarded as a technique to probe protein dynamics rather than a method for the strict determination of protein-ligand interfaces.

1.3.2 Evaluation of Thermodynamic Parameters by SUPREX

Unlike the typical H/D exchange experiments described above, SUPREX combines aspects of both equilibrium unfolding experiments and hydrogen exchange. Originally reported in 2000, this technique is capable of measuring protein folding free energies and protein-ligand binding affinities (36-46). To analyze a protein by SUPREX, the protein is introduced into a series of deuterated exchange buffers containing various concentrations of a chemical denaturant. At higher concentrations of denaturant, a larger population of the protein is unfolded, leading to an increase in the uptake of deuterons. This increase in deuteration leads to an increase in mass, which can be measured using MALDI-MS. The mass increase (i.e., Δmass) is plotted vs. denaturant concentration, and the midpoint of the resulting SUPREX curve transition ($C^{1/2}_{SUPREX}$) reflects the thermodynamic stability of the protein.

One feature of SUPREX that makes it different from conventional equilibrium unfolding experiments is that the midpoint of the SUPREX curve transition is dependent on exchange time. To measure the ΔG_f of a protein, SUPREX is carried out at a variety of different exchange times, and the apparent free energy (ΔG_{app}) at each exchange time is plotted vs. the $C^{1/2}_{SUPREX}$ value measured at that exchange time. This method is analogous to the linear extrapolation method (equation 1-3) used to analyze conventional equilibrium unfolding data (24).

$$-RT \ln K_f = m[\text{denaturant}] + \Delta G_f \qquad \text{Equation 1-3}$$

In equation 1-3, R is the universal gas constant, T is the temperature in Kelvin, K_f is the equilibrium constant for the folding reaction, and *m* is the *m*-value, which is a measure of the cooperativity of the protein folding reaction. The K_f value is given by equation 1-4 (40).

$$K_f = \frac{\langle k_{int} \rangle t}{0.693} - 1 \qquad \text{Equation 1-4}$$

In equation 1-4, $\langle k_{int} \rangle$ is the average exchange rate of an unprotected amide hydrogen atom, and t is the exchange time in seconds. The denaturant concentration ([denaturant]) in equation 1-3 is set equal to $C^{1/2}_{SUPREX}$ (the denaturant concentration at the midpoint of a SUPREX curve), and equations 1-3 and 1-4 are combined to give equation 1-5.

$$-RT \ln\left(\frac{\langle k_{int} \rangle t}{0.693} - 1\right) = mC^{1/2}_{SUPREX} + \Delta G_f \qquad \text{Equation 1-5}$$

The left side of the equality in equation 1-5 is equal to ΔG_{app}. Equation 1-5 is accurate for monomeric proteins, but to analyze multimeric proteins by SUPREX, equation 1-5 must be expanded to form equation 1-6, which accounts for the dependence of multimeric protein stability on protein concentration (41).

$$-RT \ln \frac{\left(\frac{\langle k_{int} \rangle t}{0.693} - 1\right)}{\left(\frac{n^n}{2^{n-1}}[P]^{n-1}\right)} = mC^{1/2}_{SUPREX} + \Delta G_f \qquad \text{Equation 1-6}$$

In equation 1-6, which is known as the SUPREX equation, n is the number of subunits, and [P] is the protein concentration in n-mer equivalents. Equation 1-6 can be used for

both monomeric and multimeric proteins because equation 1-6 reduces to equation 1-5 when n = 1.

Three important assumptions are made in the derivation of the SUPREX equation: 1) the protein folding reaction is assumed to be two-state (i.e., only folded and unfolded states are significantly populated during the protein folding reaction); 2) the hydrogen exchange reaction is assumed to be in the EX2 regime (i.e., the protein folding rate is significantly faster than the hydrogen exchange rate); and 3) the denatured protein is assumed to be in a random coil conformation. To obtain accurate ΔG_f values for any given protein, these assumptions must be met. However, if one or more of these assumptions is violated, relative stabilities can still be measured by comparing the SUPREX transition midpoints of a protein and its complexes. This strategy was used to compare the relative stabilities of several AGT variants as discussed in Chapter 3. Similarly, these rules can often be relaxed in ligand binding studies since the detection and quantitation of ligand binding with SUPREX relies on measuring a difference in free energy between the bound and unbound forms of the protein (43).

1.3.3 Qualitative Analysis of Ligand Binding by SUPREX

In some cases, SUPREX data for a protein-ligand complex is well-described by equation 1-6, but data on the unbound protein is not well-described by equation 1-6. In such cases, K_d value calculations are not possible because a ΔG_f value for the apoprotein cannot be determined. Consequently, the binding free energy ($\Delta\Delta G_f$) cannot be

determined. Poor fits of the SUPREX data to equation 1-6 can occur when the H/D exchange behavior of the protein is dominated by *local* unfolding events rather than *global* unfolding events (43).

The interaction of such a protein with a tight-binding ligand can often induce the H/D exchange behavior of the protein to be dominated by a global unfolding event. In such cases, the evaluation of relative binding free energies is possible if the protein has multiple known ligands. If the first ligand has a unique binding site from the other ligands, then K_d measurements are often still possible. However, in cases where the first ligand occupies the same binding site as the other ligand(s), K_d measurements are not possible, but relative binding free energies can still be measured. Bcl-x_L is an example of such a protein and will be discussed in Chapter 3.

1.4 Single-Point SUPREX

One of the unique strengths of SUPREX is its capacity for high throughput studies. Although the SUPREX experiments described above can be performed relatively quickly, the throughput can be significantly increased using a technique called single-point SUPREX (47). Single-point SUPREX is nearly identical to a conventional SUPREX experiment except that instead of collecting data over a range of denaturant concentrations, SUPREX is performed at a single denaturant concentration. Since the midpoint of a SUPREX curve shifts to the right upon ligand binding, a denaturant concentration can be chosen such that the measured Δmass value is low in the presence

of ligand and high in the absence of ligand. This difference in the Δmass values forms the basis for detecting ligand binding using single-point SUPREX.

Single-point SUPREX is not subject to many of the disadvantages of conventional screening assays that are described above. First, the technique does not require covalent labeling and thus is not susceptible to the alteration of ligand binding behavior that sometimes occurs in the presence of labels. Second, no immobilization of the compound or the target protein is necessary. Third, since this is a MALDI-MS-based technique, no time-consuming separations are required. The final advantage is that single-point SUPREX is amenable to the analysis of multiple classes of ligands (e.g., small molecules, nucleic acids, peptides, and proteins) over a wide range of binding affinities, which makes the technique more flexible than many other HTS strategies.

Single-point SUPREX was first demonstrated in a proof-of-concept study using S-protein as a model target (47). Four model peptide libraries were screened for binding to S-protein, and single-point SUPREX was found to be capable of detecting protein-peptide interactions with good efficiency. The overall goal of this dissertation project was to transform the proof-of-concept single-point SUPREX assay into a viable screening tool for protein-ligand binding.

1.5 Overview of Research Objectives

Throughout the early years of SUPREX development, the capacity of the technique for high throughput analyses has often been noted. However, this advantage

of SUPREX remained to be explored. This dissertation work has focused on exploiting the high throughput capabilities of SUPREX for the detection of protein-ligand binding. As a first step, the full SUPREX protocol was applied to two non-two-state model systems that were of interest as protein targets in future screening assays. Up to that point, only a few non-two-state folding systems had been studied by SUPREX, so these studies helped to expand the scope of SUPREX for the analysis of non-two-state proteins. The first model protein was human Bcl-x_L, an anti-apoptotic protein that has been implicated in several cancers. The identification of Bcl-x_L inhibitors for cancer therapeutics could serve to re-establish apoptosis in cancerous cells or make the cells susceptible to other therapeutic interventions (48). The second model protein was human AGT, a protein that is involved in a severe kidney stone disease called PH1 (49). Disease-causing variants of AGT are promising targets for therapeutic intervention using stabilizing ligands (i.e., pharmacological chaperones). Chapter 3 will describe the use of SUPREX to characterize the ligand binding behavior of these two proteins.

The next goal of this research project was to evaluate the throughput and efficiency of single-point SUPREX. Single-point SUPREX was first developed using a proof-of-concept study in which small libraries of peptides were screened for binding to S-protein (47). However, this proof-of-concept experiment did not involve an evaluation of the technique's throughput and efficiency for screening a real library of compounds. To evaluate these analytical parameters, this work involved performing the first

screening of a compound library against a protein target of diagnostic and therapeutic significance (50). CypA was chosen as the model target protein for this study, and although no new ligands for the protein were discovered, a blind control was successfully selected as a hit. The throughput and efficiency of single-point SUPREX is discussed in detail in Chapter 4.

The completion of an entire screening assay revealed an important limitation of the original single-point SUPREX protocol. The target protein (CypA) in this screening campaign was significantly larger than the target protein in the proof-of-concept study. MALDI-MS mass determinations for larger proteins generally have higher standard deviations than mass determinations for smaller proteins. Since the efficiency of the screening is closely related to the standard deviations of the mass measurements, the efficiency of the screening assay was significantly lower for CypA than it was for S-protein. Since this efficiency limitation would likely restrict the number of proteins that would be amenable to single-point SUPREX, an adjustment to the protocol would be required to improve its generality.

One strategy for decreasing the standard deviations of the MALDI-MS-based mass measurements is to analyze smaller proteins. While the mass of the intact protein cannot be altered, the use of an alternative SUPREX protocol in which a protease digestion step is implemented prior to MALDI-MS analysis (51) would allow for an improvement in the precision of the mass measurements. Using this protocol, the

hydrogen exchange reaction is carried out on the intact protein and thus reflects the folding properties of the full-length protein. A fast protease digestion is performed immediately before MALDI-MS analysis, and the SUPREX behavior of each fragment reports on the folding behavior of the intact protein.

Before adding a digestion step to the single-point SUPREX protocol, this work involved a full characterization of the SUPREX behavior of minor allele AGT (AGTmi) using the SUPREX-protease digestion protocol. Without the digestion protocol, AGTmi is not a well-behaved SUPREX protein (i.e., the midpoint of the SUPREX curves do not move with exchange time). However, when the SUPREX-protease digestion protocol was used, AGTmi became well-behaved, which allowed for the measurement of the dissociation constant for one of the AGTmi ligands. This full ligand binding analysis is discussed in Chapter 5.

The next step was to perform proof-of-concept experiments to evaluate the improved single-point SUPREX protocol. Two model systems were selected for this study (CypA and AGTmi), which involved screening a group of positive and negative controls for binding to each target. Chapter 6 describes the improvement in efficiency achieved by incorporating a protease digestion step into the single-point SUPREX protocol.

The final goal of this work was to address a different type of throughput challenge. Up to this point, this dissertation project has focused on screening compound

libraries against a single protein target. However, another goal of this work was to explore the multiplexing capabilities of this analytical technique. Of particular interest is the application of the technique to the thermodynamic analysis of entire proteomes. The ability to assess thermodynamic stabilities on a proteomic scale would allow for the identification of protein binding partners for ligands of interest. However, this significant challenge will require the development of new strategies for the simultaneous analysis of large numbers of proteins.

As a first step, a gel-based proteomics approach was adopted for the analysis of protein-ligand interactions. Since hydrogen exchange is reversible, SUPREX is not compatible with gel-based fractionation strategies. Overcoming this challenge involved the use of a new technique called SPROX (stability of proteins by rates of oxidation) (52) which is analogous to SUPREX except that a covalent modification (oxidation) is employed instead of hydrogen exchange.

T7 bacteriophage was chosen as a model system due to the small size of its proteome (~55 proteins). The protein product of gene 2.5 is a single stranded binding protein and is known to be a dimer (53), so this protein was chosen as a starting point for the detection of the first protein-protein interaction. Using a combination of SPROX, SDS-PAGE, and MALDI-MS, the thermodynamic stability of protein 2.5 was qualitatively analyzed. The thermodynamic stability measured using the gel-based strategy was in excellent agreement with the stability measured using a solution-based

SPROX analysis. The details of this work are discussed in Chapter 7. Finally, Chapter 8 provides a summary of this dissertation project and gives details about the benefits and limitations of SUPREX for high throughput applications. Taken in its entirety, this dissertation project resulted in the development and application of several new SUPREX- and SPROX-related protocols that significantly expand the high throughput capabilities of these analytical methodologies.

2. General Methods and Materials

2.1 Materials

The following reagents were purchased from Sigma-Aldrich (St. Louis, MO): deuterium oxide (99.9% atom D), sodium deuteroxide (40 wt% in D_2O, 99.5% atom D), deuterium chloride (35 wt% in D_2O, 99% atom D), and sinapinic acid (SA). Guanidine hydrochloride (GdmCl) was purchased from EMD (Gibbstown, NJ), and trifluoroacetic acid (TFA) was purchased from Halocarbon (River Edge, NJ). Acetonitrile (ACN) was purchased from Fisher (Fair Lawn, NJ), and deuterated phosphoric acid-d_3 (99 atom %D) was from Cambridge Isotope Laboratories (Andover, MA). Protein and ligand sources are specified in each chapter.

2.2 Mass Spectrometry Instrumentation

MALDI mass spectra were collected on either a Voyager DE-Pro Biospectrometry Workstation (PerSeptive Biosystems; Framingham, MA) or an Ultraflex II TOF/TOF mass spectrometer (Bruker Daltonics; Billerica, MA). The Voyager instrument is equipped with a nitrogen laser (337 nm) operating at approximately 3 Hz, and the Ultraflex features a Nd:YAG laser operating at 100 Hz. Mass spectra for SUPREX experiments were collected in linear and positive ion mode, and mass spectra for SPROX experiments were collected in reflectron and positive ion mode. Detailed instrument parameters are specified in each chapter.

2.3 Buffer Preparation

Deuterated H/D exchange buffers (20 mM phohsphate, pD 7.4) were prepared that contained concentrations of GdmCl ranging from 0 to ~7.4 M. Before use in SUPREX buffers, GdmCl was deuterated using four cycles of dissolution in D_2O and lyophilization. pH measurements were performed using a Jenco 6072 pH meter equipped with a Futura calomel pH electrode from Beckman Instruments (Fullerton, CA). The pD of each buffer was adjusted with sodium deuteroxide and deuterium chloride, and pH measurements were converted to pD measurements by adding 0.4 to the pH reading (54). The final concentrations of GdmCl in the buffers were measured with a refractometer (Bausch & Lomb, Rochester, NY) as previously described (24, 55).

2.4 Conventional SUPREX Protocol

Figure 1 summarizes the conventional SUPREX protocol. To perform SUPREX, a fully protonated protein or protein-ligand complex was diluted into a series of deuterated H/D exchange buffers. After a given H/D exchange time, an aliquot of the deuterated protein or protein-ligand complex was diluted into an ice-cold and low pH SA MALDI matrix solution. This step served to quench the H/D exchange reaction and prepare the sample for MALDI analysis. The saturated SA matrix solution was prepared by dissolving SA in 45% ACN/54.9% H_2O/0.1% TFA (v/v, pH ~ 2.0). A 1 μL aliquot of the protein or protein-ligand complex in the quenched H/D exchange buffer was spotted

onto a stainless steel MALDI target and allowed to dry under a gentle flow of air or N_2 before the sample was subjected to MALDI-MS analysis.

2.5 High Sensitivity SUPREX Protocol

In some cases (such as the AGT variants described in Chapter 3), a protein's ion signal is not easily detectable using MALDI-MS. This is often due to low protein concentrations or large amounts of salt in the sample that suppress the ion signal of the protein. To overcome this issue, a high sensitivity SUPREX protocol (39) was employed in which ZipTips (Millipore, Billerica, MA) were used to concentrate and desalt the samples prior to MALDI-MS analysis. As shown in Figure 2, this protocol was nearly identical to the conventional SUPREX protocol except that after the specified exchange time ended, the reaction was quenched with 1 μL of 10% TFA. Next, the proteins were extracted from the H/D exchange buffers using either C_4 or C_{18} ZipTips that were pre-equilibrated according to the manufacturer's instructions. Bound proteins were then washed twice with a solution of 0.1% TFA and eluted with the matrix solution (saturated SA in 50-75% ACN, 0.1% TFA, 24.9-49.9% water) directly onto the MALDI target. Internal standards were incorporated into the matrix solution prior to the elution step.

2.6 Single-Point SUPREX Protocol

The single-point SUPREX assay is very similar to conventional SUPREX except that instead of collecting data points at a variety of denaturant concentrations, a single

denaturant concentration is used. As shown in Figure 3, the protein sample is mixed with the library compound (or control), and SUPREX buffer is added to initiate H/D exchange. After the specified exchange time, the sample is either prepared for MALDI-MS analysis or subjected to a protease digestion prior to MALDI-MS analysis. In this work, single-point SUPREX was implemented in both a library screening (Chapter 4) and in a set of proof-of-concept experiments in which a protease digestion step was incorporated into the protocol (Chapter 6). The details of each of these protocols will be discussed in their respective chapters.

2.7 SUPREX-Protease Digestion Protocol

The SUPREX-protease digestion protocol that was employed in this work is summarized in Figure 4. The first step in performing the SUPREX-protease digestion protocol is to optimize the digestion conditions. This is typically accomplished by testing several digestion times and several protease concentrations. Optimal conditions are chosen by examining the MALDI mass spectrum of each digest; conditions were selected to produce the greatest number of fragments in the 1-10 kDa range. The specific conditions of each protease digestion are described in their respective chapters. To incorporate the protease digestion protocol into SUPREX, the H/D exchange reactions were quenched by adding an aqueous solution of HCl to each reaction mixture. A pepsin solution was added, and the digestion was allowed to proceed for a specified amount of time.

To prepare the samples for MALDI analysis, 2 μL of the digestion solution was added to 6 μL ice-cold SA matrix solution, and 1 μL of the resulting solution was spotted onto a MALDI target for mass spectral analysis. Spots were quickly dried under a gentle flow of N_2. These sample preparation steps (particularly spotting and drying the samples) were carried out as quickly as possible to minimize the back exchange of deuterons to protons.

2.8 SUPREX Data Analysis

To determine the mass-to-charge ratio for each protein or peptide peak, either a Microsoft Excel macro program or a MATLAB (The MathWorks, Inc., Natick, MA) script was used. These programs perform the following three steps: 1) a 19-point floating average smoothing of the data, 2) a one- or two-point calibration using the internal standard(s), and 3) a center of mass calculation for the peak of interest. The uptake of deuterons (i.e., Δmass) was calculated for each peptide by subtracting the undeuterated peptide mass from the partially or fully deuterated peptide mass. Ten spectra were collected at each denaturant concentration, and the Δmass values derived from these spectra were averaged. To generate SUPREX curves, the average Δmass values were plotted vs. GdmCl concentration. GdmCl concentrations were corrected to reflect the final concentration after mixing the SUPREX buffers with the equilibrated protein solutions.

SUPREX data were fit to a four-parameter sigmoidal equation using a nonlinear regression routine in SigmaPlot (Systat Software, Inc., San Jose, CA), and this fitting was used to extract a $C^{1/2}_{SUPREX}$ value (i.e., the concentration of denaturant at the transition midpoint) from each SUPREX curve. For quantitative SUPREX analyses, SUPREX data were collected at several exchange times for each protein and protein-ligand complex, and the resulting $C^{1/2}_{SUPREX}$ values were used with equation 2-1 to obtain *m*- and ΔG_f values from the SUPREX data (40).

$$-RT\left[\ln\frac{\frac{\langle k_{int}\rangle t}{0.693}-1}{\frac{n^n}{2^{n-1}}[P]^{n-1}}\right] = mC^{1/2}_{SUPREX} + \Delta G_f \qquad \text{Equation 2-1}$$

In equation 2-1, R is the gas constant, T is the temperature in Kelvin, $\langle k_{int}\rangle$ is the average intrinsic exchange rate of an unprotected amide proton, t is the H/D exchange time, *m* is defined as $\delta\Delta G_f/\delta$[Denaturant], ΔG_f is the folding free energy of the protein in the absence of denaturant, n is the number of protein subunits involved in the folding reaction (n = 1 for CypA; n = 2 for AGT), and [P] is the protein concentration expressed in n-mer equivalents. The left side of the equality in equation 2-1 will hereafter be referred to as ΔG_{app}.

Using equation 2-1, plots of ΔG_{app} vs. $C^{1/2}_{SUPREX}$ were generated for each protein and protein-ligand complex. These plots were fit using a linear regression routine, and the y-intercept and slope represent the ΔG_f and *m*-value, respectively. Values for $\langle k_{int}\rangle$ were calculated using the relationship $\langle k_{int}\rangle = 10^{pH-5}\ min^{-1}$ (36). To reduce the error of the

measurements, an averaged *m*-value (m_{avg}) was used to calculate $\Delta G_{f,avg}$ values (see Chapters 3 and 5) (43). These $\Delta G_{f,avg}$ values were ultimately used to calculate $\Delta\Delta G_{f,avg}$ values for the determination of dissociation constants.

Dissociation constants (K_d values) were calculated using equation 2-2.

$$K_d = \frac{[L]}{e^{-\Delta\Delta G_{f,avg}/NRT} - 1} \qquad \text{Equation 2-2}$$

In equation 2-2, [L] is the concentration of free ligand, $\Delta\Delta G_{f,avg}$ is the difference in $\Delta G_{f,avg}$ between the bound and unbound forms of the protein, and N is the number of independent equivalent binding sites (N = 2 for AGTmi). For all of the experiments in this work, the ligand was present in large excess relative to the protein (i.e., greater than tenfold excess); thus, [L] was taken as the total concentration of ligand in the H/D exchange reactions.

2.9 Single-Point SUPREX Data Analysis

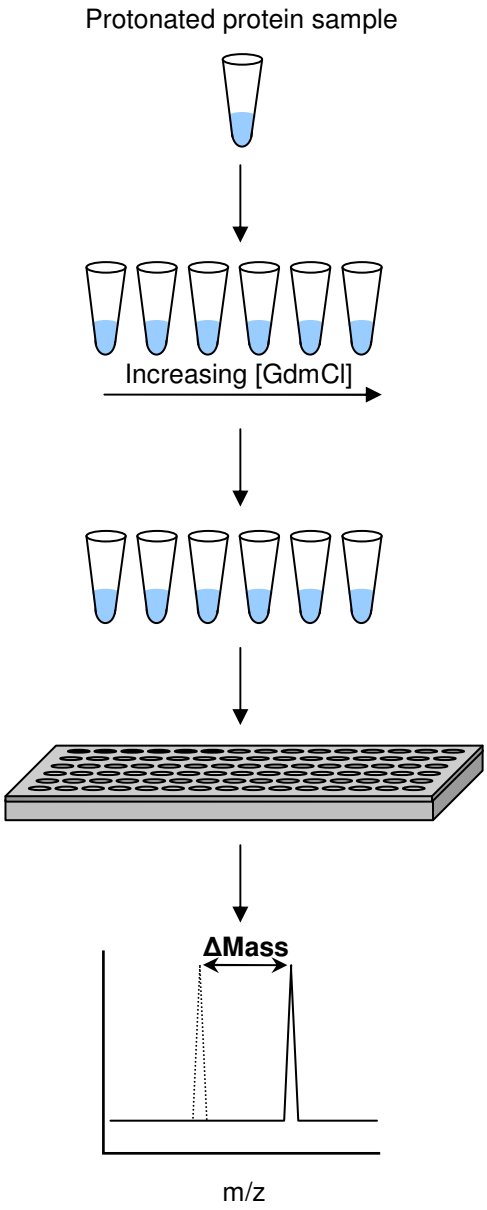
A 9-point central moving average was used to smooth the data. This manipulation involved the calculation of an unweighted average of each data point with the four previous and four subsequent data points. This smoothing step helped to adjust for day-to-day variations in experimental conditions (i.e., temperature and humidity) and helped to correct for differences in the degree of back exchange for each sample set. The averaged values were used to determine a unique cutoff value for each set of 10 library members.

The moving average strategy described above was used to calculate an average Δmass value for the negative control (i.e., a $\Delta mass_{av}$ value) in each set of 10 library compounds, and this $\Delta mass_{av}$ value was used to determine the cutoff for the corresponding data set. The cutoff was determined by subtracting a specified number of standard deviations (i.e., 3.0, 2.5, or 2.0 standard deviations) from the $\Delta mass_{av}$ value. A second MATLAB program, also developed in-house, was used to identify samples with Δmass values less than or equal to the cutoff value, and these samples were classified as hits.

Z′-factors were calculated using equation 2-3 as described in reference (56).

$$Z' = 1 - \frac{(3\sigma_{c+} + 3\sigma_{c-})}{|\mu_{c+} - \mu_{c-}|} \qquad \text{Equation 2-3}$$

In equation 2-3, σ_{c+} is the standard deviation of the positive control, σ_{c-} is the standard deviation of the negative control, μ_{c+} is the mean of the positive control, and μ_{c-} is the mean of the negative control.



1. Dilute into deuterated, denaturant-containing buffers.
2. Dilute into ice-cold, low pH matrix.
3. Spot samples on MALDI target.
4. Collect MALDI spectra.
5. Calculate Δmass.

Figure 1: Summary of the conventional SUPREX protocol.

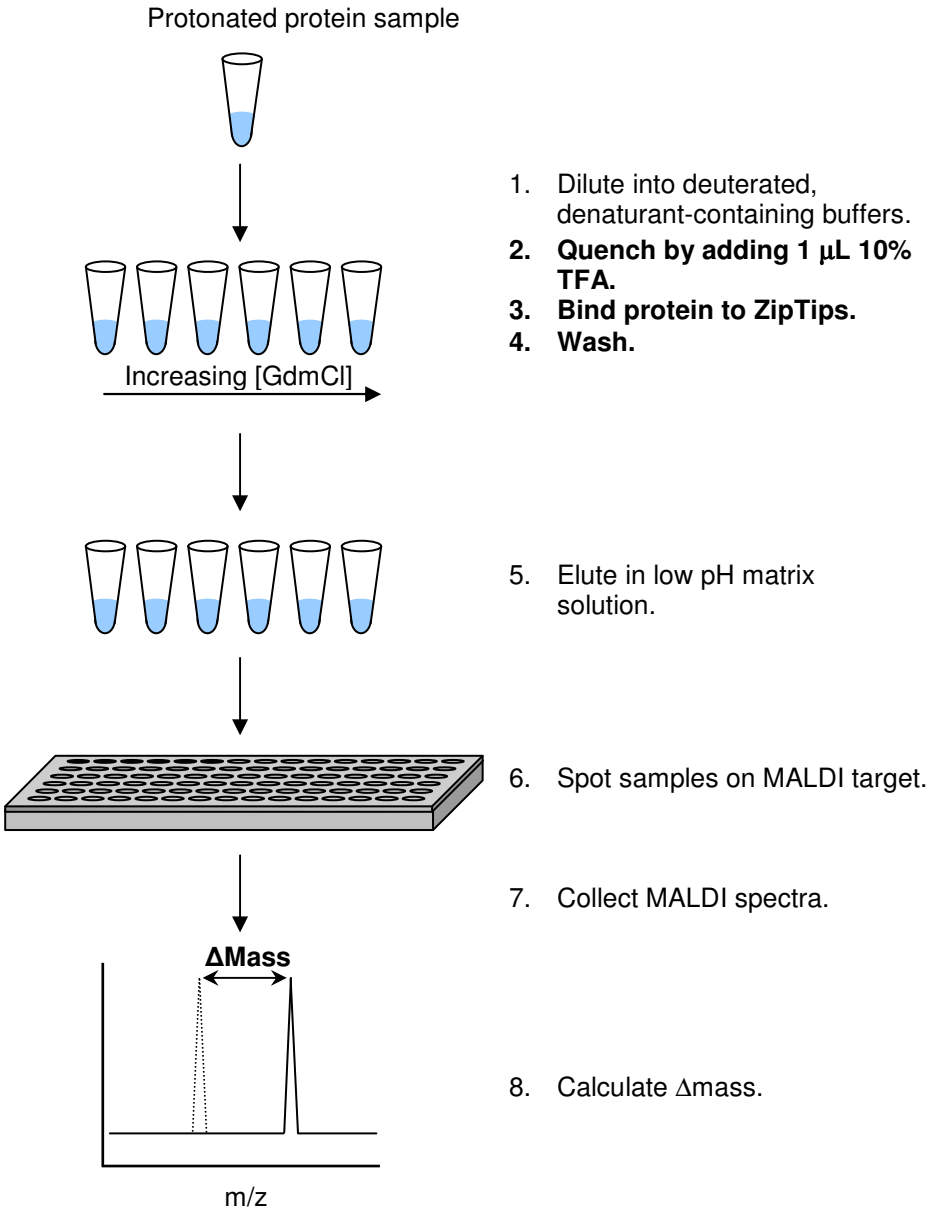


Figure 2: Summary of the high sensitivity SUPREX protocol. The text in bold indicates the differences between this protocol and the conventional SUPREX protocol.

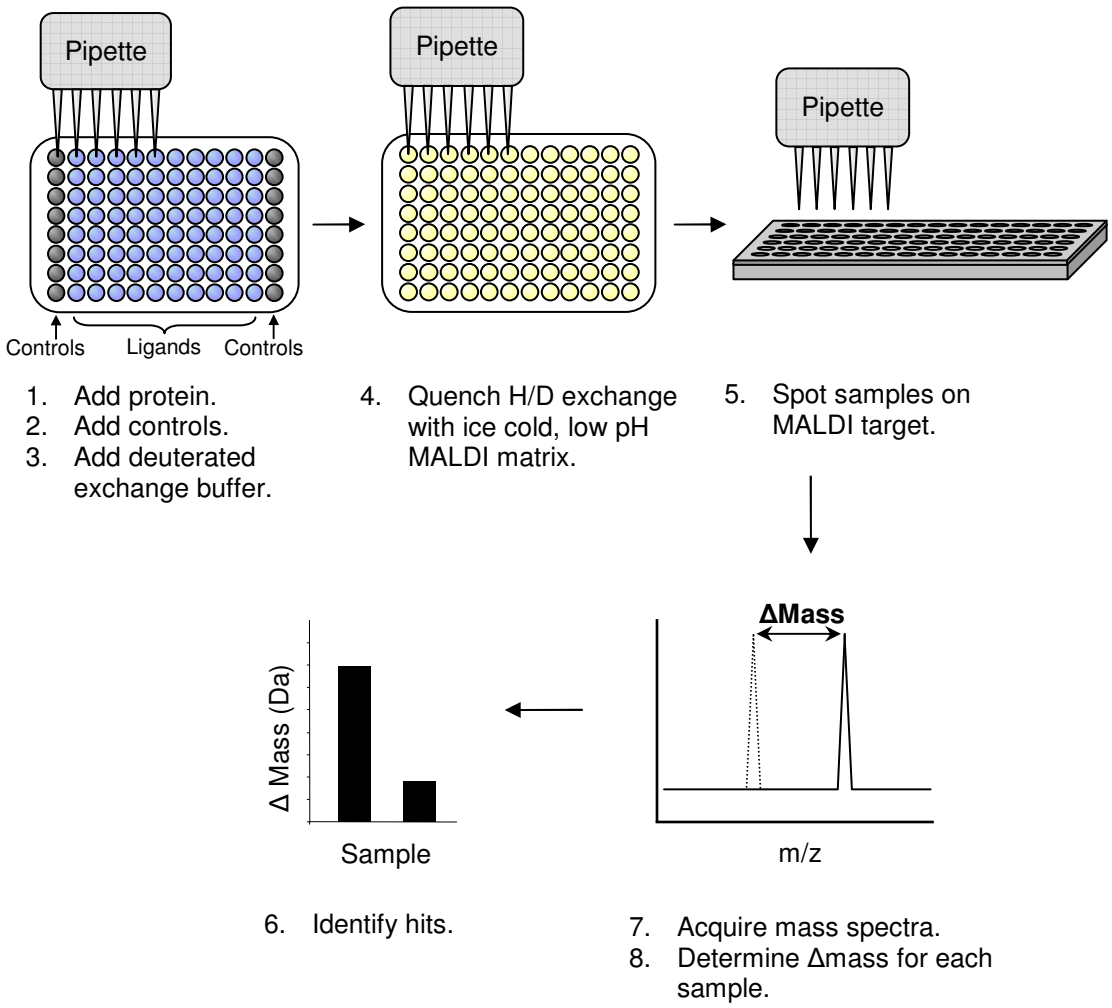


Figure 3: Summary of the single-point SUPREX protocol.

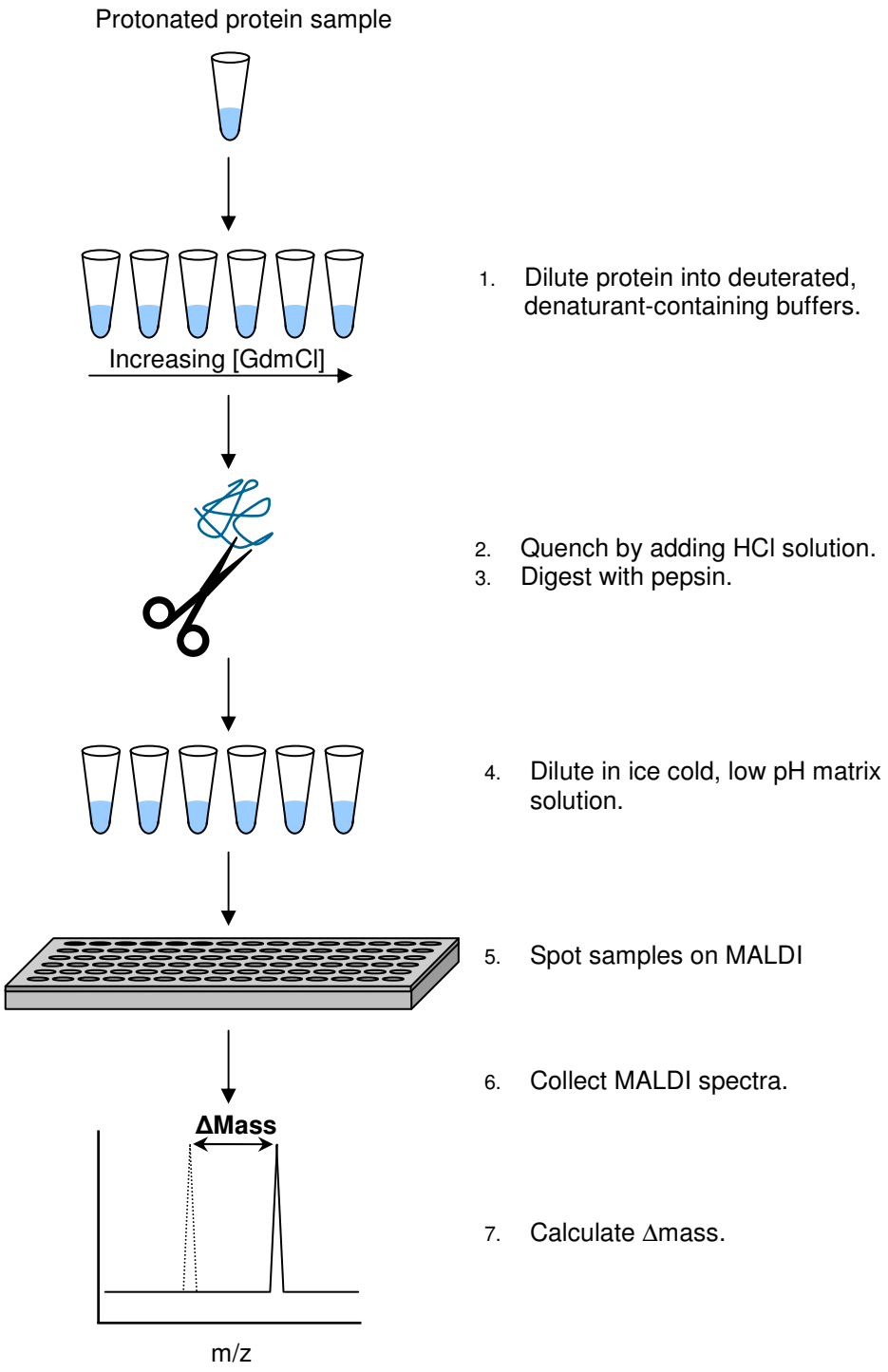


Figure 4: Summary of the SUPREX-protease digestion protocol.

3. Application of SUPREX to the Detection of Protein-Ligand Binding[1,2]

3.1 Introduction

The SUPREX technique is a relatively new method for characterizing the thermodynamic properties of protein-ligand complexes. This chapter describes work to evaluate the scope of the technique for the quantitative determination of thermodynamic parameters. One of the prerequisites for the evaluation of ΔG_f and *m*-values by any chemical denaturation-based technique is that the protein folding reaction must be well-modeled by a two-state folding mechanism (i.e., only folded or unfolded states are significantly populated during the protein folding reaction) (24). At the outset of this dissertation project, the Fitzgerald Group had previously demonstrated the utility of SUPREX for measuring the thermodynamic parameters of two-state folding proteins (36, 40, 46, 57, 58). However, the folding reactions of many proteins are much more complicated and cannot be described by a two-state folding mechanism.

From the results of this chapter and other work (45, 59), it has become apparent that for SUPREX, two-state folding is only a prerequisite when the technique is being

---

[1]Reproduced in part with permission from Tang, L.; Hopper, E. D.; Tong, Y.; Sadowsky, J. D.; Peterson, K. J.; Gellman, S. H.; Fitzgerald, M. C. H/D Exchange- and Mass Spectrometry-Based Strategy for the Thermodynamic Analysis of Protein-Ligand Binding. *Anal. Chem.* **2007**, *79*(15), 5869-5877. Copyright 2007 American Chemical Society.

[2] Reproduced in part with permission from Hopper, E. D.; Pittman, A. M. C.; Fitzgerald, M. C.; Tucker, C. L. In Vivo and in Vitro Examination of Stability of Primary Hyperoxaluria-Associated Human Alanine:Glyoxylate Aminotransferase. *J. Biol. Chem.*, **2008**, *283*(45), 30493-30502. Copyright 2008 American Society for Biochemistry and Molecular Biology.

used to determine quantitative ΔG_f and *m*-values. This rule can be relaxed when SUPREX is used for qualitative determinations of stability and for quantitative determinations of ligand binding affinities. For example, FbpA, a known non-two-state protein, and CypA, a suspected non-two-state protein, were successfully analyzed using SUPREX (45, 59), but prior to this work, the scope of the technique for the analysis of non-two-state proteins had not been established.

A further examination of the scope of SUPREX for the analysis of non-two-state proteins involved studying two additional proteins that are not expected to be two-state folders: Bcl-x_L and AGT. Both of these proteins are potential targets for HTS campaigns. Bcl-x_L is an anti-apoptotic protein that has been implicated in multiple types of cancer, and recent studies have suggested that inhibition of Bcl-x_L can restore apoptosis in some cancer types (60, 61). One goal of this work was to assess SUPREX as a technique for evaluating relative ligand binding affinities for Bcl-x_L ligands.

AGT is a target for screening campaigns due to its involvement in PH1, a disorder in which deficiencies in AGT lead to a defect in metabolism that results in a severe kidney stone disease. Deficiencies in AGT may occur for a variety of reasons, but for many patients, this deficiency arises from a small number of amino acid substitutions that result in decreased enzyme stability. Due to this decrease in stability, the enzyme may aggregate or become degraded or improperly localized (62, 63). These characteristics make AGT a promising candidate for therapeutic intervention via

pharmacological chaperones (i.e., stabilizing ligands capable of restoring native protein structure). The first goal of this study was to characterize the relative stabilities of four AGT variants. These variants included the non-disease-causing major and minor allelic forms of AGT (AGTma and AGTmi) and two disease-causing AGT mutants (AGT244 and AGT170). The ability of SUPREX to detect the binding of AGT variants to two small molecule ligands was also assessed.

The primary objective of the work in this chapter was to evaluate the ability of SUPREX to quantify ligand binding to Bcl-x_L and AGT. To test the ability of SUPREX to detect Bcl-x_L-ligand binding, the binding of two peptide ligands was assessed. These ligands were derived from Bak, a pro-apoptotic protein that interacts with Bcl-x_L. In the case of AGT, two known small molecule ligands (the cofactor pyridoxal 5′-phosphate (PLP) and the inhibitor aminooxyacetic acid (AOA)) were used to test the ability of SUPREX to analyze ligand binding to AGT. Also tested was the ability of SUPREX to detect the ligand binding interactions of several AGT variants. In the cases of both Bcl-x_L and AGT, SUPREX was an effective method for detecting ligand binding. This study helped to broaden the scope of SUPREX and to lay the groundwork for screening campaigns using single-point SUPREX.

3.2 Experimental Section

3.2.1 Materials

The following materials were purchased from Sigma-Aldrich (St. Louis, MO): trypsin inhibitor from soybean, aldolase from rabbit muscle, pyridoxal 5′-phosphate (PLP), and aminooxyacetic acid (AOA). The Microcon centrifugal filter device (YM-3, membrane NMWL 3000) and ZipTips were from Millipore (Billerica, MA).

Human Bcl-x_L was kindly provided by the Gellman Laboratory (University of Wisconsin, Madison). The Bcl-x_L construct used in this work contained residues 1-196 and lacked the C-terminal hydrophobic region that is present in the wild-type protein. The construct was expressed in *E. coli* and purified using an N-terminal 6xHis tag. The final purified Bcl-x_L sample had a concentration of 7 μM and was stored at 4 °C in a PBS buffer containing 2 mM DTT and 2 mM EDTA. The solution was concentrated with a Microcon centrifugal filter device with a 3,000 Da cutoff membrane by centrifuging for 99 min at 14,000 x *g*. The final concentration of Bcl-x_L was determined by measuring the absorbance (280 nm) of Bcl-x_L in an 8 M GdmCl solution using a molar extinction coefficient of 41940 $M^{-1}cm^{-1}$ (64). The concentration was found to be 130 μM.

The Bcl-x_L ligands are Bak-derived peptides that will be referred to hereafter as Bak1, which had an amino sequence of H_2N-GQLGRQLAIIGDDINR-$CONH_2$ and Bak2, which had an amino acid sequence of H_2N-GQLGRQLAI-Norleucine-GDDFNR-$CONH_2$. These peptides were provided by the Gellman Laboratory and were obtained by

standard solid-phase peptide synthesis as previously described (65). The Bak1 and Bak2 peptides were stored as a lyophilized powder and were reconstituted in water to a final concentration of 1.3 mM. After reconstitution, the peptides were stored at 4 °C.

The minor allele AGT variant was prepared by Adrianne M.C. Pittman from the laboratory of Dr. Chandra L. Tucker. Briefly, the AGTmi gene was cloned into the Xho I and Nde I sites of a pET-30a vector. The clone was verified by sequencing, and the vector was transformed into BL21(DE3) cells (Novagen). The resulting bacteria was grown to an OD_{600} of 0.8 and induced with 1 mM isopropyl-beta-D-thiogalactopyranoside (IPTG) for 6 hours at room temperature. Harvested cells were stored in the freezer for later use. To obtain purified AGTmi, cells were lysed with Y-PER (Pierce Scientific) and purified using the Y-PER 6xHis Fusion Protein Purification Kit (Pierce Scientific). A Bradford Assay was used to determine protein concentration, and aliquotted samples were stored in elution buffer (50 mM Tris, 300 mM NaCl, 200 mM imidazole, 10% glycerol, pH 6.8) at -80 °C.

3.2.2 Preparation of Protein Samples and SUPREX Buffers

The Bcl-x_L-peptide complexes were prepared directly in the SUPREX buffers. Aliquots of the 130 μM Bcl-x_L stock solution were diluted 11-fold into H/D exchange buffers that already contained the peptide ligand (i.e., either the Bak1 or the Bak2 peptide). The final protein and ligand concentrations in each H/D exchange buffer were

12 and 120 μM, respectively. The H/D exchange buffers for Bcl-x_L contained GdmCl concentrations ranging from 0.7 M to 6.3 M.

SUPREX experiments on AGT variants were performed using a series of deuterated H/D exchange buffers containing 0 to 7.6 M deuterated GdmCl. Prior to SUPREX analysis, AGT samples (13 μM) were preincubated for 1 hour at room temperature in PBS (pH 7.4) in the presence or absence of ligands (1.3 mM). In some cases, AGTma samples were equilibrated overnight at 4 °C, but this did not result in an observable change in the SUPREX results.

3.2.3 Instrumentation

MALDI mass spectra for Bcl-x_L and the AGT variants were acquired on an UltraFlex II TOF/TOF (Bruker Daltonics, Inc., Billerica, MA) mass spectrometer in the linear and positive ion mode. Spectra for Bcl-x_L were collected using the following parameters: 25 kV ion source 1 voltage, 23.4 kV ion source 2 voltage, 6 kV lens voltage, 100 ns pulsed ion extraction, and matrix gating to 3000 Da. Spectra for AGT variants were collected using a slightly different set of parameters, as follows: 25 kV ion source 1 voltage, 23.3 kV ion source 2 voltage, 6.75 kV lens voltage, 130 ns pulsed ion extraction, and matrix gating to 9000 Da.

3.2.4 SUPREX Data Collection

The SUPREX protocol for the work included in this chapter is described in Chapter 2 (40). For Bcl-x_L, the fully protonated protein or protein-ligand complex was

diluted into a series of deuterated H/D exchange buffers (see Section 3.2.2). After a given H/D exchange time, an aliquot of the deuterated protein or protein-ligand complex was diluted at least 4-fold into an ice-cold and low pH MALDI matrix solution. This step quenched the H/D exchange and prepared the sample for MALDI analysis. The saturated SA matrix solution was prepared by dissolving SA in 45% ACN/54.9% H_2O/0.1% TFA (v/v, pH ~ 2.0). The matrix solution was kept cold on crushed ice before use.

In the Bcl-x_L experiments, a two-layer spotting method was used for MALDI sample preparation. This method involved spotting the quenched protein or protein-ligand complex onto a pre-deposited layer of matrix crystals (66). This pre-deposited layer was formed on the MALDI target by depositing ~1 μL of a saturated SA solution prepared in a 40% acetone and 60% methanol (v/v) mixture and allowing the solvent to evaporate. A 1 μL aliquot of the protein or protein-ligand complex in the quenched H/D exchange buffer was spotted onto the first layer of matrix crystals and allowed to dry under a gentle flow of air or N_2 before the sample was introduced into the vacuum of the MALDI mass spectrometer.

For AGT variants, H/D exchange reactions were initiated by adding 7 μl of each exchange buffer to 3 μl of protein solution, resulting in a final protein concentration of 4 μM and final ligand concentrations of 400 μM. The H/D exchange reactions were carried out at room temperature for either 5 min or 1 h and quenched with 1 μl 10% TFA. The

protein was concentrated and desalted using C_4 or C_{18} ZipTips. Protein samples were eluted from the ZipTips directly onto a MALDI target using the MALDI matrix solution, which consisted of saturated SA in 0.1% TFA/50-75% ACN/24.9-49.9% water. The matrix solution also contained aldolase as an internal standard. Sample spots were dried under a gentle flow of N_2 and then analyzed by MALDI. Each mass spectrum collected for Bcl-x_L and AGTmi was the sum of the signal obtained from 100 laser shots. A total of 10 replicate mass spectra were collected for the sample in each H/D exchange buffer to obtain an average mass value for the protein in each buffer.

3.2.5 SUPREX Data Analysis

SUPREX data analysis was performed as described in Chapter 2. The concentrations of GdmCl in each SUPREX curve were adjusted to reflect the final concentration after mixing the exchange buffer with the protein solution. The standard deviations associated with the Δmass measurements were typically less than 10%. This was consistent with the ~0.05 to 0.1% MALDI-TOF mass accuracies observed for the proteins in this study.

For Bcl-x_L and its complexes, a series of SUPREX curves was generated using different exchange times. The $C^{1/2}_{SUPREX}$ values obtained at each H/D exchange time were fit to equation 2-1 (with n = 1) using a linear least-squares analysis. The y-intercept and the slope of the best-fit line were taken as the ΔG_f value and the *m*-value, respectively. The $\langle k_{int} \rangle$ values used in this work were predicted using the relationship $\langle k_{int} \rangle = 10^{pH-5}$

min^{-1} (36). For simplicity, the left side of the equality in equation 2-1 is hereafter referred to as $\Delta G_{app.}$

In theory, the ΔG_f values, as determined above, could be used to determine $\Delta\Delta G_f$ values (i.e., the change in folding free energy induced by ligand binding). However, in this study, $\Delta G_{f,avg}$ values were used to calculate $\Delta\Delta G_{f,avg}$ values (see Section 3.3) and quantify the ligand binding affinities. Calculations of $\Delta G_{f,avg}$ values involved averaging the *m*-values from multiple ΔG_{app} vs. $C^{1/2}_{SUPREX}$ value plots for a given protein system. The average *m*-value was used in equation 2-1 to directly calculate a ΔG_f value from each $C^{1/2}_{SUPREX}$ value that was collected for Bcl-x_L and its complexes. In this manner, at least four independent ΔG_f values were calculated for each protein complex. These values were ultimately averaged to generate $\Delta G_{f,avg}$ values.

For AGT, the primary goal was to compare the relative stabilities and ligand binding interactions of four variants. Since this work was qualitative rather than quantitative, a single H/D exchange time was used for each protein or protein complex. Relative stabilities were evaluated by comparing the transition midpoints (i.e., $C^{1/2}_{SUPREX}$ values) of the SUPREX curves for the variants and their complexes.

3.3 Results and Discussion

3.3.1 SUPREX Analysis of Apoproteins

SUPREX analyses were performed on apo-Bcl-x_L using H/D exchange times that ranged from 30 min to 280 min. The specific H/D exchange times and resulting $C^{1/2}_{SUPREX}$

values for each protein are included in Tables 1 and 2. The H/D exchange times used in the SUPREX experiments were different for each protein system, but in each case they were chosen such that a SUPREX transition could be detected. In theory, any H/D exchange time can be used for SUPREX analyses. However, in practice, the H/D exchange times that can be used for a given protein system are limited to those that generate experimentally accessible $C^{1/2}_{SUPREX}$ values. For example, the H/D exchange time cannot be so long as to yield a $C^{1/2}_{SUPREX}$ value at negative chemical denaturant concentrations, and it cannot be so short as to yield a $C^{1/2}_{SUPREX}$ value at a denaturant concentration at which the denaturant is not soluble. The useful range of H/D exchange times varies for different protein systems depending on the folding thermodynamics of the system. For example, relatively long exchange times can be used to generate experimentally measurable $C^{1/2}_{SUPREX}$ values if the protein has a large negative ΔG_f and/or large *m*-value associated with its folding reaction (see equation 2-1).

Typical SUPREX curves obtained for apo-Bcl-x_L are shown in Figure 5A. A single cooperative transition was detected in all of the SUPREX curves, but the SUPREX curve transition midpoints for apo-Bcl-x_L were relatively constant with H/D exchange time. This is in contrast to many of the proteins previously studied using SUPREX, in which the protein's $C^{1/2}_{SUPREX}$ values are expected to move to lower denaturant concentrations when the H/D exchange time is increased (see equation 2-1). Since Bcl-x_L does not follow this trend, it is not considered to be a well-behaved protein for SUPREX.

Well-behaved protein systems for SUPREX include those that exhibit reversible, two-state unfolding properties and EX2 exchange behavior. In the case of such well-behaved proteins, equation 2-1 can be used to describe the relationship between the $C^{1/2}_{SUPREX}$ value and the H/D exchange time used in SUPREX (40). Equation 2-1 predicts that a plot of ΔG_{app} vs. $C^{1/2}_{SUPREX}$ will be linear and that the y-intercept and slope of the resulting plot will correspond to the ΔG_f and *m*-value, respectively, of the protein's folding/unfolding reaction. Equation 2-1 clearly does not describe the SUPREX behavior of apo-Bcl-x_L because the SUPREX curve transition midpoints for this protein did not significantly change with exchange time.

Similar to Bcl-x_L, the $C^{1/2}_{SUPREX}$ values measured for apo-AGT did not vary with exchange time (as tested using AGTmi; see Figure 6). This SUPREX behavior was not surprising given the relatively large size of AGT (see below). Typical SUPREX data obtained on apo-AGTma and apo-AGTmi are shown in Figure 7A, and the average $C^{1/2}_{SUPREX}$ values obtained from three to four replicate measurements are shown in Table 2. The average $C^{1/2}_{SUPREX}$ value for AGTma was consistently greater than the average $C^{1/2}_{SUPREX}$ value for AGTmi, indicating that AGTma is more thermodynamically stable than AGTmi. This finding is in agreement with a previous study in which AGTmi was found to be more aggregation-prone than AGTma when expressed in bacteria (67). However, this result is somewhat surprising given that neither AGTma nor AGTmi is associated with PH1.

SUPREX analysis of the disease-associated mutant AGT244 (which contains an I244T mutation in the context of the minor allele) revealed that this mutant was similar in stability to AGTmi (see Table 2 and Figures 7 and 8). The similarity in stability is not particularly surprising given that this mutation appears in the context of the minor allele. However, this result was contrary to expectations; since this variant is associated with PH1, it was expected to be less stable than AGTmi. These results suggest that either 1) the stability difference is too small to be detectable by SUPREX, or 2) factors other than thermodynamic stability play a role in PH1 for patients with the I244T mutation.

Attempts to analyze an additional disease-associated mutant, AGT170 (which contains a G170R mutation in the context of the minor allele) were not successful presumably due to the rapid aggregation and precipitation of the purified protein. This observation is consistent with AGT170 being significantly destabilized compared to the wild-type protein. However, the extent of destabilization could not be determined, as a SUPREX transition was not observed for either apo-AGT170 or holo-AGT170.

3.3.2 SUPREX Analysis of Holoproteins

SUPREX analyses were performed on Bcl-x_L and AGT variants in the presence of known ligands of each protein. Bcl-x_L was analyzed in the presence of the Bak1 peptide and in the presence of the Bak2 peptide (see Figures 5B and 5C). No analysis was performed on Bcl-x_L in solution with both Bak1 and Bak2 because these peptides bind in the same location on Bcl-x_L and do not bind simultaneously. Four AGT variants were

analyzed, including the major and minor allele forms and two disease-associated variants. Each of these variants was analyzed for binding to PLP (a cofactor) and AOA (an inhibitor) except for AGT170, which precipitated too quickly to allow for SUPREX analysis, even in the presence of ligands (see above).

Tables 1 and 2 include a summary of the $C^{1/2}_{SUPREX}$ values extracted from all of the SUPREX curves generated for the different protein-ligand complexes in this work. Similar to the apoprotein SUPREX analyses, the range of H/D exchange times selected for Bcl-x_L varied for each protein-ligand complex (see above). For AGT variants, the intention was to qualitatively assess protein stabilities and ligand binding affinities, so a single exchange time was used for each complex. The $C^{1/2}_{SUPREX}$ values obtained for the holoproteins were all at higher denaturant concentrations than the $C^{1/2}_{SUPREX}$ values obtained for the respective apoproteins for a given H/D exchange time (see Figures 5 and 7 through 8 and Tables 1 and 2). Such $C^{1/2}_{SUPREX}$ value shifts are consistent with the ligand-induced stabilization of each protein (i.e., since the proteins are stabilized through ligand binding, a greater concentration of denaturant is required to unfold the holoprotein than the apoprotein).

3.3.3 Quantitative Analysis of Ligand Binding (Bcl-x_L)

Shown in Figure 5C are the ΔG_{app} vs. $C^{1/2}_{SUPREX}$ plots for the Bcl-x_L complexes in this study. In marked contrast to the data collected on the apoprotein, the $C^{1/2}_{SUPREX}$ values obtained for each of the protein-ligand complexes in Figure 5C fit well to

equation 2-1. The ligand-induced stabilization of Bcl-x_L helped to promote good SUPREX behavior, resulting in a linear correlation between ΔG_{app} vs. $C^{1/2}_{SUPREX}$ as predicted by equation 2-1. This linear relationship as shown in Figure 5C made it possible to generate a ΔG_f and *m*-value for each Bcl-x_L complex (see Table 3).

A reasonable hypothesis to explain the poor behavior of the apoprotein is that the H/D exchange behavior of this protein is dominated by local unfolding events rather than global unfolding events (43). However, in the presence of ligand, the protein is stabilized, which results in a reduction in the dominance of local unfolding events. This allows the H/D exchange behavior of the protein to be dominated by a global unfolding reaction and enables the protein to demonstrate good SUPREX behavior when it is bound to a ligand.

The accurate measurement of ΔG_f and *m*-values using equation 2-1 requires an assumption of reversible, two-state folding (40). While biophysical studies on a related Bcl-x_L construct suggest a possible two-state folding mechanism for Bcl-x_L (68), it is noteworthy that the SUPREX curve transitions recorded on the Bcl-x_L construct in this study were significantly less cooperative than would be expected for a two-state folding protein the size of Bcl-x_L. The low cooperativity of the SUPREX transition suggests that this Bcl-x_L construct is not a two-state folder. However, the motivation for deriving such ΔG_f and *m*-values for Bcl-x_L was to determine if they could be used in comparative analyses to quantify the binding affinities of different ligands to this protein (see below).

The *m*-values recorded for Bcl-x_L were not expected to change upon binding to either ligand. Indeed, significant changes in *m*-value (i.e., differences greater than 30%) were not detected in the SUPREX experiments on Bcl-x_L (see Table 3). Protein folding *m*-values have been shown to correlate with the amount of hydrophobic surface area that is buried in a protein folding reaction (69). Large proteins with more buried hydrophobic surface area in their native three-dimensional structures have larger *m*-values than small proteins with less buried hydrophobic surface area. The protein-ligand interactions described in this work do involve some burial of hydrophobic surface area. However, the amount of hydrophobic surface area that is buried in these binding interactions is small compared to the thousands of square angstroms that would be required to produce a change in *m*-value of more than 30% (69). Indeed, the *m*-value differences between the two Bcl-x_L complexes were clearly within the experimental error.

The SUPREX analyses of Bcl-x_L enabled a quantitative determination of ligand binding affinity. SUPREX-derived $\Delta\Delta G_{f,avg}$ values were determined for Bcl-x_L binding the Bak2 peptide (see Table 3). The Bak1 and Bak2 ligands in the Bcl-x_L system bind to the same site in the protein (70). Thus, the two ligands cannot simultaneously bind to Bcl-x_L during SUPREX analysis. This precludes the determination of K_d values for either of the ligands (43). However, note that the differential binding affinity of the Bak1 and Bak2 peptides for Bcl-x_L could be quantified. A comparison of the SUPREX-derived $\Delta G_{f,avg}$ values for the Bcl-x_L-Bak1 and Bcl-x_L-Bak2 complexes revealed that the binding affinity

of Bak2 is 1.6 ± 0.2 kcal/mol greater than the binding affinity of Bak1. This is in reasonable agreement with the differential binding affinity expected for the two peptides (2.0 kcal/mol) based on their relative K_i values (70).

The calculations of the above ligand binding affinities (i.e., the $\Delta\Delta G_{f,avg}$ values) employed the $\Delta G_{f,avg}$ values in Table 3 rather than the ΔG_f values. In theory, the ΔG_f values could be used to quantify protein-ligand binding affinities (i.e., to calculate $\Delta\Delta G_f$ values). Indeed, $\Delta\Delta G_f$ values calculated in this manner are very similar to the $\Delta\Delta G_{f,avg}$ values reported in Table 3. However, the error associated with such $\Delta\Delta G_f$ calculations is large (~1 kcal/mol) and close in magnitude to the $\Delta\Delta G_f$ value itself. The magnitude of this error is not consistent with the raw data collected in the SUPREX curves. For example, the two ΔG_f values obtained for the Bcl-x_L complexes are the same within experimental error (see Table 3), but the 0.6 M change in the transition midpoint between the two Bcl-x_L SUPREX curves shown in Figure 5B is clearly measurable given the error of the $C^{1/2}_{SUPREX}$ value measurements (~ ±0.1 M).

The error associated with the calculation of ΔG_f values from $C^{1/2}_{SUPREX}$ values using equation 2-1 can be significantly reduced if the *m*-value is accurately established (59). One way to establish an accurate *m*-value is to average the *m*-values obtained from multiple ΔG_{app} vs. $C^{1/2}_{SUPREX}$ value plots. For example, since the *m*-values were not expected to change with ligand (see above), the *m*-values reported for each protein complex in Table 3 can be used to obtain an average *m*-value for Bcl-x_L. With such an

established *m*-value (i.e., 1.9 kcal mol^{-1} M^{-1}), equation 2-1 could be used for the direct calculation of a ΔG_f value from each $C^{1/2}_{SUPREX}$ value. In this manner, it was possible to calculate multiple independent ΔG_f values for each protein complex and then average the values to generate $\Delta G_{f,avg}$ values (see Table 3). The standard deviations associated with these $\Delta G_{f,avg}$ values better reflect the precision of the measurements. Note also that the precision of the $\Delta G_{f,avg}$ values is comparable to that of similar measurements using conventional spectroscopy-based methods.

The quantitative analyses of ligand binding to Bcl-x_L presented above involve the use of equation 2-1 to analyze the SUPREX data on these protein-ligand complexes. The derivation of equation 2-1 (see Appendix 1 in reference (40)) assumes EX2 exchange conditions for the protein under study. The accuracy of SUPREX-derived ΔG_f and *m*-values can be compromised on protein systems when equation 2-1 is used to analyze SUPREX data obtained under non-EX2 exchange conditions (71). It has also been shown that in such cases where the accuracy of SUPREX is compromised due to non-EX2 exchange behavior, there is a characteristic non-linearity observed in the ΔG_{app} vs. $C^{1/2}_{SUPREX}$ plots at high denaturant concentrations (71). The ΔG_{app} vs. $C^{1/2}_{SUPREX}$ plots in Figure 5C were found to be linear, suggesting that the H/D exchange behavior of Bcl-x_L is in the EX2 regime. It is also important to note that the MALDI mass spectra generated in these SUPREX experiments were consistent with EX2 exchange behavior (i.e., only one population of protein ions was detected). The two populations of protein ions

typically observed in the mass spectral analysis of a protein under EX1 exchange conditions were not observed, although the resolution of the mass spectrometer used in this study was typically limited to ~500 (as defined using the full width at half maximum).

3.3.4 Qualitative Analysis of Ligand Binding (AGT)

The goal of the ligand binding experiments with AGT was to evaluate relative stabilities of several ligand-bound variants. Since this study was qualitative, a single H/D exchange time was used to evaluate the stabilities of AGT variants complexed with either PLP or PLP-AOA. As shown in Figures 7 and 8, PLP binding resulted in an increase in thermodynamic stability for AGTma, AGTmi, and AGT244. A further stabilization was observed when both PLP and AOA were added. Similar to the AGT apoproteins, the stabilities of ligand-bound AGTmi and AGT244 were quite similar. As shown in Table 2, AGTma-PLP-AOA was significantly more stable than either AGTmi-PLP-AOA or AGT244-PLP-AOA.

The Δmass values for AGT in the presence of PLP were higher than the Δmass values for AGT without added ligand (Figure 7) due to a covalent linkage between AGT and PLP. Unlike noncovalent complexes that dissociate in the MALDI matrix solution, covalent bonds generally persist throughout the sample preparation protocol. The Δmass values for AGT-PLP-AOA were similar to those for AGT without added ligand

due to the reaction between PLP and AOA, which breaks the Schiff base linkage between AGT and PLP (72).

Since PLP and AOA are known to react with each other (72, 73), it was important to evaluate whether AOA is capable of binding AGT in the absence of added PLP. No shift in the SUPREX transition midpoint was detected for AGTma when AOA was added in the absence of additional PLP, suggesting that AOA either cannot bind to AGTma without PLP or it binds too weakly to observe a detectable shift in the transition midpoint.

SUPREX analysis of AGTma, AGTmi, and AGT244 in the presence of both PLP and AOA (PLP-AOA) revealed a significant increase in the average $C^{1/2}_{SUPREX}$ values over those seen with PLP alone. Due to this considerable stabilization effect, it was not possible to observe a SUPREX transition using a 5 minute H/D exchange time as was used with PLP only. Thus, it was necessary to increase the H/D exchange time to 1 hour. By increasing the exchange time, the transition midpoint of the SUPREX curve was moved into an experimentally accessible GdmCl concentration range. Although a full quantitative analysis of AGT was not pursued in this study, the ability to move the SUPREX transition midpoint of ligand-bound AGT by adjusting exchange time suggests that, similar to Bcl-x_L, ligand binding may help AGT become amenable to quantitative ligand binding analysis by SUPREX. A full, quantitative analysis of ligand binding for one of the AGT variants is discussed in Chapter 5.

3.4 Conclusions

Taken together, the SUPREX analyses of Bcl-x_L and AGT variants demonstrate that SUPREX can be used for the qualitative and quantitative analysis of ligand binding affinities for non-two-state folding proteins. For some proteins such as Bcl-x_L that are not well-behaved in SUPREX experiments, quantitative ligand binding analysis is facilitated through the use of a stabilizing ligand. A similar strategy could likely have been used for the analysis of the AGT variants, but this dissertation work involved employing a different strategy for the quantitative analysis of AGT (see Chapter 5). These studies serve to expand the scope of SUPREX by demonstrating that this technique can be used for the quantitative and/or qualitative analysis of non-two-state proteins. Most importantly, this study demonstrated the capability of SUPREX to detect ligand binding in two medically significant proteins that are potential targets for screening campaigns to identify novel ligands. This demonstration was an important first step in the development of SUPREX as a technique for high throughput screening assays.

Table 1: Transition midpoints of SUPREX curves obtained for Bcl-x_L and its complexes.

Exchange time (min)	Apo-Bcl-x_L[a] (M, GdmCl)	Bcl-x_L+Bak1[a,b] (M, GdmCl)	Bcl-x_L+Bak2[a,b] (M, GdmCl)
30	-	3.2 (-11.5)	-
70	-	2.9 (-11.3)	-
80	2.6	-	-
90	-	-	3.8 (-13.2)
110	-	2.8 (-11.6)	-
120	2.6	-	-
130	2.5[c]	-	3.4 (-12.8)
150	-	2.8[c] (-11.6)	-
160	2.6	-	-
170	-	-	3.5 (-13.1)
210	-	-	3.3 (-12.9)
240	2.5	-	-
280	2.3	-	-
1050	-	-	2.8 (-12.9)
1320	-	-	2.9 (-13.1)
1430	-	-	3.0 (-13.3)

[a]The error associated with each value is estimated to be ~ ±0.1 M, which is the standard deviation of the $C^{1/2}_{SUPREX}$ values obtained from 3 replicate SUPREX curves for apo-Bcl-x_L at the 80-minute H/D exchange time. Transition midpoints are reported as average $C^{1/2}_{SUPREX}$ values from 2-3 SUPREX curves except where noted.

[b]Numbers in parenthesis are free energy (ΔG_f) values in kcal mol^{-1}. These free energies were calculated using equation 2-1, an average *m*-value of 1.9 kcal mol^{-1} M^{-1} (see text), and the tabulated $C^{1/2}_{SUPREX}$ value.

[c]This $C^{1/2}_{SUPREX}$ value is based on a single SUPREX curve.

Table 2: $C^{1/2}_{SUPREX}$ values for AGT variants.

	AGTma $C^{1/2}_{SUPREX}$ [GdmCl] (M)[a]	AGTmi $C^{1/2}_{SUPREX}$ [GdmCl] (M)[a]	AGT244 $C^{1/2}_{SUPREX}$ [GdmCl] (M)[a]
AGT[b]	1.4 ± 0.1	1.0 ± 0.1	1.2 ± 0.1
AGT + PLP[b]	2.7 ± 0.1	2.4 ± 0.2	2.1 ± 0.2
AGT + PLP + AOA[c]	3.2 ± 0.1	2.7 ± 0.1	2.5 ± 0.2

[a]Values are the average and standard deviation obtained from three to four replicate measurements.

[b]Exchange time = 5 min

[c]Exchange time = 1 hour

Table 3: SUPREX-derived thermodynamic parameters obtained on the model protein systems in this study.

Protein Complex	ΔG_f[a] (kcal mol^{-1})	m[a] (kcal mol^{-1} M^{-1})	$\Delta G_{f,avg}$[b] (kcal mol^{-1})	$\Delta\Delta G_{f,avg}$[c] (kcal mol^{-1})	Literature K_i[d]
Bcl-x_L-Bak1	-11.7 ± 1.2	1.9 ± 0.4	-11.5 ± 0.1	0	49 nM
Bcl-x_L-Bak2	-13.0 ± 0.8	1.8 ± 0.3	-13.1 ± 0.2	-1.6 ± 0.2	1.5 nM

[a]The ΔG_f and m-values were taken from the linear least-squares analysis of the data using equation 2-1 (see Figure 5). Errors reported for the ΔG_f and m-values are the fitting errors of the linear least-squares analyses.

[b]$\Delta G_{f,avg}$ values were the average of folding free energy values calculated using equation 2-1 and an established m-value (i.e., the m-value averaged from the m-values of the complexes of each protein). The errors reported for the $\Delta G_{f,avg}$ values are the standard deviations of the folding free energy values calculated using equation 2-1 and the established m-value.

[c]Value relative to Bcl-x_L-Bak1. Propagated error is reported.

[d]Value is the K_i from Ref (70).

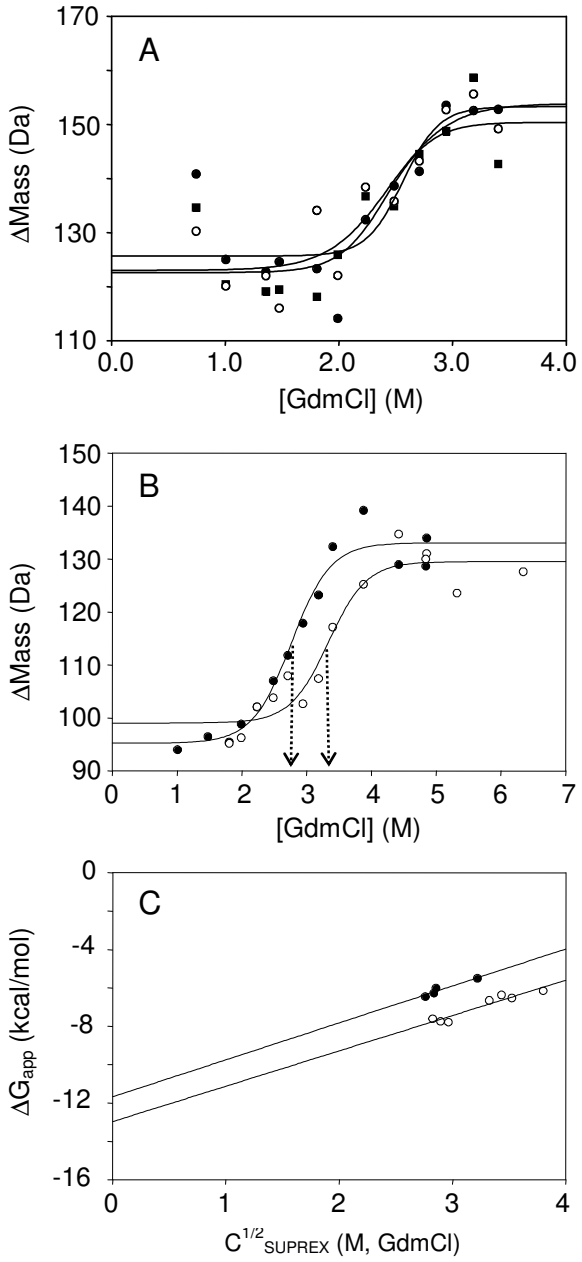


Figure 5: A) SUPREX analysis of apo-Bcl-x_L using exchange times of 80 min (filled squares), 160 min (filled circles), and 240 min (open circles). The midpoints of apo-Bcl-x_L SUPREX curves do not shift with exchange time. B) SUPREX analysis of Bcl-x_L-Bak1 (filled circles) using an exchange time of 110 minutes and Bcl-x_L-Bak2 (open circles) using an exchange time of 130 minutes. The dotted lines mark the $C^{1/2}_{SUPREX}$ value for each curve. C) Plots of ΔG_{app} versus $C^{1/2}_{SUPREX}$ for Bcl-x_L in the presence of Bak1 (filled circles) and Bcl-x_L in the presence of Bak2 (open circles). The solid lines are the results of linear least-squares fitting of the data to equation 2-1.

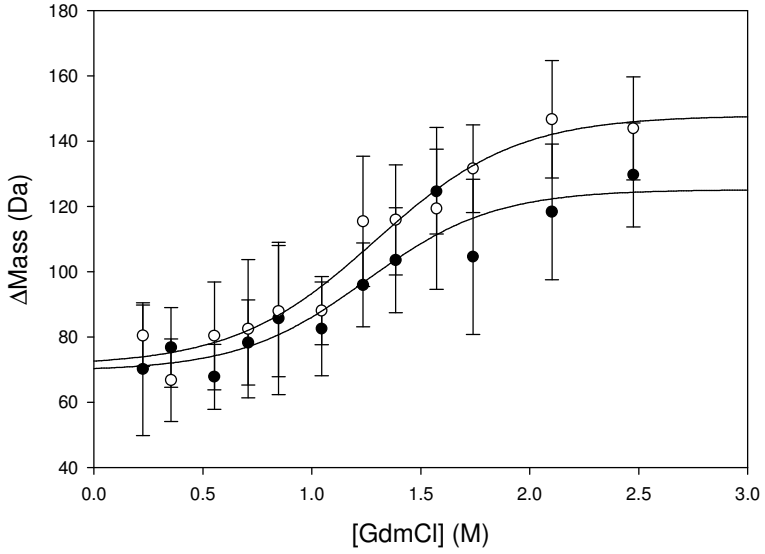


Figure 6: SUPREX curves of AGTmi using exchange times of 30 sec (filled circles) and a 5 min (open circles). In both cases, the transition midpoint was found to be 1.3 M GdmCl.

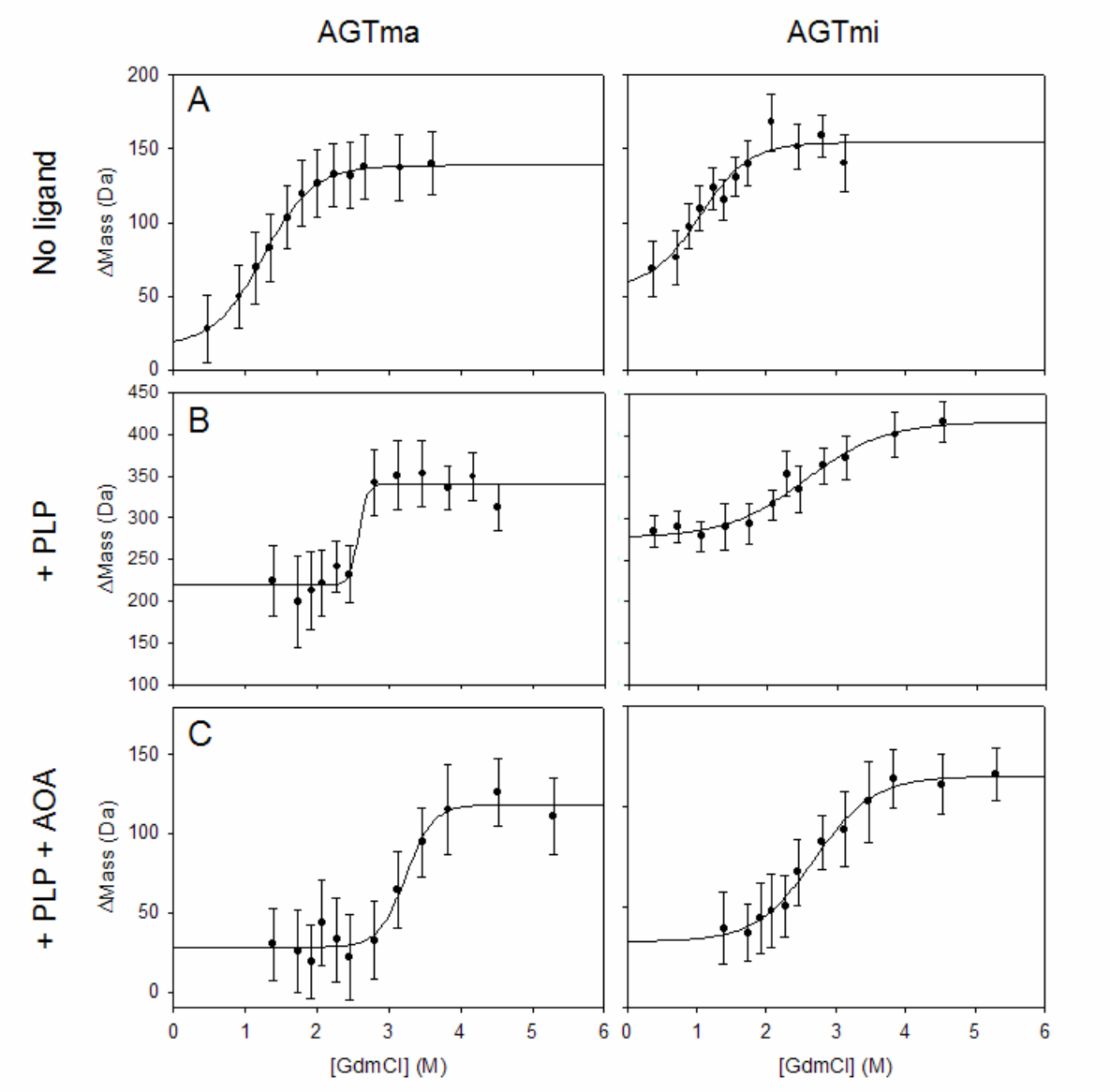


Figure 7: SUPREX analysis of AGTma and AGTmi. AGTma and AGTmi (4 μM, with C-terminal 6xHis tags) were analyzed A) in the absence of added ligand (5 minute exchange time), B) in the presence of 400 μM PLP (5 minute exchange time), and C) in the presence of 400 μM PLP and 400 μM AOA (1 hour exchange time).

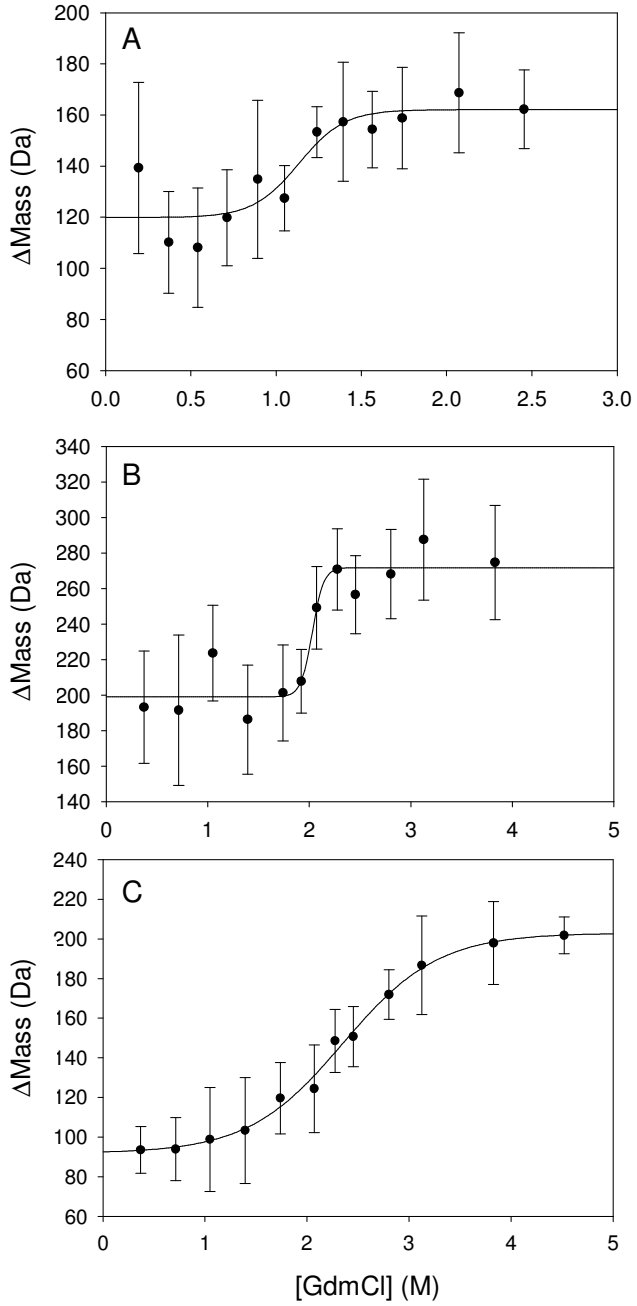


Figure 8: SUPREX analysis of AGT244 (4 μM, with a C-terminal 6xHis tag) A) in the absence of added ligand (5 minute exchange time), B) in the presence of 400 μM PLP (5 minute exchange time), and C) in the presence of 400 μM PLP and 400 μM AOA (1 hour exchange time).

4. A High Throughput Screening Application of SUPREX[1]

4.1 Introduction

The speed and sensitivity of modern mass spectrometers make them attractive tools for high throughput screening (HTS) assays, and several mass spectrometry-based HTS assays have been developed in recent years (9-11). Recently, a mass spectrometry-based HTS assay for protein-ligand binding detection was reported that utilizes an abbreviated version of SUPREX referred to hereafter as single-point SUPREX. In single-point SUPREX, binding events are detected by measuring the target protein's mass change after H/D exchange in a deuterated buffer containing a specific concentration of a chemical denaturant (74).

Several inherent advantages of SUPREX make it particularly well-suited for an HTS assay. Unlike most radiometric or fluorescence-based assays, single-point SUPREX can be performed on protein-ligand complexes directly in solution without immobilization or labeling of the target or library compounds. The technique does not require time-consuming separations or filtrations and is relatively general (i.e., it can be readily transferred to most protein-ligand systems). Additional advantages include the ability to study multi-component mixtures and to make measurements on picomole quantities of protein.

---

[1] Reproduced in part with permission from Hopper, E. D.; Roulhac, P. L.; Campa, M. J.; Patz, E. F.; Fitzgerald, M. C. Throughput and Efficiency of a Mass Spectrometry-Based Screening Assay for Protein-Ligand Binding Detection. *J. Am. Soc. Mass Spectrom.*, **2008**, *19*, 1303-1311. Copyright 2008 Elsevier.

The single-point SUPREX protocol was initially developed in a proof-of-concept study using the S-protein and a small test library of five peptides with known binding affinities for the S-protein (74). The experiments contained in this chapter represent the first application of single-point SUPREX in an HTS project designed to identify novel protein ligands, and as such it provides the first measure of the throughput and efficiency of this technique. The target protein for this study, cyclophilin A (CypA), is a protein that is overexpressed in lung tumor cells (5) and appears to be necessary for tumor growth (6). Thus, tight binding ligands for CypA could potentially be used as diagnostic imaging agents as well as lung cancer therapeutics.

Cyclosporin A (CsA), an immunosuppressive drug, is the most well-studied and tightest-binding CypA ligand identified to date (45, 75-78). Unfortunately, biodistribution studies have shown that CsA is ill-suited for use as an imaging agent as it has been shown in rats to be rapidly taken up by the liver and excreted in the GI tract (79). The immunosuppressive activities of CsA (75) also make it unattractive as a lung cancer therapeutic. One goal of the current work was to identify novel CypA ligands that might be more amenable to diagnostic imaging and therapeutic applications for lung cancer.

In an attempt to identify novel molecular scaffolds that bind to CypA, the 880-member Prestwick Chemical Library was screened. The Prestwick Chemical Library contains a variety of structurally diverse compounds with known safety and

bioavailability in humans. Over 85% of the compounds in the library are off-patent drugs that are marketed in a wide range of therapeutic areas. CsA was also present in the Prestwick Chemical Library, where it served as a blind control. This blind control was the only ligand identified in the assay as a hit. The focus of this report is to describe the analytical capabilities (i.e., the throughput and efficiency) of single-point SUPREX.

4.2 Experimental Section

4.2.1 Materials

The human CypA used in this work was kindly provided by Dr. Michael J. Campa from the laboratory of Dr. Edward F. Patz, Jr. It was obtained by recombinant DNA methods that involved the following steps: 1) overexpressing the CypA protein as a glutathione S-transferase (GST) fusion protein in *E. Coli* (BL21-DE3), 2) purifying the fusion protein using a GST-binding resin, 3) removing the GST tag in an overnight incubation with 5 units of thrombin per mg of fusion protein, and 4) removing the GST and thrombin with GST-binding resin (Promega Corporation, Madison, WI) and HiTrap Benzamidine FF (Amersham Biosciences, Piscataway, NJ), respectively. The 880-compound Prestwick Chemical Library (Prestwick Chemical, Illkirch, France) was obtained from the Small Molecule Synthesis and Screening Facility in the Center for Chemical Biology at Duke University. The library compounds were provided as 1 μL aliquots of 1 mM solutions in dimethyl sulfoxide (DMSO). CsA was purchased from LKT Laboratories (St. Paul, MN). Myoglobin (horse skeletal muscle) and trypsin

inhibitor (soybean) were from Sigma (St. Louis, MO). DMSO was purchased from Fisher (Fairlawn, NJ). The assay buffer was prepared as described in Chaper 2.

4.2.2 Theoretical SUPREX Curves

Equation 4-1, which was derived in reference (36) for the analysis of SUPREX data, was used to generate the individual theoretical SUPREX curves in this work (see Figure 10).

$$\Delta\text{Mass} = \Delta M_{\infty} + \left(\Delta M_0 - \Delta M_{\infty}\right) e^{\frac{-\langle k_{int}\rangle t}{(1+K_{fold})}} \qquad \text{Equation 4-1}$$

The variables in equation 4-1 are defined as follows: Δmass is the mass difference between the protonated and deuterated forms of the protein, ΔM_0 is the Δmass of the protein before the protein's globally protected amide protons are exchanged with solvent deuterons (i.e., the Δmass of the pre-transition baseline of a SUPREX curve), ΔM_{∞} is the Δmass of the protein after all of the protein's amide protons have exchanged with solvent deuterons (i.e., the Δmass of the post-transition baseline of a SUPREX curve), $\langle k_{int}\rangle$ is the average intrinsic exchange rate for an unprotected amide proton, t is the H/D exchange time used in the SUPREX experiment, and K_{fold} is given by equation 4-2.

$$K_{fold} = e^{\frac{-(\Delta G_f + m[\text{GdmCl}])}{RT}} \qquad \text{Equation 4-2}$$

The variables in equation 4-2 are defined as follows: R is the universal gas constant, T is the temperature in Kelvin, *m* is defined as $\delta\Delta G_f/\delta$[denaturant], [GdmCl] is

the concentration of GdmCl in the exchange buffer, and ΔG_f is the protein's folding free energy in the absence of denaturant.

In using equations 4-1 and 4-2 to construct the theoretical SUPREX curve for CypA in the absence of ligand, ΔM_0 and ΔM_∞ were assigned values of 55 and 95 Da (respectively) based on the experimental data in reference (45), and $\langle k_{int} \rangle$ was assigned a value of 7.08 s^{-1} based on a calculation using the program SPHERE, which performs the calculation using model dipeptide data along with the protein sequence and experimental temperature and pH values (80, 81). The time (t) was set to 35 minutes, and *m* and ΔG_f were assigned values of 3.7 kcal/(mol M) and -10.4 kcal/mol (respectively) based on an average of the two sets of values reported for CypA in reference (45).

Construction of the theoretical curves for CypA in the presence of the hypothetical ligands required the calculation of a $\Delta\Delta G_f$ value (i.e., the change in ΔG_f upon ligand binding) according to equation 4-3 (82).

$$\Delta\Delta G_f = -nRT \ln\left(1 + \frac{[L]}{K_d}\right) \qquad \text{Equation 4-3}$$

In equation 4-3, [L] is the concentration of free ligand in the exchange buffer, n is the number of equivalent ligand binding sites in the protein, K_d is the dissociation constant of the hypothetical ligand, and R and T are defined above. In the calculations using equation 4-3, n was set to 1 and [L] was set to 90 μM, which was the estimated ligand concentration in the SUPREX buffers. A ΔG_f value was calculated for CypA complexed with each hypothetical ligand by adding the corresponding $\Delta\Delta G_f$ value for

CypA in the presence of each hypothetical ligand to -10.4 kcal/mol (i.e., the ΔG_f value of the unbound protein – see above). Ultimately the resulting ΔG_f values were used in equation 4-1, as described above, to generate theoretical SUPREX curves for CypA in the presence of each hypothetical ligand.

4.2.3 Mass Spectrometry

The MALDI matrix was prepared as a saturated SA solution in an aqueous buffer containing 0.1% TFA and 45% ACN. Internal standards (myoglobin and trypsin inhibitor) were added directly to the matrix solution. Mass spectra were acquired on one of two different MALDI mass spectrometers, including a Voyager Biospectrometry Workstation from PerSeptive Biosystems (Framingham, MA) and an Ultraflex II TOF/TOF (Bruker Daltonics, Billerica, MA). The Voyager instrument was employed in the initial screening experiment, and the Ultraflex instrument, which was not available for the project during the initial screening, was employed in the re-screening that was performed as part of the 2-tier strategy (see below). Other than the Ultraflex's ability to collect data more rapidly, the performance characteristics (e.g., mass accuracy, precision, and resolution) were comparable.

All mass spectra were acquired in the linear and positive ion modes. Spectra collected with the Voyager instrument were a sum of 32 laser shots from a nitrogen laser operating at ~3 Hz and were collected using the following instrument parameters: an acceleration voltage of 25 kV, a grid voltage of 22.3 kV, a guide wire voltage of 37.5 V,

and a delay time of 300 nsec. Spectra collected with the Ultraflex instrument were a sum of 100 laser shots from a Nd:YAG laser operating at 100 Hz and were collected using the following instrument parameters: an ion source 1 voltage of 25 kV, an ion source 2 voltage of 23.35 kV, a lens voltage of 6 kV, and a delay time of 130 ns. For each spectrum, the mass-to-charge ratio of the CypA peak was determined using either a Microsoft Excel macro program or a MATLAB (The MathWorks, Inc., Natick, MA) script as described in Chapter 2.

4.2.4 Library Screening

Assay conditions (i.e., the H/D exchange time and denaturant concentration) for the initial library screening were chosen such that CypA ligands with K_d values in the low micromolar range would be selected. An exchange time of 35 min and a GdmCl concentration of 1.5 M were used here to select compounds with K_d values ≤ 10 μM. In theory, the H/D exchange time and denaturant concentration can be tuned to select tighter (or weaker) binding ligands. However, the goal of this study was to select for ligands with K_d values ≤ 10 μM (see Results and Discussion). In the assay, 1 μL of a 1 mM solution of each library ligand (in DMSO) was combined with 9 μL of the deuterated exchange buffer (which contained 1.5 M GdmCl). A 1-μL aliquot of a protonated stock solution of CypA (~100 μM) was added to the resulting 10-μL volume of ligand- and GdmCl-containing deuterated exchange buffer, and the H/D exchange reaction with CypA was allowed to proceed for 35 min. Each H/D exchange reaction

was quenched by adding 1 μL of the exchange reaction to 9 μL of ice-cold matrix solution (see above). Finally, 1 - 2 μL of the quenched reaction/matrix solution was spotted onto a MALDI target. Mass measurements from five replicate MALDI mass spectra were averaged, and this average was used to calculate the Δmass value. The standard deviations of the mass measurements were typically ~6 Da, or ~10% of the Δmass values. To limit the effects of differential back exchange (i.e., the loss of varying numbers of amide deuterons) from sample to sample, only 12 samples including 10 library compounds, one positive control (CsA, K_d ~ 30-200 nM (45, 75-78)), and one negative control (DMSO) were analyzed at a time. H/D exchange reactions were staggered at 20-min intervals to allow for nearly continuous data collection (i.e., MALDI-MS spectra were acquired for each sample set while the subsequent set was undergoing H/D exchange). Analysis of the single-point SUPREX data is described in Chapter 2 (see Section 2.9).

4.2.5 2-Tier Screening Strategy

Preliminary hits from the assay were subjected to an additional single-point SUPREX analysis as part of a 2-tier screening strategy. The conditions of the re-screening experiment were similar to those of the initial screening, except that C_{18} ZipTips (Millipore, Billerica, MA) were used to concentrate and desalt the samples prior to single-point SUPREX analysis. The incorporation of ZipTips into the SUPREX protocol has been reported previously (39). Briefly, this process involves quenching the H/D

exchange reaction with 1 μL of a 10% TFA solution and extracting the proteins from the H/D exchange buffers using ZipTips that were pre-equilibrated according to the manufacturer's instructions. Bound proteins were then washed with a solution of 0.1% TFA and eluted with the matrix solution (saturated SA in 50% ACN, 0.1% TFA, 49.9% water) directly onto the MALDI target. Internal standards were incorporated into the matrix solution prior to the elution step. This concentration and desalting step was not essential, but it generally improved the signal-to-noise ratio of the CypA peak.

4.3 Results and Discussion

4.3.1 General Strategy

The single-point SUPREX protocol used for the HTS assay in this work relies on the ability of SUPREX to detect ligand binding (Figure 9). In a full SUPREX analysis, aliquots of a fully protonated protein solution are diluted into a series of deuterated exchange buffers containing various concentrations of a chemical denaturant. The denaturant serves to shift the protein folding equilibrium toward the unfolded state. At higher concentrations of denaturant, a larger fraction of the protein population is unfolded, resulting in an increased uptake of deuterons and thus an increase in mass. This mass change (i.e., Δmass) is monitored by MALDI-MS. When the protein interacts with a ligand, the protein is stabilized, and a greater concentration of denaturant is required to unfold the protein. Thus, in the presence of a ligand, the midpoint of the SUPREX transition shifts toward higher denaturant concentrations.

SUPREX curves like those shown in Figure 9 can be used to derive binding free energies and to determine solution phase dissociation constants (K_d values) with reasonably high accuracy and precision. However, it would be relatively time consuming and would require large amounts of protein to perform full SUPREX analyses for every member of a chemical library. The analysis time and amount of protein required can be significantly reduced by using a single-point SUPREX protocol in which a Δmass value is recorded at a single denaturant concentration for the protein in the presence of each library compound. If appropriate conditions (i.e., denaturant concentration and H/D exchange time) are chosen, the Δmass value will be low in the presence of a binding compound and high in the presence of a nonbinding compound (see dashed line in Figure 9).

The denaturant concentration and H/D exchange time used in this work were selected prior to the assay by constructing a series of theoretical SUPREX curves (see Figure 10) based on the known SUPREX behavior of CypA using GdmCl as the denaturant (45). The effect of ligand binding on the SUPREX curve for CypA is illustrated in Figure 10, which shows a series of theoretical curves for CypA in the absence and presence of several hypothetical ligands with different binding affinities. A detailed description of the procedure used to generate the theoretical SUPREX curves is provided in the Experimental section. The theoretical curves in Figure 10 were used to select the following screening conditions: a GdmCl concentration of 1.5 M and an

exchange time of 35 min. These conditions were chosen to allow for the selection of compounds in the library that bind CypA with a $K_d \leq 10$ μM. Ligands with K_d values up to ~10 μM were expected to be useful lead compounds in a search for CypA-targeting lung cancer therapeutics and imaging agents.

One advantage of using single-point SUPREX is that the assay is flexible and can be adjusted to select for a wide range of K_d values. Longer H/D exchanges times and/or higher denaturant concentrations could be used to select for tighter binding ligands. In theory, there is no upper or lower limit to the K_d value selection in single-point SUPREX. However, in practice the upper limit is defined by the protein and ligand concentrations that are experimentally accessible. For example, one would need to increase the protein and ligand concentrations tenfold compared to those used in this work to select ligands with K_d values > 100 μM. The lower K_d value limit is not generally subject to experimental limitations.

4.3.2 Hit Identification

The 880 compounds in the Prestwick Chemical Library were individually screened for binding to CypA using single-point SUPREX. Representative data from the screen is shown in Figure 11A. Note that a positive control (CsA) and a negative control (DMSO) were analyzed with every set of 10 ligands in the library. ΔMass values were determined for all compounds in the library except for 18 compounds that appeared to suppress the MALDI ion signal for CypA. Shown in Figure 12 are the Δmass values

obtained for the positive and negative controls that were analyzed over the course of the screening, which took place in ~5-hour time blocks spread over a period of several weeks. The control data were not constant over the time course of the experiment. Figure 12 shows a random scatter to the Δmass values obtained for the controls as well as a clear trend in the $\Delta mass_{av}$ values that were calculated from the control data.

The random scatter of the points in Figure 12 can be explained by the precision of the molecular weight determinations of the CypA in this work. Standard deviations of replicate molecular weight determinations using the MALDI readout were consistently ~6 Da, whether the protein was protonated, deuterated in the presence of ligand, or deuterated in the absence of ligand. This is consistent with the mass accuracy expected for the MALDI-TOF instrument used in this study (~200 to 400 ppm). The changes observed in the $\Delta mass_{av}$ values over the course of the experiment are most likely due to variations in the degree of back exchange for each sample set. This variability in the extent of back exchange most likely resulted from day-to-day fluctuations in the temperature and humidity levels of the laboratory and from variations in the timing of sample preparation/data collection.

The variability in the control data made the selection of a constant cutoff value impractical, and it prompted the development of a new strategy for processing single-point SUPREX data that involved the use of a moving average (i.e., $\Delta mass_{av}$) to calculate cutoff values. This moving average strategy was particularly useful in correcting for the

day-to-day and set-to-set variations in back exchange that were observed. Initially, a cutoff value of three standard deviations below the $\Delta mass_{av}$ value of the negative controls (Figure 12) was used to select hits. This resulted in the selection of 41 preliminary hits.

The 41 preliminary hits were re-screened in a second single-point SUPREX analysis (Figure 11B). Four positive and four negative controls were also included in this analysis. In the re-screening, the four negative control Δmass values were averaged to generate a $\Delta mass_{av}$ value, and the cutoff was determined by multiplying the standard deviation of the negative control Δmass values by three and subtracting this product from the $\Delta mass_{av}$ value of the negative control (79 Da). This calculation resulted in a cutoff of 55 Da, and application of this cutoff resulted in a single hit. The hit was the blind control, CsA. It is not surprising that no new hits were identified in a library of this size; a conservative estimate of the average rate of confirmed HTS hits is reported to be < 0.1% (83, 84). Thus, a library of at least 1000 compounds is generally required to identify just one confirmed hit.

4.3.3 Throughput

Assay data were collected during time periods of approximately 5 hours per day over a period of several weeks. In total, it took ~3000 minutes, or about 3 min/ligand, to acquire the mass spectra for all library compounds and controls using the Voyager mass spectrometer that was exclusively employed in the initial screen. Extraction of the Δmass

data from the > 5000 mass spectra recorded in this work took a total of ~300 minutes, and subsequent analysis of the Δmass values to ascertain which compounds were hits was accomplished in ~30 minutes. However, the data analysis time could likely have been reduced ~3-fold if the MATLAB program had been exclusively employed for extracting Δmass values from the spectra. The MATLAB program was developed during the course of this work and was only employed in the hit validation component of the HTS project. The Excel-based program, which is ~10-fold slower than the MATLAB program, was employed in the early stages.

Several factors that impact the throughput of single-point SUPREX include the exchange time, the size of the sample set, and the rate at which MALDI mass spectra can be collected. Long exchange times require a larger investment of time before mass spectral analysis can begin. However, as long as H/D exchange reactions are staggered, the extra time is only invested once at the beginning of each assay. Increasing the size of each sample set could improve throughput as well, particularly if the application of a moving average is used to account for variations in back exchange. In this work, back exchange effects were minimized by keeping each sample set small (i.e., each set contained 10 library ligands, a positive control, and a negative control). This allowed for the collection of all spectra for each set within a short period of time (about 15 minutes). Small sample sizes also helped to decrease the time required to pipette samples onto the MALDI target, thereby minimizing the effects of differential back exchange. However,

note that ~50% of the deuterons that exchanged into the protein during the H/D exchange reaction were consistently back-exchanged to protons during the MALDI sample preparation and analysis step, as evidenced by the Δmass values of the negative controls, which were typically ~80 Da. This Δmass value represents ~50% of the mass gain expected if all 167 amide protons in CypA were exchanged with solvent deuterons. The use of back exchange correction factors was investigated in this work, but their use in the data analysis had no impact on the screening results.

The rate-limiting step in the initial screen using the Voyager mass spectrometer was the time required to acquire the MALDI mass spectra. Due to the low repetition rate of the nitrogen laser (~3 Hz) and the slow sample positioning mechanisms (~10 s/sample) in the instrument, ~15 minutes were required to collect the necessary mass spectra on each data set of 12 samples. The use of a high throughput MALDI instrument, like the Ultraflex instrument used in the re-screening experiments, could result in a dramatic increase in throughput. Such high throughput instruments have laser repetition rates of approximately 200 Hz and have fast sample positioning mechanisms (~2 s/sample). Additionally, the use of a high throughput MALDI instrument might allow for the analysis of larger data sets without significantly increasing the time that back exchange can occur within the mass spectrometer. The use of a high throughput MALDI instrument for single point SUPREX analyses would most likely increase the throughput of single-point SUPREX to < 20 seconds/ligand. Note that the Ultraflex

instrument was not used in the initial screening experiment in this work because it was not available for the project.

One additional approach for increasing the throughput of single-point SUPREX is to perform the analyses in a parallel fashion. In this study, each microtiter plate well contained only one ligand. However, single-point SUPREX is particularly well-suited for HTS applications because of its multiplex capabilities. Theoretically, multiple ligands could be placed in each well of the microtiter plate, allowing for the simultaneous screening of many ligands at once. For example, the simultaneous screening of 10 ligands per well has the potential to create a tenfold increase in throughput. One caveat to such multiplexing is that for every well that is identified as a hit, the compounds in the well must be individually re-tested to identify which compound(s) were responsible for the hit. Because of the need for additional analysis of each hit, the screen would need to be stringent enough to select for only the tightest binding ligands. The combined use of a high throughput MALDI and a multiplexing strategy is expected to produce up to a 100-fold increase in throughput compared to the 3 min/ligand observed here. This would make it possible to screen ~40,000 compounds in a 24-hour day using single-point SUPREX.

4.3.4 Efficiency

One goal of this work was to determine the efficiency of the single-point SUPREX protocol. The efficiency is defined here as the accuracy with which hits can be

identified. One factor that influences efficiency is the choice of the cutoff value. The cutoff value used in this work was initially taken as three standard deviations below the $\Delta mass_{av}$ values for each data set. The use of three standard deviations is customary for high throughput screens (56), but note that this value could be adjusted to make the cutoff more or less stringent. For example, the use of a cutoff value 2.5 standard deviations below the $\Delta mass_{av}$ of the negative control resulted in the selection of 54 new preliminary hits in addition to the 41 preliminary hits selected using the 3 standard deviation cutoff. However, a re-screening of the 54 new preliminary hits yielded no hits.

Two important parameters used to define efficiency are the number of false positives and false negatives that appear in the screen. The control data provide one way in which to estimate the rate of false positives and false negatives. Summarized in Table 4 are the false positive and false negative rates observed if the 2.5 and 3.0 standard deviation cutoff values described above are applied to the positive and negative control data. It is noteworthy that the false positive rate of 0% for the controls is unchanged regardless of whether the cutoff is determined using 3.0 or 2.5 standard deviations. In contrast, the false negative rate is improved by a factor of two if the cutoff value is calculated using 2.5 instead of 3.0 standard deviations. Thus, the choice of cutoff value can be tuned to optimize false positive and false negative rates.

The false positive rate can also be estimated using the library members. Out of the 41 preliminary hits identified in the initial screen, only one was found to bind CypA

when the 41 preliminary hits were re-screened. This gives a calculated false positive rate of 98%, which is high, especially considering that Hann and Oprea estimate a false positive rate of approximately 40% in typical pharmaceutical screens (83). Interestingly, the false positive rate for the initial library selection is about the same (i.e., 99%) if the cutoff value is calculated using 2.5 standard deviations. The false negative rate cannot be easily calculated using the library members. However, the selection of CsA suggests that the false negative rate is low enough for the screen to be effective.

The false positive rates observed in the initial screening of the 880-member library are quite high (see 1-tier results in Table 4). One approach that can be used to reduce false positives in HTS assays is to use a 2-tier selection strategy in which preliminary hits detected in an initial screen are subjected to a re-screening (85). Such a 2-tier selection strategy is especially useful when the assay readout is subject to random error, as is the case for the MALDI-MS readout in the single-point SUPREX protocol. If such a 2-tier selection strategy is used here to screen the 880-member library, the false positive rates observed for the single-point SUPREX protocol are reduced to 0%, regardless of how the cutoffs are calculated (i.e., using 3.0, 2.5, or a combination of 2.5 and 3.0 standard deviations). The results in Table 4 indicate that a 2-tier selection strategy using 2.5 standard deviations to calculate the cutoff values in tier 1 was optimal in this work. The results for tier 2 were the same regardless of whether 2.5 or 3.0 standard deviations were used in the calculation of the cutoff value. As shown in Table

4, the choice of 2.5 standard deviations produced the optimal false positive and false negative rates for the controls, the lowest false positive rate in the library screening, and a reasonable number of preliminary hits (i.e., 10% of the 880 ligands in the library). The use of 2.0 standard deviations to calculate the cutoff value was also investigated; however, this resulted in a large number of preliminary hits (~25% of the 880 ligands in the library), an increased false positive rate based on the controls, and only a small improvement in the false negative rate (i.e., from 9% to 7%).

Another parameter used to judge the efficiency of the assay is the Z′-factor, which can be calculated using equation 2-3 (56). Using the $\Delta mass_{av}$ and standard deviation for each control, it was possible to calculate Z′-factors for all of the control data points that were collected over the course of the screening. The Z′-factors in this assay ranged from -2.1 to 0.2, and the average Z′-factor was -0.7 over the entire screen. It is noteworthy that the control data improved over the course of the experiment. For example, Z′-factors ranged from -0.3 to 0.2 in the last 25% of the controls, and the average Z′-factor for these controls was 0.0. This improvement in the Z′-factor is attributed to an improvement in the timing of the H/D exchange reactions and mass spectrometry analyses, which decreased the variation in back exchange between data sets. It has been noted that HTS assays with Z′-factors ≥ 0 are generally useful for screening large (e.g., 100,000-member) combinatorial libraries (56). These results suggest

that the single-point SUPREX protocol is capable of screening such libraries for binding to CypA.

The Z′-factors in this work were significantly lower than the Z′-factor calculated in the single-point SUPREX proof-of-concept study involving the S-protein (Z′ = 0.77) (74). The discrepancy is most likely due to the difference in the masses of the target proteins. CypA has a mass of ~18.2 kDa, while the mass of the protein used in the proof-of-concept study (S-protein) was only ~11.5 kDa. Since the absolute precision of MALDI-TOF mass spectrometers decreases as protein mass increases, the mass measurements of CypA were less precise (i.e., standard deviations of ~ 6 Da) than the mass measurements of S-protein (i.e., standard deviations of ~2 Da). Higher standard deviations can result in lower Z′-factors. This makes the use of single-point SUPREX more challenging for larger proteins. However, the Z′-factor does not solely depend on the standard deviations of the controls; it also depends on the difference between the average positive and negative control values. Thus, the Z′-factor will also be influenced by the amplitude of the target protein's SUPREX curve. Such amplitudes are hard to predict for a given protein as they are largely defined by the number of globally protected amide protons (i.e., protons that are protected by the folded structure of the protein). Because of this dependence on the amplitude of the SUPREX curve, the Z′-factor is expected to be different for every target protein. Thus, larger proteins that have SUPREX curves with large amplitudes may still be amenable to single-point SUPREX assays. The amplitude of a SUPREX curve tends to

be slightly larger when short exchange times are used (43); therefore, for larger proteins, the Z'-factor could potentially be improved slightly by adjusting the exchange time.

Another potential strategy for improving the efficiency of single-point SUPREX for larger proteins involves the incorporation of a protease digestion step into the protocol. In a recent report, a rapid (< 2 min) protease digestion step was used to facilitate the SUPREX analyses of large multidomain proteins (51). In this SUPREX-protease digestion protocol the H/D exchange properties of the individual domains of a protein are defined when the domains are in the intact protein. However, the domain-level folding properties are evaluated using peptides generated in the protease digestion step. The standard deviations associated with peptide molecular weight determinations by MALDI-TOF mass spectrometry are typically smaller than those associated with intact protein molecular weight determinations. These smaller standard deviations are expected to result in a significant improvement in the Z'-factor of single-point SUPREX assays. Incorporation of a protease digestion protocol into the single-point SUPREX assay is discussed in Chapter 6.

4.4 Conclusions

This work demonstrated that the single-point SUPREX protocol was an effective technique for screening the 880-member Prestwick Chemical Library for binding to CypA. This application of single-point SUPREX provides the first measure of the throughput of this HTS assay, which was determined to be 3 min/ligand using a

conventional MALDI mass spectrometer equipped with a nitrogen laser operating at 3 Hz. The throughput of the single-point SUPREX protocol was found to be largely limited by the time needed to acquire the mass spectra in the MALDI readout. Thus, the use of conventional MALDI mass spectrometers that have relatively slow sample positioning mechanisms and nitrogen lasers with relatively slow repetition rates can create a bottleneck in HTS projects that utilize the single-point SUPREX protocol. The use of high throughput MALDI mass spectrometers equipped with fast sample positioning mechanisms and high repetition rate lasers could potentially increase the throughput of the protocol to less than 20 sec/ligand. The data generated in this work also permitted the best evaluation to date of the efficiency of single-point SUPREX. The results also suggest that the use of a 2-tier selection strategy is important for minimizing the false positive and false negative rates of single-point SUPREX.

Table 4: Efficiency of Single-Point SUPREX.

	False Positives	*False Negatives*
Controls		
3.0 SD Cutoff	0/86 (0%)	15/85 (18%)
2.5 SD Cutoff	0/86 (0%)	8/85 (9%)
Library Compounds (1-Tier)		
3.0 SD Cutoff	40/41 (98 %)	-
2.5 SD Cutoff	94/95 (99 %)	-
Library Compounds (2-Tier)[a]		
3.0 & 3.0 SD Cutoffs	0/1 (0%)	-
2.5 & 2.5 SD Cutoffs	0/1 (0%)	-
2.5 & 3.0 SD Cutoffs	0/1 (0%)	-

[a]In this category, the number of standard deviations used to calculate the cutoff values are listed for tiers 1 and 2, respectively.

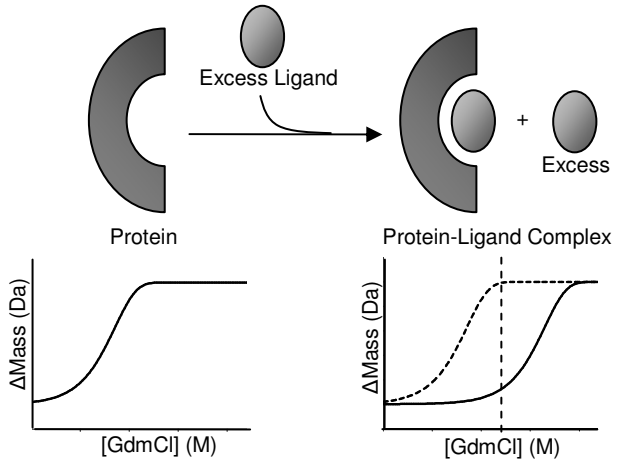


Figure 9: SUPREX-based detection of protein-ligand binding. The bottom half of the panel shows a schematic representation of SUPREX curves expected in the absence and presence of a ligand. The dashed vertical line represents an appropriate denaturant concentration at which to perform a single-point SUPREX analysis. Shown in the top half of the panel is a schematic representation of the experimental conditions used to generate the hypothetical SUPREX curves shown in the bottom half of the panel.

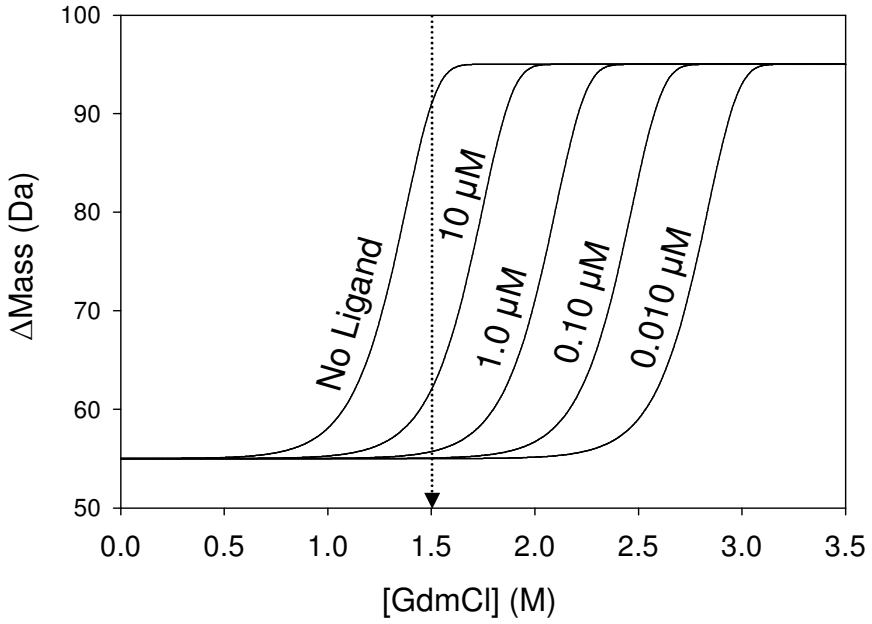


Figure 10: To detect ligand binding, single-point SUPREX exploits the difference in Δmass between the bound and unbound forms of a protein. This plot shows CypA SUPREX curves that were constructed as described in Section 4.2.2. The curves depict the SUPREX behavior of CypA in the absence and in the presence of four hypothetical ligands with various binding affinities. The curves were created using published ΔG and *m*-values along with Equations 4-1, 4-2, and 4-3. The arrow denotes the GdmCl concentration used in the single-point SUPREX analyses performed in this work; this concentration was selected to maximize the difference in the Δmass values of the bound and unbound forms of the protein.

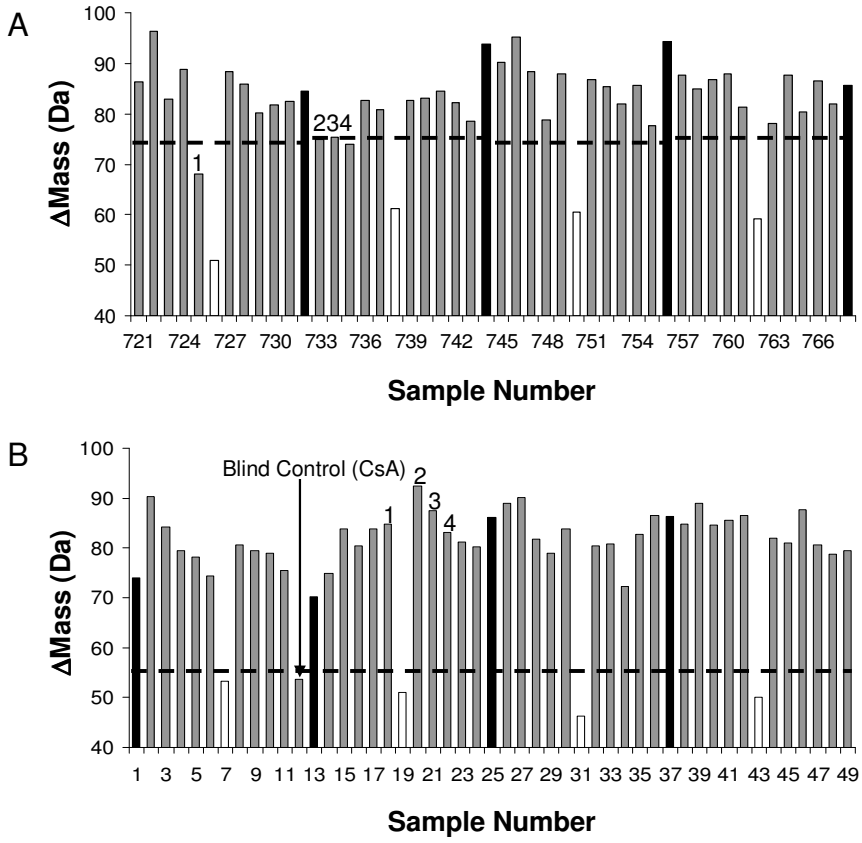


Figure 11: A) Representative data from the initial screen performed using the single-point SUPREX assay. Positive and negative controls are shown as white and black bars, respectively. The standard deviations for the mass measurements were typically ≤ 10% of the Δmass values. B) The re-screening data collected on the preliminary 41 hits identified in the initial screen using a 3.0 standard deviation cutoff. The dotted lines in both panels represent the 3.0 standard deviation cutoff. This cutoff varied slightly for each data set in panel A according to the moving average. The numbered bars represent compounds that were selected as hits in the initial screening but not in the re-screening. The arrow indicates the only hit (i.e., the blind control) that was identified using the 2-tier strategy.

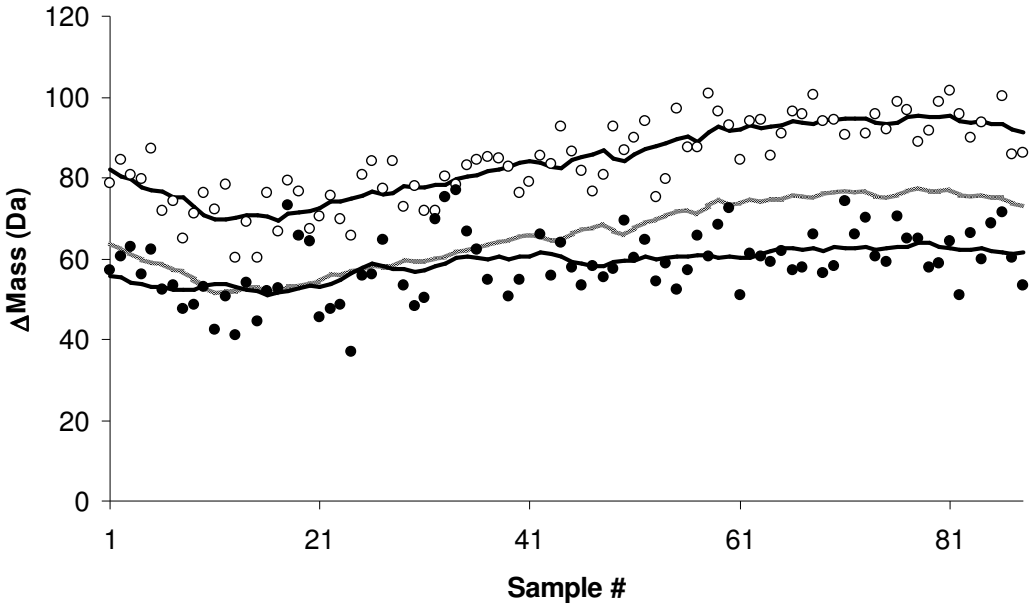


Figure 12: Summary of the positive (closed circles) and negative (open circles) control data. The upper and lower solid black lines represent the moving average ($\Delta mass_{av}$) of the negative and positive control data, respectively (see Experimental section). The gray line depicts the moving cutoff value, which was set at three standard deviations below the moving average of the negative control data.

5. Analysis of AGTmi Using the SUPREX-Protease Digestion Protocol

5.1 Introduction

For the evaluation of protein-ligand binding, techniques based on amide H/D exchange coupled with mass spectrometry have proven especially useful (28, 30, 41, 86, 87). In a typical H/D exchange experiment, a fully folded protein sample is allowed to undergo H/D exchange. After a specified amount of time, the exchange is quenched by lowering pH and temperature, and the protein is digested with pepsin, a protease with high activity under acidic conditions. The same experiment is performed on the ligand-bound protein, and comparisons are drawn between the uptake of deuterons for the apo and holo forms of the protein. Since these experiments are designed to provide site-specific information about ligand binding, the proteolytic fragments must be identified. This task is often time consuming and difficult due to the low specificity of pepsin.

While these experiments are valuable for obtaining information about protein dynamics during ligand binding, they are low throughput and do not provide any quantitative information about ligand binding affinities. A recent report outlined the development of an H/D exchange- and MALDI mass spectrometry-based technique that combines SUPREX with a protease digestion protocol (51). The SUPREX-protease digestion protocol is an extension of SUPREX in which a protease digestion step is used to allow for the thermodynamic analysis of proteins at the domain level. Unlike the conventional H/D exchange and protease digestion techniques described above, the

SUPREX-protease digestion protocol is capable of providing quantitative information about ligand binding affinities. Furthermore, in contrast to typical H/D exchange- and protease digestion-based strategies for examining protein folding and ligand binding (32, 88, 89), the SUPREX-protease digestion protocol does not require extensive peptide mapping.

This protocol can be used to obtain protein-ligand binding affinities for multiple fragments, but knowledge about the original location of the fragments is not required for the determination of dissociation constants. In addition, good peptide coverage is not a particular concern when this technique is used since thermodynamic information can be elucidated from even a single proteolytic fragment. These are distinct advantages of the SUPREX-protease digestion protocol because they allow for a significant increase in throughput relative to other H/D exchange-based strategies.

To examine the scope of this technique, the SUPREX-protease digestion protocol was used to analyze ligand binding for AGTmi, a large, multidomain protein (~44 kDa/subunit). This protein is a potential target for pharmaceutical screens, which makes it a reasonable candidate for high throughput analyses (90). The SUPREX-protease digestion protocol has previously been shown to be effective for the analysis of proteins with well-defined domains in which each domain exhibits cooperative folding behavior (51). However, this work represents the first application of this protocol for the analysis of a protein with poorly defined domains.

The work in this chapter demonstrates the ability of the SUPREX-protease digestion protocol to determine protein-ligand binding affinities for a multidomain, homodimeric protein. Since the protease digestion step takes place after the H/D exchange reaction is complete, this technique reports on the thermodynamic stability of the intact protein. This strategy is expected to be applicable to a wide variety of proteins, and the successful evaluation of the additional protein system in this work provides further evidence of the generality of the technique. In addition to providing information regarding ligand binding affinities, this technique promises to be a valuable tool for high throughput studies. Ultimately, this SUPREX-protease digestion analysis of AGTmi will lay the groundwork for the incorporation of the protease digestion protocol into the single-point SUPREX assay as discussed in Chapter 6.

5.2 Experimental Section

5.2.1 Materials

The following reagents were purchased from Sigma-Aldrich (St. Louis, MO): porcine pepsin, insulin from bovine pancreas, pyridoxal 5′-phosphate (PLP), and aminooxyacetic acid (AOA). The minor allele AGT variant was prepared by Adrianne M.C. Pittman from the laboratory of Dr. Chandra L. Tucker as described in Chapter 3. SUPREX buffers were prepared as described in Chapter 2.

5.2.2 Sample Preparation

PLP and AOA solutions were prepared in PBS (pH 7.4). AOA solutions were made fresh each day. Before use, AGTmi stock solutions were diluted with elution buffer to 40 μM. Protein-ligand complexes were prepared by combining four parts AGTmi (40 μM) with one part PLP (16 mM) and one part AOA (16 mM) and incubating at room temperature for one hour. For experiments requiring one or fewer ligands, the ligand solution was replaced with PBS. Pepsin solutions were prepared in water and stored in aliquots at -20 ˚C.

5.2.3 SUPREX-Protease Digestion Protocol

The SUPREX-protease digestion protocol is described in Chapter 2 (51). For AGTmi, H/D exchange reactions were initiated by adding 3 μL equilibrated protein complex to 7 μL of each SUPREX buffer. After the specified exchange time, the H/D exchange was quenched by adding 4 μL of the exchange reaction to 3 μL of aqueous HCl (pH 0.8). A 3-μL aliquot of 0.15 μM pepsin was immediately added to each quenched protein solution (final pepsin concentration = 0.045 μM), and the protease digestion was allowed to proceed for 30 seconds.

To prepare the samples for MALDI analysis, 2 μL of the digestion solution was added to 6 μL ice-cold matrix solution, and 1 μL of the resulting solution was spotted onto a MALDI target for mass spectral analysis. Spots were quickly dried under a gentle flow of N_2. These sample preparation steps (particularly spotting and drying the

samples) were carried out as quickly as possible to minimize the back exchange of deuterons to protons.

5.2.4 SUPREX Data Analysis

SUPREX data were analyzed as described in Chapter 2. In each SUPREX curve, GdmCl concentrations were corrected to reflect the final concentration after mixing the SUPREX buffers with the equilibrated protein solutions. Similar to the Bcl-x_L work in Chapter 3, SUPREX data were collected at several exchange times for each protein and protein-ligand complex, and the resulting $C^{1/2}_{SUPREX}$ values were used with equation 2-1 to obtain *m*- and ΔG_f values from the SUPREX data (40). The left side of the equality in equation 2-1 will hereafter be referred to as ΔG_{app}.

Using equation 2-1, plots of ΔG_{app} vs. $C^{1/2}_{SUPREX}$ were generated for each protein and protein-ligand complex. These plots were fit using a linear regression routine, and the y-intercept and slope were taken to be the ΔG_f and *m*-value (i.e., $\delta\Delta G_f/\delta$[denaturant]), respectively. Values for $\langle k_{int} \rangle$ were calculated using the relationship $\langle k_{int} \rangle = 10^{pH-5}$ min^{-1} (36). As described previously (43), an averaged *m*-value (m_{avg}) was used to calculate $\Delta G_{f,avg}$ values. These $\Delta G_{f,avg}$ values were ultimately used to calculate $\Delta\Delta G_{f,avg}$ values for determination of dissociation constants.

Dissociation constants (K_d values) were calculated using equation 2-2. For all of the experiments in this study, the ligand was present in large excess relative to the

protein (i.e., greater than tenfold excess); thus, [L] was taken as the total concentration of ligand in the H/D exchange reactions.

5.3 Results

5.3.1 Optimization of Protease Digestion Conditions

By varying pepsin concentration and digestion time, the protease digestion was optimized on AGT to obtain peptides with molecular weights in the 3-7 kDa range. Peptides within this mass range are ideal because they are large enough to give SUPREX curves with a measurable amplitude, but small enough to provide sub-global information about the stability of the intact protein and/or its domain-specific thermodynamic behaviors. For AGTmi, a 30 sec pepsin digestion at pH ~2 was found to be optimal. Using these digestion conditions, 11 peptides were observed that gave reproducible MALDI ion signals across all of the denaturant concentrations that were used in the SUPREX analyses. All 11 fragments yielded detectable SUPREX transitions, indicating that these fragments were located in regions that were at least partially protected from H/D exchange when the protein was fully folded. Each fragment is named according to its molecular weight in Daltons.

5.3.2 SUPREX-Protease Digestion Analysis of AGTmi

Figure 13 shows a single subunit of AGTma (the major allele form of AGT). This structure illustrates that the domains of AGT are closely associated with one another rather than being structured in discrete, clearly defined domains as was the case for the

proteins in the previous SUPREX-protease digestion study (51). The SUPREX-protease digestion protocol was used for the thermodynamic analysis of AGTmi alone and in complex with two of its known ligands: PLP (a cofactor) and AOA (an inhibitor). The $C^{1/2}_{SUPREX}$ values for all of the AGTmi complexes in this study are included in Tables 5 through 7. Both PLP and AOA bind in the active site of AGTmi but are able to bind simultaneously (91).

Figure 14A shows representative SUPREX-protease digestion curves obtained for AGTmi and its complexes. Note that the midpoint of the AGTmi SUPREX transition (shown here with a 5 min exchange time) shifts significantly toward the right upon PLP binding. This is indicative of the increase in stability of AGTmi as a result of ligand binding. A further shift of the midpoint was detected for the binding of AOA to AGTmi-PLP. Due to the significant stabilization of AGTmi-PLP upon AOA binding, a longer exchange time (1 h) was used to bring the SUPREX curve transition to an experimentally accessible range. Similar midpoint shifts were detected for all 11 AGTmi fragments that were analyzed in this study.

To further analyze AGTmi-ligand binding, SUPREX experiments were performed at several different exchange times for apo-AGTmi and for the two AGTmi complexes. For apo-AGTmi, exchange times varied from 0.5 min to 40 min, and for AGTmi-PLP and AGTmi-PLP-AOA, exchange times varied from 5 to 1410 min and from 60 to 1800 min, respectively. Exchange times were chosen such that a wide range of

transition midpoints could be detected, allowing for more accurate determinations of ΔG_f and *m*-values. The transition midpoint (i.e., $C^{1/2}_{SUPREX}$ value) was extracted from the SUPREX curve obtained at each exchange time (see Tables 5 through 7), and these midpoint values were used to generate plots of ΔG_{app} vs. $C^{1/2}_{SUPREX}$ as described in Chapter 2. Correlation coefficients were at least 0.92 for all plots of ΔG_{app} vs. $C^{1/2}_{SUPREX}$ for AGTmi.

Representative plots of ΔG_{app} vs. $C^{1/2}_{SUPREX}$ for one AGTmi fragment are shown in Figure 14B. Plots for all other fragments are included in Figures 15 and 16. The y-intercept and slope of each linear plot are equal to the folding free energy (ΔG_f) and *m*-value (respectively) for the protein folding reaction. The ΔG_f and *m*-values extracted from each plot of ΔG_{app} vs. $C^{1/2}_{SUPREX}$ are shown in Table 8. Although the ΔG_f values extracted from plots such as those in Figure 14B can be used to calculate $\Delta\Delta G_f$ values for ligand binding analysis, the error in such measurements is typically large due to the relatively large uncertainties associated with the extraction of the y-intercept and slope from the plots of ΔG_{app} vs. $C^{1/2}_{SUPREX}$. The uncertainty in $\Delta\Delta G_f$ measurements can be reduced by more accurately establishing an *m*-value for each protein complex (43). This strategy is useful in cases where the *m*-value is not expected to change upon ligand binding. For example, since the measured *m*-value was consistent across all 11 fragments for AGTmi-PLP and AGTmi-PLP-AOA, a more accurate *m*-value could be established by averaging the *m*-values across all fragments from these two complexes. The average *m*-

value (m_{avg}) was then used with equation 2-1 to calculate a ΔG_f value for each SUPREX curve collected at each exchange time. These multiple ΔG_f values were averaged to calculate a $\Delta G_{f,avg}$ value, which was used to calculate the $\Delta\Delta G_{f,avg}$ values to be used in the determination of the dissociation constant (see Table 9). In the case of apo-AGTmi, a significantly different slope was observed relative to AGTmi-PLP and AGTmi-PLP-AOA (see Figure 14B). This difference in *m*-value between apo-AGTmi and the AGTmi complexes precludes determination of a dissociation constant for PLP (see Discussion section).

5.4 Discussion

5.4.1 Determination of Thermodynamic Parameters

In the SUPREX-protease digestion protocol, the protease digestion step takes place under quench conditions *after* the H/D exchange reaction. This arrangement allows for the proteolytic fragments to report on the folding behavior of the intact protein since the deuterons are essentially frozen in place once the exchange reaction has been quenched. Because low pH was used to quench the H/D exchange, pepsin was used as the protease due to its high activity in acidic conditions.

To calculate accurate ΔG_f and *m*-values for a protein folding reaction, the reaction must be two-state (i.e., only the unfolded or folded states of the protein can be significantly populated during the folding reaction). The folding mechanism of AGT has not been reported, so it is unknown whether the folding reaction of AGTmi is two-state.

However, given its relatively large size and the lack of cooperativity of its folding reaction, it is unlikely that AGTmi is a two-state folder. Because two-state folding is a required assumption when determining ΔG_f and *m*-values by SUPREX, the thermodynamic parameters reported here should not be taken as an absolute representation of the folding reaction of AGTmi. However, if the folding reaction of the protein is assumed to remain the same upon ligand binding, these values can be used to calculate $\Delta\Delta G_{f,avg}$ values, which in turn can be used to calculate K_d values. Thus, the likely non-two-state nature of this protein did not preclude quantitation of ligand binding, which was the goal of this study.

5.4.2 SUPREX Analysis of AGTmi

AGT is a liver enzyme that is required for the metabolism of glyoxylate to glycine, and deficiencies in this enzyme lead to a disorder called PH1. AGT deficiencies are often due to mutations that destabilize the enzyme, leading to degradation, aggregation, and in some cases mislocalization (62). These unstable variants often retain partial enzymatic function, which makes them promising candidates for treatment using pharmacological chaperones (i.e., compounds that specifically bind and stabilize a protein's native state). Since many disease-causing mutations are in the context of the minor allele, AGTmi has been proposed as a logical target for high throughput screening campaigns (49).

Chapter 3 discussed the use of SUPREX to examine the thermodynamic stability of AGTmi relative to wild-type AGT (the major allele form, AGTma) (49). For these studies, the conventional SUPREX protocol (without the protease digestion step) was used. Although the conventional SUPREX protocol was useful for comparing relative stabilities of AGT variants and variant complexes (i.e., by comparison of $C^{1/2}_{SUPREX}$ values collected at matching exchange times), this protocol could not be used for quantitative determinations of protein-ligand binding affinities.

Quantitative analysis requires the measurement of an accurate *m*-value, which requires the use of ΔG_{app} vs. $C^{1/2}_{SUPREX}$ plots such as those shown in Figure 14B. Quantitative ligand binding analysis of AGTmi using the conventional SUPREX protocol was not possible because the $C^{1/2}_{SUPREX}$ values collected for apo-AGTmi did not vary with exchange time as predicted by equation 2-1. This behavior is not uncommon for non-two-state proteins (43). Fortunately, the SUPREX-protease digestion protocol often allows for quantitative analysis of ligand binding for proteins with this type of SUPREX behavior, as the $C^{1/2}_{SUPREX}$ values of the proteolytic fragments often do shift with exchange time as predicted by equation 2-1 (51). Thus, quantitative ligand binding analysis of AGTmi was impossible using the conventional SUPREX protocol (without the protease digestion), but the use of the SUPREX-protease digestion protocol allowed for the quantitative analysis of the protein and its complexes.

The dimer form of AGTmi consists of two identical subunits, each of which contains two domains (91). Analysis of apo-AGTmi revealed slightly different SUPREX behavior for the various fragments. As shown in Table 8, *m*-values for apo-AGTmi fragments ranged from 3.3 to 6.3 kcal/(mol M). Some variation in *m*-value is expected due to the error associated with the SUPREX-based measurements. However, differences greater than ~30% can generally be regarded as significant (43). The *m*-value is related to the amount of hydrophobic surface area that is buried upon protein folding (69). Differences in *m*-value among the various apo-AGTmi fragments is most likely a reflection of their locations in the protein (i.e., fragments came from different locations in the protein that experience variable changes in buried hydrophobic surface area upon protein folding). Because of these differences in *m*-values for the various fragments, no m_{avg} or $\Delta G_{f,avg}$ values were calculated for apo-AGTmi. Since apo-AGTmi was not used to calculate $\Delta\Delta G_f$ values (see below), the variability in *m*-value did not hinder the ability to analyze ligand binding.

Both AGTmi ligands bind in the active site, with PLP acting as a cofactor and AOA acting as an inhibitor. By keeping the digestion conditions constant for all AGTmi experiments, the same fragments were obtained for apo-AGTmi and both AGTmi complexes. However, unlike the SUPREX results for apo-AGTmi, less variability was observed in ΔG_f and *m*-values between the fragments obtained from AGTmi complexes (see Table 8; the standard deviation in the *m*-value for apo-AGTmi is over twice as large

as the standard deviations in the *m*-values for both AGTmi-PLP and AGTmi-PLP-AOA). Thus, the discussion below will address the overall behavior of all fragments rather than addressing the fragments individually.

Unlike most of the interactions previously studied using SUPREX, the AGTmi-PLP interaction involves a covalent linkage between the protein and the cofactor. When AOA binds to AGTmi-PLP, AOA reacts with PLP, resulting in the breakage of the Schiff base linkage between PLP and AGTmi. Covalent bonds generally remain intact throughout the MALDI sample preparation, and previous studies on intact AGTmi reveal that PLP remains bound throughout the sample preparation protocol (49). However, note that bound PLP was not detected on any of the fragments that we analyzed in this work using the protease digestion protocol. This suggests that none of the detected fragments contained the covalent AGTmi-PLP linkage. Interestingly, detection of PLP-induced stabilization was still possible as shown by the shift in $C^{1/2}_{SUPREX}$ upon PLP binding (see Figure 14A). This suggests that this protocol most likely can be used to detect (and often quantify) ligand binding regardless of whether the detected fragments are located at the binding site. (However, it is possible that all of the fragments are close to the binding site in the three dimensional structure of the protein despite the lack of covalent attachment to PLP.)

SUPREX analysis of AGTmi-PLP revealed a decrease in *m*-value upon ligand binding (see Table 8 and the difference in slope in Figure 14B). As mentioned above, the

m-value is related to the amount of hydrophobic surface area that is buried upon protein folding and is a reflection of the cooperativity of the folding reaction. Clearly, ligand binding can result in an increase in buried hydrophobic surface area, but to detect such changes generally requires the increase in buried surface area to be quite large (i.e., on the order of thousands of square angstroms (69)). This large of a difference would not be expected for small molecule ligands such as PLP and AOA unless ligand binding triggers a significant conformational change in the protein. Furthermore, an increase in buried hydrophobic surface area should result in an increase in *m*-value rather than a decrease as observed for AGTmi-PLP.

This change in *m*-value instead is likely due to a change in the protein folding reaction as a result of PLP binding. Generally, it is assumed that ligand binding reactions do not result in large changes in the protein folding reaction. This is often a safe assumption since the ligand is less likely to associate with partially folded intermediates than with the fully folded protein. However, if the ligand does interact with a partially folded intermediate, the interaction would stabilize the intermediate, leading to an increase in the population of that intermediate state. The stabilization of an intermediate state would be expected to decrease the cooperativity of the protein folding reaction, resulting in a decrease in *m*-value. While this scenario may not contribute significantly to most ligand binding studies, the covalent nature of the AGTmi-PLP interaction increases the chances of interaction between PLP and a partially folded

intermediate. Thus, the decrease in *m*-value upon PLP binding is not particularly surprising.

Unfortunately, the change in *m*-value upon PLP binding precludes the determination of the binding affinity of PLP. However, note in Table 8 and Figure 14B that the *m*-value of AGTmi-PLP remains the same upon AOA binding. The similarity in *m*-value for these two complexes suggests that the folding reaction of AGTmi-PLP has not been significantly altered upon AOA binding. Thus, the SUPREX-derived $\Delta G_{f,avg}$ values can be used to calculate a $\Delta\Delta G_{f,avg}$ value, which can be used for the calculation of a K_d value for AOA binding to AGTmi-PLP (see Table 9). Since the *m*-value was the same for both AGTmi-PLP and AGTmi-PLP-AOA, all *m*-values measured for both complexes were averaged to obtain an m_{avg} value. This single m_{avg} value was used to calculate the $\Delta G_{f,avg}$ values that were then used in the calculations of $\Delta\Delta G_{f,avg}$ and K_d. These results provide the first measurement of the binding affinity of AOA for AGTmi-PLP.

5.5 Conclusions

High throughput analyses of protein-ligand binding are becoming progressively more important as research in the biomedical sciences reveals increasing numbers of potential drug targets. Although the SUPREX-protease digestion protocol offers a significant improvement in throughput relative to other H/D exchange-based techniques, the throughput of this technique can likely be increased further. Chapter 4

discussed single-point SUPREX, which is an abbreviated version of SUPREX that is capable of detecting protein-ligand binding (47, 50). While single-point SUPREX is effective for proteins that are amenable to conventional SUPREX analysis, it does not appear to be amenable to the analysis of large, multidomain proteins. Chapter 6 will discuss the combination of the SUPREX-protease digestion protocol with the single-point SUPREX assay. Using AGTmi as one of the model proteins, this work will demonstrate that the protease digestion strategy results in a significant increase in the efficiency and generality of single-point SUPREX.

Table 5: Transition midpoints of SUPREX curves obtained on apo-AGTmi using the SUPREX-protease digestion protocol.

Exchange Time (min)	3221-Da Fragment	3334-Da Fragment	3531-Da Fragment	3965-Da Fragment	4078-Da Fragment	4564-Da Fragment	5288-Da Fragment	5404-Da Fragment	5951-Da Fragment	6197-Da Fragment	6941-Da Fragment
0.5	2.0	2.0	1.9	2.0	2.1	1.7	2.0	1.9	1.7	1.7	1.8
2	1.9	1.9	1.8	1.8	1.8	1.5	1.8	1.7	1.7	1.5	1.7
5	1.7	1.7	1.6	1.7	1.6	1.3	1.7	1.5	1.5	1.4	1.5
25	1.4	1.5	1.4	1.4	1.4	1.0	1.5	1.3	1.4	1.3	1.4
40	1.4	1.4	1.3	1.2	1.2	1.1	1.5	1.2	1.4	1.2	1.3

Table 6: Transition midpoints of SUPREX curves obtained on AGTmi-PLP using the SUPREX-protease digestion protocol.

Exchange Time (min)	3221-Da Fragment	3334-Da Fragment	3531-Da Fragment	3965-Da Fragment	4078-Da Fragment	4564-Da Fragment	5288-Da Fragment	5404-Da Fragment	5951-Da Fragment	6197-Da Fragment	6941-Da Fragment
5	2.5	2.5	2.4	2.5	2.5	2.1	2.5	2.4	2.4	2.5	2.3
7	2.3	2.3	2.5	—	2.5	2.0	2.3	2.3	2.1	2.5	2.2
20	2.3	—	2.3	2.3	2.2	—	2.0	1.8	1.8	1.7	—
25	—	1.9	1.9	—	—	1.7	1.9	1.8	1.9	1.8	—
100	1.7	1.7	1.6	1.5	1.6	1.6	1.6	1.6	1.5	1.5	1.6
200	1.8	1.5	1.5	1.4	1.4	1.4	1.5	1.5	1.5	1.4	1.6
450	—	—	1.2	—	1.4	—	1.1	1.4	1.2	1.2	1.3
1410	1.4	1.3	—	1.2	1.2	—	—	1.2	—	1.0	0.9

Table 7: Transition midpoints of SUPREX curves obtained on AGTmi-PLP-AOA using the SUPREX-protease digestion protocol.

Exchange Time (min)	3221-Da Fragment	3334-Da Fragment	3531-Da Fragment	3965-Da Fragment	4078-Da Fragment	4564-Da Fragment[a]	5288-Da Fragment	5404-Da Fragment	5951-Da Fragment[a]	6197-Da Fragment[a]	6941-Da Fragment[a]
60	3.0	3.0	2.8	2.8	2.9	—	2.9	2.9	—	—	—
140	2.8	2.8	2.8	2.7	2.8	—	2.8	2.8	—	—	—
360	2.5	2.5	2.4	2.3	2.7	—	2.4	2.6	—	—	—
1800	2.2	2.2	2.0	2.0	2.1	—	2.1	2.2	—	—	—

[a]The $C^{1/2}_{SUPREX}$ values for this fragment did not behave as predicted by equation 2-1.

Table 8: SUPREX-derived ΔG_f and *m*-values for AGTmi and its complexes.

AGTmi	apo-AGTmi		AGTmi + PLP		AGTmi + PLP + AOA	
Frag-ment	ΔG_f[a] (kcal/mol)	*m*[a] (kcal/mol/M)	ΔG_f[a] (kcal/mol)	*m*[a] (kcal/mol/M)	ΔG_f[a] (kcal/mol)	*m*[a] (kcal/mol/M)
3221	-18.4 ± 0.4	4.2 ± 0.3	-18.7 ± 0.7	3.0 ± 0.4	-19.6 ± 0.3	2.3 ± 0.1
3334	-18.1 ± 0.3	4.0 ± 0.2	-17.6 ± 0.6	2.7 ± 0.3	-19.8 ± 0.3	2.4 ± 0.1
3531	-17.2 ± 0.4	3.7 ± 0.6	-16.1 ± 0.4	1.9 ± 0.2	-19.0 ± 0.6	2.1 ± 0.2
3965	-16.6 ± 0.6	3.3 ± 0.4	-16.8 ± 0.7	2.2 ± 0.4	-19.1 ± 0.8	2.2 ± 0.3
4078	-16.5 ± 0.2	3.2 ± 0.2	-17.0 ± 0.4	2.3 ± 0.2	-19.2 ± 1	2.1 ± 0.4
4564	-16.3 ± 0.4	3.8 ± 0.3	-17.8 ± 0.9	3.1 ± 0.5	—[b]	—[b]
5288	-19.9 ± 0.9	5.1 ± 0.5	-16.2 ± 0.2	2.0 ± 0.1	-19.2 ± 0.7	2.2 ± 0.3
5404	-17.1 ± 0.2	3.7 ± 0.2	-17.4 ± 0.6	2.6 ± 0.3	-21.7 ± 0.7	3.1 ± 0.3
5951	-21 ± 1	6.3 ± 0.7	-16.7 ± 0.5	2.4 ± 0.3	—[b]	—[b]
6197	-19.0 ± 0.6	5.4 ± 0.6	-16.3 ± 0.4	2.1 ± 0.2	—[b]	—[b]
6941	-19 ± 1	5.0 ± 0.8	-17.1 ± 0.3	2.5 ± 0.2	—[b]	—[b]

[a]The ΔG_f and *m*-values were taken from the linear least-squares analysis of the data using equation 2-1. Errors reported for the ΔG_f and *m*-values are the fitting errors of the linear least-squares analyses.

[b]The $C^{1/2}_{SUPREX}$ values for this fragment did not behave as predicted by equation 2-1.

Table 9: SUPREX-derived thermodynamic parameters for AGTmi complexes.

Protein or Protein-Ligand Complex	m_{avg} (kcal/mol/M)	$\Delta G_{f,avg}$[a] (kcal/mol)	$\Delta\Delta G_{f,avg}$ (kcal/mol)	SUPREX-derived K_d
AGTmi	—[b]	—[c]	—[c]	—[c]
AGTmi + PLP	2.4 ± 0.4	-17.0 ± 0.4	—	—
AGTmi + PLP + AOA	2.4 ± 0.4	-19.8 ± 0.2	-2.8 ± 0.4	80 ± 20 μM[d]

[a]$\Delta G_{f,avg}$ values were the average of folding free energy values calculated using equation 2-1 and an established *m*-value. The errors reported for the $\Delta G_{f,avg}$ values are the standard deviations of the folding free energy values calculated using equation 2-1 and the established *m*-value.

[b]No m_{avg} value was calculated for apo-AGTmi because the *m*-value was not the same for every fragment.

[c]Thermodynamic parameters for apo-AGTmi are not reported due to the difference in folding behavior of the apoprotein relative to the holoprotein (see text).

[d]Propagated error is reported.

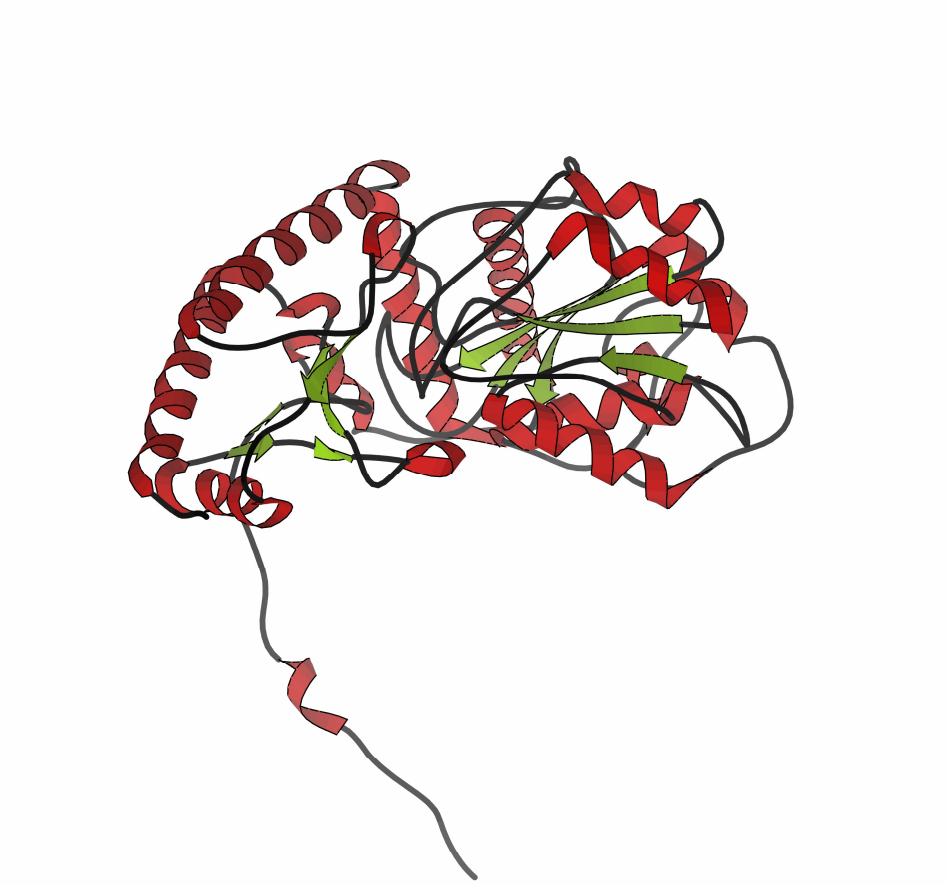


Figure 13: Crystal structure of the AGTma monomer. The domains within each monomer are closely associated and difficult to distinguish from one another. Note that AGTmi contains the following substitutions relative to AGTma: P11L and I340M. This image was created using entry 1h0c in the Protein Data Bank and the visualization program KiNG (Richardson Laboratory, Duke University).

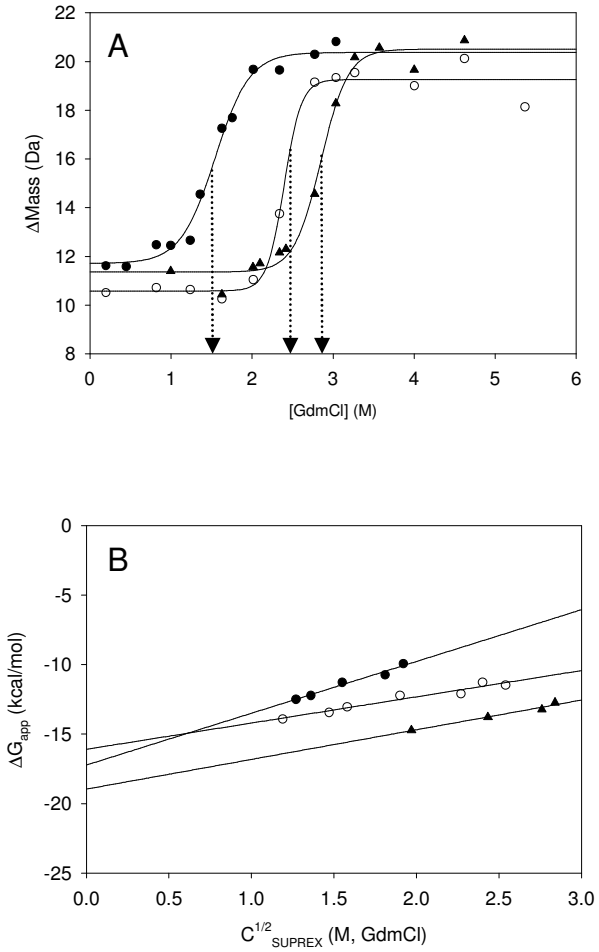


Figure 14: A) Representative SUPREX curves from fragment 5404 for apo-AGTmi (filled circles), AGTmi-PLP (open circles), and AGTmi-PLP-AOA (filled triangles). H/D exchange times were 5 min for apo-AGTmi and AGTmi-PLP. Due to its higher stability, a longer H/D exchange time (1 h) is shown for AGTmi-PLP-AOA. Dotted arrows depict the midpoint extracted from each SUPREX curve. B) Representative plots of ΔG_{app} vs. $C^{1/2}_{SUPREX}$ for fragment 3531 from apo-AGTmi (filled circles), AGTmi-PLP (open circles), and AGTmi-PLP-AOA (filled triangles). H/D exchange times ranged from 0.5 to 1800 min, and solid lines represent the linear least squares fitting of the data. Correlation coefficients for all linear fits were greater than 0.92.

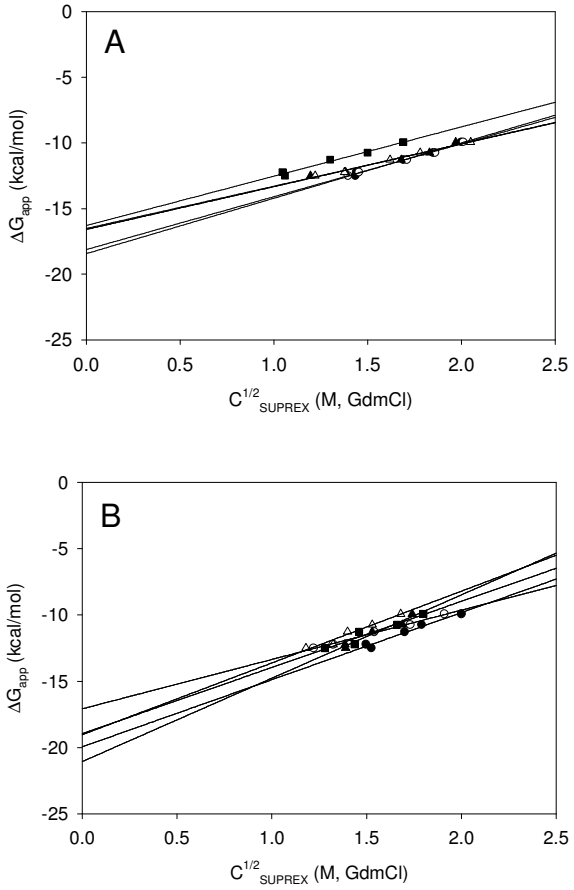


Figure 15: Plots of ΔG_{app} versus $C^{1/2}_{SUPREX}$ for peptic fragments from apo-AGTmi. (A) Peptides 3221 (filled circles), 3334 (open circles), 3965 (filled triangles), 4078 (open triangles), and 4564 (filled squares). (B) Peptides 5288 (filled circles), 5404 (open circles), 5951 (filled triangles), 6197 (open triangles), and 6941 (filled squares). The lines represent the results of the linear least-squares analysis of each data set, and the resulting y-intercept and slope were taken as the ΔG and *m*-value, respectively. The correlation coefficient obtained for each data set was greater than 0.92.

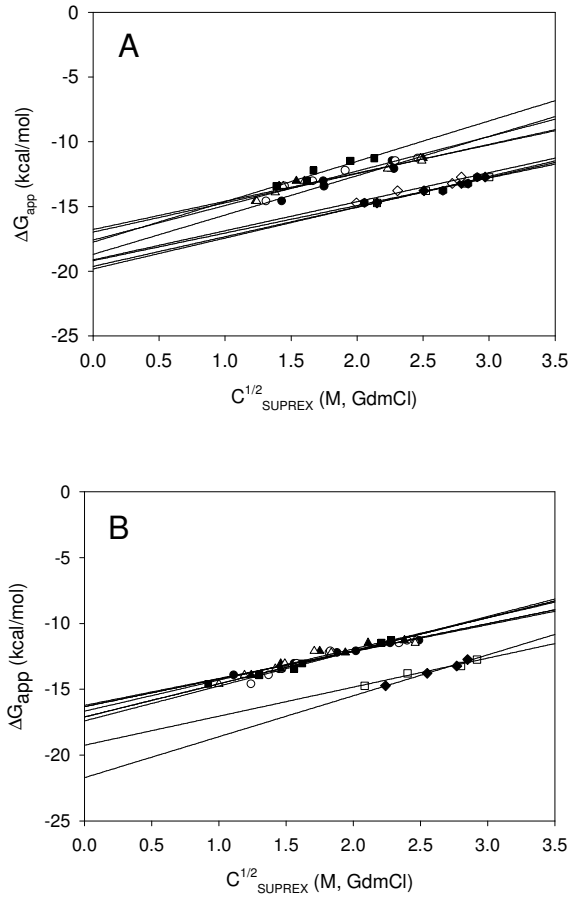


Figure 16: Plots of ΔG_{app} versus $C^{1/2}_{SUPREX}$ for peptic fragments from AGTmi complexes. (A) Plots for AGTmi-PLP (upper set of lines) peptides 3221 (filled circles), 3334 (open circles), 3965 (filled triangles), 4078 (open triangles), and 4564 (filled squares). The lower set of lines includes plots for AGTmi-PLP-AOA peptides 3221 (open squares), 3334 (filled diamonds), 3965 (open diamonds), and 4078 (filled hexagons). (B) Plots for AGTmi-PLP (upper set of lines) peptides 5288 (filled circles), 5404 (open circles), 5951 (filled triangles), 6197 (open triangles), and 6941 (filled squares). The lower set of lines includes plots for AGTmi-PLP-AOA peptides 5288 (open squares) and 5404 (filled diamonds). The lines represent the results of the linear least-squares analysis of each data set, and the resulting y-intercept and slope were taken as the ΔG and *m*-value, respectively. The correlation coefficient obtained for each data set was greater than 0.92.

6. Incorporation of a Protease Digestion Strategy into the Single-Point SUPREX Protocol

6.1 Introduction

Chapter 4 discussed the application of single-point SUPREX as a screening assay for the detection of protein-ligand binding (74, 92). This work resulted in the identification of an important limitation of the single-point SUPREX protocol. It was found that the efficiency of the assay is highly dependent on the precision of the mass measurements and on the number of globally protected amide protons in the protein. In cases where the error in the MALDI-MS mass measurement is relatively large compared to the number of globally protected amide protons, the efficiency of the assay is decreased. For example, the efficiency of the CypA screen was significantly reduced in comparison with the screening efficiency found for the S-protein in the original proof-of-concept study (74). Since CypA and S-protein have similar numbers of globally protected amide proteins (as judged by the amplitudes of their SUPREX curves), the reduced efficiency was attributed to the errors associated with the MALDI mass measurements on CypA (i.e., ± ~6 Da) being larger than those associated with the S-protein (i.e., ± 1.5 Da).

This chapter describes the introduction of a protease digestion strategy into the single-point SUPREX protocol. The protease digestion strategy employed here is directly analogous to the protease digestion strategy that was recently developed to expand the scope of the conventional SUPREX protocol (51). As in the previous work with the

SUPREX-protease digestion protocol (see Chapter 5), the SUPREX behavior of the peptides reports on the global thermodynamic properties of the intact protein. However, since the absolute uncertainty of MALDI mass measurements decreases as molecular weight decreases, the SUPREX behavior of the protein can be determined with greater accuracy and precision by reading out the peptide molecular weights.

The new protocol described here was evaluated in the context of two model protein systems, including CypA and AGTmi. CypA is a potential target for diagnostic imaging agents and pharmaceutical intervention because it is overexpressed in lung tumor cells and seems to be required for tumor growth (5, 6). This protein was chosen for this study to allow for a direct comparison to the results obtained in the previous screening experiment with CypA (see Chapter 4) (92). AGTmi is a potential target for therapeutic intervention via pharmacolgical chaperones (i.e., compounds that bind and stabilize the native state of a protein) for patients with PH1 (62, 93). AGTmi was chosen for this study because it represents a protein that is not amenable to screening using the original single-point SUPREX protocol due to its large size and relatively small number of globally protected amide protons.

6.2 Experimental Section

6.2.1 Materials

Human CypA (92) and human AGTmi (49) were expressed and purified as described in Chapters 3 and 4. The following reagents were purchased from Sigma-

Aldrich (St. Louis, MO): porcine pepsin, bovine pancreas insulin, and PLP. The internal standard ACTH clip (18-139) was from Bruker Daltonics (Billerica, MA). CsA was purchased from LKT Laboratories (St. Paul, MN).

6.2.2 General Methods and Instrumentation

Deuterated H/D exchange buffers were prepared as described in Chapter 2. For the screening assay, the H/D exchange buffers contained appropriate concentrations of deuterated GdmCl such that the denaturant concentrations during H/D exchange were 2.5 and 2 M for CypA and AGTmi, respectively (see Figure 17).

MALDI mass spectra were collected using a Bruker Daltonics (Billerica, MA) Ultraflex II TOF/TOF mass spectrometer with the following instrument parameters: an ion source 1 voltage of 25.0 kV, an ion source 2 voltage of 23.4 kV, a lens voltage of 6.5 kV, and a delay time of 100 ns. Gating was used to suppress peptides with molecular weights of less than 1 kDa for CypA and less than 3 kDa for AGTmi. Mass spectra were a sum of 100 laser shots. A total of 5 (for the screening assay) or 10 (for the SUPREX-protease digestion protocol) mass spectra were collected for each data point.

6.2.3 Conventional SUPREX-Protease Digestion Protocol

Full SUPREX curves for CypA were collected using the SUPREX-protease digestion protocol as discussed in Chapter 5 (51). This protocol involved performing a rapid protease digestion immediately after the H/D exchange reaction had been

quenched. The digestion was performed using the protease pepsin due to its high activity under acidic quench conditions.

SUPREX analyses performed on CypA using the protease digestion protocol involved combining 1 μL of a 50 μM CypA stock solution with 0.5 μL of either DMSO or 1 mM CsA in DMSO. H/D exchange was initiated by adding 3.5 μL of SUPREX buffer to each protein solution. After a five minute exchange time, the reaction was quenched by adding 4 μL of the reaction mixture to 3 μL of aqueous HCl (pH ~ 0.8). A short pepsin digestion was performed by adding 3 μL of a 1 μM pepsin solution to the quenched protein mixture. After a 1 min digestion, a 2 μL aliquot of the digested protein mixture was combined with 6 μL of matrix solution (saturated SA in 50% ACN/49.9% water/0.1% TFA) containing an internal standard (ACTH clip (18-139)), and 1 μL of the final matrix solution was spotted onto the MALDI target. Spots were dried quickly using a stream of nitrogen gas.

For AGTmi, H/D exchange reactions were initiated by adding 3 μL equilibrated protein complex to 7 μL of each SUPREX buffer following the protocol in Chapter 5. The remainder of the SUPREX protocol used to analyze AGTmi was very similar to that described above for the CypA analysis with the following exceptions: 1) the AGTmi stock solution was 40 μM; 2) the ligand (PLP) was dissolved in phosphate buffered saline (PBS) rather than DMSO; 3) AGTmi was equilibrated with PBS or PLP in PBS for 1

h at room temperature prior to SUPREX analysis; 4) the stock pepsin concentration was ~0.2 μM; 5) the digestion time was 30 sec; and 6) the internal standard was insulin.

6.2.4 Single-Point SUPREX-Protease Digestion Protocol

The single-point SUPREX assay was performed by distributing positive and negative controls (CsA and DMSO, respectively, for CypA; PLP and PBS, respectively, for AGTmi) in alternating wells of a microtiter plate. Exchange reactions were performed in sets of 10 (i.e., 5 positive and 5 negative controls per set), and MALDI analysis of each set was performed before the exchange reaction for the subsequent set was initiated. For AGTmi, the protein-ligand complexes were equilibrated as described in Section 6.2.3, and H/D exchange reactions were scaled down such that the final volume of the H/D exchange solution was 5 μL instead of 10 μL. This helped to reduce protein consumption. The remainder of the protocol is identical to the SUPREX protocol described above except that instead of using a series of SUPREX buffers containing a range of GdmCl concentrations, a single GdmCl concentration was used for each protein. Six sets of 10 controls were analyzed for each model protein.

6.2.5 Data Analysis

The mass-to-charge ratio for each peptide peak was determined using an in-house Matlab (The MathWorks, Inc., Natick, MA) program as described in Chapter 2. Using this program, the data were smoothed using a 9-point (CypA) or 19-point (AGTmi) floating average smoothing of the data. In the case of the conventional

SUPREX-protease digestion protocol, SUPREX curves were plotted and analyzed as described in Chapter 2. For the control screening assay, the Δmass values were subjected to a three-point moving average smoothing. The smoothed Δmass values will hereafter be called $\Delta mass_{av}$ values (92). Z′-factors were calculated using equation 2-3 (56). The use of the moving average allowed for the calculation of a Z′-factor for each $\Delta mass_{av}$ value, and each Z′-factor was averaged to determine a single average Z′-factor for each peptide (see Tables 11 and 12).

6.3 Results

6.3.1 The SUPREX-Protease Digestion Protocol

Before performing either the conventional SUPREX-protease digestion protocol or the single-point SUPREX protocol, the protease digestion was optimized for both CypA and AGTmi. To optimize the digestion, a variety of pepsin concentrations and digestion times were tested to obtain as many fragments as possible in the 1.5 to 10 kDa mass range. Once the protease digestion conditions were optimized, the number of fragments suitable for SUPREX analysis was evaluated. Suitable fragments for SUPREX are those that produce reliable ion signals across a range of denaturant concentrations and that yield observable SUPREX transitions. Under the conditions used in this work, 10 suitable fragments for CypA and 11 suitable fragments for AGTmi were identified (see Tables 10-12 for a list of fragments for each protein).

This work involved the first application of the SUPREX-protease digestion protocol to CypA. Although a variety of exchange times could have been used, a relatively short exchange time (5 min) was chosen to minimize the uptake of deuterons due to local fluctuations in protein structure. This helped to maximize the difference in Δmass between the positive and negative controls. The $C^{1/2}_{SUPREX}$ values collected for apo- and holo-CypA are listed in Table 10. The collection of full SUPREX curves for CypA allowed for the selection of an appropriate denaturant concentration for the screening assay.

Chapter 3 discussed the qualitative analysis of the intact form of AGTmi using SUPREX with an exchange time of 5 min (49), and Chapter 5 described the analysis of AGTmi using the SUPREX-protease digestion protocol and a variety of exchange times. The work in Chapter 5 using the SUPREX-protease digestion protocol included the collection of SUPREX curves using a 5 min exchange time, and these curves were used to determine an appropriate denaturant concentration for the screening assay in this chapter.

Figure 17 shows representative CypA and AGTmi SUPREX curves that were used for the selection of appropriate experimental conditions for the screening assays. At the chosen denaturant concentrations for the assays (see dashed lines in Figure 17), a significant difference in Δmass was observed between the bound and unbound forms of each protein. The full SUPREX curves for CypA and AGTmi helped to define the

SUPREX behavior of the fragments and to facilitate the selection of an appropriate denaturant concentration for each screening assay. In the cases of CypA and AGTmi, the SUPREX behavior of each fragment was within the experimental error typically observed in SUPREX experiments (i.e., SUPREX transition midpoints were within 0.4 M of each other). Furthermore, all of the CypA and AGTmi fragments that were analyzed were sensitive to ligand binding (i.e., a shift in the transition midpoint was observed for each fragment when the proteins were analyzed in the presence of a ligand). Thus, these sets of fragments were monitored (10 for CypA and 11 for AGTmi) during the screening assays.

6.3.2 Evaluation of the Single-Point SUPREX-Protease Digestion Assay

For both CypA and AGTmi, the assay was performed on 6 sets of 10 controls (5 positive controls and 5 negative controls). Figures 18 and 19 contain assay data for the most and least ideal peptides for single-point SUPREX. The best peptides for screening were those that showed a clear separation between the positive and negative controls (Figures 18A and 19A), whereas the least ideal peptides for screening were those that did not show a clear separation between the controls (Figures 18B and 19B). Screening data for the remaining CypA and AGTmi peptides are included in Figures 20 through 25. The dashed lines in these figures represent the cutoff values that would typically be used to select for hits when performing a single-point SUPREX assay. These cutoff

values were determined by subtracting three standard deviations from the $\Delta mass_{av}$ values as discussed in Chapter 4 (92).

The screening efficiency is defined as the accuracy with which the assay can select ligands for the target protein. One way to evaluate efficiency is to examine the rates of false positives and false negatives; this is easily accomplished using the control data collected for CypA and AGTmi along with the calculated cutoff values described above. As shown in Tables 11 and 12, the false positive and false negative rates were 0% for many of the peptides (CypA peptides 1-5 and 7-9 and AGTmi peptides 1-3, 5, and 7-8). The false negative rate is a significant improvement over the screening of intact CypA (see Chapter 4), in which a false negative rate of 18% was observed (92). Even for CypA peptides that had Z′-factors equal to the average Z′-factor for the intact CypA screening (-0.2), the false negative rate was lower (0-3%) when the protease digestion protocol was employed.

Another method for evaluating screening efficiency involves an analytical parameter called the Z′-factor. This parameter gives a measure of the separation between the positive and negative controls in an assay (56). It is generally accepted that the Z′-factor must be at least zero for an assay to be viable, and a Z′-factor equal to zero is acceptable for a "yes/no" assay. However, Z′-factors greater than zero ($0 < Z' \leq 1$) are better for screening assays, with larger values (up to $Z' = 1$) indicating improved separation between controls and thus improved efficiency. Using the protease digestion

protocol described here, a significant improvement in Z′-factors was achieved for CypA (see Table 11 and Figures 18 and 20-22). As mentioned above, the maximum Z′-factor achieved when screening the intact protein was 0. In contrast, when the protease digestion protocol was applied to CypA, Z′-factors as high as 0.4 were observed, and 7 of the 10 peptides yielded Z′-factors ≥ 0 (see Table 11).

The results of the control assay for AGTmi were similar to those of CypA. The maximum Z′-factor observed for AGTmi was 0.5, indicating an excellent separation between positive and negative controls. Z′-factors of at least zero were observed for 9 of the 11 AGTmi peptides (see Table 12). As expected, the false positive and false negative rates for both CypA and AGTmi are lowest for peptides with the highest Z′-factors (see Tables 11 and 12) since a higher Z′-factor indicates better separation between controls.

6.4 Discussion

The improved screening protocol described here is a combination of single-point SUPREX (74, 92) and the SUPREX-protease digestion protocol (see Chapters 4 and 5) (51). Similar to single-point SUPREX, this protocol is an abbreviated form of SUPREX in which ligand binding is detected by monitoring the uptake of deuterons at a single denaturant concentration. However, in this work a protease digestion step was incorporated prior to mass spectral analysis. Protease digestion has previously been incorporated into the SUPREX protocol to obtain domain-specific thermodynamic parameters (51) and to facilitate the quantitative thermodynamic analysis of proteins

that are not amenable to the conventional SUPREX protocol (see Chapter 5). In the SUPREX-protease digestion protocol, the digestion is performed after H/D exchange but before MALDI-MS analysis. This allows for the H/D exchange behavior of each protein fragment to reflect the H/D exchange behavior of the intact protein.

In Chapter 4, a single-point SUPREX protocol was used to screen for binding to CypA (92). This study permitted a calculation of the throughput and efficiency of single-point SUPREX for the first time. The efficiency of single-point SUPREX was found to be maximized when two factors are optimized: 1) the standard deviation of the Δmass values should be as small as possible, and 2) the difference in Δmass values between the positive and negative controls should be as large as possible. Since variability in the mass measurement of larger proteins by MALDI becomes greater as the molecular weight of the protein increases, the standard deviations of the Δmass measurements also increase as protein molecular weight increases. In contrast, the difference in Δmass between the positive and negative controls depends on the number of globally protected amide protons and thus is largely protein-dependent.

In some cases, the difference in Δmass between the positive and negative controls is large enough to compensate for large standard deviations, but for many proteins such as AGTmi, the large standard deviations will likely preclude screening by single-point SUPREX. Since the number of globally protected amide protons is protein-dependent, a more reasonable approach for increasing efficiency involves decreasing the

standard deviations of the mass measurements. It was hypothesized that the simplest way to decrease the standard deviations was to fragment the protein into smaller pieces, which is easily accomplished using a protease digestion strategy.

Since the protease digestion is performed after the H/D exchange reaction, the SUPREX behavior of each peptide reports on the global stability of the intact protein. Thus, unlike other common H/D exchange and protease digestion techniques, extensive peptide maps of the protein are not required to perform single-point SUPREX. Extensive peptide coverage is not important for this assay since the detection of ligand binding does not require the analysis of a peptide that is located at the binding site. The basic requirement for the assay is that at least one peptide can be detected that will report on ligand binding. For multidomain proteins, this may require the detection of a peptide in the ligand binding domain. In some proteins (particularly large, multidomain proteins), different fragments may demonstrate variable SUPREX behaviors (51). In these cases, multiple fragments may be monitored to increase the probability of selecting for compounds that influence protein stability in diverse ways.

In the previous study on CypA (see Chapter 4), it was found that under ideal conditions, a successful screen for CypA ligand binding could be accomplished with a Z′-factor of 0. The goal of the work in this chapter was to improve assay efficiency (i.e., increase the Z′-factor) and increase the generality of the assay. It was found that the incorporation of a protease digestion strategy indeed resulted in increased Z′-factors.

Unlike CypA, intact AGTmi is not amenable to single-point SUPREX. Due to its larger size (~44 kDa per subunit), the standard deviations for AGTmi mass measurements were significantly larger than those for CypA. Using data from the SUPREX analysis of intact AGTmi (49), the predicted maximum Z'-factor for the intact protein is -0.2, indicating that screening the intact protein using a single-point SUPREX protocol without protease digestion would not yield useful results. In contrast, when a protease digestion step was incorporated into the protocol, AGTmi became amenable screening using single-point SUPREX, and screening efficiencies comparable to those of CypA (using the protease digestion protocol) were achieved. Interestingly, a relationship between Z'-factor and peptide molecular weight was not observed in our studies on CypA and AGTmi; instead, the Z'-factors appeared to be completely peptide-dependent.

In addition to the improvement in Z'-factor, the single-point protease digestion protocol described here has the added advantage of allowing for replicate measurements within a single screening. For example, in the screening of CypA, one could choose to monitor fragments 1,2,3, and 8 (see Table 11). This allows for a built-in redundancy within each screening. However, it is noteworthy that some sources of error (e.g., back exchange) may influence different peptides in a similar way, causing the error to be observed for each peptide. Even if this is the case, the improved efficiency of this assay should lead to a significant reduction in false positives and false negatives compared to screening using the intact protein.

The new protocol described in this chapter is expected to significantly improve the generality of single-point SUPREX. It is reasonable to hypothesize that the incorporation of the protease digestion protocol should diminish the protein molecular weight limitations described above (see Section 6.1). In addition to reducing the standard deviations of the MALDI mass measurements, the single-point protease digestion protocol could allow for the screening of very large proteins that are typically difficult to detect using MALDI-MS. For example, recent work in the Fitzgerald Laboratory has shown that glycogen phosphorylase b (~98 kDa/subunit) is amenable to analysis using the SUPREX-protease digestion protocol, even though detection of the intact protein using MALDI-MS is difficult. Proteins that can be analyzed by the SUPREX-protease digestion protocol are expected to be amenable to screening using the single-point protease digestion protocol.

The addition of the protease digestion step did not significantly alter the throughput of the protocol. The control assays in this work were performed at a rate of 1.8 min/ligand as compared to 3 min/ligand in the previous screening described in Chapter 4. This increase in throughput was observed despite the fact that an extra step was added to the protocol and that the screening was performed by one researcher instead of two. The increased throughput in this work is attributed to the use of a high throughput MALDI mass spectrometer. It is also predicted that the throughput of single-point SUPREX and the single-point SUPREX with protease digestion protocol can

be improved further as described in Chapter 4 (92). The goal of the work in this chapter was to improve the efficiency rather than the throughput, and the incorporation of a protease digestion strategy into the single-point SUPREX protocol is a simple and effective means with which to accomplish this goal.

6.5 Conclusions

The efficiency of single-point SUPREX was improved by incorporating a protease digestion protocol into the assay. The new protocol was evaluated using two model proteins, CypA and AGTmi. With this strategy, Z'-factors of up to 0.5 were achieved, and false positive and false negative rates of 0% were observed for many of the peptides. Because this strategy served to reduce the standard deviations of the mass measurements, it is expected to be applicable to a wide range of proteins, including large proteins that are not easily detected in their intact form using MALDI-MS. In addition, the approach described here should be applicable to proteins that require protease digestion for quantitative analysis by SUPREX. Since multiple peptides can be monitored simultaneously, this assay allows for replicate measurements within a single screening. This development promises to reduce the rates of false positives and false negatives in single-point SUPREX assays and to make the assay applicable to a wider variety of protein targets.

Table 10: $C^{1/2}_{SUPREX}$ values for CypA fragments.

Fragment Number	Molecular Weight	apo-CypA $C^{1/2}_{SUPREX}$ (M, GdmCl)[a]	CypA + CsA $C^{1/2}_{SUPREX}$ (M, GdmCl)[a]
1	1623	1.9	3.7
2	1865	1.9	3.8
3	2375	1.6	3.4
4	2427	1.6	3.6
5	2574	1.6	3.7
6	2671	1.7	3.9
7	2742	1.7	3.9
8	2839	1.6	3.4
9	3083	1.9	3.8
10	6046	2.1	—[b]

[a]Exchange time = 5 min

[b]No transition detected.

Table 11: Screening efficiency for CypA.

Fragment number	Molecular Weight (Da)	Z′-Factor	False Positives	False Negatives
1	1623	0.4	0	0
2	1865	0.4	0	0
3	2375	0.4	0	0
4	2427	0.2	0	0
5	2574	0.1	0	0
6	2671	-0.2	0	1/30 (3%)
7	2742	-0.2	0	0
8	2839	0.4	0	0
9	3083	0.2	0	0
10	6046	-0.9	3/30 (10%)	0

Table 12: Screening efficiency for AGTmi.

Fragment number	Molecular Weight (Da)	Z′-Factor	False Positives	False Negatives
1	3221	0.4	0	0
2	3334	0.5	0	0
3	3531	0.3	0	0
4	3965	0.0	0	1/30 (3%)
5	4078	0.1	0	0
6	4564	-0.3	0	1/30 (3%)
7	5288	0.2	0	0
8	5404	0.3	0	0
9	5951	0.0	0	1/30 (3%)
10	6197	0.1	1/30 (3%)	0
11	6941	-1.3	0	5/30 (17%)

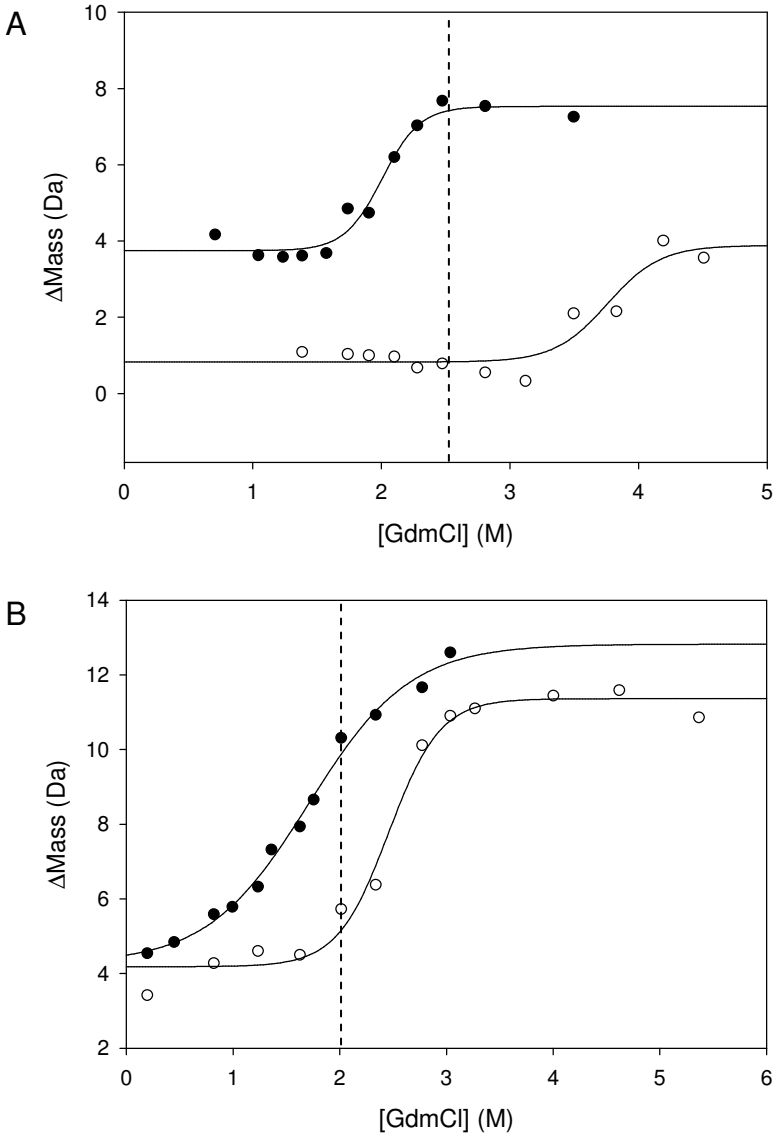


Figure 17: Representative SUPREX curves for A) apo-CypA fragment 1865 (filled circles) and CypA-CsA fragment 1865 (open circles) and B) apo-AGTmi fragment 3334 (filled circles) and AGTmi-PLP fragment 3334 (open circles). In all cases, the exchange time was 5 min. The dashed lines indicate the chosen denaturant concentration for the screening assay.

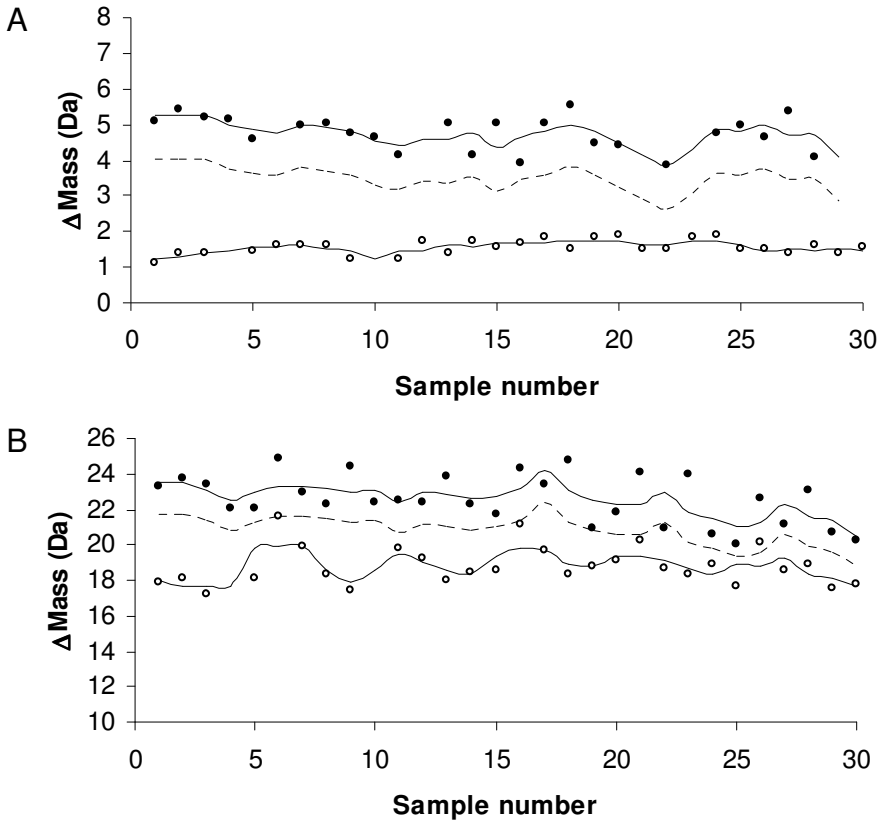


Figure 18: Representative control screening for CypA peptides with A) the best (fragment 2) and B) the worst (fragment 10) observed Z′-factors. In each panel the filled circles represent the negative controls, and the open circles represent the positive controls. The upper and lower solid lines represent the $\Delta mass_{av}$ of the negative and positive controls, respectively, and the dashed line represents the cutoff values, which are three standard deviations below the negative control $\Delta mass_{av}$ values.

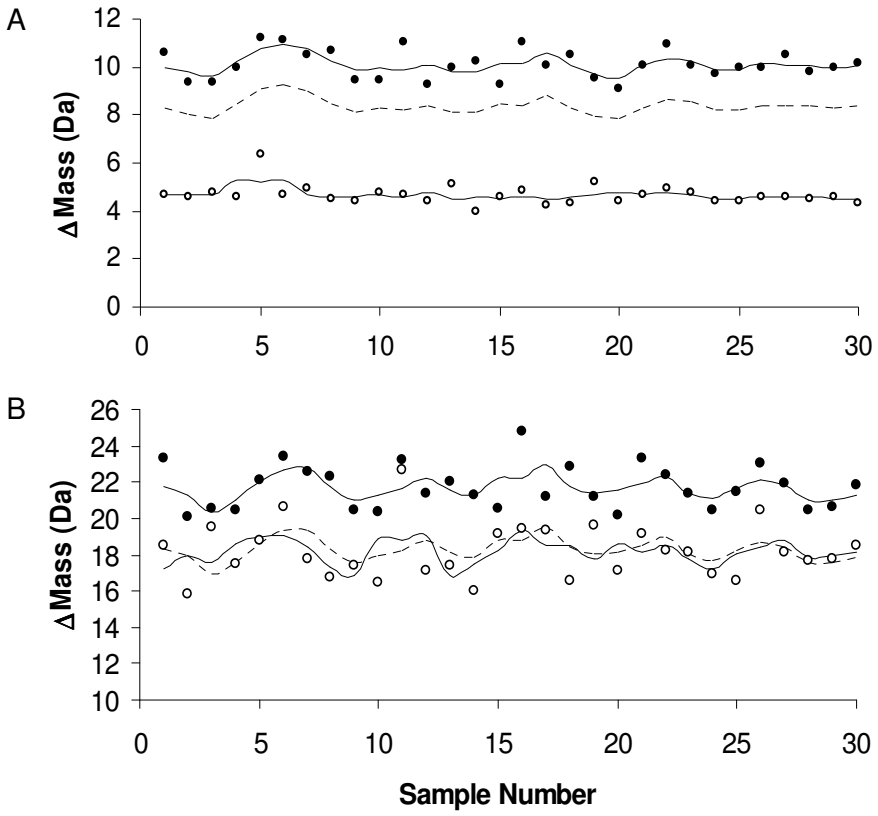


Figure 19: Representative control screening for AGTmi peptides with A) the best (fragment 2) and B) the worst (fragment 11) observed Z′-factors. In each panel the filled circles represent the negative controls, and the open circles represent the positive controls. The upper and lower solid lines represent the $\Delta mass_{av}$ of the negative and positive controls, respectively, and the dashed line represents the cutoff values, which are three standard deviations below the negative control $\Delta mass_{av}$ values.

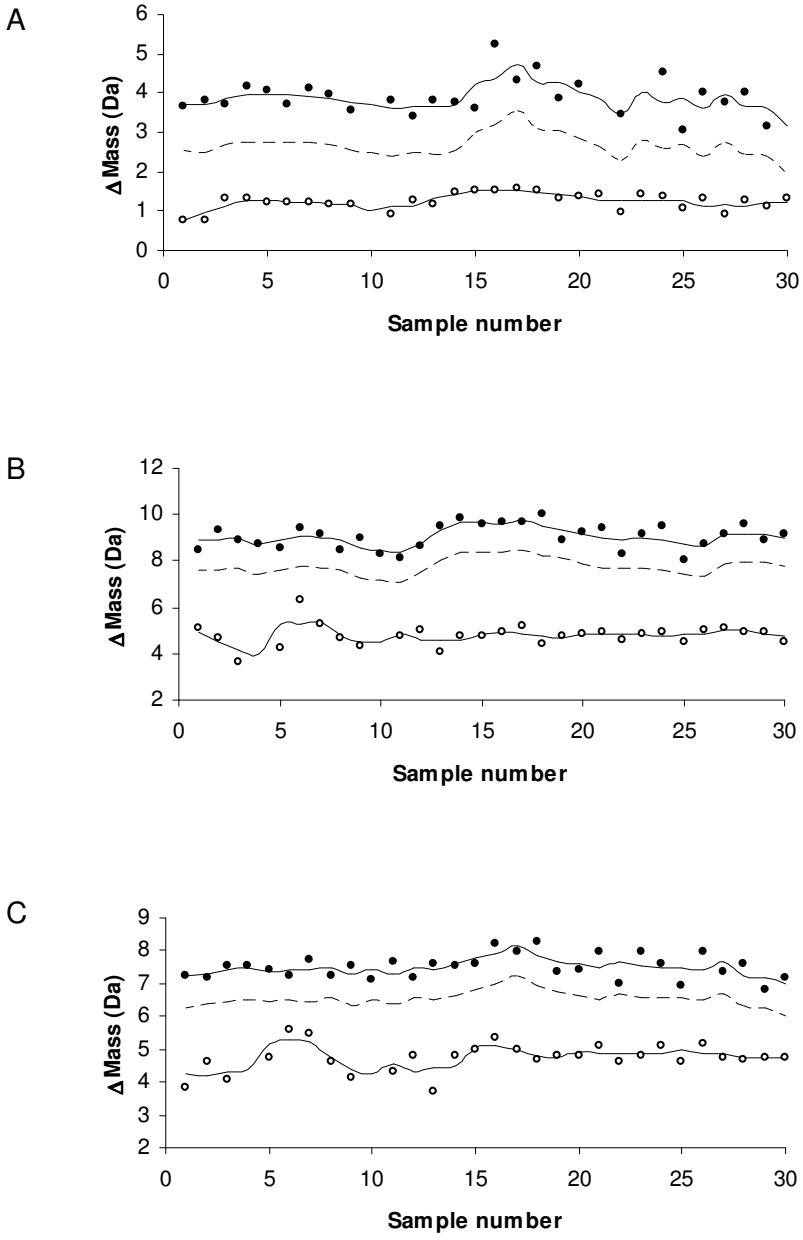


Figure 20: Control screening data for CypA peptides A) 1, B) 3, and C) 4. ΔMass values are plotted for negative controls (filled circles) and positive controls (open circles), and the moving average for each set of controls is represented by a solid line. The dashed line indicates the suggested cutoff values for each peptide.

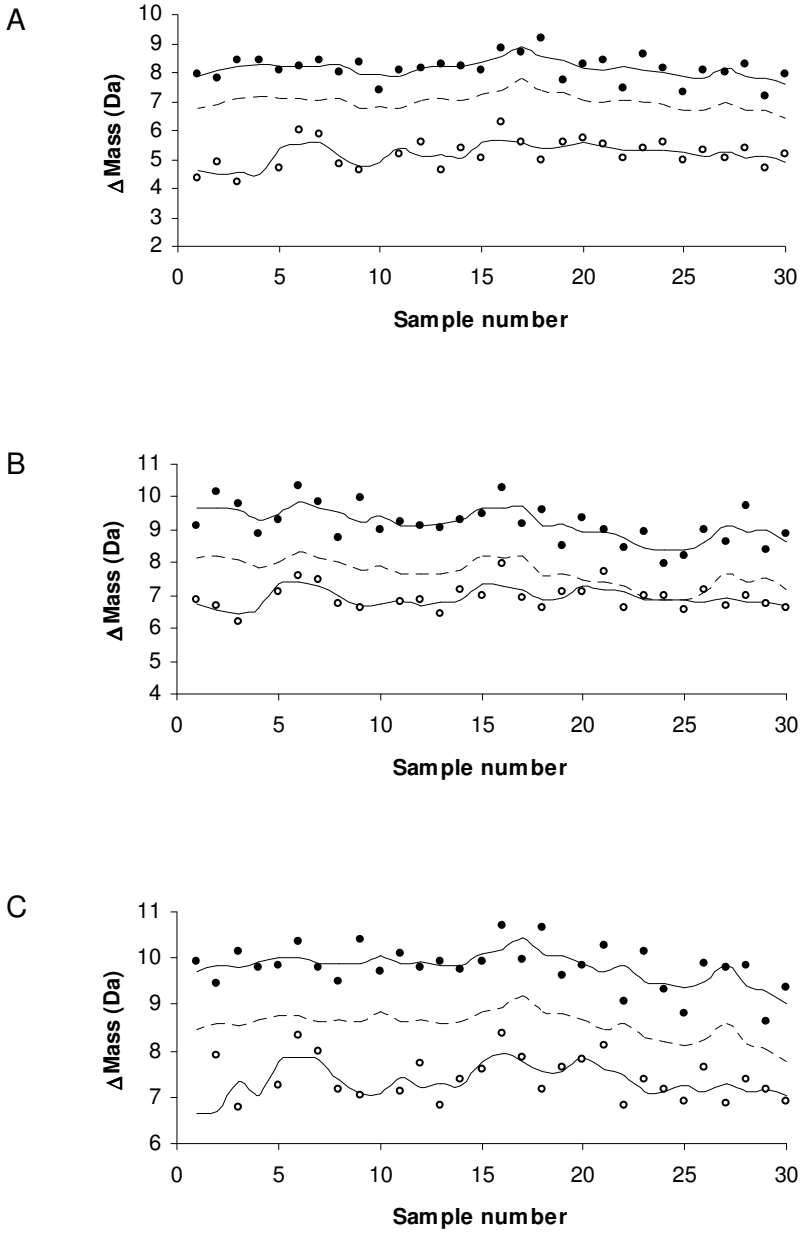


Figure 21: Control screening data for CypA peptides A) 5, B) 6, and C) 7. ΔMass values are plotted for negative controls (filled circles) and positive controls (open circles), and the moving average for each set of controls is represented by a solid line. The dashed line indicates the suggested cutoff values for each peptide.

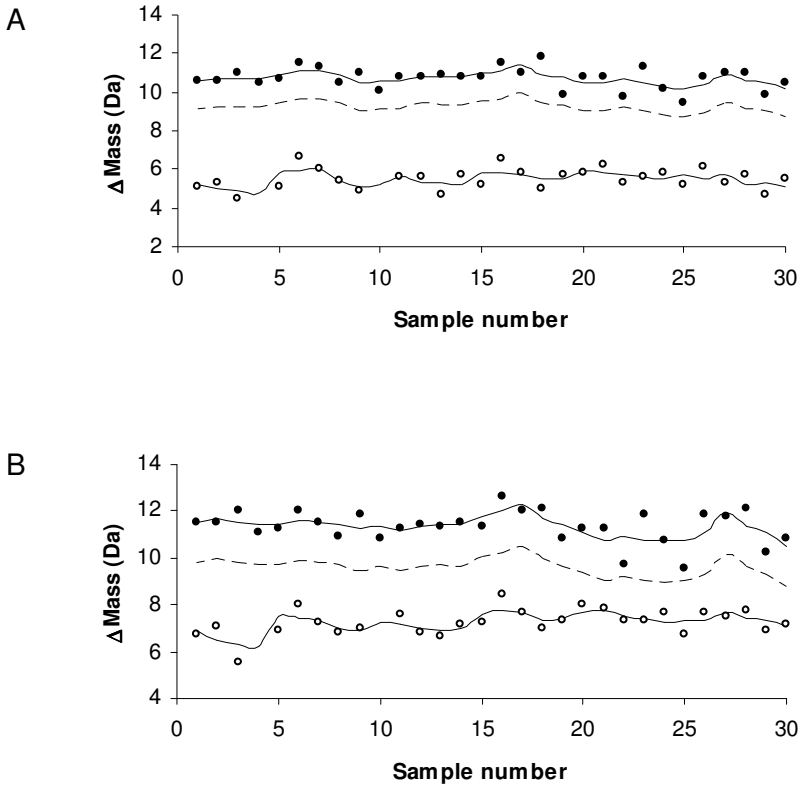


Figure 22: Control screening data for CypA peptides A) 8 and B) 9. ΔMass values are plotted for negative controls (filled circles) and positive controls (open circles), and the moving average for each set of controls is represented by a solid line. The dashed line indicates the suggested cutoff values for each peptide.

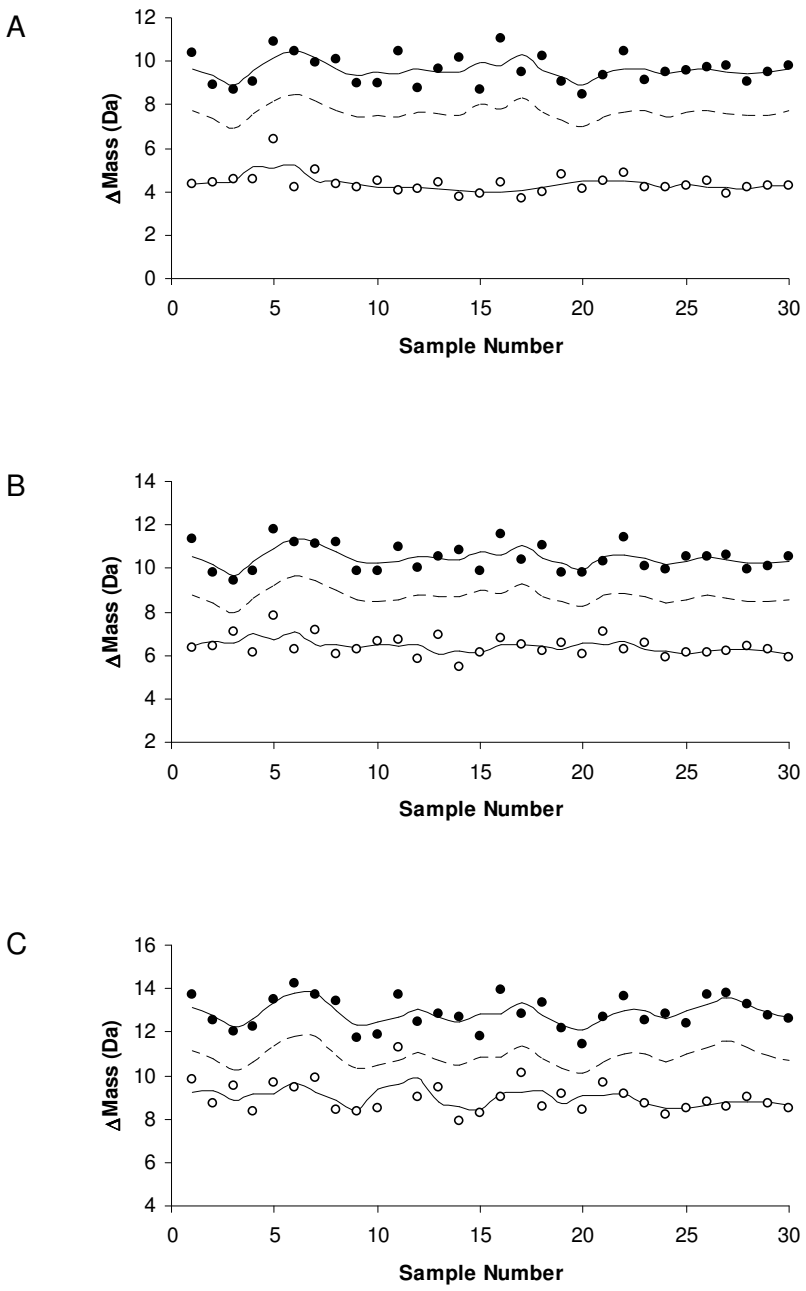


Figure 23: Control screening data for AGTmi peptides A) 1, B) 3, and C) 4. ΔMass values are plotted for negative controls (filled circles) and positive controls (open circles), and the moving average for each set of controls is represented by a solid line. The dashed line indicates the suggested cutoff values for each peptide.

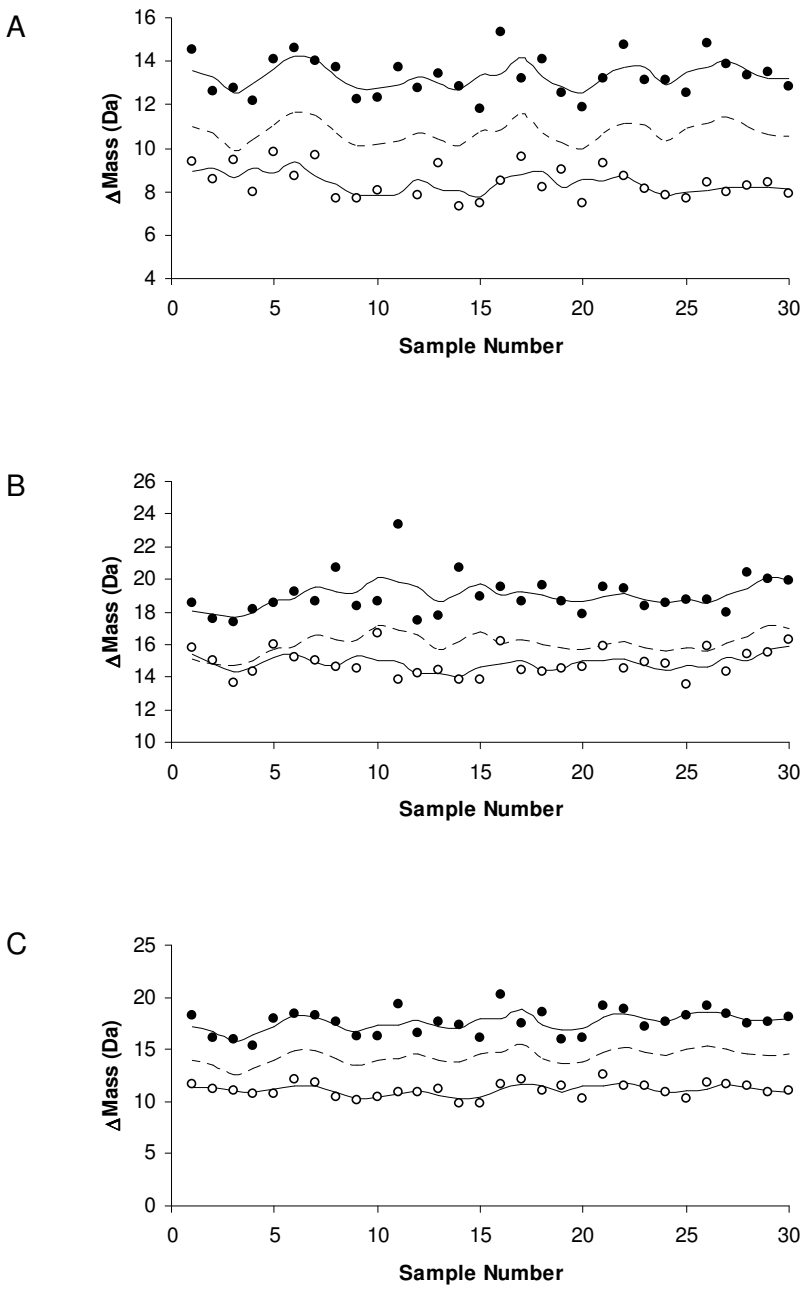


Figure 24: Control screening data for AGTmi peptides A) 5, B) 6, and C) 7. ΔMass values are plotted for negative controls (filled circles) and positive controls (open circles), and the moving average for each set of controls is represented by a solid line. The dashed line indicates the suggested cutoff values for each peptide.

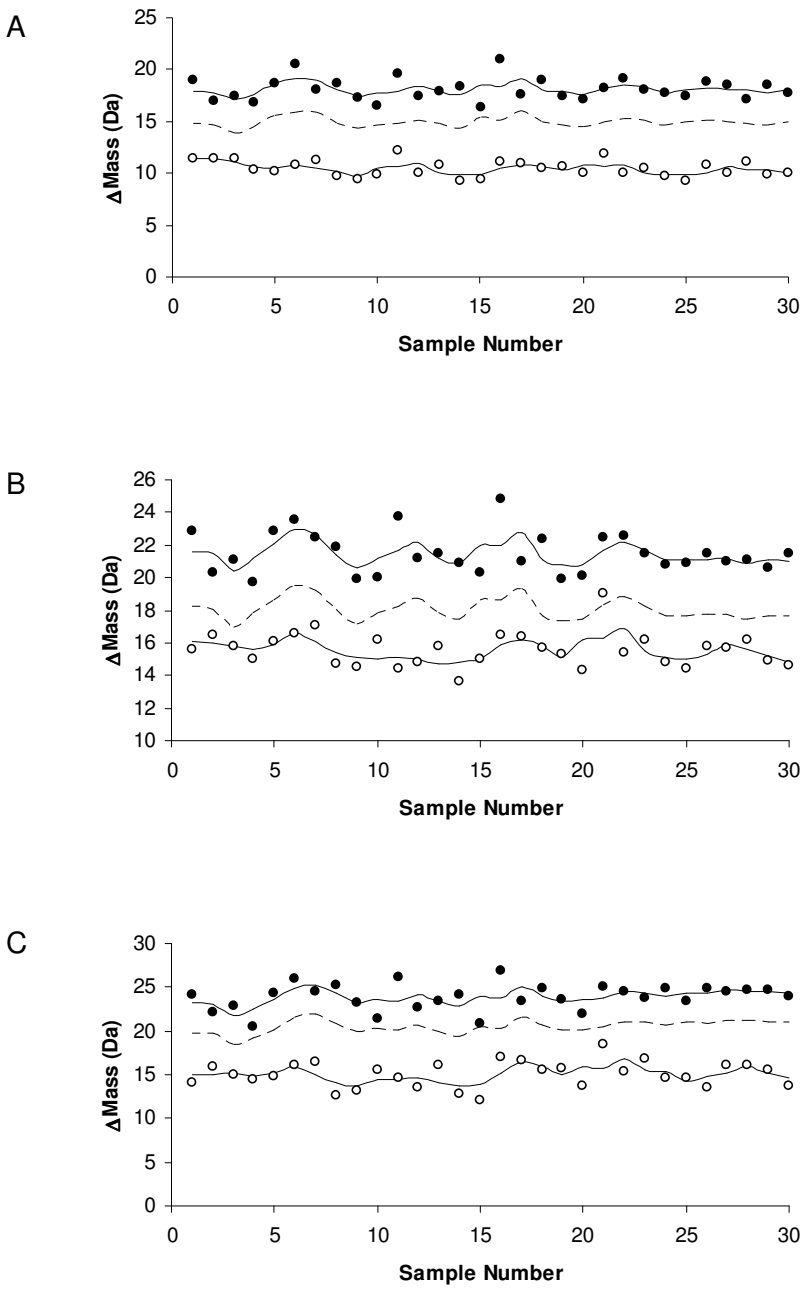


Figure 25: Control screening data for AGTmi peptides A) 8, B) 9, and C) 10. ΔMass values are plotted for negative controls (filled circles) and positive controls (open circles), and the moving average for each set of controls is represented by a solid line. The dashed line indicates the suggested cutoff values for each peptide.

Chapter 7. Thermodynamic Analysis of Protein Folding in Complex Mixtures using a Gel-Based Fractionation Strategy

The majority of this dissertation project has centered on the high throughput applications of SUPREX. In particular, this work has focused on the use of SUPREX to screen libraries of compounds for binding to a single target protein. However, another goal of this work involves the use of chemical modification techniques such as SUPREX to perform the reverse experiment: to screen multiple proteins for binding to a single target compound. While this idea may sound straightforward, performing this task is a significant challenge that will require the ability to perform chemical modification techniques on entire proteomes in the presence and absence of a target compound. Furthermore, this strategy requires efficient protein separation techniques to reduce the complexity of these complicated protein mixtures prior to mass spectral analysis. This study involved the combination of a recently developed chemical modification technique called SPROX with a common gel-based proteomic strategy.

7.1 Introduction

Many of the biological roles of proteins are modulated through protein-ligand interactions, but identifying the protein interaction partners of a given ligand is a challenging task. One approach to this challenge involves combining a technique for the detection of protein-ligand binding with a proteomics approach to analyze and identify large numbers of proteins. Although SUPREX is capable of detecting protein-ligand

interactions in a high throughput fashion, the transient nature of the protein deuteration prevents SUPREX from being easily coupled to the time-consuming protein fractionation techniques that are commonly used in proteomics.

Due to the incompatibility of SUPREX with conventional proteomics strategies, another technique called SPROX was used for the evaluation of the thermodynamic stability of a model protein. SPROX is a technique similar to SUPREX that uses oxidation rather than H/D exchange to report on protein stability (52). Since oxidation is a more persistent chemical modification than H/D exchange, the protein remains chemically modified throughout subsequent protein fractionation protocols. This feature makes SPROX more suitable than SUPREX for proteome-scale studies. Just like SUPREX, SPROX involves equilibrating a protein of interest in a series of buffers containing a range of denaturant concentrations. Hydrogen peroxide is added to each protein mixture to initiate the oxidation reaction, and after a specified time, the reaction is quenched with catalase, an enzyme that consumes the excess hydrogen peroxide. The extent of oxidation is then evaluated using mass spectrometry.

SPROX has been performed on a number of purified proteins (52), but its application to protein mixtures is just beginning to be explored. To assess the applicability of SPROX to mixture analysis, a gel-based fractionation strategy was incorporated into the SPROX protocol. Since the ultimate goal of this work involves performing thermodynamic analyses on a proteomic scale, T7 bacteriophage was chosen

as the model system. The T7 proteome contains only ~55 proteins, and each gene has been cloned into *E. coli*, allowing for the convenient analysis of individual T7 proteins. This work describes the development of a protocol that combines SPROX analysis with an SDS-PAGE-based protein fractionation approach. This approach was applied to the analysis of T7 bacteriophage gene 2.5 protein. The results of this proof-of-concept study indicate that SPROX can be readily coupled to SDS-PAGE-based protein fractionation without any change in the qualitative protein stability measurement. The gene 2.5 protein was also analyzed by SUPREX to assess the effects of oxidation on the stability of the protein. It was found that the relative stability of the unoxidized protein was higher than the stability of the oxidized protein. While this oxidation-induced destabilization may decrease the accuracy of the SPROX results, it is not expected to preclude the detection of protein-ligand interactions.

7.2 Experimental Section

7.2.1 Materials

The *E. coli* strain was kindly provided by F. William Studier (Brookhaven National Laboratory, NY). This strain contained bacteriophage T7 gene 2.5 in a pBR322 plasmid. To facilitate expression and purification of the protein product of gene 2.5, the gene was cloned into a pET-30a vector as described below. Hydrogen peroxide was purchased from ScienceLab.com, and catalase was purchased from Sigma. Vendors for the remaining materials are described in Chapter 2.

7.2.2 Protein Cloning and Expression

Freezer stocks of the original *E. coli* strain were maintained in MDG non-inducing media (25 mM Na_2HPO_4/25 mM KH_2PO_4/50 mM NH_4Cl/5 mM Na_2SO_4/2 mM $MgSO_4$/10 μM Fe + trace metals/28 mM glucose/19 mM aspartate) containing ampicillin and chloramphenicol (94). A QIAprep Spin Miniprep kit (Qiagen) was used to isolate plasmid DNA from a 10 mL overnight culture grown in MDG media. To clone gene 2.5 into a pET-30a plasmid for more facile expression and purification, NcoI and BamHI sites were added to gene 2.5 using PCR (Phusion; New England Biolabs). PCR products were digested with NcoI and BamHI and ligated into pET-30a digested with NcoI and BamHI. The clone was sequenced to check for errors and was transformed into BL21(DE3) *E. coli* cells (Novagen).

E. coli cells containing the affinity-tagged clone of T7 bacteriophage gene 2.5 were grown at 37 °C in LB medium to an OD_{600} of 0.7. They were then induced at 25 °C for 5 hours with 1 mM isopropryl-beta-D-thiogalactopyranoside (IPTG; Fisher Bioreagents) to produce proteins with N-terminal affinity purification tags. The affinity tag included a 6xHis region, a thrombin cleavage site, an S-tag, and an enterokinase cleavage site. Cells were harvested by centrifugation and stored at -80 °C overnight. To extract proteins, the cells were lysed with B-PER (Pierce Scientific) using the manufacturer's protocol. The insoluble fraction of the lysate was discarded, and Benzonase Nuclease (Novagen) was added to the soluble fraction according to the manufacturer's instructions. This nuclease

helped to reduce the viscosity of the lysate by degrading DNA and RNA. To remove the large affinity tag, the lysate was incubated with 0.1 μL/mL of bovine enterokinase (New England Biolabs) for 3 hours at room temperature.

7.2.3 Mass Spectrometry

MALDI mass spectra were acquired on an UltraFlex II TOF/TOF mass spectrometer (Bruker Daltonics, Inc., Billerica, MA). For the intact protein, spectra were collected in linear and positive ion mode using a Smartbeam Nd:YAG laser (355 nm). The following instrument parameters were used: 25 kV ion source 1 voltage, 23.25 kV source 2 voltage, 6 kV lens voltage, 130 ns pulsed ion extraction, and matrix gating to 3000 Da. Spectra were calibrated using soybean trypsin inhibitor as an internal standard. This internal calibrant was incorporated directly into the matrix solution, which consisted of saturated SA in 0.1% TFA/75% acetonitrile/24.9% water. This solution was used in conjunction with ZipTips as described below.

For peptide fragments, spectra were collected in reflectron and positive ion mode. The following instrument parameters were used: 25.0 kV ion source 1 voltage, 21.6 kV ion source 2 voltage, 9.6 kV lens voltage, 26.4 kV reflector voltage, 13.6 kV reflector 2 voltage, 30 ns pulsed ion extraction, and matrix deflection to 450 Da. Spectra were externally calibrated using the Peptide Calibration Standard II mixture from Bruker Daltonics. For MALDI analysis of the peptide fragments, the matrix solution consisted of saturated SA in 0.1% TFA/45% ACN/54.9% water.

Peptides were mapped by performing a theoretical digest of the T7 gene 2.5 protein using the ExPASy PeptideMass program (http://ca.expasy.org/tools/peptide-mass.html). Peptide molecular weights (determined by mass spectrometry) were compared to the molecular weights of the peptides in the theoretical digest to identify the sequence of the detected peptides. Peptide identifications were further confirmed by checking for methionine residues in identified peptides that were oxidized during SPROX.

7.2.4 SPROX Protocol

SPROX analyses on the cell lysate were performed using a set of buffers composed of 20 mM sodium phosphate (pH 7.4) and concentrations of GdmCl that varied between 0 and 8 M. The protocol used in this study was very similar to the previously reported SPROX protocol (52). Briefly, the lysate was equilibrated in the buffers for 30 min prior to SPROX analysis. To initiate the oxidation reaction, H_2O_2 was added to each lysate/buffer solution to reach a final concentration of 1.3 M. Oxidation was allowed to proceed for 2 min, and each reaction was quenched by adding an equal volume of an aqueous catalase solution (12.5 units/μL), which served to decompose any remaining hydrogen peroxide in each solution. The purpose for using an equal volume of catalase solution was to sufficiently dilute the denaturant in the buffers to prevent denaturation of the catalase.

Protein samples were subjected to gel-based SPROX, solution-based SPROX, and solution-based SUPREX. All solution-based analyses were carried out on the intact form of the protein. For SPROX and SUPREX analyses of the intact form, the protein was concentrated and desalted using C_4 ZipTips (Millipore) according to the manufacturer's instructions (see Section 2.5). Proteins were eluted directly onto the MALDI target using the MALDI matrix solution, which contained soybean trypsin inhibitor as an internal standard (see above). The gel-based analysis required trypsin digestion to extract the protein fragments from the gel. Analysis of the digested protein is described in the next section.

7.2.5 Gel-Based Fractionation

A gel-based proteomic strategy was used to fractionate the lysate that had been subjected to SPROX analysis. To concentrate the proteins in the lysate and remove excess GdmCl, a trichloroacetic acid (TCA; J.T. Baker) precipitation was used. Briefly, a solution of 100% TCA was added to each sample to reach a final concentration of 20%. The sample was incubated on ice for 30 min, then microcentrifuged for 1 min. The pellets were washed with sufficient PBS to remove any additional GdmCl and again microcentrifuged. The resulting pellets were resuspended in Laemmli buffer (Boston BioProducts) in preparation for SDS-PAGE, and sodium hydroxide was used to increase the pH of the resulting solution to compensate for residual TCA in the samples.

One dimensional SDS-PAGE was performed using precast 8-16% Tris-HCl polyacrylamide Ready Gels (BioRad). 30 μL of the sample was loaded in each lane. Electrophoresis was carried out using an initial voltage of 100 V in the stacking gel and then increasing the voltage to 170 V for the resolving gel. The gels were stained with Bio-Safe Coomassie stain (Bio-Rad) following the manufacturer's protocol and then destained overnight in deionized water.

Bands corresponding to protein 2.5 were cut from the gel and cleaved into small pieces. The protein bands were destained using 25 mM NH_4HCO_3 in 50% acetonitrile, dehydrated with acetonitrile, and rehydrated with a solution of trypsin in 25 mM NH_4HCO_3 (pH 8). The digestion was allowed to proceed for 1 h at 37 °C and was then quenched with TFA. Peptides were extracted from the gel using three cycles of addition of extraction buffer (60% acetonitrile/35% water/5% formic acid), shaking for 10 min, and collection of the resulting solution. The extracted peptide solution was evaporated to dryness using a SpeedVac, and peptides were reconstituted in 10 μL deionized water prior to MALDI analysis.

7.2.6 SUPREX Protocol

The gene 2.5 protein was analyzed using the SUPREX protocol described in Chapter 2 (40). The H/D exchange buffers contained concentrations of deuterated GdmCl ranging from 0 to 4 M. H/D exchange reactions were initiated by combing 1 μL of the cell lysate with 9 μL of exchange buffer. To bring the transition midpoint into an

experimentally accessible GdmCl range, an exchange time of 27 hours was used. The H/D exchange reaction was quenched with 1 μL of 10% TFA, and the protein was concentrated and desalted using C_4 ZipTips as described above.

7.2.7 Data Analysis

After subjecting the lysate to SPROX analysis, a weighted average mass was calculated for the ion signals of the unoxidized and oxidized species. The change in mass due to oxidation ($\Delta mass_{wt,av}$) was calculated by subtracting the mass of the unoxidized protein from the weighted average mass of the protein after SPROX. These $\Delta mass_{wt,av}$ values were plotted vs. [GdmCl] to generate SPROX curves. ΔMass values were extracted from the SUPREX data as discussed in Chapter 2.

To estimate the thermodynamic stability of the gene 2.5 protein, the measurement of an *m*-value was required. This value was calculated by fitting several SUPREX and SPROX curves to equation 7-1.

$$\Delta Mass = \Delta M_{\infty} + \left(\Delta M_0 - \Delta M_{\infty}\right) e^{\frac{-\langle k \rangle t}{(1+K_{fold})}} \qquad \text{Equation 7-1}$$

In equation 7-1, the variables are as described in Chapter 4 (see equation 4-1) except that for SUPREX, k is equal to k_{int}, and for SPROX, k is equal to k_{ox} (i.e., the rate of oxidation). In this study, $\langle k_{int} \rangle$ was estimated using the relationship $\langle k_{int} \rangle = 10^{pH-5}$ min^{-1} (36), and k_{ox} was estimated based on the concentration of H_2O_2 present during the oxidation reaction and the estimated second order rate constant for the oxidation of unprotected

methionine residues (for the reaction conditions in this study, k_{ox} was set to 0.016 sec^{-1}). In equation 7-1, K_{fold} is defined as described in equation 7-2.

$$K_{fold} = e^{\frac{-(\Delta G_f + m[GdmCl])}{RT}} \qquad \text{Equation 7-2}$$

The variables in equation 7-2 are as described in Chapter 4 (see equation 4-2). To calculate an average *m*-value, several SPROX and SUPREX curves were fit to equation 7-1, and the *m*-values obtained from each curve fitting were averaged. The average *m*-value (m_{av}) was multiplied by two to reflect the multimeric state of the protein, which is a dimer. Since equations 7-1 and 7-2 do not account for the multimeric state of the protein, each SPROX and SUPREX curve was fit to a four-parameter sigmoidal equation using a nonlinear regression routine in SigmaPlot (SYSTAT Software, Inc., San Jose, CA). This allowed for the extraction of a transition midpoint from each curve. These transition midpoints were used in conjunction with equation 7-3 to determine a ΔG_f value for each data set.

$$-RT\left[\ln\frac{\frac{\langle k\rangle t}{0.693}-1}{\frac{n^n}{2^{n-1}}[P]^{n-1}}\right] = mC^{1/2}_{SUPREX/SPROX} + \Delta G_f \qquad \text{Equation 7-3}$$

In equation 7-3, the variables are as described in Chapter 2 (see equation 2-1) except that for SUPREX, k is equal to k_{int}, and for SPROX, k is equal to k_{ox} (see above). The protein concentration of the stock solution (expressed in dimer equivalents) was estimated to be 100 μM, and [P] in equation 7-3 was calculated according to the dilution

factors used in each SPROX and SUPREX experiment. This procedure allowed for the calculation of a ΔG_f value for each SUPREX or SPROX curve, including the curves obtained from the gel-based experiments.

7.3 Results

SPROX is a chemical modification- and mass spectrometry-based technique in which a chemical denaturant is used to perturb the folding equilibrium of a protein in the presence of hydrogen peroxide. As the amount of chemical denaturant is increased, a larger population of the protein unfolds, and the extent of unfolding is evaluated by measuring the extent of oxidation. Although hydrogen peroxide is capable of oxidizing cysteine residues (95), under the conditions of SPROX, methionine residues are preferentially oxidized (52). Thus, performing SPROX at the peptide level, as is required in this gel-based protocol, requires the ability to extract and identify methionine-containing peptides from the gel.

As shown in Figure 26, gel-based SPROX requires the execution of the following steps: 1) subjecting a protein mixture to SPROX analysis, 2) separating the protein mixture on a polyacrylamide gel, 3) extracting the band of interest from the gel, 4) performing an in-gel digestion, 5) extracting the peptides from the gel, and 6) analyzing the peptides by MALDI-MS. As a first step, the in-gel digestion protocol was evaluated to determine whether methionine-containing peptides could be extracted from the gel and detected using MALDI-MS. Using a digestion time of 1 hour, two reproducible

methionine-containing fragments were detected that could be mapped back to the T7 gene 2.5 protein. The two fragments had molecular weights of 2790 and 3916 Da, and each fragment contained a single methionine residue.

In many proteomics protocols, proteins are digested completely (generally overnight) to allow for more facile peptide identification. In this study, a shorter digestion time was chosen to increase the chances of detecting methionine-containing peptides. However, for comparison to the typical proteomics approach, an overnight digestion was performed. This longer digestion time did not increase the coverage of methionine residues, but the two methionine residues in fragments 2790 and 3916 were still detected in fragments produced using an overnight digestion. Thus, a variety of trypsin digestion times would likely be compatible with gel-based SPROX analysis.

Once the digestion protocol had been optimized, a gel-based SPROX analysis was performed on the lysate. SPROX analysis of the two methionine-containing peptides yielded clear SPROX transitions (Figure 27A), suggesting that the methionine residue in each of these fragments is protected from oxidation in the native three-dimensional structure of the protein. Both peptides were mapped using the masses of the peptides, the sequence of the protein, and the expected trypsin cleavage sites (see Section 7.2.3). The presence of a single methionine residue in each peptide sequence provides further evidence that the identified peptide sequences are correct.

Since the transition midpoints of the SPROX and SUPREX curves cannot be compared directly, the midpoints were instead used to calculate a ΔG_f value from each curve. While equations 7-1 and 7-2 can be used to extract a ΔG_f value for each SPROX and SUPREX curve, ΔG_f values can be determined more accurately by establishing a more accurate *m*-value (43). Fitting several SPROX and SUPREX curves to equation 7-1 (in conjunction with equation 7-2) allowed for the extraction of an *m*-value for each curve. These *m*-values were averaged to obtain a single *m*-value (m_{av}) and multiplied by two to reflect the multimeric state of the protein (dimer). Since equations 7-1 and 7-2 do not accurately describe multimeric proteins, the transition midpoint was extracted from each SPROX and SUPREX curve, and these transition midpoints were used in conjunction with equation 7-3 to determine a ΔG_f value for each data set.

Transition midpoints and estimated ΔG_f values extracted from the SPROX and SUPREX data are shown in Table 13. The transition midpoints extracted for the two methionine-containing fragments using the gel-based SPROX protocol are in good agreement with one another, even though the two fragments contain different methionine residues. To more thoroughly evaluate the gel-based SPROX protocol, a solution-based SPROX analysis was performed on the intact protein using the same cell lysate. A representative SPROX curve from this analysis is shown in Figure 27B. As shown in Table 13, the average transition midpoint and ΔG_f value extracted from the

solution-based SPROX curves is in good agreement with the values extracted from the gel-based SPROX curves.

To examine the effects of oxidation on the stability of the gene 2.5 protein, SUPREX was performed on the protein in the context of the cell lysate. A representative SUPREX curve is shown in Figure 27C. Table 13 shows the average transition midpoint and ΔG_f value extracted from SUPREX curves collected using a 27-hour H/D exchange time. Comparison of the SUPREX data with the SPROX data reveals that the oxidation of gene 2.5 protein has reduced its stability by approximately 3 kcal/mol.

7.4 Discussion

The gel-based SPROX protocol differs from the solution-based SPROX protocol in two important ways: 1) the gel-based protocol requires a significant amount of sample cleanup prior to SDS-PAGE (i.e., TCA precipitation), and 2) the gel-based protocol requires a trypsin digestion step prior to mass spectral analysis. Although desalting columns are often used to remove salts from protein mixtures, in these experiments such columns did not remove the denaturant effectively enough to allow for successful electrophoresis runs. A TCA precipitation, however, more effectively removed the salt and allowed for improved electrophoresis. One other option for gel-based SPROX is to use urea as a denaturant instead of GdmCl. Since urea is not a salt, it does not interfere with electrophoresis and does not need to be removed from the samples. However,

GdmCl is a stronger denaturant than urea, and the high stability of the gene 2.5 protein would most likely have precluded SPROX analysis using urea as the denaturant.

One potential problem with the gel-based SPROX protocol is that protein oxidation commonly occurs within SDS-PAGE gels (96). In fact, performing SDS-PAGE analysis of a lysate that had not been subjected to SPROX resulted in oxidation of many of the gene 2.5 protein fragments, including fragments that do not contain methionine. As expected, when SDS-PAGE analysis was performed on the lysate that had been subjected to SPROX, in-gel oxidation was observed. However, no denaturant dependence to this oxidation was observed. Thus, in-gel oxidation could be distinguished from the denaturant-dependent SPROX-based oxidation and did not prevent SPROX analysis of the oxidized peptides.

The quantitative determination of ΔG_f values using any chemical denaturation-based technique requires that the protein folding reaction must be well-described by a two-state mechanism (24). As described in Chapter 3, however, this requirement can be relaxed for qualitative protein stability studies and for quantitative determinations of ligand binding affinities. The folding mechanism of T7 bacteriophage gene 2.5 protein has not been reported. However, this protein is not expected to be a two-state folder given the relatively low cooperativity of the folding reaction observed in the SPROX and SUPREX experiments.

Given the assumption that the gene 2.5 protein is not a two-state folder, the ΔG_f values calculated in this study cannot be taken to be a true representation of the thermodynamic stability of this protein, but the SPROX- and SUPREX-derived values can be used as a gauge of relative stability. The SPROX- and SUPREX-derived relative stabilities of the T7 gene 2.5 protein were examined, and it was found that the stability of the protein decreased upon oxidation. This change in stability, which has been observed in previous SPROX studies (52), affects the accuracy of the SPROX results, but does not necessarily preclude analysis of relative stabilities by SPROX. Such a qualitative assessment of relative protein stabilities still allows for the detection of ligand binding since protein-ligand interactions increase the thermodynamic stability of the protein. Since relative stabilities can still be evaluated using gel-based SPROX, this decrease in stability due to oxidation is not expected to prevent the use of gel-based SPROX to detect protein-ligand interactions as long as the oxidation does not interfere with ligand binding.

7.5 Conclusions

The use of oxidation as a chemical modification strategy is preferable over H/D exchange in situations that require lengthy purification or sample preparation protocols. The ease with which SPROX was coupled to SDS-PAGE suggests that this approach could also be used with other sample fractionation protocols. This work with the T7 gene 2.5 protein suggests that the oxidation reaction destabilizes the protein, but this

destabilization did not preclude the measurement of qualitative protein stabilities. Thus, as long as the oxidation reaction does not interfere with ligand binding, the gel-based SPROX approach should allow for the detection of protein-ligand binding interactions. This proof-of-concept study has shown that SPROX is compatible with mixture analysis and protein fractionation strategies and suggests that SPROX is a promising approach for the analysis of ligand binding on a proteomic scale.

Table 13: Summary of SPROX and SUPREX results for the protein product of T7 bacteriophage gene 2.5.

Protein Samples	Residue Number[a]	Peptide Sequence	Transition Midpoint (M, GdmCl)[b]	Estimated ΔG_f (kcal/mol)[c]
Fragment 2790	83-107	GKKPLKPYEGDMPFFDNGDG TTTFK	3.1 ± 0.1	-17.5 ± 0.4
Fragment 3916	170-206	LQLESVMLVELATFGGGEDD WADEVEENGYVASGSAK	2.9 ± 0.1	-16.7 ± 0.4
Intact SPROX	—	—	2.97 ± 0.06	-17.0 ± 0.2
Intact SUPREX	—	—	1.8 ± 0.1	-20.6 ± 0.5

[a]Peptides were identified by comparing their molecular weights with a theoretical digest (see text).

[b]Error is the standard deviation of the transition midpoints extracted from at least three curves.

[c]Error is the standard deviation of the free energies calculated from at least three curves.

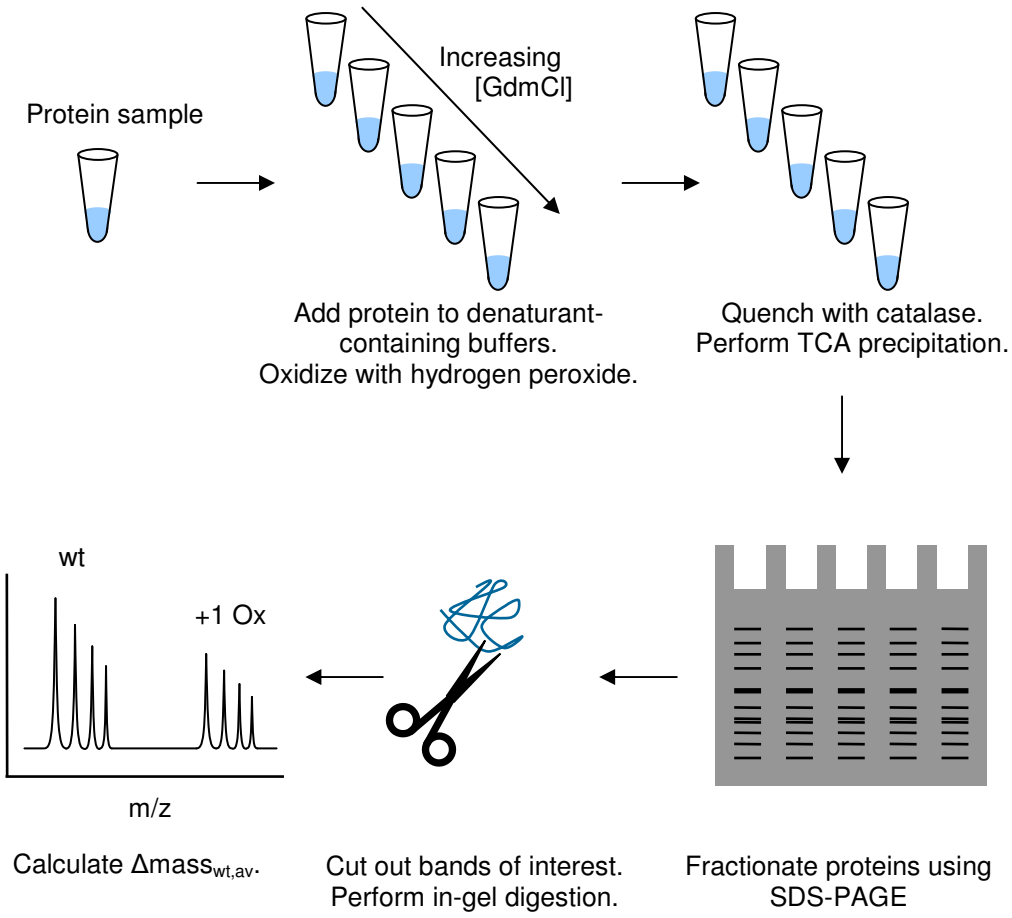


Figure 26: Summary of the gel-based SPROX protocol.

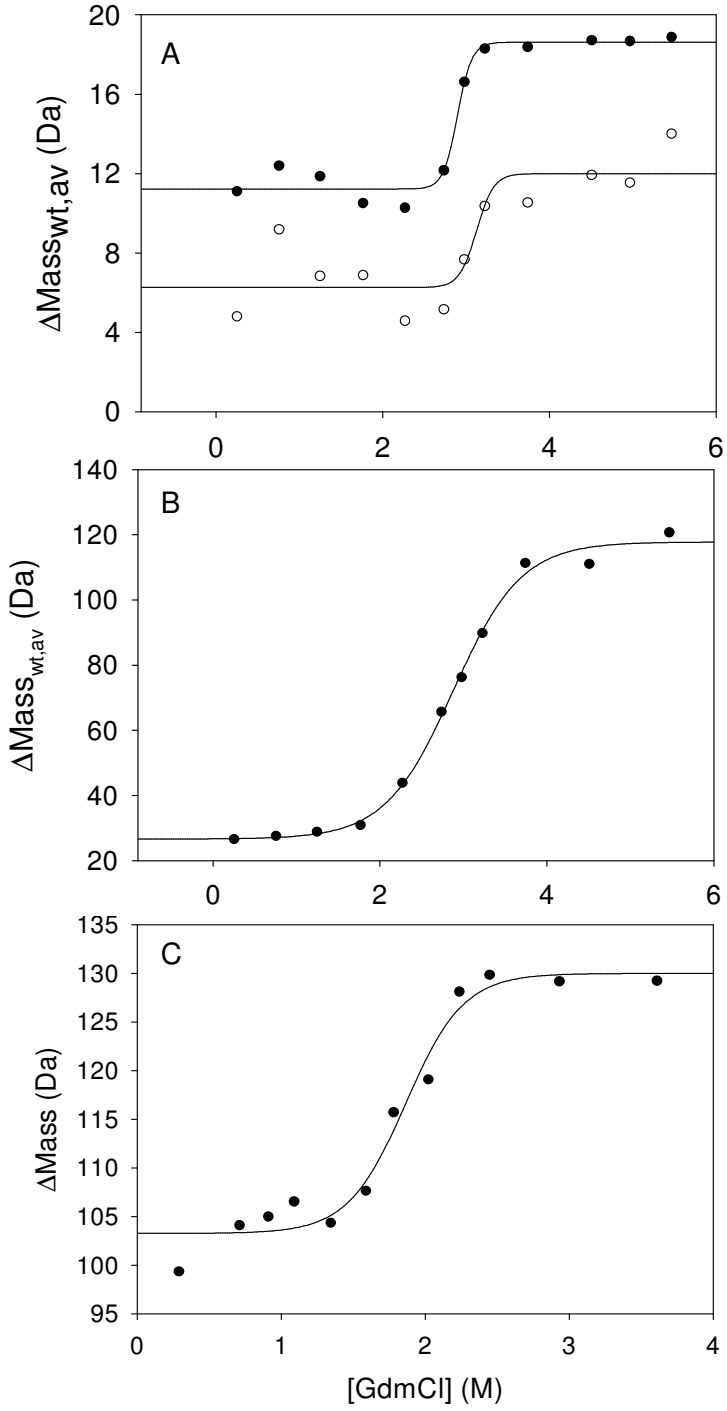


Figure 27: Representative SUPREX and SPROX curves. A) Gel-based SPROX curves of fragment 2790 (open circles) and 3916 (filled circles). B) Solution-based SPROX curve of the intact protein. C) Solution-based SUPREX curve of the intact protein using an exchange time of 27 hours.

Chapter 8. Benefits and Limitations of SUPREX for the High Throughput Detection of Protein-Ligand Binding.

This dissertation work has focused on the development of SUPREX into an assay for the high throughput selection of novel ligands that bind to target proteins of interest. Chapter 7 also discussed efforts toward the use of a similar technique (SPROX) to identify protein binding partners of a target ligand of interest. Both of these techniques have strengths and weaknesses. This chapter serves to summarize the work in this dissertation and place the strengths and weaknesses of the methodologies that have been developed in a broader context.

8.1 Summary of Dissertation Research

8.1.1 SUPREX Analysis of Non-Two-State Proteins

At the beginning of this dissertation project, SUPREX was a well established technique for the quantitative analysis of two-state folding proteins (36, 40, 41). The experimental protocols and data analysis methods were in place to evaluate the thermodynamic parameters of such two-state folding protein folding reactions and to evaluate the thermodynamic parameters associated with the ligand binding properties of such proteins. However, many proteins are not two-state folders, and the applicability of SUPREX for the analysis of this category of proteins was untested. The work in Chapter 3 describes the SUPREX analysis of two model non-two-state proteins, Bcl-x_L and AGT.

Apo-Bcl-x_L was found to be poorly behaved in the conventional SUPREX experiment; when this protein was analyzed at various exchange times, the transition midpoint did not move in accordance with equation 2-1. However, when Bcl-x_L was analyzed in the presence of ligands, the midpoint of the SUPREX curve did shift to higher denaturant concentrations as expected, indicating that SUPREX could be used to detect ligand binding. Even more interestingly, the midpoint of the SUPREX curves for the Bcl-x_L complexes moved with exchange time as predicted by equation 2-1. Although the K_d values for the complexes could not be determined due to the lack of thermodynamic information for the apoprotein, the relative affinities of the two ligands could be quantitatively measured.

In Chapter 3, SUPREX was also used to compare the relative stabilities of several AGT variants. Similar to Bcl-x_L, apo-AGT was not well-behaved in SUPREX, but SUPREX could still be used to compare the relative stabilities of the apoproteins. SUPREX was also found to be capable of detecting the binding of two AGT ligands (PLP and AOA). These experiments with Bcl-x_L and AGT demonstrated that SUPREX could be used to evaluate ligand binding for non-two-state proteins. They also helped to expand the scope of SUPREX to include a wider variety of proteins. This work was an important step in the development of a SUPREX-based assay for protein-ligand binding because it provided some of the first evidence that the assay could be used for the analysis of non-two-state folding proteins.

8.1.2 Application of Single-Point SUPREX for the Screening of a Compound Library

The work in Chapter 4 described the first use of single-point SUPREX for the screening of a compound library. The primary motivation for this screening was to find new ligands for CypA. However, from an analytical method development perspective, this screening enabled the first evaluation of the throughput and efficiency of single-point SUPREX. This work also revealed an important limitation of single-point SUPREX: the efficiency of the screening was highly dependent on the standard deviations of the mass measurements. Such standard deviations generally increase as protein size increases. Thus, the single-point SUPREX assay was not expected to be generally applicable to a wide variety of proteins. This finding motivated the development of a new single-point SUPREX protocol, which is described in Chapter 6.

8.1.3 Quantitative Analysis of AGTmi using the SUPREX-Protease Digestion Protocol

As mentioned above (see Section 8.1.2), the original single-point SUPREX protocol was found to have an important limitation related to the standard deviations of the mass measurements. The efficiency of single-point SUPREX is related to both the standard deviations of the mass measurements and the difference in Δmass value between the bound and unbound forms of the protein. Since the difference in Δmass between these forms of the protein is largely controlled by the number of protected amide protons in the folded protein structure, this factor could not be changed by

adjusting the protocol. However, the standard deviations of the mass measurements could be reduced by performing a protease digestion prior to mass spectral analysis.

The SUPREX-protease digestion protocol was developed previously to facilitate the analysis of large and/or multidomain proteins (51). However, in this dissertation work, it was hypothesized that this same strategy could be used to improve the efficiency and expand the generality of the single-point SUPREX assay. In preparation for the development of a new single-point SUPREX protocol, the SUPREX-protease digestion strategy was applied to AGTmi. These experiments were critical because screening AGTmi with the original single-point SUPREX protocol would not have been possible due to the large standard deviations associated with the AGTmi mass measurements. Although apo-AGTmi was not amenable to full SUPREX analysis in its intact form (see Section 8.1.1), the protease digestion facilitated the analysis of this apoprotein. Furthermore, this approach allowed for the successful analysis of AGTmi-PLP and AGTmi-PLP-AOA and was used to make the first measurement of the binding affinity of AOA for AGTmi-PLP (see Chapter 5).

8.1.4 Incorporation of a Protease Digestion Protocol into the Single-Point SUPREX Assay

The experiments described in the previous section (8.1.3) were the first step toward incorporating the protease digestion strategy into the single-point SUPREX protocol. The data from these experiments suggested that the combination of these two protocols could significantly increase the efficiency of single-point SUPREX. The new

protocol was evaluated by performing control screenings on CypA and AGTmi. In both cases, a significant improvement in efficiency was observed. For CypA, the maximum Z'-factor associated with screening the intact protein was 0, which is acceptable for screening, but not ideal. However, Z'-factors of up to 0.4 were achieved for CypA when the protease digestion protocol was used. Even more significantly, AGTmi would not have been amenable to single-point SUPREX using the original protocol to analyze the intact protein, but the protease digestion strategy allowed for the successful screening of this protein with Z'-factors of up to 0.5 (see Chapter 6). The experiments in Chapter 6 demonstrate that the new single-point SUPREX protease digestion protocol offers a significant improvement in efficiency in comparison to the original protocol and suggest that the new protocol will also increase the generality of the assay.

8.1.5 Gel-Based SPROX for the Analysis of Protein Stabilities in Complex Mixtures

A major part of this dissertation was focused on the development of SUPREX-based assays for the detection of protein-ligand binding. However, a part of this dissertation also contributed to the early development of a similar chemical modification-based strategy to identify protein binding partners for a ligand of interest. Reaching this goal required the analysis of large numbers of proteins, so it was clear that a proteomics approach would be a logical strategy for meeting this challenge. One of the limitations of SUPREX is that this technique is difficult to couple with protein fractionation strategies such as SDS-PAGE and LC that are commonly used in

proteomics. The reason for this difficulty is that SUPREX relies on a reversible chemical modification (H/D exchange) that is not stable throughout protein fractionation.

One strategy for overcoming this limitation of SUPREX is to use a more permanent chemical modification instead of using H/D exchange. It was hypothesized that an oxidation-based strategy termed SPROX (52) could be successfully coupled to protein fractionation techniques. To test this hypothesis, a gel-based proteomics approach was used to fractionate a cell lysate after SPROX analysis (Chapter 7). The results of this experiment demonstrated that SPROX could be coupled to SDS-PAGE protein separations. Although the oxidation was found to perturb the model protein's stability, this perturbation is not expected to prevent the detection of ligand binding by SPROX.

8.2 Benefits and Limitations of SUPREX-Based Assays for Protein-Ligand Binding

8.2.1 General Considerations

One of the most significant advantages of the single-point SUPREX assay is that ligand binding interactions are identified by detecting the increase in protein stability that occurs upon ligand binding. Because ligand-induced stabilization is a general phenomenon, the single-point SUPREX assay is expected to be relatively general. Thus, unlike enzymatic assays for protein-ligand binding, a new assay does not need to be developed for each new protein that is screened using single-point SUPREX. However, this assay does require some knowledge of the biophysical properties of the protein (i.e.,

ΔG_f and *m*-values). Thus, if this information has not been previously published for the target protein, conventional SUPREX analysis may be required to allow for the selection of appropriate screening conditions for single-point SUPREX. While the flexibility of single-point SUPREX for the screening of non-enzymatic proteins is generally an advantage, one drawback to this aspect of the assay is that single-point SUPREX only selects for ligand binding; the assay does not provide any information about whether the ligand modulates protein function.

The basic requirements for the single-point SUPREX assay are that the target protein must be amenable to analysis by SUPREX and that the thermodynamic properties of the protein folding reaction must be known (either using SUPREX-based measurements or values obtained from the literature). Since SUPREX has been found to be a fairly general technique for the analysis of protein folding (36, 39, 41, 43, 44, 46, 49, 51, 59), single-point SUPREX is expected to be applicable for the screening of a wide variety of proteins. Single-point SUPREX can be used to screen a wide variety of ligands, including small molecules, nucleic acids, and peptides. Screening a library of proteins is also theoretically possible, although it may be more challenging because the *m*-value for the protein folding reaction may change in the presence of protein-protein interactions (44).

A final advantage of single-point SUPREX is that it can be used to screen unpurified proteins due to the capacity of MALDI for the analysis of protein mixtures.

This is a particular advantage for proteins that are difficult to purify using standard techniques. The primary consideration for the analysis of unpurified proteins is that the protein ion signal must be detectable (with an acceptable signal-to-noise ratio) and well-resolved in the MALDI readout. The use of the single-point SUPREX-protease digestion protocol would be more difficult to implement for unpurified proteins since the MALDI spectrum will contain peptides from all of the proteins in the lysate. However, if the target protein is present in high enough quantity and acceptable peptides can be mapped to the target protein, the protease digestion protocol could theoretically be used for unpurified proteins as well.

8.2.3 Theoretical and Practical Constraints

Although the single-point SUPREX protocol is expected to be relatively general, especially when the protease digestion protocol is employed, several theoretical and practical constraints must be considered before applying this technique. One important consideration is related to the dynamic range for the detection of protein-ligand binding. In theory there is no upper or lower limit to the range of K_d values that can be selected using the single-point SUPREX protocol. In practice, however, the ability to detect weak binding ligands depends on the ability to achieve sufficiently high concentrations of the library members such that the ligand binding equilibrium is sufficiently shifted toward the bound form of the protein.

The ability to detect tight-binding ligands is determined by the appropriate selection of screening conditions (i.e., longer exchange times and higher denaturant concentrations). These concepts are discussed in more detail in Chapter 4. To select appropriate screening conditions for a target protein, the ΔG_f and *m*-value for the protein can be used along with the K_d value corresponding to the desired screening stringency to construct a theoretical SUPREX curve that represents the desired binding affinity for the screening hits. This theoretical curve can be used with the SUPREX curve for the apoprotein to determine a set of screening conditions that maximizes the difference in Δmass between the bound and unbound forms of the protein.

Another significant consideration for the single-point SUPREX assay is related to back exchange (i.e., the loss of deuteration during sample preparation and mass spectral analysis). In the screens described in Chapters 4 and 6, back exchange was kept to a minimum by keeping the sample sets small and analyzing the sample sets immediately after the sample preparation step. This strategy has proven effective for the screening of small libraries, but to maximize throughput, the ability to analyze larger sample sets would be advantageous.

Two primary strategies have been used with single-point SUPREX to minimize effects due to back exchange. The first strategy involves chilling the MALDI target after sample spotting and performing a mathematical back exchange correction to account for variable back exchange for the various samples (47). The second strategy, which was

developed as a part of this dissertation project, involves using a moving average to calculate the cutoff values for the screening assay as described in Chapter 4 (50). The moving average method has the advantage of accounting for variations that occur over the course of an extended screening campaign. Furthermore, when the moving average was used, the mathematical back exchange correction was not found to improve the efficiency of the screening. The incorporation of appropriate automation is also expected to minimize variation from sample to sample, thus reducing the effects of back exchange.

One approach that has yet to be explored is the use of SPROX (see Chapter 7) as a screening tool. Since the chemical modification used in SPROX (oxidation) is more chemically stable than the modification used in SUPREX (H/D exchange), SPROX could theoretically be used to screen libraries for ligand binding without the need to consider back exchange. Although this is an attractive approach, it will require the assumptions that oxidation does not interfere with ligand binding and that it does not modify the library ligands. Thus, this approach may be less general than single-point SUPREX.

8.2.4 Throughput Considerations

The throughput of the original single-point SUPREX protocol was established for the first time in this dissertation work and was found to be 3 min/ligand. This measure was based on the time required to screen an 880-member compound library (see Chapter 4) (50). During the course of this screening, the primary factor limiting the throughput

was found to be the speed of the MALDI mass spectrometer. Using a higher throughput mass spectrometer along with the single-point SUPREX protease digestion protocol was found to increase the throughput to 1.8 min/ligand even though the protocol was more complex and only one researcher was performing the screening instead of two (see Chapter 6).

While the use of a higher throughput mass spectrometer was successful in improving the throughput of the technique, a multiplexing strategy could also be used to make more significant increases in throughput. For example, if the compound library was manipulated such that each sample well contained 10 library compounds instead of one, the throughput of the technique could be increased by approximately tenfold. The drawback of this approach is that for each hit in the screen, each of the ten compounds associated with the hit must be analyzed individually. However, the high throughput nature of the screening makes this re-screening fairly straightforward. This approach is currently being explored in the Fitzgerald Laboratory and seems to be a promising strategy for improving throughput.

One additional possibility for improving throughput is to screen multiple proteins simultaneously. This approach takes advantage of the capacity of MALDI-MS for mixture analysis. Using this strategy, the target proteins of interest would be combined and screened simultaneously against a single compound library. This approach could result in a significant increase in throughput (e.g., the throughput

would be expected to double if two proteins were screened simultaneously). However, it should be noted that this method could result in an increase in false positives or false negatives since the presence of multiple proteins could influence ligand binding for one or more of the target proteins.

Several factors would need to be considered before adopting this multiple protein approach. For example, only one set of screening conditions can be used in a given single-point SUPREX assay. Thus, the target proteins must be amenable to screening using a single denaturant concentration and exchange time. The simplest scenario would involve screening multiple proteins that have similar ΔG_f and *m*-values, as would likely be the case if the target proteins are variants of a single protein. In this case, a single set of screening conditions could easily be chosen to suit all of the target proteins.

Another important consideration involves the choice of whether or not to use the protease digestion strategy in combination with the multiple protein approach. If no protease digestion is used, the signals of the target proteins would need to be well-resolved to facilitate monitoring the SUPREX behavior of each protein. If the protease digestion strategy is used, a set of digestion conditions that is suitable for all of the target proteins must be chosen. The digestion conditions should be chosen so that at least one fragment (but preferably multiple fragments) can be detected for each target protein. The fragments of interest must exhibit a good signal-to-noise ratio and must be well-

resolved from one another. They also must be suitable for SUPREX (i.e., a SUPREX transition midpoint can be detected at the selected exchange time). This optimization of the digestion conditions will likely involve some peptide mapping so that the each fragment can be matched with its parent protein. Any hits from this type of screen should be validated using the purified target protein if possible.

8.3 Overall Conclusions

This dissertation project was largely focused on the use of a SUPREX-based assay for the detection of protein-ligand binding. Before implementing a single-point SUPREX assay, a number of factors should be considered, including the following: 1) whether the original or protease digestion protocol should be used; 2) whether the protein will be purified or unpurified; 3) what set of conditions would be best for selecting ligands with the desired binding affinity; 4) which strategy will be used to minimize the effects of back exchange; 5) whether a multiplexing strategy will be employed; and 6) if a multiplexing strategy is desirable, which approach is most suitable for the target protein(s). Although some aspects of the single-point SUPREX protocol (such as the multiplexing strategies) remain to be explored, this dissertation project has resulted in a significant advancement in the high throughput capabilities of SUPREX.

References

1. Voet, D.; Voet, J. G.; Pratt, C. W., *Fundamentals of Biochemistry*. John Wiley & Sons, Inc.: New York, 2002.

2. Hopkins, A. L.; Groom, C. R. The Druggable Genome. *Nat. Rev. Drug Discovery* **2002**, *1*, 727-730.

3. Berger, F.; Gambhir, S. S. Recent Advances in Imaging Endogenous or Transferred Gene Expression Utilizing Radionuclide Technologies in Living Subjects: Applications to Breast Cancer. *Breast Cancer Res.* **2001**, *3*, 28-U1.

4. Blasberg, R. G.; Tjuvajev, A. G. Molecular-Genetic Imaging: Current and Future Perspectives. *J. Clin. Invest.* **2003**, *111*, 1620-1629.

5. Campa, M. J.; Wang, M. Z.; Howard, B.; Fitzgerald, M. C.; Patz, E. F. Protein Expression Profiling Identifies Macrophage Migration Inhibitory Factor and Cyclophilin A as Potential Molecular Targets in Non-Small Cell Lung Cancer. *Cancer Res.* **2003**, *63*, 1652-1656.

6. Howard, B. A.; Furumai, R.; Campa, M. J.; Rabbani, Z. N.; Vujaskovik, Z.; Wang, X. F.; Patz, E. F. Stable RNA Interference-Mediated Suppression of Cyclophilin A Diminishes Non-Small-Cell Lung Tumor Growth in Vivo. *Cancer Res.* **2005**, *65*, 8853-8860.

7. Sundberg, S. A. High-Throughput and Ultra-High-Throughput Screening: Solution- and Cell-Based Approaches. *Curr. Opin. Biotechnol.* **2000**, *11*, 47-53.

8. Hertzberg, R. P.; Pope, A. J. High-Throughput Screening: New Technology for the 21st Century. *Curr. Opin. Chem. Biol.* **2000**, *4*, 445-451.

9. Papac, D. I.; Shahrokh, Z. Mass Spectrometry Innovations in Drug Discovery and Development. *Pharm. Res.* **2001**, *18*, 131-145.

10. Shin, Y. G.; van Breemen, R. B. Analysis and Screening of Combinatorial Libraries Using Mass Spectrometry. *Biopharm. Drug Dispos.* **2001**, *22*, 353-372.

11. Sussmuth, R. D.; Jung, G. Impact of Mass Spectrometry on Combinatorial Chemistry. *J. Chromatogr. B* **1999**, *725*, 49-65.

12. Annis, D. A.; Nickbarg, E.; Yang, X.; Ziebell, M. R.; Whitehurst, C. E. Affinity Selection-Mass Spectrometry Screening Techniques for Small Molecule Drug Discovery. *Curr. Opin. Chem. Biol.* **2007**, *11*, 518-526.

13. Deng, G. J.; Sanyal, G. Applications of Mass Spectrometry in Early Stages of Target Based Drug Discovery. *J. Pharm. Biomed. Anal.* **2006**, *40*, 528-538.

14. Greis, K. D.; Zhou, S. T.; Burt, T. M.; Carr, A. N.; Dolan, E.; Easwaran, V.; Evdokimov, A.; Kawamoto, R.; Roesgen, J.; Davis, G. F. MALDI-TOF MS as a Label-Free Approach to Rapid Inhibitor Screening. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 815-822.

15. Shen, Z. X.; Go, E. P.; Gamez, A.; Apon, J. V.; Fokin, V.; Greig, M.; Ventura, M.; Crowell, J. E.; Blixt, O.; Paulson, J. C.; Stevens, R. C.; Finn, M. G.; Siuzdak, G. A Mass Spectrometry Plate Reader: Monitoring Enzyme Activity and Inhibition with a Desorption/Ionization on Silicon (DIOS) Platform. *Chembiochem* **2004**, *5*, 921-927.

16. Raji, M. A.; Frycak, P.; Beall, M.; Sakrout, M.; Ahn, J. M.; Bao, Y.; Armstrong, D. W.; Schug, K. A. Development of an ESI-MS Screening Method for Evaluating Binding Affinity between Integrin Fragments and RGD-Based Peptides. *Int. J. Mass Spectrom.* **2007**, *262*, 232-240.

17. Chan, N. W. C.; Lewis, D. F.; Hewko, S.; Hindsgaul, O.; Schriemer, D. C. Frontal Affinity Chromatography for the Screening of Mixtures. *Comb. Chem. High Throughput Screen.* **2002**, *5*, 395-406.

18. Ng, E. S. M.; Yang, F.; Kameyama, A.; Palcic, M. M.; Hindsgaul, O.; Schriemer, D. C. High-Throughput Screening for Enzyme Inhibitors Using Frontal Affinity Chromatography with Liquid Chromatography and Mass Spectrometry. *Anal. Chem.* **2005**, *77*, 6125-6133.

19. Muckenschnabel, I.; Falchetto, R.; Mayr, L. M.; Filipuzzi, I. Speedscreen: Label-Free Liquid Chromatography-Mass Spectrometry-Based High-Throughput Screening for the Discovery of Orphan Protein Ligands. *Anal. Biochem.* **2004**, *324*, 241-249.

20. Johnson, B. M.; Nikolic, D.; van Breemen, R. B. Applications of Pulsed Ultrafiltration-Mass Spectrometry. *Mass Spectrom. Rev.* **2002**, *21*, 76-86.

21. van Breemen, R. B.; Huang, C. R.; Nikolic, D.; Woodbury, C. P.; Zhao, Y. Z.; Venton, D. L. Pulsed Ultrafiltration Mass Spectrometry: A New Method for Screening Combinatorial Libraries. *Anal. Chem.* **1997**, *69*, 2159-2164.

22. Dill, K. A. Dominant Forces in Protein Folding. *Biochemistry* **1990**, *29*, 7133-7155.

23. Pace, C. N.; Vanderburg, K. E. Determining Globular Protein Stability - Guanidine-Hydrochloride Denaturation of Myoglobin. *Biochemistry* **1979**, *18*, 288-292.

24. Pace, C. N. Determination and Analysis of Urea and Guanidine Hydrochloride Denaturation Curves. *Methods Enzymol.* **1986**, *131*, 266-280.

25. Englander, S. W.; Sosnick, T. R.; Englander, J. J.; Mayne, L. Mechanisms and Uses of Hydrogen Exchange. *Curr. Opin. Struct. Biol.* **1996**, *6*, 18-23.

26. Sivaraman, T.; Arrington, C. B.; Robertson, A. D. Kinetics of Unfolding and Folding from Amide Hydrogen Exchange in Native Ubiquitin. *Nat. Struct. Biol.* **2001**, *8*, 331-333.

27. Bai, Y. W.; Englander, S. W. Future Directions in Folding: The Multi-State Nature of Protein Structure. *Proteins: Struct., Funct., Genet.* **1996**, *24*, 145-151.

28. Englander, J. J.; Del Mar, C.; Li, W.; Englander, S. W.; Kim, J. S.; Stranz, D. D.; Hamuro, Y.; Woods, V. L. Protein Structure Change Studied by Hydrogen-Deuterium Exchange, Functional Labeling, and Mass Spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 7057-7062.

29. Englander, S. W. Hydrogen Exchange and Mass Spectrometry: A Historical Perspective. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1481-1489.

30. Lam, T. T.; Lanman, J. K.; Emmett, M. R.; Hendrickson, C. L.; Marshall, A. G.; Prevelige, P. E. Mapping of Protein : Protein Contact Surfaces by Hydrogen/Deuterium Exchange, Followed by on-Line High-Performance Liquid Chromatography-Electrospray Ionization Fourier-Transform Ion-Cyclotron-Resonance Mass Analysis. *J. Chromatogr. A* **2002**, *982*, 85-95.

31. Mandell, J. G.; Falick, A. M.; Komives, E. A. Identification of Protein-Protein Interfaces by Decreased Amide Proton Solvent Accessibility. *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 14705-14710.

32. Mandell, J. G.; Falick, A. M.; Komives, E. A. Measurement of Amide Hydrogen Exchange by MALDI-TOF Mass Spectrometry. *Anal. Chem.* **1998**, *70*, 3987-3995.

33. Raschke, T. M.; Marqusee, S. Hydrogen Exchange Studies of Protein Structure. *Curr. Opin. Biotechnol.* **1998**, *9*, 80-86.

34. Wales, T. E.; Engen, J. R. Hydrogen Exchange Mass Spectrometry for the Analysis of Protein Dynamics. *Mass Spectrom. Rev.* **2006**, *25*, 158-170.

35. Mayne, L.; Paterson, Y.; Cerasoli, D.; Englander, S. W. Effect of Antibody-Binding on Protein Motions Studied by Hydrogen-Exchange Labeling and 2-Dimensional NMR. *Biochemistry* **1992**, *31*, 10678-10685.

36. Ghaemmaghami, S.; Fitzgerald, M. C.; Oas, T. G. A Quantitative, High-Throughput Screen for Protein Stability. *Proc. Natl. Acad. Sci. U. S. A.* **2000**, *97*, 8296-8301.

37. Ghaemmaghami, S.; Oas, T. G. Quantitative Protein Stability Measurement in Vivo. *Nat. Struct. Biol.* **2001**, *8*, 879-882.

38. Ma, L. Y.; Fitzgerald, M. C. A New H/D Exchange- and Mass Spectrometry-Based Method for Thermodynamic Analysis of Protein-DNA Interactions. *Chem. Biol.* **2003**, *10*, 1205-1213.

39. Powell, K. D.; Fitzgerald, M. C. Measurements of Protein Stability by H/D Exchange and Matrix-Assisted Laser Desorption Ionization Mass Spectrometry Using Picomoles of Material. *Anal. Chem.* **2001**, *73*, 3300-3304.

40. Powell, K. D.; Fitzgerald, M. C. Accuracy and Precision of a New H/D Exchange- and Mass Spectrometry-Based Technique for Measuring the Thermodynamic Properties of Protein-Peptide Complexes. *Biochemistry* **2003**, *42*, 4962-4970.

41. Powell, K. D.; Ghaemmaghami, S.; Wang, M. Z.; Ma, L.; Oas, T. G.; Fitzgerald, M. C. A General Mass Spectrometry-Based Assay for the Quantitation of Protein-Ligand Binding Interactions in Solution. *J. Am. Chem. Soc.* **2002**, *124*, 10256-10257.

42. Powell, K. D.; Wang, M. Z.; Silinski, P.; Ma, L.; Wales, T. E.; Dai, S. Y.; Warner, A. H.; Yang, X.; Fitzgerald, M. C. The Accuracy and Precision of a New H/D Exchange- and Mass Spectrometry-Based Technique for Measuring the Thermodynamic Stability of Proteins. *Anal. Chim. Acta* **2003**, *496*, 225-232.

43. Tang, L.; Hopper, E. D.; Tong, Y.; Sadowsky, J. D.; Peterson, K. J.; Gellman, S. H.; Fitzgerald, M. C. H/D Exchange- and Mass Spectrometry-Based Strategy for the Thermodynamic Analysis of Protein-Ligand Binding. *Anal. Chem.* **2007**, *79*, 5869-5877.

44. Tong, Y.; Wuebbens, M. M.; Rajagopalan, K. V.; Fitzgerald, M. C. Thermodynamic Analysis of Subunit Interactions in Escherichia Coli Molybdopterin Synthase. *Biochemistry* **2005**, *44*, 2595-601.

45. Wang, M. Z.; Shetty, J. T.; Howard, B. A.; Campa, M. J.; Patz, E. F.; Fitzgerald, M. C. Thermodynamic Analysis of Cyclosporin A Binding to Cyclophilin A in a Lung Tumor Tissue Lysate. *Anal. Chem.* **2004**, *76*, 4343-4348.

46. Powell, K. D.; Wales, T. E.; Fitzgerald, M. C. Thermodynamic Stability Measurements on Multimeric Proteins Using a New H/D Exchange- and Matrix-Assisted Laser Desorption/Ionization (MALDI) Mass Spectrometry-Based Method. *Protein Sci.* **2002**, *11*, 841-851.

47. Powell, K. D.; Fitzgerald, M. C. High-Throughput Screening Assay for the Tunable Selection of Protein Ligands. *J. Comb. Chem.* **2004**, *6*, 262-269.

48. Huang, Z. W. Bcl-2 Family Proteins as Targets for Anticancer Drug Design. *Oncogene* **2000**, *19*, 6627-6631.

49. Hopper, E. D.; Pittman, A. M. C.; Fitzgerald, M. C.; Tucker, C. L. In Vivo and in Vitro Examination of Stability of Primary Hyperoxaluria-Associated Human Alanine:Glyoxylate Aminotransferase. *J. Biol. Chem.* **2008**, *283*, 30493–30502.

50. Hopper, E. D.; Roulhac, P. L.; Campa, M. J.; Patz, E. F.; Fitzgerald, M. C. Throughput and Efficiency of a Mass Spectrometry-Based Screening Assay for Protein-Ligand Binding Detection. *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1303-1311.

51. Tang, L. J.; Roulhac, P. L.; Fitzgerald, M. C. H/D Exchange and Mass Spectrometry-Based Method for Biophysical Analysis of Multidomain Proteins at the Domain Level. *Anal. Chem.* **2007**, *79*, 8728-8739.

52. West, G. M.; Tang, L.; Fitzgerald, M. C. Thermodynamic Analysis of Protein Stability and Ligand Binding Using a Chemical Modification- and Mass-Spectrometry Based Strategy. *Anal. Chem.* **2008**, *80*, 4175-4185.

53. Kim, Y. T.; Richardson, C. C. Bacteriophage T7-Gene 2.5-Protein - an Essential Protein for DNA-Replication. *Proc. Natl. Acad. Sci. U. S. A.* **1993**, *90*, 10173-10177.

54. Glasoe, P. K.; Long, F. A. Use of Glass Electrodes to Measure Acidities in Deuterium Oxide. *J. Phys. Chem.* **1960**, *64*, 188-190.

55. Nozaki, Y. The Preparation of Guanidine Hydrochloride. *Methods Enzymol.* **1972**, *26*, 43-50.

56. Zhang, J. H.; Chung, T. D. Y.; Oldenburg, K. R. A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays. *J. Biomol. Screen.* **1999**, *4*, 67-73.

57. Dai, S. Y.; Gardner, M. W.; Fitzgerald, M. C. Protocol for the Thermodynamic Analysis of Some Proteins Using an H/D Exchange- and Mass Spectrometry Based Technique. *Anal. Chem.* **2005**, *77*, 693-697.

58. Yang, X. Y.; Fitzgerald, M. C. Total Chemical Synthesis of the B1 Domain of Protein L from Peptostreptococcus Magnus. *Bioorg. Chem.* **2006**, *34*, 131-141.

59. Roulhac, P. L.; Powell, K. D.; Dhungana, S.; Weaver, K. D.; Mietzner, T. A.; Crumbliss, A. L.; Fitzgerald, M. C. SUPREX (Stability of Unpurified Proteins from Rates of H/D Exchange) Analysis of the Thermodynamics of Synergistic Anion Binding by Ferric-Binding Protein (FbpA), a Bacterial Transferrin. *Biochemistry* **2004**, *43*, 15767-15774.

60. Leech, S. H.; Olie, R. A.; Gautschi, O.; Simoes-Wust, A. P.; Tschopp, S.; Haner, R.; Hall, J.; Stahel, R. A.; Zangemeister-Wittke, U. Induction of Apoptosis in Lung-Cancer Cells Following Bcl-Xl Anti-Sense Treatment. *Int. J. Cancer* **2000**, *86*, 570-576.

61. Simoes-Wust, A. P.; Olie, R. A.; Gautschi, O.; Leech, S. H.; Haner, R.; Hall, J.; Fabbro, D.; Stahel, R. A.; Zangemeister-Wittke, U. Bcl-Xl Antisense Treatment Induces Apoptosis in Breast Carcinoma Cells. *Int. J. Cancer* **2000**, *87*, 582-590.

62. Danpure, C. J. Molecular Etiology of Primary Hyperoxaluria Type 1: New Directions for Treatment. *Am. J. Nephrol.* **2005**, *25*, 303-310.

63. Danpure, C. J. Primary Hyperoxaluria Type 1: AGT Mistargeting Highlights the Fundamental Differences between the Peroxisomal and Mitochondrial Protein Import Pathways. *Biochim. Biophys. Acta Mol. Cell Res.* **2006**, *1763*, 1776-1784.

64. Manion, M. K.; O'Neill, J. W.; Giedt, C. D.; Kim, K. M.; Zhang, K. Y. Z.; Hockenbery, D. M. Bcl-x_L Mutations Suppress Cellular Sensitivity to Antimycin A. *J. Biol. Chem.* **2004**, *279*, 2159-2165.

65. Sadowsky, J. D.; Fairlie, W. D.; Hadley, E. B.; Lee, H.-S.; Umezawa, N.; Nikolovska-Coleska, Z.; Wang, S.; Huang, D. C. S.; Tomita, Y.; Gellman, S. H. (α/β+α)-Peptide Antagonists of Bh3 Domain/Bcl-xL Recognition: Toward General Strategies for Foldamer-Based Inhibition of Protein-Protein Interactions. *J. Am. Chem. Soc.* **2007**, *129*, 139-154.

66. Dai, Y. Q.; Whittal, R. M.; Li, L. Two-Layer Sample Preparation: A Method for MALDI-MS Analysis of Complex Peptide and Protein Mixtures. *Anal. Chem.* **1999**, *71*, 1087-1091.

67. Coulter-Mackie, M. B.; Lian, Q. Consequences of Missense Mutations for Dimerization and Turnover of Alanine : Glyoxylate Aminotransferase: Study of a Spectrum of Mutations. *Molec. Genet. Metab.* **2006**, *89*, 349-359.

68. Thuduppathy, G. R.; Hill, R. B. Acid Destabilization of the Solution Conformation of Bcl-x_L Does Not Drive Its pH-Dependent Insertion into Membranes. *Protein Sci.* **2006**, *15*, 248-257.

69. Myers, J. K.; Pace, C. N.; Scholtz, J. M. Denaturant *m* Values and Heat Capacity Changes: Relation to Changes in Accessible Surface Areas of Protein Unfolding. *Protein Sci.* **1995**, *4*, 2138-2148.

70. Sadowsky, J. D.; Peterson, K. J.; Fairlie, W. D.; Huang, D. C. S.; Tomita, Y.; Gellman, S. H. Universitiy of Wisconsin, Madison, WI. To be submitted for publication, 2007.

71. Dai, S. Y.; Fitzgerald, M. C. Accuracy of SUPREX (Stability of Unpurified Proteins from Rates of H/D Exchange) and MALDI Mass Spectrometry-Derived Protein Unfolding Free Energies Determined under Non-EX2 Exchange Conditions. *J. Am. Soc. Mass Spectrom.* **2006**, *17*, 1535-1542.

72. Liu, W. S.; Peterson, P. E.; Carter, R. J.; Zhou, X. Z.; Langston, J. A.; Fisher, A. J.; Toney, M. D. Crystal Structures of Unbound and Aminooxyacetate-Bound Escherichia Coli Gamma-Aminobutyrate Aminotransferase. *Biochemistry* **2004**, *43*, 10896-10905.

73. Beeler, T.; Churchich, J. E. Reactivity of Phosphopyridoxal Groups of Cystathionase. *J. Biol. Chem.* **1976**, *251*, 5267-5271.

74. Powell, K. D.; Fitzgerald, M. C. High-Throughput Screening Assay for the Tunable Selection of Protein Ligands. *J. Comb. Chem.* **2004**, *6*, 262-269.

75. Handschumacher, R. E.; Harding, M. W.; Rice, J.; Drugge, R. J. Cyclophilin - a Specific Cytosolic Binding-Protein for Cyclosporin-A. *Science* **1984**, *226*, 544-547.

76. Harding, M. W.; Handschumacher, R. E. Cyclophilin, a Primary Molecular Target for Cyclosporine - Structural and Functional Implications. *Transplantation* **1988**, *46*, S29-S35.

77. Holzman, T. F.; Egan, D. A.; Edalji, R.; Simmer, R. L.; Helfrich, R.; Taylor, A.; Burres, N. S. Preliminary Characterization of a Cloned Neutral Isoelectric Form of the Human Peptidyl Prolyl Isomerase Cyclophilin. *J. Biol. Chem.* **1991**, *266*, 2474-2479.

78. Liu, J.; Albers, M. W.; Chen, C. M.; Schreiber, S. L.; Walsh, C. T. Cloning, Expression, and Purification of Human Cyclophilin in Escherichia-Coli and Assessment of the Catalytic Role of Cysteines by Site-Directed Mutagenesis. *Proc. Natl. Acad. Sci. U. S. A.* **1990**, *87*, 2304-2308.

79. Taylor, D. M. Radioiodinated Cyclosporin A - Preparation and Biodistribution. *Int. J. Appl. Rad. Iso.* **1980**, *31*, 192-193.

80. Bai, Y. W.; Milne, J. S.; Mayne, L.; Englander, S. W. Primary Structure Effects on Peptide Group Hydrogen-Exchange. *Proteins: Struct., Funct., Genet.* **1993**, *17*, 75-86.

81. Zhang, Y. Z. *Structural Biology and Molecular Biophysics*. PhD Thesis, University of Pennsylvania, 1995.

82. Schellman, J. A. Macromolecular Binding. *Biopolymers* **1975**, *14*, 999-1018.

83. Hann, M. M.; Oprea, T. I. Pursuing the Leadlikeness Concept in Pharmaceutical Research. *Curr. Opin. Chem. Biol.* **2004**, *8*, 255-263.

84. Oprea, T. I. Current Trends in Lead Discovery: Are We Looking for the Appropriate Properties? *J. Comput. Aided Mol. Des.* **2002**, *16*, 325-334.

85. Zhang, J. H.; Wu, X.; Sills, M. A. Probing the Primary Screening Efficiency by Multiple Replicate Testing: A Quantitative Analysis of Hit Confirmation and False Screening Results of a Biochemical Assay. *J. Biomol. Screen.* **2005**, *10*, 695-704.

86. Neubert, T. A.; Walsh, K. A.; Hurley, J. B.; Johnson, R. S. Monitoring Calcium-Induced Conformational Changes in Recoverin by Electrospray Mass Spectrometry. *Protein Sci.* **1997**, *6*, 843-850.

87. Zhu, M. M.; Rempel, D. L.; Du, Z. H.; Gross, M. L. Quantification of Protein-Ligand Interactions by Mass Spectrometry, Titration, and H/D Exchange: PLIMSTEX. *J. Am. Chem. Soc.* **2003**, *125*, 5252-5253.

88. Engen, J. R.; Gmeiner, W. H.; Smithgall, T. E.; Smith, D. L. Hydrogen Exchange Shows Peptide Binding Stabilizes Motions in Hck Sh2. *Biochemistry* **1999**, *38*, 8926-8935.

89. Zhang, Z. Q.; Smith, D. L. Determination of Amide Hydrogen-Exchange by Mass-Spectrometry - a New Tool for Protein-Structure Elucidation. *Protein Sci.* **1993**, *2*, 522-531.

90. Lumb, M. J.; Birdsey, G. M.; Danpure, C. J. Correction of an Enzyme Trafficking Defect in Hereditary Kidney Stone Disease in Vitro. *Biochem. J.* **2003**, *374*, 79-87.

91. Zhang, X. X.; Roe, S. M.; Hou, Y. W.; Bartlam, M.; Rao, Z. H.; Pearl, L. H.; Danpure, C. J. Crystal Structure of Alanine:Glyoxylate Aminotransferase and the Relationship between Genotype and Enzymatic Phenotype in Primary Hyperoxaluria Type 1. *J. Mol. Biol.* **2003**, *331*, 643-652.

92. Hopper, E. D.; Roulhac, P. L.; Campa, M. J.; Patz, E. F.; Fitzgerald, M. C. Throughput and Efficiency of a Mass Spectrometry-Based Screening Assay for Protein-Ligand Binding Detection. . *J. Am. Soc. Mass Spectrom.* **2008**, *19*, 1303-1311.

93. Lumb, M. J.; Danpure, C. J. Functional Synergism between the Most Common Polymorphism in Human Alanine : Glyoxylate Aminotransferase and Four of the Most Common Disease-Causing Mutations. *J. Biol. Chem.* **2000**, *275*, 36415-36422.

94. Studier, F. W. Protein Production by Auto-Induction in High-Density Shaking Cultures. *Prot. Expr. Purif.* **2005**, *41*, 207-234.

95. Winterbourn, C. C.; Metodiewa, D. Reactivity of Biologically Important Thiol Compounds with Superoxide and Hydrogen Peroxide. *Free Radic. Biol. Med.* **1999**, *27*, 322-328.

96. Sun, G.; Anderson, V. E. Prevention of Artifactual Protein Oxidation Generated During Sodium Dodecyl Sulfate-Gel Electrophoresis. *Electrophoresis* **2004**, *25*, 959-965.

Biography

Erin Hopper was born in Boone, North Carolina on November 12, 1981. She graduated from Watauga High School in May, 2000 and was awarded a Spencer Fellowship to attend the University of North Carolina at Chapel Hill. During her time at UNC-CH, she performed undergraduate research in the laboratory of Professor Mark H. Schoenfisch and received a Smallwood Foundation Summer Undergraduate Research Fellowship during the summer of 2003. She was married to Steven D. Hopper on May 25, 2002. In May, 2004, Erin graduated summa cum laude with a B.S. in chemistry, and she entered graduate school at Duke University the following fall. Erin joined the laboratory of Professor Michael C. Fitzgerald and was awarded a Center for Biomolecular and Tissue Engineering (CBTE) Fellowship sponsored by an NIH Training Grant to Duke University. She was also the recipient of several travel grants, including two Duke University Conference Travel Fellowships, a CBTE Student Travel Award, and an American Society for Mass Spectrometry Travel Stipend. Erin will receive her PhD in analytical chemistry in May of 2009.

Publications

1. **Hopper, E. D.**; Pittman, A. M. C.; Tucker, C. L.; Campa, M. J.; Patz, E. F.; Fitzgerald, M. C. An H/D Exchange- and Protease Digestion-Based Screening Assay for Protein-Ligand Binding Detection. In preparation.

2. **Hopper, E. D.**;* Xu, Y.;* Fitzgerald, M. C. Investigation of an H/D Exchange and Protease Digestion Protocol for the Thermodynamic Analysis of Protein-Ligand Binding. In preparation. *Equally contributing authors

3. Anbalagan, V.; Dangkulwanich, M.; **Hopper, E. D.**; Fitzgerald, M. C. Thermodynamic Analysis of Protein Folding and Ligand Binding in Complex Mixtures using a Gel-Based Fractionation Strategy. In preparation.

4. Fitzgerald, M. C.; Tang, L.; **Hopper, E. D.** Thermodynamic Analysis of Protein Folding and Ligand Binding by SUPREX. In *Protein Mass Spectrometry*; Whitelegge, J. P., Ed.; Comprehensive Analytical Chemistry, Vol. 52; Elsevier: New York, 2009; pp. 127–149.

5. **Hopper, E. D.**;* Pittman, A. M. C.;* Fitzgerald, M. C.; Tucker, C. L. In Vivo and in Vitro Examination of Stability of Primary Hyperoxaluria-Associated Human Alanine:Glyoxylate Aminotransferase. *J. Biol. Chem.*, **2008**, *283*(45), 30493-30502. *Equally contributing authors

6. **Hopper, E. D.**; Roulhac, P. L.; Campa, M. J.; Patz, E. F.; Fitzgerald, M. C. Throughput and Efficiency of a Mass Spectrometry-Based Screening Assay for Protein-Ligand Binding Detection. *J. Am. Soc. Mass Spectrom.*, **2008**, *19*, 1303-1311.

7. Tang, L.; **Hopper, E. D.**; Tong, Y.; Sadowsky, J. D.; Peterson, K. J.; Gellman, S. H.; Fitzgerald, M. C. H/D Exchange- and Mass Spectrometry-Based Strategy for the Thermodynamic Analysis of Protein-Ligand Binding. *Anal. Chem.* **2007**, *79*(15), 5869-5877.

8. Robbins, M. E.; Oh, B. K.; **Hopper, E. D.**; Schoenfisch, M. H. Nitric Oxide-Releasing Xerogel Microarrays Prepared with Surface-Tailored Poly(dimethylsiloxane) Templates. *Chem. Mater.* **2005,** *17*(12), 3288-3296.

9. Robbins, M. E.; **Hopper, E. D.**; Schoenfisch, M. H. Synthesis and Characterization of Nitric Oxide-Releasing Sol-Gel Microarrays. *Langmuir* **2004,** *20*(23), 10296-10302.