Published in final edited form as:

J Am Soc Mass Spectrom. 2014 January ; 25(1): . doi:10.1007/s13361-013-0754-2.

False Positive Rate Determination of Protein Target Discovery using a Covalent Modification- and Mass Spectrometry-Based Proteomics Platform

Erin C. Strickland, M. Ariel Geer, Jiyong Hong, and Michael C. Fitzgerald*

Department of Chemistry, Duke University, Durham, North Carolina 27708

Abstract

Detection and quantitation of protein-ligand binding interactions is important in many areas of biological research. The Stability of Proteins from Rates of Oxidation (SPROX) technique is an energetics-based technique for identifying the proteins targets of ligands in complex biological mixtures. Knowing the false positive rate of protein target discovery in proteome-wide SPROX experiments is important for the correct interpretation of results. Reported here are the results of a control SPROX experiment in which chemical denaturation data is obtained on the proteins in two samples that originated from the same yeast lysate, as would be done in a typical SPROX experiment except that one sample would be spiked with the test ligand. False positive rates of 1.2–2.2% and <0.8% are calculated for SPROX experiments using Q-TOF and orbitrap mass spectrometer systems, respectively. Our results indicate that the false positive rate is largely determined by random errors associated with the mass spectral analysis of the isobaric mass tag (e.g., iTRAQ[®]) reporter ions used for peptide quantitation. Our results also suggest that technical replicates can be used to effectively eliminate such false positives that result from this random error, as is demonstrated in a SPROX experiment to identify yeast protein targets of the drug, manassantin A. The impact of ion purity in the tandem mass spectral analyses and of background oxidation on the false positive rate of protein target discovery using SPROX is also discussed.

INTRODUCTION

The ability to detect and quantify the strength of protein-ligand binding interactions is critical to many areas of biological research, from the discovery of novel protein-protein interactions to understanding drug mode-of-action. Many techniques for the analysis of protein-ligand binding interactions (e.g., isothermal titration and differential scanning calorimetry(1), surface plasmon resonance(2), yeast two- or three-hybrid assay (3–5), and

*Address reprint requests to: Professor Michael C. Fitzgerald, Department of Chemistry, Box 90346, Duke University, Durham, North Carolina 27708-0346, Tel: 919-660-1547, Fax: 919-660-1605, michael.c.fitzgerald@duke.edu.

protein microarrays(6)) require large amounts of highly purified protein, and/or are limited to binary and direct protein-ligand interactions. Several energetics-based approaches such as Stability of Unpurified Proteins from Rates of H/D Exchange (SUPREX),(7–9) Stability of Proteins from Rates of Oxidation (SPROX) (10–14), and pulse proteolysis methods (15–17) have proven useful for the detection and quantitation of protein-ligand binding interactions in complex biological mixtures such as cell lysates. Such energetics-based methods are attractive because they do not require large amounts of purified protein and are amenable to multiplexed analyses.

Recently, we described several proteome-wide applications of the SPROX technique (11, 13), in which the thermodynamic properties of proteins and protein-ligand complexes are assessed by evaluating the chemical denaturant dependence of the hydrogen peroxide-mediated oxidation of methionine side chains in proteins. The proteome-wide applications of SPROX reported to date have relied on the use of isobaric mass tags to quantify the denaturant dependence of the oxidation reaction in SPROX using a quantitative bottom-up shotgun proteomics strategy. In this strategy, the isobaric mass tag reporter ion intensities obtained for the methionine-containing peptides are used to generate chemical denaturation data. Ultimately, a transition midpoint ($C_{1/2}$ value) is determined from the data (see Figure 1), and used to evaluate a protein's thermodynamic stability (10).

An important application of the SPROX technique is the detection and quantitation of protein-ligand binding interactions. Protein-ligand binding applications involve subjecting a protein sample to a SPROX analysis both in the absence and in the presence of the test ligand, and comparing the chemical denaturation data obtained from each analysis (Figure 1). A shift in the $C_{1/2}$ value to higher or lower denaturant concentration can indicate ligand-induced stabilization or destabilization of the protein, respectively. In ideal cases, these measurements can be used to calculate binding affinities for the protein-ligand interaction. Proteome-wide applications of SPROX have enabled the ligand-binding properties of hundreds of proteins to be simultaneously assayed in the context of complex biological samples (11–13).

While SPROX is a powerful tool for the detection and quantitation of protein-ligand binding interactions in complex biological mixtures, understanding the false positive rate associated with ligand binding detection in proteome-wide SPROX experiments is crucial for the correct interpretation of results. As part of this work, we report on the results of a control SPROX experiment, which enables an evaluation of the false positive rate associated with protein target discovery in proteome-wide SPROX applications using isobaric mass tags. The false positive rate is also examined in a second set of experiments in which the proteins in a yeast cell lysate were assayed for interactions with the drug manassantin A. Also discussed is the impact of ion purity in the tandem mass spectral analyses and of background oxidation on the precision with which $C_{1/2}$ value assignments can be made in proteome-wide SPROX applications using isobaric mass tags.

Materials and Methods

Materials

The *Saccharomyces cerevisiae* open reading frame strain used in this experiment (Y258) was purchased from Open Biosystems (Waltham, MA). Guanidine hydrochloride (GdmCl), 30% hydrogen peroxide, trichloroacetic acid (TCA), trifluoroacetic acid (TFA), triethylammonium bicarbonate (TEAB), s-methyl methanethiosulfonate (MTS), ammonium bicarbonate (AmBic), L-methionine, L-tryptophan, L-histidine, L-leucine, sodium dodecyl sulfate (SDS), and proteomics grade porcine pancreas trypsin were purchased from Sigma-Aldrich (St. Louis, MO). Ethanol and methanol were purchased from

VWR International (Radnor, PA). Solvents for the LC mobile phases, acetonitrile (ACN) and high purity water were purchased from Honeywell Burdick & Jackson (Morristown, NJ). Sodium phosphate dibasic, anhydrous was purchased from EM Science (Gardena, CA) and acetic acid was purchased from Mallinckrodt (Hazelwood, MO). The Pi³TM - Methionine Reagent Kit and the macro C-18 spin columns were purchased from The Nest Group (Southboro, MA). The iTRAQ[®] 8-plex reagents were purchased from ABSciex (Framingham, MA). The Bradford reagent, formic acid (FA), tris (2-carboxyethyl) phosphine hydrochloride (TCEP HCl), and the protease inhibitors (AEBSF, bestatin, E-64, leupeptin, and pepstatin A) were purchased from Thermo Scientific (Waltham, MA). A 20X inhibitor cocktail was prepared in which stock solutions of the inhibitors were mixed to give a final concentration of 0.2 mM pepstatin A, 0.4 mM leupeptin, 0.3 mM E-64, 1 mM bestatin, and 20 mM AEBSF. Manassantin A was prepared by total chemical synthesis and purified as described elsewhere (18).

SPROX Analysis: Control Experiment

A yeast cell lysate solution containing 20 µg/µL total protein was prepared as described elsewhere (12). Two 180 µL aliquots of the lysate were each combined with 20 µL of 20 mM phosphate buffer (pH 7.4) to generate the two (–) ligand samples in the control experiment (i.e., Controls A and B). After a 30 min equilibration, 20 µL aliquots of the (–) ligand samples were combined with 25 µL volumes of the SPROX buffers, which were comprised of 20 mM phosphate buffer (pH 7.4) and increasing amounts of GdmCl such that the final [GdmCl] in each SPROX oxidation reaction was 0, 0.6, 1.0, 1.4, 1.7, 2.0, 2.4, and 3.0 M. The resulting two sets of eight SPROX samples were equilibrated for 30 min before 5 µL of 30% H₂O₂ was added to initiate the methionine oxidation reaction in each SPROX sample. After 3 min, the oxidation reaction in each SPROX sample was quenched with 1 mL of 300 mM L-methionine. A 300 µL volume of a TCA solution (1 g/ml) was added to each sample. The samples were kept on ice at 4°C overnight and centrifuged at 8,000 rcf for 30 min at 4°C. The supernatant was decanted and the protein pellets were washed 3 times with 300 µL of ice-cold ethanol prior to the proteomic sample preparation and analysis described below.

SPROX Analysis: Manassantin A Binding Experiment

A yeast cell lysate solution containing 10 µg/µL total protein was prepared as described elsewhere (12), and divided into two 180 µL portions. A 20 µL aliquot of a 3 mM manassantin A solution prepared in DMSO was added to one portion to create a (+) ligand sample, and 20 µL of DMSO was added to the other portion to create a (–) ligand sample. The samples were incubated for 1 hr at room temperature before 20 µL aliquots of the (–) and (+) samples were combined with 25 µL volumes of the SPROX buffers, which were comprised of 20 mM phosphate buffer (pH 7.4) and increasing amounts of GdmCl such that the final [GdmCl] in each SPROX oxidation reaction was 0.5, 1.0, 1.3, 1.5, 1.8, 2.0, 2.5, and 3.0 M. The resulting two sets of eight samples were equilibrated for 30 min at room temperature before 5 µL of 30% H₂O₂ was added to initiate the SPROX oxidation reaction. After 3 min the SPROX oxidation reactions were each quenched with 1 mL of 300 mM L-methionine. The proteins in each SPROX sample were precipitated by addition of 200 µL of 1 g/mL TCA on ice for 3 hr. Samples were centrifuged at 8000 rcf for 30 min at 4°C. The supernatant was decanted, and the protein pellets were washed four times with 300 µL of ice-cold ethanol. Samples were placed in a SpeedVac concentrator for 6 min to remove the ethanol prior to the proteomic sample preparation described below.

Proteomic Sample Preparation and Analysis

The protein pellets generated in the control experiment and in the manassantin A binding experiment were each dissolved in 35 µL of 0.5 M TEAB buffer, pH 8.5, containing 0.1%

SDS. Samples were vortexed, sonicated, and heated at 60 °C for 10–15 min each in order to facilitate the dissolution of the protein pellets. A 5 µL aliquot of each sample was removed and used in a Bradford Assay to determine the total protein concentration in each sample, which ranged from 0.3 to 1.7 mg/mL. Samples were reacted with 5 mM TCEP at 60°C for 1 hr to reduce disulfide bonds, reacted with 10 mM MMTS for 10 min at room temperature to block sulfhydryl groups in cysteine side chains, and digested with 1 µg of trypsin for ~16 hr at 37 °C.

The four sets of eight SPROX samples (i.e., the two sets of (–) ligand samples in the control experiment and the (–) and (+) samples in the manassantin A binding experiment) were each labeled with a set of iTRAQ® 8-plex reagents. Each sample was labeled with 0.5 units of iTRAQ® reagent dissolved in 50 µL isopropanol. The (–) ligand samples from the 0, 0.6, 1.0, 1.4, 1.7, 2.0, 2.4, and 3.0 M GdmCl-containing SPROX samples in the control experiment were labeled with the 113, 114, 115, 116, 117, 118, 119, and 121 iTRAQ® tags, respectively. The (–) and (+) ligand samples from the 0.5, 1.0, 1.3, 1.5, 1.8, 2.0, 2.5, and 3.0 M GdmCl-containing SPROX samples in the manassantin A binding experiment were labeled with the 113, 117, 115, 119, 116, 114, 121, and 118 iTRAQ® tags, respectively. The pH was adjusted to between pH 7.5 and 8.5, as necessary. The labeling reaction was allowed to proceed for 2 hr at room temperature before the samples were stored at –20°C until further use.

Aliquots, corresponding to the same amount of total peptide (~10–30 µg), from the eight different iTRAQ® labeled samples from each SPROX analysis were combined to generate two (–) ligand samples from the control experiment and a (–) and (+) ligand sample from the manassantin A binding experiment. Each of the resulting four samples was split into two portions. One portion of each sample was desalted using C18 resin (The Nest Group) to generate “non-enriched” samples, and the other portion of each sample was enriched for methionine-containing peptides using the Pi³™ Methionine Reagent (The Nest Group) according to the manufacturer’s protocols to generate “enriched” samples. The samples used for Controls 1A and 2A were from the same C18-clean-up and the same methionine enrichment, as were the samples used for Controls 1B and 2B. The samples used for Controls 3A and 3B were from additional C18 treatments and methionine enrichments. The samples for Man A Replicate 1 and 2 were from the same C18 clean-up and methionine enrichment.

LC-MS/MS Sample Analysis: Control Experiment

The two (–) ligand samples in the control experiment were analyzed on a Q-TOF LC-MS/MS system from Agilent Technologies, Inc. (Santa Clara, CA) that was comprised of a 1200 series Nano and Cap pump with a micro-well plate sampler and a 6520B Q-TOF with a Chip Cube interface. The HPLC Chip used in this experiment contained a 160 nL trapping column and 75 µm × 150 mm column with 300 Å Zorbax C18 packing (5 µM). The flow rate was 400 nL/min. The two solvents used to establish the gradient were solvent A, which was 0.1% formic acid in H₂O and solvent B, which was 0.1% formic acid in ACN. The gradient increased linearly from 3–15% solvent B in 2.5 min, 15–45% solvent B in 78 min, and 45–100% solvent B in 10 min. The drying gas flow rate was set to 4 or 6 L/min at a temperature of 350°C. The capillary voltage, fragmentor, skimmer, and octapole RF were set to 1,900–1,975 V, 175 V, 65.0 V, and 750 V, respectively. Precursor ion scans were collected at a rate of 3 scans/sec and product ion scans were collected at a rate of 2 scans/sec. Product ion scans were collected for the top 4 most intense peaks in a given precursor scan with an intensity threshold of 1,000, and dynamic exclusion of a given m/z ratio was set at 2 scans in a 0.6 min window. The precursor isolation width was 4 m/z, the scan range for the precursor scan was 300–3,000 m/z and 100–3,000 m/z for the product ion scan. Collision induced dissociation was achieved with an energy of 3.9 V/100 m/z and an offset

of +2.9 V. For the LC-MS/MS runs in Control 3, an additional setting of preferred ions from the previous controls' hits were added to increase the chance of collecting additional quantitative data for those peptides.

The two (–) ligand samples, A and B, generated in the control experiment were each subjected to total of 22 LC-MS/MS analyses. The LC-MS/MS data were searched on Spectrum Mill Workbench Software B.04 Rev. B against the SwissProt *Saccharomyces cerevisiae* database with trypsin digestion allowing only tryptic termini and up to 3 missed cleavages. Fixed modifications were set to MMTS on cysteines and iTRAQ® 8-Plex on the N-terminus and lysines. Oxidation was set as a variable (0–1) modification for methionine. The mass tolerance was set to 20 ppm for the precursor ion and 50 ppm for the product ions. In the methionine enriched samples, methionine was set as a required residue.

The peptide and protein identifications and iTRAQ® reporter ion intensities from the first three LC-MS/MS analyses of the enriched and non-enriched samples of A and B constituted the data in Control 1. The data resulting from the next three LC-MS/MS analyses of the enriched and non-enriched samples of A and B constituted the data for Control 2, and the data resulting from the last 5 LC-MS/MS analyses of enriched and non-enriched samples of A and B constituted the data in Control 3. Thus, only Control samples 1 and 2 are true technical replicates.

LC-MS/MS Sample Analysis: Manassantin A Binding Experiment

The (–) and (+) ligand samples in the manassantin A binding experiment were analyzed at the Proteomics Resource at the Fred Hutchinson Cancer Research Center on an Orbitrap Elite ETD mass spectrometer equipped with an EASY-nLC system (ThermoFisher Scientific Inc., Waltham, MA). The trapping column was a 100 µm × 2 cm Integragrit column (New Objective) packed with 200 Å Magic C18 AQ 5 µm material (Michrom), and the column was a 75 µm × 25 cm PicoFrit column (New Objective) packed with 100 Å Magic C18 AQ 5 µm material (Michrom). The flow rate was 400 nL/min. The LC gradient was 5 to 7% solvent B over 2 min, 7 to 35% solvent B over 90 min, 35 to 50% solvent B over 1 min, isocratic at 50% solvent B for 9 min, increased from 50 to 95% B over 1 min, and finally isocratic at 95% solvent B for the last 8 min. Product ion scans (resolution 15,000) were collected for the top 10 most intense peaks in a given precursor scan (resolution 60,000) with an intensity threshold of 5,000, and dynamic exclusion of a given m/z ratio was set at 1 scan in a 0.75 min window. The precursor isolation width was 1.2 m/z. The scan range for the precursor scan was 400–1,800 m/z and 100–2,000 m/z for the product ion scan. Collision induced dissociation was achieved using HCD with normalized collision energy of 40% and an HCD activation time of 0.1 ms.

The LC-MS/MS data were searched using Proteome Discoverer with fixed modifications of MMTS on cysteine and iTRAQ® 8-Plex on N-terminus and lysine residues. Proteome Discoverer was set with a variable modification of oxidation on methionine residues. For data analysis, only peptides identified with high and medium confidence (i.e., false discovery rates (FDR) <1% and 5%, respectively) were used. The resulting peptide and protein identifications and iTRAQ® reporter ion intensities from the first LC-MS/MS analyses of the enriched and non-enriched samples of the (–) and (+) sample comprised manassantin A replicate 1, and the resulting data from the second LC-MS/MS analyses of the enriched and non-enriched samples of the (–) and (+) samples comprised manassantin A replicate 2. Thus, the two manassantin A replicate analyses were true technical replicates.

SPROX Data Analysis

The SPROX data analysis was performed as described previously (12). Briefly, the identified peptide sequences and their corresponding iTRAQ[®] reporter ion intensities were exported into Excel along with the filename, identification, score, charge state, modifications, isolation purity, retention time, protein accession number, and protein name. Only peptides with high quality quantitative data (i.e., iTRAQ[®] reporter ion intensities at m/z 113–121 that summed to >1,000) were used in subsequent analyses. The iTRAQ[®] reporter ion intensities for an individual peptide were normalized as we have previously described. (12) Detailed information is provided in the Supporting Material section.

The normalized iTRAQ[®] reporter ion intensities were used to generate chemical denaturation data sets for the peptides identified and successfully quantified in the two (–) ligand samples in the control experiment and in the two (–) and (+) ligand samples in the manassantin A binding experiment. In each experiment a set of hit peptides with normalized iTRAQ[®] reporter ion differences that resulted in transition midpoint shifts of >0.5 M were identified by a visual inspection of the data or by visual inspection of the data after a difference analysis as described in the Supporting Material. All the peptides with significant transition midpoint shifts (i.e., shifts >0.5 M) and consistent retention times (i.e., within 3 min) in the control experiments were labeled as false positives, and all the peptides with significant transition midpoint shifts and consistent retention times (i.e., within 3 min) in the manassantin A mode-of-action experiment were labeled as potential hits.

RESULTS AND DISCUSSION

Control Experiment

In order to directly assess the false positive rate of the SPROX technique, a control experiment was performed using the SPROX protocol outlined in Figure 1. Two equal portions of the same yeast cell lysate sample were subjected to parallel SPROX analyses, as would be done in a typical protein-ligand binding experiment using SPROX with the exception that neither protein sample was incubated with ligand. The proteomic results obtained in replicate LC-MS/MS analyses of the control SPROX samples are summarized in Table 1. The LC-MS/MS data in each control replicate were that from 3 – 5 LC-MS/MS runs of the non-enriched peptide sample and 3 – 5 LC-MS/MS runs of the methionine-enriched peptide sample. Also, shown in Table 1 is a summary of the proteomic results obtained when the isobaric mass tag data from all three control experiments were analyzed together. The iTRAQ data from each of the control experiments and in the combined data set were analyzed and false positive peptides and proteins were identified (see Table S-1 in the Supporting Material). All the identified false positive peptides resulted from normalized reporter ion intensity differences that yielded transition midpoint shifts between 0.5 and 1.0 M [GdmCl]. Shown in Figure 2 are the chemical denaturation data obtained for a typical non-hit and false positive peptide.

The results shown in Table 1 reveal that the false positive rate of protein-ligand discovery using SPROX is in the range of 1.2 – 2.2%. This rate is slightly smaller than the 5% false positive rate previously reported in a study that compared the results obtained on control samples generated in two different protein-ligand binding studies performed on two different instrument platforms (13). This earlier study did not employ a retention time filter. An inspection of the 8 false positive peptides and proteins identified in the three control experiments in this work revealed that none of the 8 peptide hits were detected as a hit in more than one control experiment, despite the fact that all but 1 of the peptide hits were effectively assayed in at least two of the control experiments (Table S-1). The lack of common false positive peptides among the control experiments suggests that the false

positives observed in these control experiments are the result of random error associated with the mass spectral readout of the isobaric mass tag reporter ion intensities.

The variation in reporter ion intensities obtained on the non-methionine-containing peptides in the SPROX experiment can be used to assess the expected variability of the isobaric mass tag reporter ion intensities in the mass spectral readout. The normalized reporter ion intensities for all of the non-methionine containing peptides in the SPROX experiment were generally distributed around 1.0 (see Figure 3). The standard deviations associated with the normalized reporter ion intensities of the non-methionine containing peptides in the technical replicates of the control experiment ranged from 0.28 – 0.56 with an average standard deviation of 0.41 (see Table S-2). Our results suggest that with such variances the false discovery rate of hit peptides in the SPROX experiment is on the order of 1.2 – 2.2% (Table 1).

Protein-Ligand Binding Experiment with Manassantin A

The SPROX protocol outlined in Figure 1 was utilized in a drug mode-of-action study to assess the reproducibility of potential hits in technical replicates. In this analysis, the endogenous proteins in a yeast cell lysate were analyzed for binding to manassantin A, a natural product that has been shown to have anti-cancer activity in cell-based assays but that has an unknown mode-of-action (18–20). The LC-MS/MS data for this ligand-binding experiment were acquired on an orbitrap mass spectrometer with HCD. The proteomic results obtained from replicate LC-MS/MS analyses of the (–) and (+) ligand samples generated in this manassantin A binding experiment are summarized in Table 1. Representative chemical denaturation data obtained in the manassantin A binding experiment are shown in Figure 4.

The increased speed and sensitivity of the orbitrap mass spectrometer system over the Q-TOF system not only increased the peptide and protein coverage in our SPROX experiments, but also significantly reduced the number of LC-MS/MS runs needed to attain such coverage. We also found that the raw iTRAQ reporter ion intensities were significantly higher on the orbitrap mass spectrometer system than on the Q-TOF platform. For example, the median iTRAQ reporter ion intensities for the LC-MS/MS analyses of the non-enhanced samples analyzed on the Q-TOF were between 250 and 500 relative intensity units, whereas they were between 12,000 and 18,000 relative intensity units for similar samples analyzed on the orbitrap instrument.

It is noteworthy that the variation in reporter ion intensities obtained for the non-methionine-containing peptides in the manassantin A binding experiment was significantly smaller than that observed in the control experiments. The standard deviations associated with the normalized reporter ion intensities of the non-methionine containing peptides in the technical replicates of the manassantin A binding experiments ranged from 0.12–0.24, with an average standard deviation of 0.15 (Table S-2). This variation is approximately 2.5-fold smaller than that observed in the control experiment described above, which was performed on an Agilent Q-TOF mass spectrometer (see Figure 3 and Table S-2). The lower variation better distinguishes the pre- and post-transition baselines of a SPROX curve, which facilitates the assignment of transition midpoints. We attribute the lower variation observed on the orbitrap mass spectrometer system to the increased ion signal intensities of the iTRAQ reporter ions recorded on this instrument.

In order to assess the false positive rate associated with our SPROX experiments on the orbitrap mass spectrometer system, we compared the chemical denaturation data obtained in the two technical replicates of the (–) ligand samples from the manassantin A binding experiment. A total of 8 out of the 1,057 methionine-containing peptides analyzed in the two

technical replicates were identified as hits even though the two data sets originated from the same (–) ligand sample (Tables 1 and S-3). From this analysis we calculated a false positive rate of protein target discovery of 0.8 % for SPROX data acquired on the orbitrap instrument used in this work.

The chemical denaturation data sets obtained in the two replicate analyses of the(–) and (+) ligand samples in the manassantin A binding experiment were analyzed to identify potential protein targets of manassantin A. In replicate 1 of the manassantin A binding study, 7 out of the 1,082 methionine-containing peptides analyzed were identified as potential hits. In replicate 2, 6 out of the 1,269 methionine-containing peptides analyzed were identified as potential hits (see Table S-4). Although 6 of the 13 potential hit peptides were assayed in both replicates, none of these 6 potential hit peptides consistently appeared as a hit in both replicates (Table S-4). In theory, these 6 peptides could be false positives in the replicate in which they were detected as hits or false negatives in the replicate in which they were not detected as hits. Unfortunately, the false negative rate of protein target discovery in the SPROX experiment is difficult to determine. However, given our finding that the false positive rate in SPROX experiments is relatively small (<1%) and largely due to random error, it is likely that the 6 potential hit peptides yielding inconsistent results in the two replicate analyses are false positives. The false positives observed in the control experiments described above displayed the same such variability. Shown in Figure 4 are the chemical denaturation data obtained for a non-hit peptide and for one of the 6 peptides that only appeared as a hit in one technical replicate. The 7 potential hit peptides that were only assayed in one technical replicate (see Table S-4) require further validation (e.g., using additional biological and/or technical replicates of SPROX and other assays) to determine if the observed $C_{1/2}$ value shifts are indicative of true interactions with manassantin A or if they are also false positives.

Impact of Ion Purity in Tandem Mass Spectral Analyses

One source of error in quantitative proteomics experiments using isobaric mass tags is that co-eluting isobaric peptides can interfere with quantitation (21). In order to determine if such a phenomenon was impacting the false discovery of protein-ligand interactions in the SPROX experiment, we examined the ion purity of the hit peptide ions when they were selected for CID (or HCD) in the LC-MS/MS analyses (Figure 5). Global analyses of the data in the control and manassantin A binding experiments revealed that the ion purities of the hit peptides in these experiments were not dramatically different from those of the non-hit peptides (Figure 5). However, an inspection of the individual product ion mass spectra used to generate the chemical denaturation data of the hit peptides revealed that a number of the peptide hits in our experiments did indeed have at least one product ion mass spectra generated from peptide ions of relatively low purity (i.e., ion purities less than 50% and 70% in the control and manassantin A binding experiments, respectively). In both cases these ion purities include the bottom ~20% of the data (see Figure 5).

The fraction of peptide hits with product ion mass spectra from low purity ions varied in the experiments described. The fraction was 2 out of 8 in the control experiments on the Q-TOF instrument (see Table S-1), 6 out of 8 in the control comparison on the orbitrap instrument (see Table S-3), 1 out of 7 in replicate 1 of the manassantin A binding experiment (see Table S-4), and 3 out of 6 in replicate 2 of the binding experiment (see Table S-4). In general, the chemical denaturation data of hit peptides with product ion mass spectra from low purity ions tended to have no transitions, very little distinction between the pre- and post-transition baselines, or one reporter ion intensity that was relatively high. In theory, peptide hits resulting from chemical denaturation data compromised by product ion mass spectra from low purity ions can be identified and filtered out to reduce the false positive rate in SPROX experiments. Interestingly, the peptide hits with chemical denaturation data that were

compromised by low ion purities tended to be from oxidized peptides from the non-enriched sample runs, which were more complex and largely comprised of non-methionine-containing peptides. Interferences from non-methionine-containing peptides, which have no denaturant dependence to their iTRAQ reporter ion intensities, are likely responsible for the poor separation of the pre- and post-transition baselines that ultimately lead to incorrect $C_{1/2}$ value determinations.

Impact of Background Oxidation

A potential complication in proteome-wide SPROX analyses is the oxidation of methionine-containing peptides before and/or after the methionine oxidation reaction in the SPROX buffers is quenched. Such pre- and/or post-SPROX oxidation is difficult to identify and can compromise the accuracy of $C_{1/2}$ value determinations using oxidized peptides. If the extent of such pre- and post-SPROX oxidation is different in the (–) and (+) ligand samples, then the precision of $C_{1/2}$ value determinations will be compromised, and false positives may result. Clearly, the impact of background oxidation on the false positive rate in SPROX will be experiment dependent. It is also noteworthy that the impact of such background oxidation on the false positive rate was small. For example, the oxidized peptide hits in the control and manassantin A binding experiments in this work constituted small fractions (i.e., between 0 and 2.5%) of the total number of oxidized peptides assayed in each experiment.

One source of background oxidation is the oxidation of methionine-containing peptides at the electrospray interface between the LC column and the mass spectrometer (22). In contrast to other sources of background oxidation in the SPROX experiment, this source of oxidation can be identified in the SPROX experiment. Methionine-containing peptides oxidized at the electrospray interface are identified in the LC-MS/MS analysis as oxidized, but they have the retention time and chemical denaturation data (i.e., iTRAQ[®] reporter ion intensities that decrease as a function of denaturant) of the corresponding un-oxidized peptide. Oxidized peptides displaying such behavior can be easily identified and eliminated from the analysis. In this work, 7 oxidized peptides identified in the manassantin A binding experiments performed on the orbitrap instrument displayed this behavior. These peptides were eliminated from the analysis and are not included in the hit peptide lists summarized in Table S-3 or 4. Interestingly, no such hit peptides were identified in the control experiments performed on the Q-TOF instrument. Presumably, this was a result of the decreased sensitivity and/or different source design of the Q-TOF.

Conclusions

The false discovery rate of peptide hits observed in SPROX experiments was ~2% in a control experiment performed on a Q-TOF mass spectrometer and ~0.8 % in a ligand-binding study performed on an orbitrap mass spectrometer system equipped with HCD. If peptide hits with low ion purity are filtered out, the false positive rates determined in our experiments on the Q-TOF and orbitrap instruments decrease to ~1.2% and 0.2%, respectively. We conclude that the false discovery rate of peptide “hits” in the SPROX experiment is ultimately limited by the inherent variability of the reporter ion intensities generated in the LC-MS/MS analyses. The technical variability defined here sets a lower bound to the overall false positive rate of protein target discovery in SPROX experiments, as the overall false positive rate may also be impacted by biological variability, which was not assessed here.

In the control experiment performed on the Q-TOF mass spectrometer, the variability in the iTRAQ[®] reporter ion intensities was determined to be $\pm 40\%$, while the variability of the iTRAQ intensities was $\pm 15\%$ in the manassantin A binding study performed on the orbitrap mass spectrometer. This 2.5-fold decrease in iTRAQ[®] reporter ion variation is likely

responsible for the reduced false positive rate we observed in our iTRAQ® SPROX experiments performed on the orbitrap instrument. We find that ion purity in tandem mass spectral analyses and background oxidation can affect the false positive rate of protein ligand discovery in SPROX analyses, but such effects are not widespread and in many cases false positives resulting from such effects can be identified and filtered out. Our results suggest that technical replicates of SPROX samples provide an effective and efficient way to eliminate false positives obtained in proteome-wide SPROX studies of protein-ligand interactions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant from the National Institutes of Health (GM084174) to M.C.F. and by a Research Scholar Grant (122057-RSG-12-045-01-CDD) from the American Cancer Society to J.H.

References

1. Jelesarov I, Bosshard HR. Isothermal titration calorimetry and differential scanning calorimetry as complementary tools to investigate the energetics of biomolecular recognition. *J Mol Recognit*. 1999; 12:3–18. [PubMed: 10398392]
2. Oshannessy DJ, Brighamburke M, Sonesson KK, Hensley P, Brooks I. Determination of Rate and Equilibrium Binding Constants for Macromolecular Interactions Using Surface-Plasmon Resonance - Use of Nonlinear Least-Squares Analysis-Methods. *Anal Biochem*. 1993; 212:457–468. [PubMed: 8214588]
3. Fields S, Song OK. A Novel Genetic System to Detect Protein Protein Interactions. *Nature*. 1989; 340:245–246. [PubMed: 2547163]
4. Fashena SJ, Serebriiskii I, Golemis EA. The continued evolution of two-hybrid screening approaches in yeast: how to outwit different preys with different baits. *Gene*. 2000; 250:1–14. [PubMed: 10854774]
5. SenGupta DJ, Zhang BL, Kraemer B, Pochart P, Fields S, Wickens M. A three-hybrid system to detect RNA-protein interactions in vivo. *Proc Natl Acad Sci USA*. 1996; 93:8496–8501. [PubMed: 8710898]
6. Templin MF, Stoll D, Schrenk M, Traub PC, Vohringer CF, Joos TO. Protein microarray technology. *Trends Biotechnol*. 2002; 20:160–166. [PubMed: 11906748]
7. Ghaemmaghami S, Fitzgerald MC, Oas TG. A quantitative, high-throughput screen for protein stability. *Proc Natl Acad Sci U S A*. 2000; 97:8296–8301. [PubMed: 10890887]
8. Ghaemmaghami S, Oas TG. Quantitative protein stability measurement in vivo. *Nat Struct Biol*. 2001; 8:879–882. [PubMed: 11573094]
9. Siburt CJP, Roulhac PL, Weaver KD, Noto JM, Mietzner TA, Cornelissen CN, Fitzgerald MC, Crumbliss AL. Hijacking transferrin bound iron: protein-receptor interactions involved in iron transport in *N-gonorrhoeae*. *Metallomics*. 2009; 1:249–255. [PubMed: 20161024]
10. West GM, Tang L, Fitzgerald MC. Thermodynamic analysis of protein stability and ligand binding using a chemical modification- and mass-spectrometry based strategy. *Anal Chem*. 2008; 80:4175–4185. [PubMed: 18457414]
11. West GM, Tucker CL, Xu T, Park SK, Han XM, Yates JR, Fitzgerald MC. Quantitative proteomics approach for identifying protein-drug interactions in complex mixtures using protein stability measurements. *Proc Natl Acad Sci USA*. 2010; 107:9078–9082. [PubMed: 20439767]
12. Strickland EC, Geer MA, Tran DT, Adhikari J, West GM, DeArmond PD, Xu Y, Fitzgerald MC. Thermodynamic analysis of protein-ligand binding interactions in complex biological mixtures using the stability of proteins from rates of oxidation. *Nat Protoc*. 2013; 8:148–161. [PubMed: 23257983]

13. DeArmond PD, Xu Y, Strickland EC, Daniels KG, Fitzgerald MC. Thermodynamic Analysis of Protein-Ligand Interactions in Complex Biological Mixtures using a Shotgun Proteomics Approach. *J Prot Res*. 2011; 10:4948–4958.
14. DeArmond PD, West GM, Huang HT, Fitzgerald MC. Stable Isotope Labeling Strategy for Protein-Ligand Binding Analysis in Multi-Component Protein Mixtures. *J Am Soc Mass Spectrom*. 2011; 22:418–430. [PubMed: 21472561]
15. Liu PF, Kihara D, Park C. Energetics-Based Discovery of Protein-Ligand Interactions on a Proteomic Scale. *J Mol Biol*. 2011; 408:147–162. [PubMed: 21338610]
16. Chang Y, Schleich JP, VerHeul RA, Park C. Simplified proteomics approach to discover protein-ligand interactions. *Protein Sci*. 2012; 21:1280–1287. [PubMed: 22733688]
17. Park CW, Marqusee S. Pulse proteolysis: A simple method for quantitative determination of protein stability and ligand binding. *Nat Methods*. 2005; 2:207–212. [PubMed: 15782190]
18. Kim H, Kasper AC, Moon EJ, Park Y, Wooten CM, Dewhirst MW, Hong JY. Nucleophilic Addition of Organozinc Reagents to 2-Sulfonyl Cyclic Ethers: Stereoselective Synthesis of Manassantins A and B. *Org Lett*. 2009; 11:89–92. [PubMed: 19111058]
19. Hossain CF, Kim YP, Baerson SR, Zhang L, Bruick RK, Mohammed KA, Agarwal AK, Nagle DG, Zhou YD. Saururus cernuus lignans - Potent small molecule inhibitors of hypoxia-inducible factor-1. *Biochem Biophys Res Com*. 2005; 333:1026–1033. [PubMed: 15967416]
20. Kasper AC, Moon EJ, Hu XQ, Park Y, Wooten CM, Kim H, Yang WT, Dewhirst MW, Hong JY. Analysis of HIF-1 inhibition by manassantin A and analogues with modified tetrahydrofuran configurations. *Bioorg Med Chem Lett*. 2009; 19:3783–3786. [PubMed: 19423348]
21. Christoforou AL, Lilley KS. Isobaric tagging approaches in quantitative proteomics: the ups and downs. *Anal Bioanal Chem*. 2012; 404:1029–1037. [PubMed: 22580419]
22. Morand K, Talbo G, Mann M. Oxidation of Peptides during Electrospray-Ionization. *Rapid Commun Mass Spectrom*. 1993; 7:738–743. [PubMed: 8374164]

Figure 1. Schematic representation of the SPROX protocol using iTRAQ[®] quantitation that was investigated in this work.

Figure 2.

Representative chemical denaturation data obtained in the control experiment. Data obtained for a non-hit and false positive peptide from Control 1 are shown in (A) and (B), respectively. The dark and light shaded bars indicate the data generated in Control 1A and 1B, respectively. The dotted lines indicate the normalized reporter ion value that best separates the pre- and post-transition baselines. The dark and light arrows indicate the denaturant concentration at the transition midpoint. In cases where data from multiple product ion mass spectra were averaged, the average values are shown. The error bars represent one standard deviation.

Figure 3.

Representative distributions of the N1-normalized iTRAQ[®] reporter ion intensities (see Supporting Material for normalization procedure) generated using the non-methioinine-containing peptides in the control experiment (A) and in the manassantin A binding experiment (B). The data in (A) are from the 115 iTRAQ reporter ion in Control 3B. The data in (B) are from the 115 iTRAQ reporter ion in the (+) manassantin A sample in replicate 2. The averages and standard deviations for the values in (A) and (B) are 0.91 ± 0.30 and 1.10 ± 0.13 , respectively.

Figure 4.

Representative chemical denaturation data obtained in the manassantin A binding experiment. Data obtained on a non-hit peptide that were assayed in both replicates are shown in (A). Data obtained from a hit peptide in Replicate 1 that was not a hit in Replicate 2 are shown in (B). The dark and light shaded bars indicate the data generated in the absence and in the presence of manassantin A, respectively. The dotted lines indicate the normalized reporter ion value that separates the pre- and post-transition baseline. The dark and light arrows indicate the denaturant concentration at the transition midpoint obtained in the absence and in the presence of manassantin A, respectively. In cases where data from at least two product ion mass spectra were averaged, the average values are shown. The error bars represent one standard deviation.

Figure 5.

Distribution of peptide ion purities observed in the control and manassantin A binding experiments. (A) Distribution of peptide ion purities from the LC-MS/MS runs on the Q-TOF instrument used in Control Experiments 1, 2, and 3. (B) Distribution of peptide ion purities from the LC-MS/MS runs on the orbitrap instrument used in the manassantin A binding experiments. In each case, the dotted black line and left scale is the data from all the peptides. The dashed line and left scale are for the data from all the methionine-containing peptides. The solid black line and right scale are for the data from the hit peptides identified in each experiment. Peptide ion purities in (A) were taken directly from Spectrum Mill, and peptide ion purities in (B) are reported as $[100 - (\text{the "peptide isolation interference value (\%)"} \text{ reported in Proteome Discoverer})]$. The vertical line marks the ion purity below which includes the bottom 20% of the data from all the peptides.

Table 1

Summary of proteomic results obtained in the control and manassantin A (Man A) binding experiments performed in this work using the SPROX protocol. See the **LC-MS/MS Sample Analysis** sections in the **EXPERIMENTAL** for description of data that comprise each sample set.

Sample	Methionine-Containing Peptides (Proteins) Assayed	Hit Peptides (Proteins)	False Positive Rate
Control 1A vs. 1B	231 (144)	5 (5)	2.2%
Control 2A vs. 2B	145 (104)	2 (2)	1.4%
Control 3A vs. 3B	83 (65)	1 (1)	1.2%
All Control A vs. B*	400 (292)	9 (9)	2.2%
Man A Replicate 1	1082 (415)	7 (7)	0.6% **
Man A Replicate 2	1269 (447)	6 (6)	0.5% **
Man A Control 1 vs. 2	1057 (412)	8 (8)	0.8%

* The number of methionine-containing peptides (proteins) assayed is significantly increased when the data from all 22 LC-MS/MS runs of the A and B samples is used in the data analysis, because more peptides are identified and quantified in both the A and B samples. The smaller number of LC-MS/MS runs in the Control 1, 2, and 3, samples limits the number of overlapping peptides identified in samples A and B.

** Reported value is an upper bound as it assumes that all identified hits are false positives.