Published in final edited form as:

J Proteome Res. 2010 December 3; 9(12): 6696–6704. doi:10.1021/pr100843x.

Comparison of Strong Cation Exchange and SDS/PAGE Fractionation for Analysis of Multi-Protein Complexes

Sudipto Das^{#,1}, Allen D. Bosley^{#,1}, Xiaoying Ye¹, King C. Chan¹, Isabel Chu², Jeffery E. Green², Haleem J. Issaq¹, Timothy D. Veenstra¹, and Thorkell Andresson^{*,1}

¹Laboratory of Proteomics and Analytical Technologies, Advanced Technology Program, SAIC-Frederick, Inc., NCI-Frederick, Frederick, Maryland 21702

²Laboratory of Cancer Biology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA

Abstract

Affinity purification of protein complexes followed by identification using liquid chromatography/mass spectrometry (LC-MS/MS) is a robust method to study the fundamental process of protein interaction. While affinity isolation reduces the complexity of the sample, fractionation prior to LC-MS/MS analysis is still necessary to maximize protein coverage. In this study, we compared the protein coverage obtained via LC-MS/MS analysis of protein complexes pre-fractionated using two commonly employed methods, SDS-PAGE and strong cation exchange chromatography (SCX). The two complexes analyzed focused on the nuclear proteins Bmi-1 and GATA3 that were expressed within the cells at low and high levels, respectively. Pre-fractionation of the complexes at the peptide level using SCX consistently resulted in the identification of approximately 3-fold more proteins compared to separation at the protein level using SDS-PAGE. The increase in the number of identified proteins was especially pronounced for the Bmi-1 complex, where the target protein was expressed at a low level. The data shows that pre-fractionation of affinity isolated protein complexes using SCX prior to LC-MS/MS analysis significantly increases the number of identified proteins and individual protein coverage, particularly for target proteins expressed at low levels.

Keywords

strong cation exchange; immuno-precipitation; protein complex isolation; mass spectrometry; FLAG

Introduction

The present systems biology era is focused on addressing the need to comprehensively analyze cells, tissues, and organisms at multiple biomolecular levels. The datasets are then integrated to enhance our understanding of how alterations in individual components can affect the entire system.^{1,2} An integral part of any systems biology view is insight into the plethora of dynamic protein complexes that form under specific cellular conditions. Identifying changes in protein complex formation and assembling this information into

*Corresponding author: Thorkell Andresson, Laboratory of Proteomic sand Analytical Technologies, Advanced Technology Program, SAIC-Frederick, Inc., NCI-Frederick, Frederick, Maryland 21702., Tel: (301) 846- 7190. Fax: (301) 846-6037. andressont@mail.nih.gov.

[#]These authors contributed equally to this work.

biomolecular networks that enhance our understanding of biological systems under interrogation is critical if systems biology is to mature.

Several approaches have been employed to identify “interactomes” both at the individual protein level and on a global scale.³ Frequently used methods involve affinity purification of the protein complex followed by characterization using mass spectrometry (MS)-based shotgun sequencing.^{4,5} This method has proven highly successful in characterizing protein complexes from entire cell lysates as well as different subcellular compartments. This method, however, still results in relatively high rates of false positive and negative discoveries that adversely impacts post discovery validation of potentially interacting proteins. The wide dynamic range of protein concentrations in the cell and even within the same protein complex can lead to robust identification of high abundant components but under-sampling of proteins at the lowest concentration. Furthermore, significant differences in the binding affinities between proteins within the same complex can result in a disproportional loss of individual components during complex isolation. A recent study identifying proteins that bind to a variety of commonly used affinity reagents during complex isolation effectively illustrates another challenge in characterizing protein complexes; the inherent protein background that is inevitably part of every isolated protein complex.⁶ Hence, optimization of the different steps of the procedure is needed to maximize the number of proteins identified (and their individual coverage) within the complex.

While the ability of mass spectrometers to identify proteins both in terms of accuracy and throughput continues to improve, it is still necessary to separate protein complexes prior to tandem MS (MS/MS) analysis if maximum coverage is to be obtained. The benefit of pre-fractionation is that it minimizes potential ion suppression and the under sampling associated with data dependent data acquisition.^{7,8} Several methods have been used to separate protein complexes prior to MS/MS analysis. Two methods are used more readily than others. The first method involves one-dimensional sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of intact proteins followed by reversed-phase liquid chromatography (RPLC) of peptides resulting from in-gel proteolytic digestion of the protein bands cut from the gel. In the second method, the entire protein lysate is proteolytically digested into peptides, which are then separated by strong cation exchange (SCX) and RPLC.^{7,9,10} During protein complex isolation, affinity purification serves to reduce the complexity of the sample allowing for more robust identification of the proteins in the sample. However, even though the complexity of the sample is reduced it is still high enough to cause significant under sampling without any pre-fractionation prior to LC-MS/MS analysis.

In this study, we compare the ability to characterize affinity purified protein complexes fractionated using either SDS-PAGE or SCX in the first dimension prior to RPLC-MS/MS analysis. Immunoprecipitation was performed on two different nuclear proteins; Bmi-1, a member of the Polycomb protein complex 1 (PcG1) that is involved in transcriptional repression of specific genes through regulation of chromatin structure¹¹ and GATA3, a transcription factor involved in regulating cell differentiation.¹² To compare the performance of the separation strategies, the number of proteins identified within the Bmi-1 and GATA3 immunoprecipitates was evaluated. The data demonstrates that separation by SCX increases the number of identified proteins compared to separation by SDS PAGE. Furthermore, it suggests that the separation by SCX increases the ability of the mass spectrometer in identifying low abundant novel binding factors.

Materials and Methods

Materials

Dulbecco's modified Eagle's medium (DMEM), 0.25% trypsin, penicillin/streptomycin were obtained from Invitrogen (Carlsbad, CA). Ammonium bicarbonate (NH_4HCO_3) was obtained from Sigma (St. Louis, MO). Sequencing grade trypsin was obtained from Promega (Madison, WI). Trifluoroacetic acid (TFA) and formic acid (FA) were purchased from Fluka (Milwaukee, WI). HPLC grade acetonitrile (ACN) and methanol (CH_3OH) were obtained from EM Science (Darmstadt, Germany). Affinity purification reagents, M2 FLAG Beads and 3x FLAG peptide were purchased from Sigma (St. Louis, MO).

Generation of 3x-FLAG Fusion Proteins

Human Bmi-1 and GATA3 were expressed in human embryonic kidney 293T cells as a fusion protein with a 3x-FLAG peptide. Bmi-1 and GATA3 were cloned into Gateway Entry clone pDonr253 (Invitrogen, Carlsbad, CA) by PCR using either a gene optimized Bmi-1 template (Geneart) or a pcDNA3-GATA3 template. The Bmi-1/GATA3 ORFs were subcloned by Gateway LR recombination using the manufacturer's protocols (Invitrogen, Carlsbad, CA) into pDest-737, a vector containing an amino-terminal 3xFLAG tag and a CMV promoter. Expression clones were verified using agarose gel electrophoresis and restriction digest. Transfection-ready DNA for the final clones was prepared using the GenElute XP Maxiprep kit (Sigma, St. Louis, MO).

Cell Culture and Transfection

HEK293T cells were cultured in DMEM supplemented with 10% fetal bovine serum, 100 $\mu\text{g}/\text{ml}$ penicillin and 100 $\mu\text{g}/\text{ml}$ streptomycin. The cultures were kept in a 5% $\text{CO}_2/95\%$ air humidified incubator maintained at 37°C . Prior to transfection 293T cells were grown to ~60% confluency and then transfected with a complex of polymer PEI (Polysciences Inc, Warrington, PA) and DNA at a ratio of 5:2. The cells were grown for a further 48 hrs post-transfection.

Affinity Purification of Protein Complexes

HEK293T cells expressing the protein of interest were washed twice with cold PBS. Cells were harvested on ice in the presence of 1 mL/100 mm dish of lysis buffer (50 mM Tris, pH 7.4, 150 mM NaCl, 0.5 mM EDTA, 0.1% NP-40, and protease inhibitor cocktail 100 $\times$). Cells were scraped and lysed through one freeze thaw. Lysates were centrifuged at 14 000 rpm for 10 minutes and the supernatant was incubated with 40 μL of anti-FLAG M2 agarose (Sigma, St. Louis, MO) for 2 hrs. The resin was washed three times with wash buffer (50 mM Tris, pH 7.4, 150 mM NaCl) and centrifuged at 4000 rpm for 30 sec. Following the final wash the bait protein interactomes were eluted via a 3x-FLAG (N-Met-Asp-Tyr-Lys-Asp-His-Asp-Gly-Asp-Tyr-Lys-Asp-His-Asp-Ile-Asp-Tyr-Lys-Asp-Asp-Asp-Lys-C) peptide challenge in 100 μL of elution buffer (6.25mM NH_4HCO_3 , pH 8.4, 150 $\mu\text{g}/\text{mL}$ 3x-FLAG peptide) for 20 min. After two rounds of elution, the samples were pooled, aliquoted based upon the experimental design, and lyophilized.

Electrophoresis and In-Gel Tryptic Digestion

Samples were incubated at room temperature for 15 min in 2x LDS-Reducing agent (Invitrogen, Carlsbad, CA) containing reducing agent. Protein samples were resolved on 4–12% Bis-Tris gels and stained using SimplyBlue Safe Stain (Invitrogen, Carlsbad, CA) as per manufacturer instructions. Each lane of the gel was cut into 12 bands that were placed in individual 1.5 ml eppendorf tubes. Each gel band was then further cut into small pieces and destained using 50% ACN/25 mM NH_4HCO_3 , pH 8.4. After removal of the organic solvent,

gel pieces were dried by vacuum centrifugation for 1 hr. Trypsin (15 ng/ μ L) in 25 mM NH_4HCO_3 , pH 8.4, was added to each sample (50 μ L) and incubated on ice for 30 min. Excess trypsin solution was removed, and replaced with 30 μ L of 25 mM NH_4HCO_3 , pH 8.4. The resulting samples were incubated overnight at 37°C, and the mixture of tryptic peptides extracted with 70% (v/v) ACN/5% FA ($2 \times 150 \mu$ L). The combined extracts were lyophilized, reconstituted in 20 μ L of 0.1% TFA and desalted using C_{18} ZipTips (Millipore Corporation, Bedford, MA). The C_{18} ZipTips were conditioned in 20 μ L of 50% CAN, followed by 3x equilibration in 20 μ L of 0.1% TFA. The binding of the peptides to the ZipTips was performed by pipetting the sample 10 times over the bed followed by 3x washing in 20 μ L of 0.1% TFA. The peptides were eluted of the resin with 10 μ L of 80% ACN. The eluate was lyophilized and reconstituted in 0.1% TFA (12 μ L) and aliquots (5 μ L) analyzed by nanoRPLC-MS/MS.

In-Solution Tryptic Digestion

The lyophilized eluate obtained from the anti-FLAG M2 affinity purification was resuspended in 25 μ L of 25 mM NH_4HCO_3 , pH 8.4. The samples were subjected to overnight digestion with trypsin (1 μ g) at 37°C followed by lyophilization. The tryptic digest was reconstituted in 25% ACN/0.1% FA (100 μ L) and fractionated using strong cation exchange (SCX) liquid chromatography (LC).

Strong Cation Exchange Liquid Chromatography

Separation of the tryptic peptide fragments was performed using a SCX/LC column (5 μ m, 200Å, 1×50 mm, polysulfoethyl A; PolyLC, Columbia, MD). The solvent system was composed of mobile phase A, 25% ACN and mobile phase B, 25% ACN/0.5M ammonium formate (pH 3). Elution of the peptides was performed by following a multistep gradient: 0–2 min/3%B, 27 min/30%B, 28–45 min/100%B at a flow rate of 50 μ L/min. Forty-five fractions were collected over a 45 minute time interval using a fraction collector (Isco, Lincoln, NE). The first 36 SCX fractions were pooled into 12 fractions based on the peptide chromatography profile, lyophilized, reconstituted in 0.1% TFA (12 μ L) and aliquots (5 μ L) analyzed using nRPLC-MS/MS.

Nanoflow Reversed Phase Liquid Chromatography (nanoRPLC) - Tandem Mass Spectrometry

Nanoflow RPLC-MS/MS analyses were performed using an Agilent 1100 nanoflow LC system coupled online with a linear ion trap (LIT) mass spectrometer (LTQ from ThermoElectron, San Jose, CA). Microcapillary RPLC column (75 μ m i.d. $\times$ 10 cm fused silica capillary with a flame pulled tip) was in-house slurry-packed with 5 μ m, 300 Å pore size, Jupiter C-18 stationary phase (Phenomenex, Torrance, CA). After sample injection, the column was washed for 30 min with 98% mobile phase A (0.1% FA) at 0.5 μ L/min and peptides were eluted using a linear gradient of 2% mobile phase B (0.1% FA in ACN) to 42% B in 40 minutes at 0.25 μ L/min, then to 98% B for an additional 10 min. The LIT-MS was operated in a data dependent mode in which each full MS scan was followed by seven MS/MS scans wherein the seven most abundant molecular ions were dynamically selected for collision-induced dissociation using normalized collision energy of 35%.

Data Analysis and Interpretation

Acquired MS/MS spectra were searched against a human protein database (March, 2009) using the SEQUEST algorithm implemented in BioWorks 3.2 package (ThermoElectron, San Jose, CA) using 1.5 Da for precursor ion tolerance and 0.5 Da for fragment ions tolerance with methionine oxidation included as dynamic modification. Only fully tryptic peptides with up to two miscleavages with charge state dependent cross correlation X_{corr}

2.1 for $[M+H]^1+$, 2.5 for $[M+2H]^2+$ and 3.2 for $[M+3H]^3+$ and delta correlation (ΔC_n) 0.08 were considered as positive identification with a 1% false discovery rate.

Results

The Pre-fractionation Method Significantly Affects the Identification of Immunoprecipitated Proteins

Regulated formation of protein complexes is a fundamental process in all cells. A clear understanding of how protein complexes function is necessary to fully understand basic processes such as cell division and the underlying mechanisms of diseases. Immunoprecipitation of a target protein followed by sample fractionation and LC-MS/MS analysis is frequently used to identify protein-protein interactions. A major challenge is the limited protein concentration available after immunoprecipitation. This limitation makes achieving the goal of identifying low abundant proteins a challenge and requires careful optimization of each experimental step.

One critical step in analyzing isolated protein complexes is fractionation prior to MS analysis. The predominant method is separation of intact proteins using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS PAGE). The separated proteins are then in-gel digested (usually with trypsin) and peptides extracted from the gel. The recovered peptides are desalted using C₁₈ ZipTips followed by RPLC-MS/MS analysis. Another fractionation method involves tryptic digestion of the protein complex followed by peptide separation using SCX-LC followed by RPLC-MS/MS analysis. In this study, immunoprecipitated protein complexes were characterized using both SDS-PAGE and SCX fractionation. The immunoprecipitation was performed targeting Bmi-1, a member of the Polycomb protein complex 1 (PcG1) that is involved in transcriptional regulation of specific genes through regulation of chromatin modification. The core PcG1 complex is composed of several known proteins including Ring1A, Ring1B, Cbx2, and Phc1 in addition to Bmi-1.¹¹ These known interacting proteins serve as quality controls allowing the efficiency of the immunoprecipitation analysis to be monitored.

A schematic outline of the procedures is shown in Figure 1A. To mimic endogenous cell conditions, the expression of N-terminal FLAG tagged Bmi-1 in 293T cells was adjusted to a level in which it was barely detectable using Coomassie blue staining (Figure 1 B, lane 3). To keep the comparison as valid, a single master immunoprecipitation was performed and equivalent amounts of eluate were separated using SDS-PAGE or SCX (post-tryptic digestion) (Figure 1A). The gel lane was cut into 12 slices followed by in-gel tryptic digestion of the proteins. Thirty-six SCX fractions were collected off-line and pooled into 12 aliquots. All gel and SCX peptide fractions were analyzed using a Thermo LTQ iontrap mass spectrometer. Identical RPLC conditions were used to separate the peptides prior to MS/MS analysis and the raw data was processed using identical search parameters.

The MS/MS data was evaluated using three different criteria; total number of peptides identified, proteins identified, and proteins identified by 2 peptides (Figure 2A). As shown in figure 2A, SCX identified 861 unique peptides corresponding to 377 proteins with 1 peptide compared to 368 peptides corresponding to 210 proteins identified using SDS-PAGE fractionation (Gel 1). Using 2 peptide as a criterion for protein identification, 168 proteins were identified within the SCX fractions compared to 52 proteins identified using SDS-PAGE fractionation (Gel1). This represents a 3.2 fold increase in the number of proteins identified by 2 peptides using SCX fractionation (Figure 2A). Since Bmi-1 was expressed at a low level, the amount of immunoprecipitate loaded onto the SDS-PAGE gel was doubled (Figure 1A, Gel 2). Doubling the amount of material resulted in a 1.8 fold increase in the number of proteins identified by 2 peptides (93 vs. 52), however, gel

fractionation still resulted in significantly fewer protein identifications compared to the SCX (168 proteins) experiment (Figure 2A).

To confirm the observation made with Bmi-1, an immunoprecipitation experiment was performed using GATA3 as the target protein (Figure 1A). GATA-3 belongs to a family of transcription factors involved in regulating cell differentiation in a variety of tissues.¹² Since GATA3 is expressed at significantly higher levels than Bmi-1 (Figure 1B, lane 2) it allowed us to test the method's reproducibility and the affect of higher target protein levels. Fractionation of the isolated GATA3 complex using SCX (973 proteins/2583 peptides) identified more proteins compared to SDS-PAGE (521 proteins/1275 peptides) (Figure 2B). Using a 2 peptide threshold, 552 proteins were identified using SCX fractionation compared to 252 from the SDS-PAGE gel. Doubling the amount of material loaded on the gel (as done for Bmi-1) did not significantly change the number of proteins identified by 2 peptide (Fig 2B). Beyond substantiating the greater number of proteins identified using SCX fractionation, these results show that for more abundant proteins the SDS-PAGE gel can become saturated (even with immunoprecipitated material from 1×10^7 cells) and no benefit is gained by increasing the amount of material loaded.

Control Subtraction Leads to Significantly More Specific Interactions Identified using SCX Fractionation Compared to SDS-PAGE

A constant issue regarding the analysis of affinity isolated protein complexes is the presence of nonspecifically interacting proteins. These proteins are believed to interact with either protein A/G coded sepharose beads or the antibody directed to the targeted protein. The most common way of processing the MS/MS data is to compile a list of background proteins found in control experiments and subtract these identifications from proteins found in the experimental sample. The remaining proteins are generally considered potential candidates that specifically bind to the protein of interest. To test if SCX generates a more comprehensive list of "background" proteins compared to SDS-PAGE after subtractive analysis, two control samples (transfected with empty vector) were analyzed in an identical manner used for the experimental cells (i.e. transfected with Bmi-1 and GATA-3 constructs) (Figure 1A). As expected, SCX fractionation led to more protein identifications than SDS-PAGE (data not shown).

Originally a separation method-specific subtraction was performed, however analysis of the gel separation data revealed that several proteins assigned as a Bmi-1 and GATA3 specific interacting proteins were also identified in the SCX control. Therefore, even though these proteins are not subtracted using the gel control they should be considered as non-specific binders and omitted from the list of potential interacting proteins. To circumvent this result a master list of all background proteins identified using both methods was compiled and subtracted from the proteins identified in the analysis of the affinity purified Bmi-1 and GATA3 complexes. When the subtractive analysis is performed, the advantages of the SCX fractionation method over SDS-PAGE became even more evident. Using a 2 peptide threshold for protein identification of the Bmi-1 complex, only 10 proteins remained on the list of those identified using gel fractionation (starting from 1×10^7 cells; Gel1), while 85 proteins remained on the list generated using SCX fractionation (Figure 3A). As doubling the amount of material loaded onto the SDS-PAGE gel increased the number of proteins identified, we compared the number of specific proteins isolated from the gel and SCX. Not surprisingly, even with increasing the amount of material loaded on the gel (Gel 2), 2.8 fold more proteins were identified using SCX fractionation (30 vs. 85 proteins for SDS-PAGE and SCX, respectively). The same trend was observed in the analysis of the GATA3 protein complex (Figure 3B), as 3.3 fold more proteins were identified using the SCX fractionation versus SDS-PAGE (Figure 3B). In addition, no benefit was gained when the amount of GATA3 complex loaded onto the gel was doubled.

The overlap in proteins identified by 2 peptides for the Bmi-1 and GATA3 complexes using the two different pre-fractionation methods is shown in Figures 3C–F. SCX fractionation resulted in the identification of 84–95% of the total number of proteins in the individual immunoprecipitation experiments. In fact, 70–89% of the proteins were identified only using SCX fractionation, while only 5–16% were unique to the SDS-PAGE gel datasets. Since both GATA3 and Bmi-1 are nuclear proteins, any non-nuclear proteins were assumed to be interacting non-specifically and were excluded from the list. The list of identified proteins (2 peptides) was filtered using gene ontology (GO), selecting proteins under the categories of nuclear localization and nucleic acid interaction (Terms: GO:0003676 and GO:0005634). This list was reanalyzed comparing the overlap between the SCX and SDS-PAGE methods (Figure 4). This filtering significantly reduced the number of total proteins, especially for Bmi-1, which is expressed at a relatively low level. Only three nuclear proteins were identified from 1×10^7 cells using the SDS-PAGE approach compared to 25 using SCX (Figure 4A) with Bmi-1 being the only common protein between the two datasets. Some benefit was gained from loading material from 2×10^7 cells on the gel, as the number of identified nuclear proteins increased to 13 with 7 of these overlapping with the SCX dataset (Figure 4B). A similar trend was observed for GATA3, as 76 proteins were identified using SDS-PAGE fractionation from both 1×10^7 and 2×10^7 cells (Figure 4C and D). Doubling the amount of material loaded onto the SDS-PAGE gel did not increase the number of proteins identified, nor did it have a significant effect on the overlap with the proteins identified using SCX fractionation (Figure 4C and D).

Common protein contaminants can also significantly affect the amount of information obtained from affinity purification experiments. These contaminants were especially evident in the datasets acquired via SDS-PAGE pre-fractionation. For example, keratin proteins showed a 30-fold higher abundance level in the gel pre-fractionated data (Table I). Several other background proteins such as hornerin (Table I) were also in much higher abundance in the gel compared to the SCX fractions. Highly abundant contaminating proteins can suppress ionization of peptides originating from specifically bound proteins or prevent their selection for MS/MS analysis in a data-dependent experiment.

Pre-fractionation by SCX Identifies More Known Bmi-1 Interacting Proteins than SDS-PAGE

Bmi-1 is known to interact with several proteins involved in chromatin remodeling. Using SCX to pre-fractionate the Bmi-1 protein complex resulted in the identification of 4 proteins previously shown to be associated with Bmi-1. Two of these proteins, the E3 ubiquitin-protein ligases RING1 and RING2, were each identified by 10 peptides. In addition, YY1-binding protein (RYBP) and YY1-associated factor 2 (YAF2) were identified by 2 and 3 peptides, respectively (Table II). RING1 was identified using SDS-PAGE fractionation, by two peptides from Gel2 (2×10^7 cells) and a single peptide of RING2 was identified from Gel1 (1×10^7 cells). The same trend was seen for Bmi-1 itself, which was identified by 4, 67 and 119 peptides in Gel1, Gel2 and SCX, respectively (Table II). The number of unique Bmi-1 peptides identified also increased from 3 in Gel1 to 7 using SCX.

Bmi-1 has been associated with a number of cancers including those of the breast. To compare the two fractionation methods for identifying Bmi-1 interacting proteins under a more physiological setting, a stable human mammary epithelial cell line (MCF-7) transfected with a 3x FLAG-tagged Bmi-1 construct was generated. As shown in lane 2 of Figure 5A, exogenous Bmi-1 is expressed at a similar level as the endogenous protein. A master immunoprecipitate was isolated from 2×10^8 cells using a control cell line and the stable 3x FLAG Bmi-1 MCF-7 cell line. A master immunoprecipitate isolated from 2×10^8 cells was split into equal aliquots; one fractionated into 12 gel slices using SDS-PAGE and the other into 12 fractions using SCX. Consistent with the previous experiments, the LC-

MS/MS data showed that significantly more information is obtained using SCX pre-fractionation. Following control subtraction, seven proteins were identified by 2 more peptides using SDS-PAGE versus 34 using SCX. This result agrees with the analysis of the protein complexes generated from HEK293 cells, with about 90% of the potential interacting proteins identified using SCX fractionation (Figure 5B). Ten known Bmi-1 interacting proteins were identified by 2 peptides in the SCX pre-fractionate sample. Only 2 proteins were identified using the same criteria in the data set obtained using SDS-PAGE pre-fractionation, with two additional known interactors identified by a single peptide (Table III).

Discussion

The strengths of the interactions between proteins within a multi-component complex can vary considerably. There are several examples of protein complexes containing a nucleus of proteins that are tightly bound with additional components that transiently associate with the core, depending on the physiological status of the cells.^{13–16} These transiently bound proteins are of interest since they can represent factors that regulate the complex's molecular function. The transient nature of interactions associated with very low abundance and binding affinity presents a challenge in isolating and identifying these proteins. Capturing these proteins requires careful experimental optimization and in some cases the study must be tailored to a specific physiological state to maximize the chances of identifying these components.

In this study, two commonly and widely used separation methods for characterizing various protein complexes were compared. While several studies have compared different separation methods;^{17–19} these typically focus on global proteomics using relatively high concentrations of starting material. No empirical studies comparing the affect of different separation schemes on protein identification following protein complex isolation, where the concentration of the starting material is typically in the range of 100 ng to 1 µg, have been reported. The principle behind pre-fractionation is to reduce the complexity of the samples prior to MS/MS analysis resulting in an increase in the information obtained. Even though immunoprecipitation significantly reduces the complexity of the sample, the number of non-specifically bound and background proteins still saturates the capacity of the mass spectrometer if samples are directly analyzed by LC-MS/MS.

The preferred way of conducting protein complex isolation is to target the endogenous protein within a physiologically relevant cell line. However, the lack of quality antibodies, low endogenous expression of the target protein, and the co-extraction of background contaminants can make this strategy extremely challenging. Therefore the most common method is to exogenously express the target protein in either heterologous cells, such as HEK293T cells, or by generating stable cell lines. The downside to this approach is that the heterologous protein is often over-expressed compared to the endogenous protein leading to co-extraction of non-specifically bound proteins. Keeping this observation in mind, the expression of Bmi-1 was adjusted to mimic conditions in which the abundance of the target protein is the limiting factor. In addition, the second gene, GATA3, was selected because of its relatively high expression in HEK293T cells. We also generated a cell line stably expressing an epitope tagged version of Bmi-1 at a level similar to the endogenous protein expression. By selecting these different conditions, the overall performance of the SDS-PAGE and SCX fractionations could be effectively evaluated.

The data clearly demonstrate that pre-fractionation at the peptide level using SCX results in considerably more protein identifications using RPLC-MS/MS compared to fractionation at the protein level using SDS-PAGE. In general, 2.5–3-fold more proteins were identified

using SCX, however, after bioinformatically filtering the data to only include nuclear proteins, the difference increased to ~7-fold. This difference becomes even more striking when the recovery of known Bmi-1 interacting proteins is considered¹¹. Analysis of the Bmi-1 complex from HEK293 cells by SDS-PAGE identified only RING2 (1 peptide in Gel 1) and RING1 (2 peptides in Gel 2). In contrast RING1 (10 peptides), RING2 (10 peptides), RYBP (2 peptides) and YAF2 (3 peptides) were readily identified when the complex was fractionated by SCX. When the immunoprecipitation was performed using a lysate from MCF-7 cells stably expressing Bmi-1, only four known interacting proteins were identified by 15, 6, 1, and 1 peptides using SDS-PAGE fractionation. Using SCX fractionation resulted in the identification of 10 known Bmi-1 interacting proteins, identified by anywhere from 3 to 19 peptides. Figure 6 illustrates the difference in the known Bmi-1 interacting proteins identified following fractionation by the SDS-PAGE or SCX. The figure visually shows the difference in the performance between the two methods and drives home our main conclusion that SCX is advantageous in identifying the low abundant components of the protein complex. The data also demonstrates that for low abundant interacting partners it is important to maximize the number of proteins identified even though the true interacting proteins might be buried in a long list of no-specific background proteins causing high false discovery rate. Under these circumstances it is invaluable to have a robust bioinformatics' system that can help with prioritizing those proteins that will be further validated using methods such as co-immunoprecipitation and co-localization.

Currently SDS-PAGE is the most frequently used method to reduce the complexity of immunoprecipitated samples prior to RPLC-MS/MS analysis. Usually the control and experimental samples are fractionated in adjacent lanes and protein bands exclusively seen in the experimental lane are excised and analyzed along with the corresponding area within the control lane. An alternative strategy involves cutting the entire lanes into defined segments and analyzing all of the fractionated proteins. Examining the overall SDS-PAGE process reveals possible reasons for the identification of a substantially greater number of peptides/proteins using SCX fractionation. Processing gels requires significant sample handling increasing the chances of introducing contaminating proteins. Examination of Table I shows that the number of keratin-related peptides identified within the SDS-PAGE fractions was between 8 and 26 times the number identified within the SCX fractions. This large proportion of contaminating proteins could prevent the MS/MS selection and identification of proteins that selectively bind to Bmi-1 and GATA3. In addition, some peptides are not efficiently extracted from the gel following tryptic digestion. However, analysis of all peptides identified following affinity purification of GATA3 using the two separation methods did not reveal any difference in their overall physical properties such as hydrophilicity, hydrophobicity, charge or size distribution (data not shown). The greater individual protein coverage found using SCX suggests that proteins specific to the complexes comprised a greater proportion of the total sample analyzed using RPLC-MS/MS using this pre-fractionation method compared to SDS-PAGE. Therefore the chances of a peptide from a complex-specific proteins being selected for MS/MS are greater and its chance of suffering ion suppression are lower.

In this study a commonly utilized experimental approach for isolating protein complexes was used to target Bmi-1 and GATA3 and the separation methods were challenged at two different protein concentrations. Regardless of the expression level, the overall protein complex coverage showed the same trend with SCX fractionating resulting in greater coverage. These results suggest that this observation applies to the purification of most protein complexes identified by LC-MS/MS, with the highest benefit seen for low expressing proteins.

Acknowledgments

This project has been funded in whole or in part with federal funds from the National Cancer Institute, National Institutes of Health, under Contract HHSN261200800001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the United States Government.

Reference

1. Bader S, Kuhner S, Gavin AC. Interaction networks for systems biology. *FEBS Lett.* 2008; 582:1220–1224. [PubMed: 18282471]
2. Russell RB, Aloy P. Targeting and tinkering with interaction networks. *Nat. Chem. Biol.* 2008; 4:666–673. [PubMed: 18936751]
3. Shoemaker BA, Panchenko AR. Deciphering protein-protein interactions. Part I. Experimental techniques and databases. *PLoS Comput. Biol.* 2007; 3:e42. [PubMed: 17397251]
4. Gingras AC, et al. Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell. Biol.* 2007; 8:645–654. [PubMed: 17593931]
5. Collins MO, Choudhary JS. Mapping multiprotein complexes by affinity purification and mass spectrometry. *Curr. Opin. Biotechnol.* 2008; 19:324–330. [PubMed: 18598764]
6. Trinkle-Mulcahy L, et al. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J. Cell. Biol.* 2008; 183:223–239. [PubMed: 18936248]
7. Motoyama A, Yates JR 3rd. 3rd, Multidimensional LC separations in shotgun proteomics. *Anal. Chem.* 2008; 80:7187–7193. [PubMed: 18826178]
8. Yates JR 3rd, Ruse CI, Nakorchevsky A. Proteomics by mass spectrometry: approaches, advances, and applications. *Annu. Rev. Biomed. Eng.* 2009; 11:49–79. [PubMed: 19400705]
9. Ishihama Y, Rappsilber J, Mann M. Modular stop and go extraction tips with stacked disks for parallel and multidimensional Peptide fractionation in proteomics. *J. Proteome Res.* 2006; 5:988–994. [PubMed: 16602707]
10. Washburn MP, Wolters D, Yates JR 3rd. Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 2001; 19:242–247. [PubMed: 11231557]
11. Simon JA, Kingston KE. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat. Rev. Mol. Cell. Biol.* 2009; 10:697–708. [PubMed: 19738629]
12. Patient RK, McGhee JD. The GATA family (vertebrates and invertebrates). *Curr. Opin. Genet. Dev.* 2002; 12:416–422. [PubMed: 12100886]
13. Mousson F, et al. Quantitative proteomics reveals regulation of dynamic components within TATA-binding protein (TBP) transcription complexes. *Mol. Cell. Proteomics.* 2008; 7:845–852. [PubMed: 18087068]
14. Wang X, Huang L. Identifying dynamic interactors of protein complexes by quantitative mass spectrometry. *Mol. Cell. Proteomics.* 2008; 7:46–57. [PubMed: 17934176]
15. Scott JD, Pawson T. Cell signaling in space and time: where proteins come together and when they're apart. *Science.* 2009; 326:1220–1224. [PubMed: 19965465]
16. Sardiù ME, et al. Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. *Proc. Natl. Acad. Sci. U.S.A.* 2008; 105:1454–1459. [PubMed: 18218781]
17. Fang Y, Robinson DP, Foster LJ. Quantitative analysis of proteome coverage and recovery rates for upstream fractionation methods in proteomics. *J. Proteome Res.* 2010; 9:1902–1912. [PubMed: 20078137]
18. Hubner NC, Ren S, Mann M. Peptide separation with immobilized pI strips is an attractive alternative to in-gel protein digestion for proteome analysis. *Proteomics.* 2008; 8:4862–4872. [PubMed: 19003865]
19. Slebos RJ, et al. Evaluation of strong cation exchange versus isoelectric focusing of peptides for multidimensional liquid chromatography-tandem mass spectrometry. *J. Proteome Res.* 2008; 7:5286–5294. [PubMed: 18939861]

Figure 1. Schematic illustrating comparison of SDS-PAGE and SCX fractionation for the analysis of immunoprecipitated protein complexes

(A) Immunoprecipitation was performed from HEK293T cells transfected with 3x-FLAG-Bmi-1 or 3x-FLAG-GATA3. Cells transfected with the vector alone was used as a negative control. Half of the eluate was separated on a 4–12% SDS-PAGE gel [Gel 2] and the proteins were stained with Coomassie Blue. The second half of the eluate was further divided into two halves; one half being loaded onto a 4–12% SDS-PAGE gel [Gel 1] and other digested with trypsin followed by separation using strong cation exchange [SCX] into 12 aliquots. Each lane of the SDS-PAGE gel was cut into 12 pieces, which were subjected to in-gel digestion and peptide extraction. The peptides present within the 12 SDS-PAGE bands and 12 SCX fractions were analyzed using identical RPLC-MS/MS conditions. (B) Eluate from the immuno-precipitate was separated onto a 4–12% SDS-PAGE gel and stained with a coomassie blue. The lane was cut into 12 bands and in-situ trypsinization and extraction of the peptides was performed. The peptides were analyzed using RPLC-MS/MS. Arrow heads show bands corresponding to the transfected proteins.

Figure 2. Quantitative comparison of peptides and proteins identified using SDS-PAGE and strong cation exchange fractionation

The proteins obtained after immunopurification were fractionated using SDS-PAGE, followed by in-gel digestion and peptide extraction or tryptically digested fractionated using SCX. A total of 12 fractions were collected and analyzed by RPLC-MS/MS. Identifications from the all the 12 fractions were combined and represented in terms of the total of peptides, total number of proteins and number of proteins identified by ≥ 2 peptides. (A) Bmi-1. (B) GATA3.

Figure 3. Comparison of the Bmi-1 and GATA3 specific interacting proteins

Bmi-1 (A) and GATA3 (B) specifically interacting proteins were obtained by subtraction of the non-specific proteins identified from the affinity purification of vector transfected HEK293T cells. Bar graphs indicate the total number of specific proteins and also proteins identified with ≥ 2 peptides within the different experiments. For Gel 1 and SCX identical amounts of immunoprecipitate were fractionated, while twice as much protein was fractionated on Gel 2. (C–F) Venn diagrams showing the percentage of the overlap of the Bmi-1 (C–D) and GATA3 (E–F) specifically interacting proteins identified with ≥ 2 peptides obtained using SCX and SDS-PAGE fractionation. Extent of overlap between the lists of

proteins when equal amount of immunoprecipitate was used for fractionation (**C & E**).
Extent of overlap between the lists of proteins when double the amount of immuno-
precipitate compared to SCX was used for SDS-PAGE (**D & F**).

Figure 4. Overlap of the nuclear proteins identified from the fractionation performed by SDS-PAGE and strong cation exchange

Exclusive list of nuclear proteins were generated by using GeneOntology (GO) by selecting proteins under the categories of nuclear localization and nucleic acid interaction (Terms: GO:0003676 and GO:0005634). Venn diagrams depicts the percentage of the overlap of the Bmi-1 (A–B) and GATA3 (C–D) nuclear specific interacting proteins identified with 2 peptides obtained from strong cation exchange and SDS-PAGE. Extent of overlap between the lists of proteins when equal amount of immuno-precipitate was used for fractionation (A & C). Extent of overlap between the lists of proteins when double the amount of immuno-precipitate compared to SCX was used for SDS-PAGE (B & D).

Figure 5. Comparison of the proteins identified from affinity purified 3x-FLAG Bmi-1 from MCF7 cells using SDS/PAGE and strong cation exchange

(A) Western analysis of 3x-FLAG-Bmi-1 expression in stable cell lines. Lane 1, MCF7 vector control and lane 2, MCF7 3xFLAG Bmi-1 were analyzed by western blotting using mouse anti-Bmi-1 antibody. A protein band of ~45 kDa was observed for 3x-FLAG-Bmi-1 transfected cells (lane 2). A secondary band ~35 kDa was observed representing the native Bmi-1 protein in lanes MCF7 cells. (B) Overlap of proteins identified from the fractionation of the eluate using SDS/PAGE and SCX. All proteins identified with 2 peptides were used for the analysis.

Figure 6. Model showing the disparity of the complex isolation for the same protein under different fractionation method

Affinity purification of Bmi-1 was performed from MCF7 cells stably expressing 3xFLAG Bmi-1. Fractionation of the immuno-precipitate using SDS-PAGE led to the detection of only 4 known interacting proteins, while the fractionation of the same material using SCX detected 10 known interactors. The interacting proteins identified by SDS-PAGE are colored in yellow, while the interactors detected specifically by SCX are shown in green. Bmi-1, the bait protein is shown in red. Pentagons depict the transcription factors, while other proteins are shown in oval shapes. Direct interactions between the proteins are shown by a straight line.

Common protein contaminants are observed in higher frequency in SDS-PAGE gel bands compared to SCX fractions.

Table 1

Protein name	Uniprot	Number of peptides					
		Control gel	Control SCX	Bmi-1 gel	Bmi-1 SCX	GATA3 gel	GATA3 SCX
Keratin, type II cytoskeletal 1	P04264	353	11	255	31	305	17
Keratin, type II cytoskeletal 2 epidermal	P35908	262	5	144	18	220	9
Keratin, type I cytoskeletal 10	P13645	252	13	178	23	191	14
Keratin, type I cytoskeletal 9	P35527	202	5	131	14	164	7
Keratin, type II cytoskeletal 5	P13647	81	7	33	6	34	1
Keratin, type I cytoskeletal 14	P02533	62	5	21	3	18	0
Hornerin	Q86YZ3	43	0	13	0	18	0

Table 2

Comparison of known Bmi-1 interacting proteins identified by affinity purification in HEK293T cells by SCX and SDS-PAGE.

Protein name	Uniprot	<u>Number of peptides</u>		
		SCX	Gel 1	Gel 2
Bmi-1	P35226	119	4	67
RING1	Q06587	10	0	2
RING2	Q99496	10	1	0
RYBP	Q8N488	2	0	0
YAF2	Q8IY57	3	0	0

SCX: Affinity purification of Bmi-1 from 1×10^7 HEK293T cells transfected with 3x-FLAG Bmi-1 was performed and the eluate trypsinized, fractionated by strong cation exchange and run on LC-MS/MS.

Gel 1: Affinity purification of Bmi-1 from 1×10^7 HEK293T cells transfected with 3x-FLAG Bmi-1 was performed and the eluate was ran on a 4–12% SDS-PAGE gel. The lanes were cut, trypsinized and run on LC-MS/MS.

Gel 2: Affinity purification of Bmi-1 from 2×10^7 HEK293T cells transfected with 3x-FLAG Bmi-1 was performed and the eluate was ran on a 4–12% SDS-PAGE gel. The lanes were cut, trypsinized and run on LC-MS/MS.

Table 3

Comparison of known Bmi-1 interacting proteins identified by affinity purification in MCF7 cells by SCX and SDS-PAGE.

Protein name	Uniprot	<u>Number of peptides</u>	
		SDS-PAGE	SCX
Bmi-1	P35226	5	16
RING1	Q06587	15	19
RING2	Q99496	0	9
RYBP	Q8N488	1	8
YAF2	Q8IY57	1	4
Cbx 2	Q14781	0	5
Cbx 4	O00257	0	7
Cbx 8	Q9HC52	6	9
PHC 1	P78364	0	4
PHC 2	Q8IXK0	0	3
PHC 3	Q8NDX5	0	9

SDS-PAGE: Affinity purification of Bmi-1 from 1×10^8 MCF7 cells stably transfected with 3x-FLAG Bmi-1 was performed and the eluate was ran on a 4–12% SDS-PAGE gel. The lanes were cut, trypsinized and run on LC-MS/MS.

SCX: Affinity purification of Bmi-1 from 1×10^8 MCF7 cells stably transfected with 3x-FLAG Bmi-1 was performed and the eluate trypsinized, fractionated by strong cation exchange and run on LC-MS/MS.