Biochemistry

pubs.acs.org/biochemistry

From the Bench

Analysis of Tagged Proteins Using Tandem Affinity-Buffer Exchange Chromatography Online with Native Mass Spectrometry

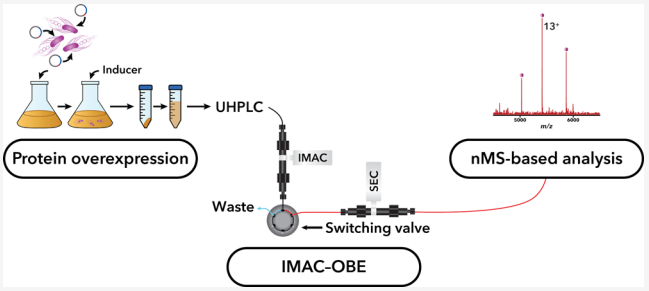
Florian Busch,[‖] Zachary L. VanAernum,[‖] Stella M. Lai,[‖] Venkat Gopalan, and Vicki H. Wysocki*

Cite This: *Biochemistry* 2021, 60, 1876−1884

Read Online

ACCESS | Metrics & More | Article Recommendations

ABSTRACT: Protein overexpression and purification are critical for *in vitro* structure−function characterization studies. However, some proteins are difficult to express in heterologous systems due to host-related (e.g., codon usage, translation rate) and/or protein-specific (e.g., toxicity, aggregation) challenges. Therefore, it is often necessary to test multiple overexpression and purification conditions to maximize the yield of functional protein, particularly for resource-heavy downstream applications (e.g., biocatalysts, tertiary structure determination, biotherapeutics). Here, we describe an automatable liquid chromatography−mass spectrometry-based method for direct analysis of target proteins in cell lysates. This approach is facilitated by coupling immobilized metal affinity chromatography (IMAC), which leverages engineered poly-histidine tags in proteins of interest, with size exclusion-based online buffer exchange (OBE) and native mass spectrometry (nMS). While we illustrate a proof of concept here using relatively straightforward examples, the use of IMAC−OBE−nMS to optimize conditions for large-scale protein production may become invaluable for expediting structural biology and biotherapeutic initiatives.



Protein purification is a critical prerequisite for biochemical and biophysical characterization of proteins and protein−ligand complexes *in vitro*. Therefore, well-characterized, cost-effective overexpression systems like *Escherichia coli* are commonly used to generate large amounts of recombinant protein. The method involves overexpression of the protein of interest in an appropriate host and subsequent chromatography-based purification to obtain the pure protein of interest (Figure 1A,B). Typically, the gene encoding the target protein is cloned into an expression vector that is, in turn, transformed into host cells. The target gene is usually placed under the control of an inducible promoter,[1−3] which allows for regulated overexpression of the protein of interest. A chemical inducer is typically added during the log phase of cell growth to induce target gene expression, and the culture is grown post-induction at a specific temperature for a defined time period to promote maximal protein synthesis. Recombinant protein production can usually be scaled up relatively easily and cost-effectively by simply increasing the cell culture volume. Following confirmation of successful overexpression, the target protein is isolated from host cell proteins using column-based chromatographic purification methods that exploit intrinsic properties of the protein of interest [e.g., isoelectric point (pI), hydrophobicity].[4,5] Alternatively, genetically engineered affinity tags can be fused to either terminus of the protein to enable affinity-based purification.[3]

Unfortunately, obtaining large amounts of soluble, correctly folded, and active recombinant protein is frequently not straightforward. Some of the major challenges that undermine expression of certain recombinant proteins arise from issues ranging from host cell translation (e.g., codon usage differences, translation rate) to target protein attributes (e.g., poor solubility, propensity to aggregate, and toxicity[6−8]). Additionally, specific incompatibilities between the protein of interest and the selected heterologous host system may culminate in misfolding or premature translation termination, generating truncated proteins. Therefore, different expression constructs are often evaluated in small-scale pilot tests to assess total overexpression and solubility. To this end, cells are harvested and analyzed by sodium dodecyl sulfate−polyacrylamide gel electrophoresis (SDS−PAGE).[9] The absence or presence of the target protein, as inferred from a band of the expected molecular weight, is then used to make a rapid and qualitative assessment of protein overexpression.

Received: February 24, 2021
Revised: April 25, 2021
Published: June 8, 2021

https://doi.org/10.1021/acs.biochem.1c00138
Biochemistry 2021, 60, 1876−1884

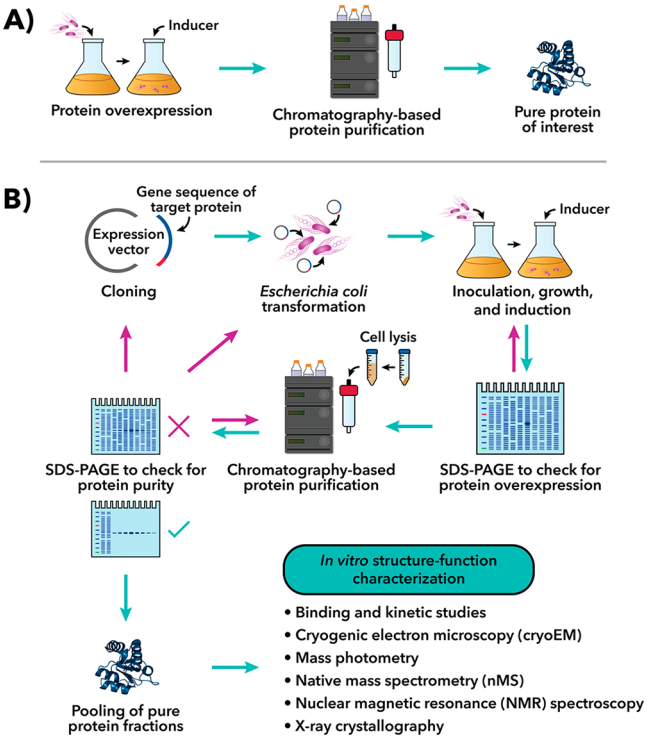


Figure 1. (A) General overview of recombinant protein preparation. (B) Workflow for overexpression and purification of recombinant proteins of interest. Teal arrows indicate the expected experimental workflow, while magenta arrows denote steps that may require troubleshooting and optimization (L7Ae structure: Protein Data Bank entry 3NVI[13]).

SDS−PAGE, which is inexpensive and routinely used, has been a cornerstone of modern biochemistry. However, there may be ambiguity when assessing the extent of protein overexpression. First, determining whether the target protein has been successfully overexpressed in its full-length form may be occasionally complicated by aberrant migration of proteins in SDS−PAGE, which is caused by compositional eccentricities (e.g., above-average presence of charged residues) of the protein of interest.[10,11] Second, the target protein may be obscured by co-migrating host proteins[12] and therefore scored as absent despite being present at low levels. Consequently, opportunities to isolate even modest amounts of the target protein may be lost. Thus, confirming the overexpression and solubility of a target protein by SDS−PAGE may not reliably forecast its subsequent successful purification.

The limitations of SDS−PAGE, coupled with potential challenges associated with heterologous gene expression, could necessitate laborious and time-consuming rounds of screening, troubleshooting, and optimization to maximize the yield of pure recombinant protein (Figure 1B). Therefore, methods that can expedite the process of identifying suitable overexpression and purification conditions are a necessary and ever-evolving part of the biochemist's toolbox. Advances in sample preparation (e.g., separation strategies) and mass analyzers have broadened the scope and utility of mass spectrometry (MS)-based methods for studying macromolecules and their noncovalent interactions under "native" conditions while also providing information about mass and subunit composition. One such method uses offline direct-infusion nanoelectrospray ionization (nESI) to rapidly characterize intact proteins in crude cell lysates to obtain information about their mass, solubility, oligomeric state, post-translational modifications, and binding to other biomolecules.[14−18]

In this study, we sought to expand upon the offline direct-infusion MS method by using liquid chromatography (LC)-driven separation strategies to remove endogenous host cell proteins, enrich for target species, and buffer exchange samples online immediately prior to MS analysis. Specifically, we coupled immobilized metal affinity chromatography (IMAC),[19] which leverages engineered poly-histidine tags (poly-His tags) in proteins of interest, with our recently developed method for performing online size exclusion-based buffer exchange with native MS (OBE−nMS)[20] to streamline assessment of cell lysates for successful protein overexpression and purification into a single workflow with minimal sample preparation. Here, we selected five proteins to illustrate the utility of the IMAC−OBE−nMS method for (i) selectively capturing and enriching for His_6-tagged target proteins, (ii) determining the oligomeric state of proteins of interest, and (iii) assessing the ability of solubility tags to prevent protein aggregation and precipitation.

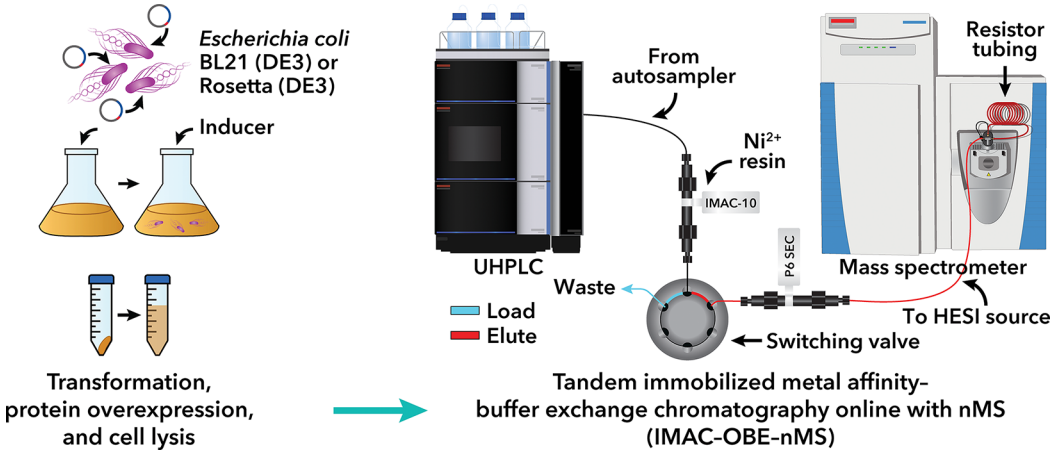


Figure 2. Recombinant protein overexpression and characterization using IMAC−OBE−nMS. The switching valve (gray) has two positions: "Load" (blue) and "Elute" (red).

https://doi.org/10.1021/acs.biochem.1c00138
Biochemistry 2021, 60, 1876−1884

MATERIALS AND METHODS

Protein Overexpression and Lysis. Purified His_6-tagged *Salmonella* YidA and overexpression plasmids encoding His_6-tagged *Methanococcus maripaludis* L7Ae, *Salmonella* FrlB, *Salmonella* FraR, and *Salmonella* FraR fused to maltose-binding protein (MBP-FraR) were used in this study. The motivation for this selection of plasmids was twofold: (i) the in-house availability of purified protein and clones (V. Gopalan, unpublished data) and (ii) the desire to exploit our method to advance ongoing biochemical characterization. Overexpression conditions and purification protocols for these proteins have not been previously reported and will be described elsewhere. Briefly, *E. coli* BL21 (DE3) or Rosetta (DE3) cells were transformed with plasmids encoding the His_6-tagged proteins mentioned above. Overnight seed cultures were used to inoculate 5 mL of lysogeny broth (LB) medium containing the appropriate antibiotics, which were subsequently grown at 37°C with shaking to an OD_{600} of ∼0.6. While the scale of overexpression may be adjusted as needed, the 5-mL culture volume used here was chosen to allow for efficient probe-based sonication and to ensure sufficient protein amounts for analysis. Protein overexpression was induced with appropriate amounts of isopropyl β-D-thiogalactoside (IPTG). Post-induction, these cultures were allowed to grow further under previously optimized conditions (V. Gopalan, unpublished data). Cells were then harvested by centrifugation (20000*g*, 2 min, ∼22°C) and lysed by sonication in 400 μL of 1× phosphate-buffered saline (PBS) supplemented with Halt protease inhibitor (Thermo Scientific). Following centrifugation (20000*g*, 30 min, 4°C) to remove cellular debris and insoluble components, supernatants were transferred to 96-well plates for MS analysis. Untransformed BL21 (DE3) cells were cultured, induced, and processed as described above to prepare the cell lysate used in the YidA dilution experiments.

LC-Based IMAC−OBE−nMS. For all experiments, a mobile phase of 200 mM ammonium acetate at pH 7.5 (pH adjusted using ammonium hydroxide) was maintained at a flow rate of 100 μL/min. A diagram of the instrument setup described below is shown in Figure 2. A Vanquish Duo Ultra-High-Performance LC (UHPLC) system (Thermo Scientific) equipped with a dual pump and autosampler was used to load crude cell lysates onto an IMAC column (ProPac IMAC-10, 1 mm × 50 mm, 1.7 μm; Thermo Scientific, 063617). The column was charged with Ni^{2+} at the beginning of this study, and we did not find it necessary to strip and recharge the column because there did not appear to be a decrease in protein binding capacity during successive runs.

A six-port switching valve downstream of the IMAC column was used to direct flow: unbound species were directed to waste, while bound proteins were eluted using 3 μL of 5 M imidazole (pH 7.5), buffer exchanged into the 200 mM ammonium acetate (pH 7.5) mobile phase, and separated from imidazole using a size exclusion chromatography (SEC) column made of P6 gel medium (Bio-Rad) that was self-packed in PEEK tubing (0.03 in. internal diameter, 12 cm length).[20] Desalted samples were then analyzed using a Q Exactive Ultra-High Mass Range (UHMR) Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Scientific) that was modified with a customized device for performing surface-induced dissociation (SID)[21,22] and equipped with an Ion Max ion source (Thermo Scientific) and a heated ESI (HESI-II) probe (Thermo Scientific). The presence of the SID device did not influence selection of standard settings for this method. The instrument was tuned to optimize ion transmission over a wide m/z range (see Table 1 for specific ESI and MS tune settings). Because nMS analysis was carried out in a mobile phase of moderately high ionic strength (200 mM), PEEK tubing (0.005 in. internal diameter, 10 ft length) was placed between the HESI-II probe and the Ion Max source's stainless-steel grounding union as "resistor tubing" to ensure that the electrospray current did not exceed the maximum limit set by the instrument software. Exceeding this threshold would cause a loss of electrospray and decrease sensitivity.[20]

Table 1. Tune Settings for the Q Exactive UHMR Hybrid Quadrupole-Orbitrap Mass Spectrometer

setting	value
scan range (m/z)	1000−14000
resolution (at m/z 400)	12500
no. of microscans	5
max inject time (ms)	200
sheath gas (psi)	50
auxiliary gas (psi)	0
sweep gas (psi)	0
spray voltage (kV)	4.0
capillary temperature (°C)	320
S-lens RF level (V)	200
in-source dissociation (V)	10
in-source trapping (V)	variable
source DC offset (V)	21
injection flatapole DC (V)	5
inter flatapole lens (V)	4
bent flatapole DC (V)	2
trapping gas pressure	3

Data Analysis. All data analyses were conducted using the Intact Mass software (version w2.15-294-gba5daea4b; Protein Metrics Inc., San Carlos, CA)[23] and the following deconvolution parameters: minimum difference between mass peaks, 10; charge vector spacing, 1; baseline radius (m/z), 15; smoothing σ (m/z), 0.02; spacing (m/z), 0.04; mass smoothing σ, 3; mass spacing, 0.5; iteration maximum, 10.

RESULTS

Instrument Setup for Tandem Affinity-Buffer Exchange Chromatography Online with nMS. For implementation of IMAC−OBE−nMS, an analytical flow LC system equipped with an autosampler was coupled to a mass spectrometer (Figure 2). Located between the LC autosampler and the mass spectrometer were a Ni^{2+}-charged IMAC column; a six-port switching valve, which was used to direct flow as described below; and an SEC column for removal of nonvolatile components and buffer exchange into ammonium acetate for nMS analysis.

The switching valve had two positions: "Load" and "Elute" (Figure 2). When the autosampler injected crude cell lysate onto the IMAC column, the switching valve was set to the "Load" position to direct the flow-through toward waste, allowing for the removal of unbound proteins and nonvolatile species from the sample. In contrast, when the autosampler injected imidazole onto the column to elute the poly-His-tagged protein, the switching valve was set to the "Elute" position. Subsequently, the eluate was directed to the size exclusion column, which separated the sample from imidazole

https://doi.org/10.1021/acs.biochem.1c00138
Biochemistry 2021, 60, 1876−1884

Table 2. IMAC−OBE−nMS Method[a]

step	time (min)	autosampler	LC pump	switching valve	MS
load	4	1 μL injection of crude cell lysate	100 μL/min	position 1-6 (to waste)	tune-file2 (no source gas, no ESI voltage)
wash	4	5 μL injection of mobile phase	100 μL/min	position 1-2 (to MS)	tune-file2 (no source gas, no ESI voltage)
elute	4	3 μL injection of 5 M imidazole (pH 7.5)	100 μL/min	position 1-2 (to MS)	tune-file1 (source gas, ESI voltage) during protein elution; switch to tune-file2 (no source gas, no ESI voltage) before imidazole elutes after ∼1.5 min

[a]The minimum total run time for each sample is 12 min. An additional 4-min wash step post-elution was typically included to minimize the risk of protein carryover.

and nonvolatile salts prior to detection by MS. Note: to analyze unbound proteins (e.g., host cell proteins), the switching valve can remain in the "Elute" position during the load step. However, if cellular debris is not carefully removed after cell lysis or protein aggregation occurs, the performance of the size exclusion column may be compromised.

A mobile phase of 200 mM ammonium acetate (pH 7.5) was selected for two primary reasons: (i) it is a volatile electrolyte, which is required to transfer samples from the solution phase to the gas phase without extensive adduction for MS analysis; and (ii) ammonium acetate is generally suitable and commonly used for preserving noncovalent inter- and intramolecular interactions for a wide range of macromolecular samples and complexes.[24,25] In all cases, the pH of the mobile phase should be neutral to slightly alkaline (pH 7−8) to ensure that the poly-histidine tag on the protein of interest binds the Ni^{2+}-charged resin.

General Workflow for IMAC−OBE−nMS Experiments. Following overexpression of poly-His-tagged proteins of interest, cells were resuspended in an appropriate buffer and then lysed. While we used 1× PBS for all experiments described here, inclusion of a switching valve that can direct nonvolatile salts to waste as well as an online buffer exchange step should permit the use of any buffer, regardless of its compatibility with MS analysis. With the switching valve set to the "Load" position, the autosampler was used to inject the sample onto the Ni^{2+}-charged IMAC column. Poly-His-tagged proteins remained bound to the column, while unbound species were directed to waste. A subsequent injection of the mobile phase minimized carryover from the previous injection. The valve was then switched to the "Elute" position, and imidazole was injected to elute bound species from the IMAC column. The eluate was directed toward the SEC column, where proteins were separated from imidazole and concomitantly buffer exchanged into the ammonium acetate mobile phase prior to nMS analysis.

Mass spectral acquisition was timed to permit ionization of proteins, but not imidazole, upon elution from the SEC column; due to the large size discrepancy, imidazole elutes later than proteins during size exclusion-based separation. To this end, the MS method was comprised of two separate tune files (Table 2): the first file turned on the source gas and applied electrospray voltage as proteins were eluted from the SEC column, while the second turned both off before imidazole began to elute at ∼1.5 min. This strategy allowed for efficient detection of proteins while imidazole and remaining nonvolatile salts dripped out of the ESI needle and into a waste tube connected to the ion source housing drain.

To evaluate the contribution of the added IMAC purification step, we also performed OBE−nMS-only experiments on the same samples subjected to IMAC−OBE−nMS analysis. For OBE−nMS, the IMAC column was removed from the instrument setup described above, and the autosampler was used to direct samples onto the SEC column via the switching valve. As with IMAC−OBE−nMS, two separate tune files were used to ensure that nonvolatile salts were not electrosprayed and the instrument was kept clean for subsequent experiments.

IMAC−OBE−nMS as a Tool for Analyzing Proteins in Cell Lysates. We used several proteins to highlight the ability of IMAC−OBE−nMS to selectively capture and enrich for overexpressed poly-His-tagged proteins of interest in crude lysate. First, His_6-tagged *M. maripaludis* L7Ae, a soluble protein that ionizes well under a variety of conditions, was used as a proof of concept (Figure 3). When a cell lysate containing overexpressed L7Ae was subjected to OBE−nMS (Figure 3A), L7Ae was the dominant species, though the spectrum revealed the presence of other, less intense peaks corresponding to *E. coli* host cell proteins. In contrast, combining the IMAC purification step and OBE−nMS removed almost all of the contaminating species, demonstrating that IMAC−OBE−nMS effectively enriches for His_6-tagged proteins (Figure 3B).

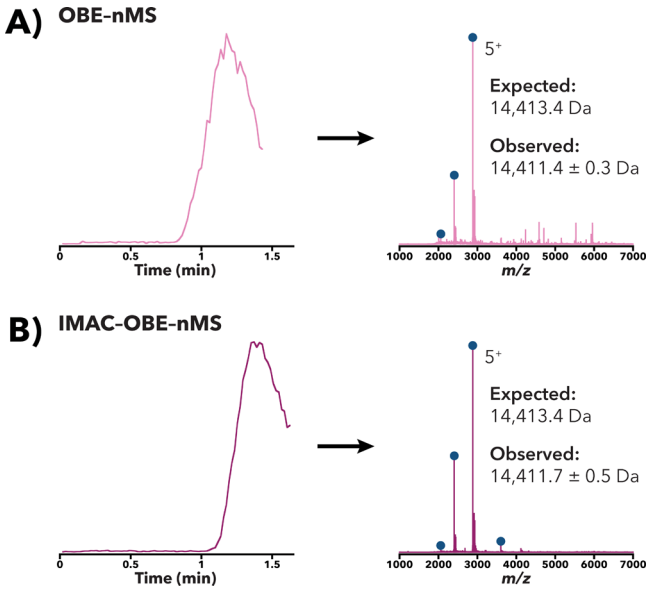


Figure 3. Total ion chromatograms and mass spectra of *M. maripaludis* L7Ae analyzed using (A) OBE−nMS or (B) IMAC−OBE−nMS. The expected mass indicated above accounts for loss of the N-terminal methionine, which is a common post-translational modification.[26−28] Charge state distributions for L7Ae are indicated with blue circles, and the main charge state is labeled. In all plots, the *y*-axis (not shown) represents relative intensity.

https://doi.org/10.1021/acs.biochem.1c00138
Biochemistry 2021, 60, 1876−1884

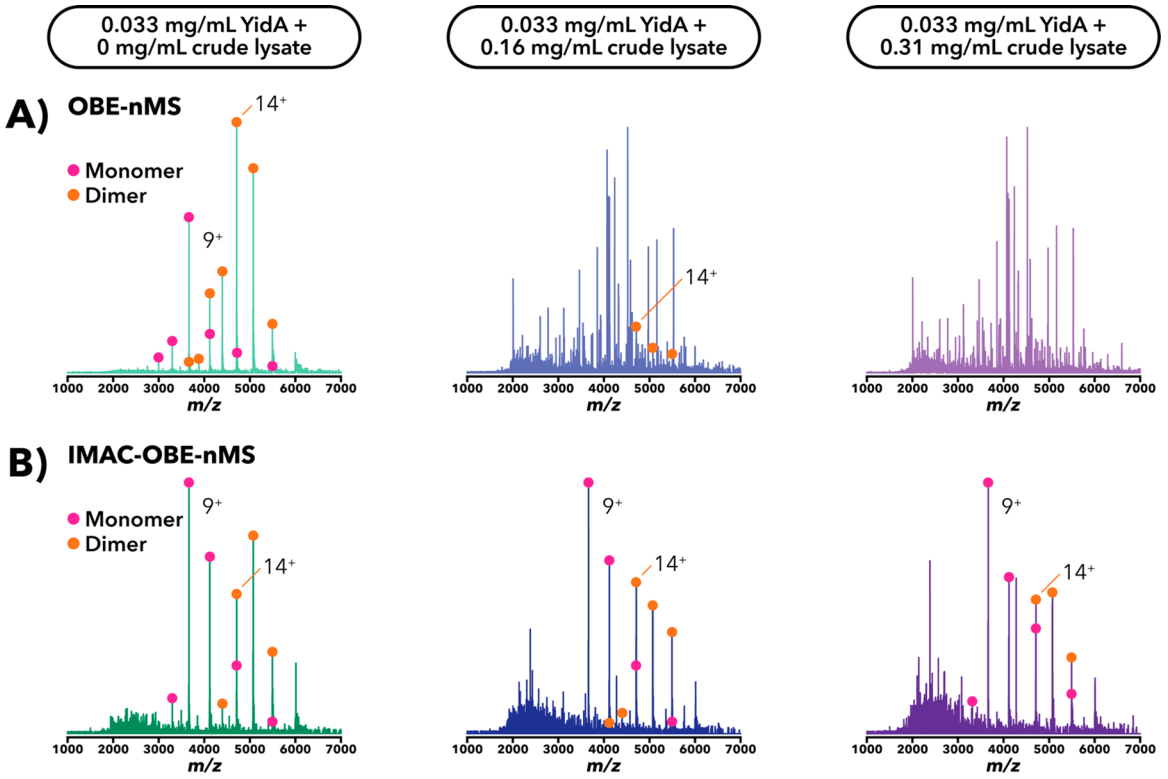


Figure 4. Mass spectra of purified *Salmonella* YidA, either alone or diluted as indicated with *E. coli* crude lysate, analyzed using (A) OBE−nMS or (B) IMAC−OBE−nMS. Charge state distributions for monomeric and dimeric species are indicated with pink and orange circles, respectively, and the main charge state for each species is labeled. In all plots, the *y*-axis (not shown) represents relative intensity.

However, due to the relatively high ionization efficiency and overexpression level of *M. maripaludis* L7Ae, the benefit of our new method is not particularly striking here. Therefore, to better establish the merits of IMAC−OBE−nMS, we intentionally spiked a pure His_6-tagged protein into untransformed cell lysate and compared the results of OBE−nMS and IMAC−OBE−nMS (Figure 4).

For this next set of experiments, we used His_6-tagged *Salmonella* YidA, a sugar phosphatase, as the target protein because its ionization efficiency is lower than that of *M. maripaludis* L7Ae, so it is likely to be more representative of an average protein of interest. A fixed amount of purified YidA (0.033 mg/mL), either alone or in combination with varying amounts of untransformed cell lysate (0.16 or 0.31 mg/mL), was subjected to OBE−nMS and IMAC−OBE−nMS. Consistent with the expectation that both methods should be gentle enough to preserve and screen for the oligomeric state of overexpressed proteins of interest,[20] a mixture of monomeric and dimeric YidA species was observed when YidA alone was analyzed using either method. However, addition of crude lysate to the sample clearly distinguished the two methods: whereas the YidA signal for OBE−nMS was lost with increasing amounts of cell lysate (Figure 4A), YidA was still detectable at a 1:10 YidA:crude lysate dilution with IMAC−OBE−nMS (Figure 4B). Loss of the YidA signal with OBE−nMS is characteristic of increased charge competition and ion suppression from the higher abundance of host cell proteins, which, in extreme cases, can prevent detection of target proteins even if they are present in the sample.[29] Thus, the results of this experiment suggest that the addition of the IMAC step increases the dynamic range of detection by selectively enriching for tagged proteins while depleting endogenous species.

Nearly 40% of all enzymes with known structures function as dimers.[30,31] Structure−function studies of homodimers, in particular, are complicated by the fact that introducing any mutations that target the active site could potentially disrupt oligomerization. Therefore, prior to the initiation of large-scale purification and functional characterization efforts, it is useful to carry out preliminary small-scale studies to determine whether introduced mutations impair homodimer formation and lead to a loss of activity. Because SDS denatures proteins and disrupts protein−protein interactions, SDS−PAGE, in contrast to native MS-based methods, is not suitable for assessing the oligomeric state of overexpressed proteins. To demonstrate the utility of IMAC−OBE−nMS in this regard, we selected His_6-tagged *Salmonella* FrlB deglycase as a test case because previous nMS experiments had shown that two monomers of wild-type FrlB assemble to form a functional homodimeric enzyme (S. Kovvali, A. Di Capua, V. Gopalan, and V.H. Wysocki, unpublished data; see also refs 32 and 33). Our IMAC−OBE−nMS results showed that a FrlB E224Q/H240N mutant mirrors the wild type in preferentially forming a homodimer, as there were no detectable peaks corresponding to a FrlB monomer (Figure 5). However, the presence of a bimodal charge state distribution with lower intensity, higher charge state peaks indicated that there was some partially unfolded homodimer in the sample. Therefore, the appearance of multimodal and/or broad charge state distributions during IMAC−OBE−nMS analysis may be useful for assessing how overexpression conditions affect the distribution between folded and unfolded populations of overexpressed target proteins,[34−37] though we cannot rule out the possibility that

https://doi.org/10.1021/acs.biochem.1c00138
Biochemistry 2021, 60, 1876−1884

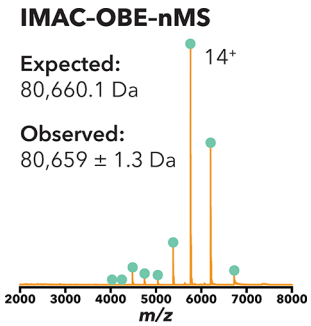


Figure 5. Mass spectrum of *Salmonella* FrlB E224Q/H240N, as analyzed using IMAC−OBE−nMS. The charge state distribution for the homodimer is indicated with green circles, and the main charge state is labeled. The lower intensity, higher charge state peaks in the bimodal charge state distribution are indicative of a minor population of partially unfolded homodimer. The observed mass is consistent with the expected E224Q and H240N substitutions. The *y*-axis (not shown) represents relative intensity.

LC and/or MS conditions may contribute, at least in part, to any observed protein unfolding.

One notable challenge of using heterologous systems for recombinant protein production is that a significant number of proteins are not expressed in a soluble form, accumulating instead in inclusion bodies and precluding the use of nondenaturing purification.[6−8] While there are many potential reasons for protein insolubility, the primary cause is misfolding, which may result from suboptimal translation rates or a lack of post-translational modifications, chaperones, or binding partners required for proper folding. Some of these insoluble proteins can be rescued by fusing a solubility tag such as maltose-binding protein (MBP) to one terminus of the polypeptide chain, which can significantly increase the yield of overexpressed, soluble protein.[38,39] To test whether IMAC−OBE−nMS can be used to assess the ability of solubility tags to prevent protein aggregation and precipitation, overexpression samples containing either His_6-tagged wild-type *Salmonella* FraR or a variant that is fused to MBP (MBP-FraR) were analyzed (Figure 6). Because wild-type FraR is largely

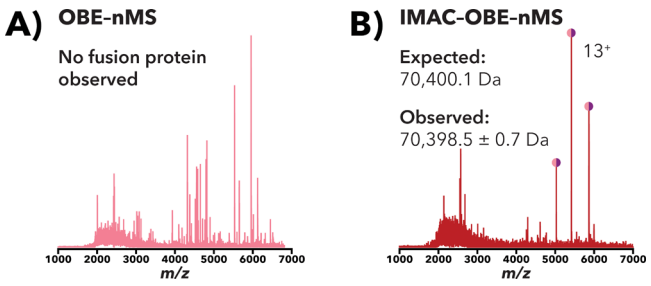


Figure 6. Mass spectra of MBP-*Salmonella* FraR analyzed using (A) OBE−nMS or (B) IMAC−OBE−nMS. The charge state distribution for MBP-FraR is indicated with pink and purple circles, and the main charge state is labeled. In both plots, the *y*-axis (not shown) represents relative intensity.

expressed as an insoluble protein in *E. coli* (S. M. Lai and V. Gopalan, unpublished data), it was not detected by either OBE−nMS or IMAC−OBE−nMS (data not shown). Even as a fusion construct (B.E. Szkoda and V. Gopalan, unpublished data), though, the apparent low abundance and/or ionization efficiency of MBP-FraR prevented detection via OBE−nMS, and no peaks corresponding to the target protein were observed (Figure 6A). However, IMAC−OBE−nMS successfully enriched for MBP-FraR while simultaneously depleting *E. coli* host cell proteins (Figure 6B), thereby facilitating detection of the protein of interest and revealing that addition of the MBP tag helped promote the production of soluble FraR. While this analysis cannot provide a quantitative assessment of protein yield, it provides a useful indication that a solubility tag-based strategy could be successful for overexpression and purification of a target protein that is prone to misfolding.

DISCUSSION

We have developed and validated a laboratory-scale approach for analyzing cell lysates for successful overexpression and IMAC-based purification of poly-His-tagged proteins of interest without extensive sample preparation. By adding an upstream IMAC-based enrichment step to our recently described OBE−nMS approach,[20] we have developed an automatable method for determining whether target proteins (i) have been overexpressed in a soluble form, (ii) are able to bind and elute from an IMAC column (i.e., the poly-His tag is solvent-accessible), (iii) exist as monomers or higher-order oligomers, and (iv) importantly, have the expected mass. The high mass accuracy typically achieved with nMS is greater than that of other methods, including SDS−PAGE, which affords a decisive advantage when verifying the identity of an overexpressed protein of interest, especially if there is amino acid misincorporation and/or proteolysis during recombinant protein production.[40,41]

Our approach is particularly effective for analyzing protein overexpression because it can be used for a variety of different samples, including those with a low level of overexpression and poor protein solubility. In the future, we plan to show that the method may be adapted for use with an even wider range of samples by, for example, substituting the IMAC resin with other media (e.g., Protein A, Protein G, glutathione, and streptavidin) that exploit different affinity tags. Coupling online fragmentation, either by collision-induced dissociation (CID) or SID, with IMAC−OBE−nMS would permit subunit mass and complex composition/stoichiometry measurements, which may be particularly useful for characterizing protein and protein−ligand complexes that are co-expressed within a single host.[42,43] Finally, while *E. coli* overexpression cultures were used for this study, IMAC−OBE−nMS should, in principle, be compatible with cell lysates or secretomes (through analysis of the growth medium) from any bacterial, archaeal, or eukaryotic host.

Because of their dependence on electrospray ionization for transferring samples into the gas phase, OBE−nMS and IMAC−OBE−nMS are primarily suited for analyzing soluble species in volatile electrolytes. However, the incorporation of a switching valve and OBE allows inclusion of moderate amounts of nonvolatile solutes, such as monovalent and divalent cations necessary for nucleic acid−protein assemblies or detergents for membrane protein solubility. However, if exceptionally high concentrations of nonvolatile buffer components are necessary for species stability and solubility, solution-based methods like mass photometry,[44−48] small-angle X-ray scattering (SAXS),[49,50] or SEC-multiangle light scattering (MALS)[51,52] may be more suitable, though they provide lower-resolution mass information compared to MS and the target protein would need to be purified before analysis. It should also be noted that the ability to use MS to

https://doi.org/10.1021/acs.biochem.1c00138
Biochemistry 2021, 60, 1876−1884

detect a particular species of interest within a mixture is dependent on its relative abundance and ionization efficiency. Differences in ionization efficiencies make it challenging to quantitatively measure protein concentration based on observed peak intensities. Therefore, the IMAC−OBE−nMS method described in this study is best suited for providing qualitative confirmation of the presence of a specific protein within a single sample, albeit with sensitivity higher than those of other methods that rely on direct infusion of cell lysates, due to its ability to enrich for target proteins.

Notwithstanding these limitations, the proof-of-concept results shown here suggest that future development of IMAC−OBE−nMS has the potential to help expedite the overall process of optimizing overexpression and purification conditions for large-scale protein production, making it both a powerful stand-alone technique and a complement to other MS-based approaches[17,53−56] for studying protein structure and function.

■ ASSOCIATED CONTENT

Accession Codes

The following proteins were used in this study: *M. maripaludis* L7Ae (UniProt entry P62426), *Salmonella* YidA (UniProt entry A0A0F6B8Z9), *Salmonella* FrlB (UniProt entry A0A0F6BB82), and *Salmonella* FraR (UniProt entry A0A0F6B872).

■ AUTHOR INFORMATION

Corresponding Author

Vicki H. Wysocki − *Department of Chemistry and Biochemistry, Resource for Native Mass Spectrometry-Guided Structural Biology, and Campus Chemical Instrument Center, Mass Spectrometry and Proteomics, The Ohio State University, Columbus, Ohio 43210, United States;* orcid.org/0000-0003-0495-2538; Email: wysocki.11@osu.edu

Authors

Florian Busch − *Department of Chemistry and Biochemistry, Resource for Native Mass Spectrometry-Guided Structural Biology, and Campus Chemical Instrument Center, Mass Spectrometry and Proteomics, The Ohio State University, Columbus, Ohio 43210, United States;* orcid.org/0000-0002-4324-6065

Zachary L. VanAernum − *Department of Chemistry and Biochemistry and Resource for Native Mass Spectrometry-Guided Structural Biology, The Ohio State University, Columbus, Ohio 43210, United States;* orcid.org/0000-0002-0956-3956

Stella M. Lai − *Department of Chemistry and Biochemistry and Resource for Native Mass Spectrometry-Guided Structural Biology, The Ohio State University, Columbus, Ohio 43210, United States*

Venkat Gopalan − *Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, United States;* orcid.org/0000-0002-3016-2719

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.biochem.1c00138

Author Contributions

‖F.B., Z.L.V., and S.M.L. contributed equally to this work.

Funding

This work was supported by the National Institutes of Health (Grants P41 GM128577 to V.H.W., GM120582 to V.G. and V.H.W., AI116119 to V.G. and V.H.W., and AI140541 to V.G.) and Ohio Eminent Scholar funds (to V.H.W.).

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors are grateful to members of Dr. Venkat Gopalan's laboratory for overexpression plasmids and protocols (Dr. I-Ming Cho for L7Ae, Sravya Kovvali for mutant FrlB, and Blake E. Szkoda for FraR with or without MBP), purified protein (Eleanor Bashian for YidA), sequencing results (Dr. Pankajavalli Thirugnanasambantham for YidA), and valuable comments on the manuscript (Dr. Walter J. Zahurancik). The authors also thank Julia Baek, Terry Zhang, Dr. Kyle L. Fort, and Dr. Rosa Viner (Thermo Fisher Scientific) for providing the ProPac IMAC-10 columns used in this study and their assistance with the IMAC−OBE−nMS setup and QE UHMR mass spectrometer.

■ REFERENCES

(1) Tabor, S., and Richardson, C. C. (1985) A bacteriophage T7 RNA polymerase/promoter system for controlled exclusive expression of specific genes. *Proc. Natl. Acad. Sci. U. S. A. 82*, 1074−1078.

(2) Tabor, S. (1990) Expression using the T7 RNA polymerase/promoter system. *Curr. Protoc. Mol. Biol. 11*, 16.12.11−16.12.11.

(3) Structural Genomics Consortium, Architecture et Fonction des Macromolécules Biologiques, Berkeley Structural Genomics Center, China Structural Genomics Consortium, Integrated Center for Structure and Function Innovation, Israel Structural Proteomics Center, Joint Center for Structural Genomics, Midwest Center for Structural Genomics, New York Structural GenomiX Research Center for Structural Genomics, Northeast Structural Genomics Consortium, Oxford Protein Production Facility, Protein Sample Production Facility, Max Delbrück Center for Molecular Medicine, RIKEN Structural Genomics/Proteomics Initiative, and SPINE2-Complexes. (2008) Protein production and purification. *Nat. Methods 5*, 135−146.

(4) Jungbauer, A., and Hahn, R. (2009) Ion-exchange chromatography. *Methods Enzymol. 463*, 349−371.

(5) McCue, J. T. (2009) Theory and use of hydrophobic interaction chromatography in protein purification applications. *Methods Enzymol. 463*, 405−414.

(6) Francis, D. M., and Page, R. (2010) Strategies to optimize protein expression in E coli. *Curr. Protoc. Protein Sci. 61*, 5.24.21−25.24.29.

(7) Lebendiker, M., and Danieli, T. (2014) Production of prone-to-aggregate proteins. *FEBS Lett. 588*, 236−246.

(8) Rosano, G. L., and Ceccarelli, E. A. (2014) Recombinant protein expression in Escherichia coli: advances and challenges. *Front. Microbiol. 5*, 172.

(9) Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature 227*, 680−685.

(10) Shi, Y., Mowery, R. A., Ashley, J., Hentz, M., Ramirez, A. J., Bilgicer, B., Slunt-Brown, H., Borchelt, D. R., and Shaw, B. F. (2012) Abnormal SDS-PAGE migration of cytosolic proteins can identify domains and mechanisms that control surfactant binding. *Protein Sci. 21*, 1197−1209.

(11) Guan, Y., Zhu, Q., Huang, D., Zhao, S., Jan Lo, L., and Peng, J. (2015) An equation to estimate the difference between theoretically predicted and SDS PAGE-displayed molecular weights for an acidic peptide. *Sci. Rep. 5*, 13370.

(12) Bolanos-Garcia, V. M., and Davies, O. R. (2006) Structural analysis and classification of native proteins from E. coli commonly

co-purified by immobilised metal affinity chromatography. *Biochim. Biophys. Acta, Gen. Subj. 1760*, 1304−1313.
(13) Xue, S., Wang, R., Yang, F., Terns, R. M., Terns, M. P., Zhang, X., Maxwell, E. S., and Li, H. (2010) Structural basis for substrate placement by an archaeal box C/D ribonucleoprotein particle. *Mol. Cell 39*, 939−949.
(14) Catcott, K. C., Yan, J., Qu, W., Wysocki, V. H., and Zhou, Z. S. (2017) Identifying unknown enzyme-substrate pairs from the cellular milieu with native mass spectrometry. *ChemBioChem 18*, 613−617.
(15) Gan, J., Ben-Nissan, G., Arkind, G., Tarnavsky, M., Trudeau, D., Noda Garcia, L., Tawfik, D. S., and Sharon, M. (2017) Native mass spectrometry of recombinant protein from crude cell lysates. *Anal. Chem. 89*, 4398−4404.
(16) Ben-Nissan, G., Vimer, S., Warszawski, S., Katz, A., Yona, M., Unger, T., Peleg, Y., Morgenstern, D., Cohen-Dvashi, H., Diskin, R., Fleishman, S. J., and Sharon, M. (2018) Rapid characterization of secreted recombinant proteins by native mass spectrometry. *Commun. Biol. 1*, 213.
(17) Vimer, S., Ben-Nissan, G., and Sharon, M. (2020) Direct characterization of overproduced proteins by native mass spectrometry. *Nat. Protoc. 15*, 236−265.
(18) Vimer, S., Ben-Nissan, G., and Sharon, M. (2020) Mass spectrometry analysis of intact proteins from crude samples. *Anal. Chem. 92*, 12741−12749.
(19) Sulkowski, E. (1985) Purification of proteins by IMAC. *Trends Biotechnol. 3*, 1−7.
(20) VanAernum, Z. L., Busch, F., Jones, B. J., Jia, M., Chen, Z., Boyken, S. E., Sahasrabuddhe, A., Baker, D., and Wysocki, V. H. (2020) Rapid online buffer exchange for screening of proteins, protein complexes, and cell lysates by native mass spectrometry. *Nat. Protoc. 15*, 1132−1157.
(21) VanAernum, Z. L., Gilbert, J. D., Belov, M. E., Makarov, A. A., Horning, S. R., and Wysocki, V. H. (2019) Surface-induced dissociation of noncovalent protein complexes in an extended mass range Orbitrap mass spectrometer. *Anal. Chem. 91*, 3611−3618.
(22) Harvey, S. R., VanAernum, Z. L., Kostelic, M. M., Marty, M. T., and Wysocki, V. H. (2020) Probing the structure of nanodiscs using surface-induced dissociation mass spectrometry. *Chem. Commun. (Cambridge, U. K.) 56*, 15651−15654.
(23) Bern, M., Caval, T., Kil, Y. J., Tang, W., Becker, C., Carlson, E., Kletter, D., Sen, K. I., Galy, N., Hagemans, D., Franc, V., and Heck, A. J. R. (2018) Parsimonious charge deconvolution for native mass spectrometry. *J. Proteome Res. 17*, 1216−1226.
(24) Heck, A. J. R., and van den Heuvel, R. H. H. (2004) Investigation of intact protein complexes by mass spectrometry. *Mass Spectrom. Rev. 23*, 368−389.
(25) Hernández, H., and Robinson, C. V. (2007) Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. *Nat. Protoc. 2*, 715−726.
(26) Sherman, F., Stewart, J. W., and Tsunasawa, S. (1985) Methionine or not methionine at the beginning of a protein. *BioEssays 3*, 27−31.
(27) Ben-Bassat, A., Bauer, K., Chang, S.-Y., Myambo, K., Boosman, A., and Chang, S. (1987) Processing of the initiation methionine from proteins: properties of the Escherichia coli methionine aminopeptidase and its gene structure. *J. Bacteriol. 169*, 751−757.
(28) Giglione, C., Boularot, A., and Meinnel, T. (2004) Protein N-terminal methionine excision. *Cell. Mol. Life Sci. 61*, 1455−1474.
(29) Annesley, T. M. (2003) Ion suppression in mass spectrometry. *Clin. Chem. 49*, 1041−1044.
(30) Marianayagam, N. J., Sunde, M., and Matthews, J. M. (2004) The power of two: protein dimerization in biology. *Trends Biochem. Sci. 29*, 618−625.
(31) Matthews, J. M., and Sunde, M. (2012) Dimers, oligomers, everywhere. *Adv. Exp. Med. Biol. 747*, 1−18.
(32) Sengupta, A., Wu, J., Seffernick, J. T., Sabag-Daigle, A., Thomsen, N., Chen, T.-H., Capua, A. D., Bell, C. E., Ahmer, B. M. M., Lindert, S., Wysocki, V. H., and Gopalan, V. (2019) Integrated use of biochemical, native mass spectrometry, computational, and genome-editing methods to elucidate the mechanism of a Salmonella deglycase. *J. Mol. Biol. 431*, 4497−4513.
(33) Graf von Armansperg, B., Koller, F., Gericke, N., Hellwig, M., Jagtap, P. K. A., Heermann, R., Hennig, J., Henle, T., and Lassak, J. (2021) Transcriptional regulation of the Nδ-fructoselysine metabolism in Escherichia coli by global and substrate-specific cues. *Mol. Microbiol. 115*, 175−190.
(34) Chowdhury, S. K., Katta, V., and Chait, B. T. (1990) Probing conformational changes in proteins by mass spectrometry. *J. Am. Chem. Soc. 112*, 9012−9013.
(35) Konermann, L., and Douglas, D. J. (1998) Unfolding of proteins monitored by electrospray ionization mass spectrometry: a comparison of positive and negative ion modes. *J. Am. Soc. Mass Spectrom. 9*, 1248−1254.
(36) Kaltashov, I. A., and Mohimen, A. (2005) Estimates of protein surface areas in solution by electrospray ionization mass spectrometry. *Anal. Chem. 77*, 5370−5379.
(37) Xia, Z., DeGrandchamp, J. B., and Williams, E. R. (2019) Native mass spectrometry beyond ammonium acetate: effects of nonvolatile salts on protein stability and structure. *Analyst 144*, 2565−2573.
(38) Fox, J. D., and Waugh, D. S. (2002) Maltose-binding protein as a solubility enhancer. *Methods Mol. Biol. 205*, 99−117.
(39) Tropea, J. E., Cherry, S., Nallamsetty, S., Bignon, C., and Waugh, D. S. (2007) A generic method for the production of recombinant proteins in Escherichia coli using a dual hexahistidine-maltose-binding protein affinity tag. *Methods Mol. Biol. 363*, 1−19.
(40) Tsai, L. B., Lu, H. S., Kenney, W. C., Curless, C. C., Klein, M. L., Lai, P.-H., Fenton, D. M., Altrock, B. W., and Mann, M. B. (1988) Control of misincorporation of de novo synthesized norleucine into recombinant interleukin-2 in E. coli Biochem. *Biochem. Biophys. Res. Commun. 156*, 733−739.
(41) Rozkov, A., and Enfors, S.-O. (2004) Analysis and control of proteolysis of recombinant proteins in Escherichia coli. *Adv. Biochem. Eng./Biotechnol. 89*, 163−195.
(42) Chen, W.-Y., Pulukkunat, D. K., Cho, I.-M., Tsai, H.-Y., and Gopalan, V. (2010) Dissecting functional cooperation among subunits in archaeal RNase P, a catalytic ribonucleoprotein complex. *Nucleic Acids Res. 38*, 8316−8327.
(43) Perederina, A., Berezin, I., and Krasilnikov, A. S. (2018) In vitro reconstitution and analysis of eukaryotic RNase P RNPs. *Nucleic Acids Res. 46*, 6857−6868.
(44) Young, G., Hundt, N., Cole, D., Fineberg, A., Andrecka, J., Tyler, A., Olerinyova, A., Ansari, A., Marklund, E. G., Collier, M. P., Chandler, S. A., Tkachenko, O., Allen, J., Crispin, M., Billington, N., Takagi, Y., Sellers, J. R., Eichmann, C., Selenko, P., Frey, L., Riek, R., Galpin, M. R., Struwe, W. B., Benesch, J. L. P., and Kukura, P. (2018) Quantitative mass imaging of single biological macromolecules. *Science 360*, 423−427.
(45) Soltermann, F., Foley, E. D. B., Pagnoni, V., Galpin, M., Benesch, J. L. P., Kukura, P., and Struwe, W. B. (2020) Quantifying protein-protein interactions by molecular counting with mass photometry. *Angew. Chem., Int. Ed. 59*, 10774−10779.
(46) Sonn-Segev, A., Belacic, K., Bodrug, T., Young, G., VanderLinden, R. T., Schulman, B. A., Schimpf, J., Friedrich, T., Dip, P. V., Schwartz, T. U., Bauer, B., Peters, J.-M., Struwe, W. B., Benesch, J. L. P., Brown, N. G., Haselbach, D., and Kukura, P. (2020) Quantifying the heterogeneity of macromolecular machines by mass photometry. *Nat. Commun. 11*, 1772.
(47) Li, Y., Struwe, W. B., and Kukura, P. (2020) Single molecule mass photometry of nucleic acids. *Nucleic Acids Res. 48*, No. e97.
(48) Olerinyova, A., Sonn-Segev, A., Gault, J., Eichmann, C., Schimpf, J., Kopf, A. H., Rudden, L. S. P., Ashkinadze, D., Bomba, R., Frey, L., Greenwald, J., Degiacomi, M. T., Steinhilper, R., Killian, J. A., Friedrich, T., Riek, R., Struwe, W. B., and Kukura, P. (2021) Mass photometry of membrane proteins. *Chem. 7*, 224−236.
(49) Blanchet, C. E., and Svergun, D. I. (2013) Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution. *Annu. Rev. Phys. Chem. 64*, 37−54.

https://doi.org/10.1021/acs.biochem.1c00138
Biochemistry 2021, 60, 1876−1884

(50) Gräwert, T. W., and Svergun, D. I. (2020) Structural modeling using solution small-angle X-ray scattering (SAXS). *J. Mol. Biol. 432*, 3078−3092.
(51) Some, D., Amartely, H., Tsadok, A., and Lebendiker, M. (2019) Characterization of proteins by size-exclusion chromatography coupled to multi-angle light scattering (SEC-MALS). *J. Visualized Exp. 148*, No. e59615.
(52) Some, D., and Razinkov, V. (2019) High-throughput analytical light scattering for protein quality control and characterization. *Methods Mol. Biol. 2025*, 335−359.
(53) Morris, J. H., Knudsen, G. M., Verschueren, E., Johnson, J. R., Cimermancic, P., Greninger, A. L., and Pico, A. R. (2014) Affinity purification-mass spectrometry and network analysis to understand protein-protein interactions. *Nat. Protoc. 9*, 2539−2554.
(54) Zhang, Y., Sun, H., Zhang, J., Brasier, A. R., and Zhao, Y. (2017) Quantitative assessment of the effects of trypsin digestion methods on affinity purification-mass spectrometry-based protein-protein interaction analysis. *J. Proteome Res. 16*, 3068−3082.
(55) Ehkirch, A., D'Atri, V., Rouviere, F., Hernandez-Alba, O., Goyon, A., Colas, O., Sarrut, M., Beck, A., Guillarme, D., Heinisch, S., and Cianferani, S. (2018) An online four-dimensional HICxSECx-IMxMS methodology for proof-of-concept characterization of antibody drug conjugates. *Anal. Chem. 90*, 1578−1586.
(56) Takano, K., Arai, S., Sakamoto, S., Ushijima, H., Ikegami, T., Saikusa, K., Konuma, T., Hamachi, I., and Akashi, S. (2020) Screening of protein-ligand interactions under crude conditions by native mass spectrometry. *Anal. Bioanal. Chem. 412*, 4037−4043.

https://doi.org/10.1021/acs.biochem.1c00138
Biochemistry 2021, 60, 1876−1884