nature protocols

PROTOCOL

https://doi.org/10.1038/s41596-019-0281-0

Rapid online buffer exchange for screening of proteins, protein complexes and cell lysates by native mass spectrometry

Zachary L. VanAernum[1,2,9], Florian Busch[1,2,9], Benjamin J. Jones[1,2], Mengxuan Jia[1,2], Zibo Chen[3,4,6], Scott E. Boyken[3,4,7], Aniruddha Sahasrabuddhe[1,8], David Baker[3,4,5] and Vicki H. Wysocki[1,2*]

It is important to assess the identity and purity of proteins and protein complexes during and after protein purification to ensure that samples are of sufficient quality for further biochemical and structural characterization, as well as for use in consumer products, chemical processes and therapeutics. Native mass spectrometry (nMS) has become an important tool in protein analysis due to its ability to retain non-covalent interactions during measurements, making it possible to obtain protein structural information with high sensitivity and at high speed. Interferences from the presence of non-volatiles are typically alleviated by offline buffer exchange, which is time-consuming and difficult to automate. We provide a protocol for rapid online buffer exchange (OBE) nMS to directly screen structural features of pre-purified proteins, protein complexes or clarified cell lysates. In the liquid chromatography coupled to mass spectrometry (LC-MS) approach described in this protocol, samples in MS-incompatible conditions are injected onto a short size-exclusion chromatography column. Proteins and protein complexes are separated from small molecule non-volatile buffer components using an aqueous, non-denaturing mobile phase. Eluted proteins and protein complexes are detected by the mass spectrometer after electrospray ionization. Mass spectra can inform regarding protein sample purity and oligomerization, and additional tandem mass spectra can help to further obtain information on protein complex subunits. Information obtained by OBE nMS can be used for fast (<5 min) quality control and can further guide protein expression and purification optimization.

Introduction

Protein production and purification have become increasingly accessible to researchers in all biomedical disciplines due to the rise of cost-efficient gene synthesis methods, standardized vectors and expression systems and the routine use of protein purification tags[1–3]. Commonly, proteins are overexpressed with an affinity tag in a suitable host cell system, for instance a derivative of *Escherichia coli* BL21(DE3). Cells are subsequently lysed in a well-buffered, high–ionic strength solution to preserve the initial structure of the protein of interest[1]. Reducing agents, stabilizing agents and ligands are frequently added to minimize protein oxidation and stabilize proteins, thereby also preventing them from aggregating[4,5]. Soluble proteins can be directly purified from the supernatant by affinity chromatography, typically resulting in a preparation with a relatively low host cell protein contamination level[6]. It is often desirable to determine key protein properties at this point to make an informed decision on whether a sample is, for example, suitable for in-depth biophysical and structural analysis. Commonly, intact protein molecular weight and sample purity are estimated based on SDS-PAGE, or are assessed after denaturation by MALDI- or ESI–mass spectrometry (MS)[1]. As those measurements are performed under denaturing conditions, no information on protein quaternary structure, a key indicator of protein functionality, is obtained. Here, we demonstrate that online buffer exchange (OBE) coupled to native mass spectrometry (nMS) can be readily implemented to obtain information on tertiary and quaternary structure, thus rapidly assessing protein and

[1]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, OH, USA. [2]Resource for Native Mass Spectrometry Guided Structural Biology, The Ohio State University, Columbus, OH, USA. [3]Department of Biochemistry, University of Washington, Seattle, WA, USA. [4]Institute for Protein Design, University of Washington, Seattle, WA, USA. [5]Howard Hughes Medical Institute, University of Washington, Seattle, WA, USA. [6]Present address: Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA, USA. [7]Present address: Lyell Immunopharma, Inc., Seattle, WA, USA. [8]Present address: Amgen Inc., Thousand Oaks, CA, USA. [9]These authors contributed equally: Zachary L. VanAernum, Florian Busch. *e-mail: wysocki.11@osu.edu

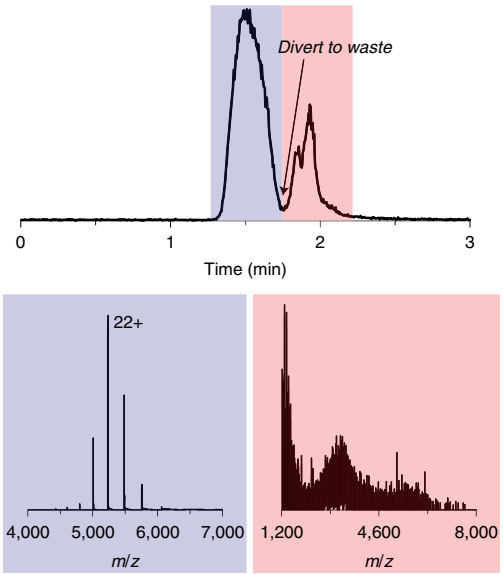

Fig. 1 | Separation of protein from non-volatile buffer components. Total ion chromatogram and mass spectra of CRP pentamer (blue) separated from non-volatile PBS components (red) using the OBE nMS method. A mobile phase of 200 mM ammonium acetate was delivered at a flow rate of 100 µl/min to a Yarra SEC-3000 column (290 Å pore size, 3.0 µm, 2.1 mm × 50 mm). The y-dimension of each spectrum represents relative intensity.

protein complex integrity of large numbers of samples, in an automated fashion, using small sample quantities. This procedure consists of four stages: in the first stage (Steps 1–11), buffer exchange columns are prepared; in the second stage (Step 12), the protein concentration in each sample is measured; in the third stage (Steps 13–23), the OBE method timing is established, and the samples are analyzed by nMS; finally, in the fourth stage (Step 24), the data are analyzed and reported.

Development of the protocol

OBE nMS was first described by Cavanagh et al.[7], with further development and potential use for drug discovery being reported by Waitt et al.[8] As any non-volatile components (salts, electrolytes and additives) present in a sample are not compatible with downstream MS analysis, OBE was introduced as a faster alternative to manual buffer exchange before MS analysis. More recently, OBE has been implemented as a fast desalting step after hydrophobic interaction chromatography separation coupled online with nMS[9]. The separation of proteins from non-volatile small molecules is accomplished by a short size-exclusion column, typically polyether ether ketone (PEEK) tubing filled with a porous stationary phase. We have improved upon and implemented OBE nMS to accommodate aqueous mobile phases containing enough ammonium acetate to provide sufficient ionic strength to maintain native protein structure and prevent interactions between analytes and the stationary phase. A typical chromatogram from the OBE method is shown in Fig. 1, demonstrating the efficient removal of non-volatile salts from a protein complex and subsequent MS detection. We have recently used this method for the high-throughput characterization of de novo designed proteins, allowing for unprecedented speed of nMS analysis to guide protein design and purification[10]. The procedure can be used for a variety of protein and protein complex samples and can help with efficient removal of non-volatiles before MS.

Suitable columns for OBE nMS

The main purpose of the stationary phase in OBE is to separate proteins from small non-volatiles within a short amount of time at a given flow rate, thereby limiting sample dilution and the extent to which biomolecular interactions with high k_{off} rates dissociate. For optimal OBE performance, a column should be chosen that has an exclusion limit below the mass of the proteins to be buffer-exchanged, by a factor of 2–3. This allows the buffer-exchanged protein to be rapidly eluted in the void volume, followed by the non-volatile salts. We have found that Bio-Gel P6 material (Bio-Rad) can easily be packed in 0.03 inch i.d. PEEK tubing to manufacture disposable gel filtration columns at very low cost. The self-packed P6 columns efficiently separate proteins from non-volatile salts with favorably short elution times. A column length of 12 cm generally provides enough capacity to

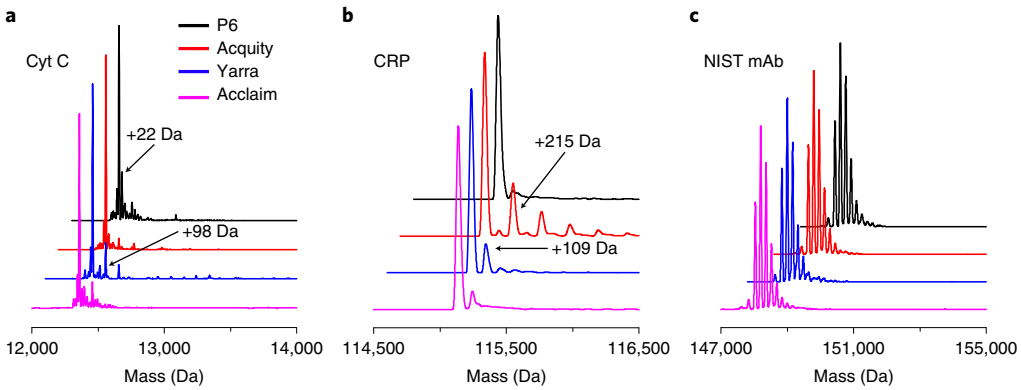

Fig. 2 | Comparison of OBE nMS using different size-exclusion columns. An Acquity UPLC BEH SEC (Waters, 125 Å, 1.7 µm, 4.6 mm × 30 mm), Acclaim SEC-300 (Thermo Scientific, 300 Å, 5 µm, 4.6 mm × 33 mm), and Yarra SEC-3000 (Phenomenex 290 Å, 3.0 µm, 2.1 mm × 50 mm) column were compared to the self-packed P6 Bio-Gel columns. Deconvoluted mass spectra of an (5 µl 4 µM) injection of cytochrome C (Cyt C; **a**), CRP (**b**) and NIST mAb (**c**) exchanged from PBS into 200 mM ammonium acetate using different columns (shown in legend). Common mass adducts are sodium (+22 Da) and phosphoric acid (+98 Da), in addition to two unknown adducts (+215 Da and +109 Da). Additional adduct peaks are primarily combinations of these masses. All spectra were acquired on an Exactive Plus EMR instrument and deconvoluted using Intact Mass software. The y-dimension of each spectrum represents relative intensity.

efficiently separate proteins from even high concentrations of non-volatiles, and in most cases it is likely possible to use even shorter columns. Alternatively, short size-exclusion chromatography (SEC) columns are available from several commercial manufacturers and can also be used for OBE nMS. A comparison of OBE nMS using commercial and self-packed columns is shown in Fig. 2. Cytochrome C (12-kDa monomeric protein), C-reactive protein (CRP, 115-kDa pentameric protein complex) and National Institute of Standards reference material 8671 (NIST mAb, 148 kDa) prepared in 1× PBS were buffer-exchanged using different columns. The desalting performance of each commercial column was comparable to that of the self-packed P6 column, with a few minor exceptions. The Yarra column resulted in less efficient non-volatile removal from cytochrome C (Fig. 2a), as cytochrome C is close to the lower working range of this column and is not as well separated from the non-volatile salts compared to larger proteins. CRP retained noticeable ~215-Da mass adducts when buffer-exchanged using the Acquity column. The origin of these adducts is not known and will require further investigation; however, it may be responsible for the shift to the lower charge-state distribution shown in Supplementary Fig. 1. The elution times of protein varied between the columns we investigated. The elution time of BSA was determined for each column by injecting 5 µl of 4 µM BSA with a mobile phase composition of 200 mM ammonium acetate and flow rate of 100 µl/min. The self-packed P6 column had the shortest elution time of all the columns, while the Acclaim column had the longest (Supplementary Table 1), demonstrating the advantage of using a column with an exclusion limit below the mass of the protein of interest. Each column generally exhibits efficient removal of non-volatile salts from the protein of interest, so the next most valuable figure of merit for a column used for sample screening is likely speed. Under these conditions, the self-packed P6 column would allow for the highest throughput. Although mobile phase flow rate can be modified to make up for the increased retention time for some of the columns, one should take caution in increasing the flow rate too much as too high flow rates and pressure can induce protein structure changes due to frictional heating[11]. In our experience, a backpressure of less than 400 p.s.i. at a flow rate of 0.1 ml/min makes self-packed P6 columns a good general choice, in particular for temperature- and pressure- sensitive proteins and protein complexes. Commercial columns with tighter packing and/or smaller particle sizes can, however, be sometimes advantageous, for instance in cases where some extent of separation between eluting proteins is desired.

Established mass range for OBE nMS

During the development of the OBE protocol, nine proteins and protein complexes ranging in size from 12 to 150 kDa were used to optimize MS tuning conditions (Supplementary Fig. 2). However, it should be noted that there is no reason that OBE cannot be used for larger analytes. For instance, we have recently successfully analyzed the 800-kDa tetradecamer bacterial chaperonin GroEL on a Q Exactive UHMR instrument without any changes to the OBE method aside from the MS tuning

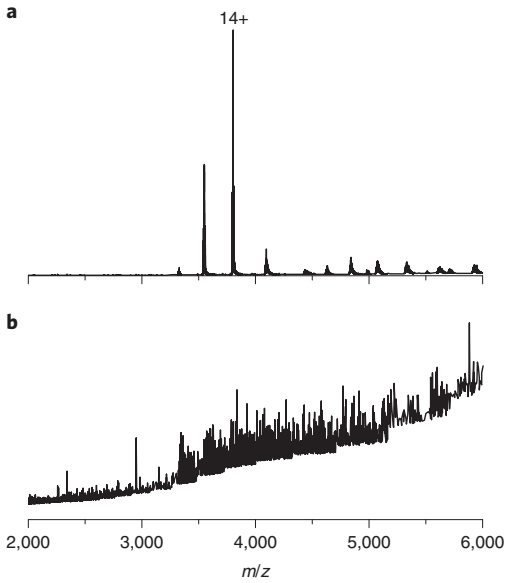

Fig. 3 | Effect of OBE on protein spectral quality. Mass spectra of streptavidin tetramer in PBS collected on a Solarix 15 T FT-ICR with a P6 OBE column (**a**) and without the use of a buffer exchange column (**b**). The experimental setup and all variables (MS tune settings, LC settings, etc.) were identical except that the P6 column for **a** was replaced with tubing for **b**. Minor peaks to higher *m/z* in **a** are due to non-specific 8-mer and 12-mer. The y-dimension of each spectrum represents relative intensity.

parameters (Supplementary Fig. 3). The proteins were dissolved or diluted in 1× PBS, desalted by OBE using a self-packed column with Bio-Rad P6 resin at an injection concentration of 4 μM protein or protein complex and recorded on a Thermo Scientific Exactive Plus EMR Orbitrap instrument. At a flow rate of 100 μl/min, the buffer-exchanged proteins are detected between 0.7 and 1.3 min, followed by the non-volatile salts between 1.3 and 2.3 min. The elution time was observed to shift by up to 0.05–0.1 min between different columns, presumably due to slight differences in column packing efficiency. Importantly, the elution time for an individual column remained constant over hundreds of runs. Because all proteins used here are above the exclusion limit of the resin (~6 kDa), all proteins elute from the column in the void volume, which allows for the development of a single liquid chromatography (LC)-MS method regardless of the size of the protein or protein complex being analyzed. Desalting efficiency of all nine proteins via the OBE approach was comparable to or better than offline buffer exchange via P6 spin columns (a direct comparison is shown in Supplementary Fig. 4). In all cases, the most abundant signal corresponded to adduct-free protein, with only minor adduction occurring in a few of the samples. Some samples also show multiple proteoforms present in minor abundances. A zoomed-in, deconvoluted spectrum of each buffer-exchanged protein and protein complex is available in Supplementary Fig. 5. The minor adducts present in each spectrum are due to non-volatile salts such as sodium (+22 Da) and phosphoric acid (+98 Da). Some of the peaks to high and low mass of the main peak are also due to proteoforms present in the sample, such as in the case of NIST mAb, which has multiple different glycoforms present, and streptavidin, which has the N-terminal methionine removed on a fraction of subunits present in each tetramer. In the cases where sodium adducts could not be resolved from the adduct-free ion at the resolution setting used (i.e., CRP and NIST mAb), the mass accuracy of the adduct-free signal was not sufficiently affected, indicating that only small amounts of sodium adduction are likely present. A comparison of streptavidin in PBS analyzed with and without the buffer exchange column is shown in Fig. 3, demonstrating the performance of the P6 column and the necessity of non-volatile removal before MS analysis.

Removal of non-volatiles from samples

A variety of buffers are used during protein expression and purification. A buffer is generally chosen based on the pH range of interest, ionic strength and chemical properties to stabilize the native structure of the protein or protein complex of interest. In addition to the wide range of buffers, solution additives such as preservatives, metal chelators and cryoprotectants are often included into the biomolecule purification workflow and storage process to further stabilize and protect the protein

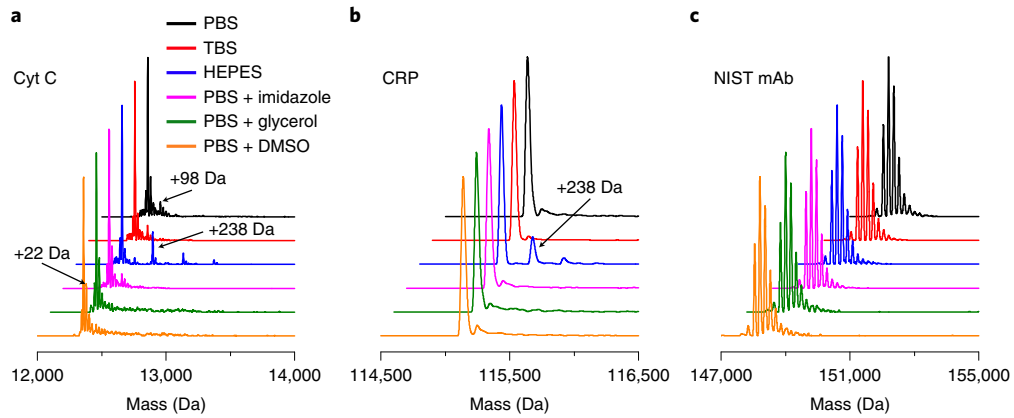

Fig. 4 | Deconvoluted mass spectra demonstrating the removal of non-volatile components from proteins in common biological buffers by OBE. Deconvoluted mass spectra of an (5 µl 4 µM) injection of cytochrome C (**a**), CRP (**b**) and NIST mAb (**c**) exchanged from various non-volatile buffers into 200 mM ammonium acetate. All spectra were acquired on an Exactive Plus EMR instrument after removal of small-molecular-weight non-volatiles using a self-packed P6 column. Common mass adducts identified are sodium (+22 Da), phosphoric acid (+98 Da) and HEPES (+238 Da). Additional adduct peaks are primarily combinations of these masses. The heterogeneity in **c** is due to the presence of various glycoforms. The y-dimension of each spectrum represents relative intensity. Spectra were deconvoluted using Intact Mass software.

of interest. Here, we demonstrate the removal efficiency by OBE of three different common buffers mimicking physiological conditions (PBS, TBS and HEPES buffer) and three different commonly used additives (glycerol, imidazole and dimethylsulfoxide (DMSO)). Cytochrome C, CRP and NIST mAb were diluted or dissolved in PBS, TBS or HEPES buffer, or in PBS with 200 mM imidazole, 20% glycerol or 20% DMSO added. The samples were buffer-exchanged online using a self-packed P6 column, and data were acquired on an Exactive Plus EMR instrument (Fig. 4). The dominant peak in each spectrum is the adduct-free protein or protein complex, with only minor signals due to small mass adducts such as sodium (+22 Da) and phosphoric acid (+98 Da). The extent of adducting on the samples prepared in TBS, as well as PBS with imidazole, glycerol and DMSO, is similar to the level of adducting present on the ions prepared in PBS only and is comparable to what would be expected for samples prepared by offline buffer exchange. The main adducts from these buffers were also sodium and phosphoric acid. No distinct adducts corresponding to Tris, imidazole, glycerol or DMSO were observed. The samples that contained 200 mM imidazole exhibit ions that are shifted to lower charge states (higher mass-to-charge ratio, *m/z*), which is consistent with imidazole having been previously reported as a charge-reducing reagent in electrospray ionization[12–15]. Interestingly, the samples in HEPES buffer displayed +238-Da mass adducts, indicating that HEPES is not as efficiently removed compared to the other buffers and additives. However, it should be noted that even in the case of HEPES, the adducted protein ions are in far lower abundance than the adduct-free protein ions, and sensitivity does not seem to be significantly lower. Mass spectra containing all charge states are shown in Supplementary Fig. 6. Overall, these experiments demonstrate that the OBE method is useful for analyzing protein samples directly from common expression, purification and storage buffer conditions.

Analysis of cell lysates

While the previous results have demonstrated the OBE method's utility for pre-purified proteins and protein complexes, in the case where a protein of interest is overexpressed, we have found that it is not necessary to carry out pre-purification steps such as affinity, size-exclusion or ion exchange chromatography. Here, we have directly analyzed a clarified cell lysate containing a protein of interest using the OBE method (Fig. 5). The results show the protein of interest in both the monomeric (32 kDa) and dimeric (64 kDa) form as the most abundant signals in the spectrum. In this case, it is clear that the protein of interest was overexpressed and is a good candidate to be screened by OBE nMS without prior purification steps. This method allows the determination of molecular weight, proteoforms and oligomeric state in <5 min after clarification of the cell lysate and could possibly be extended to use top-down tandem MS (MS/MS) to determine sequence information of the protein of interest.

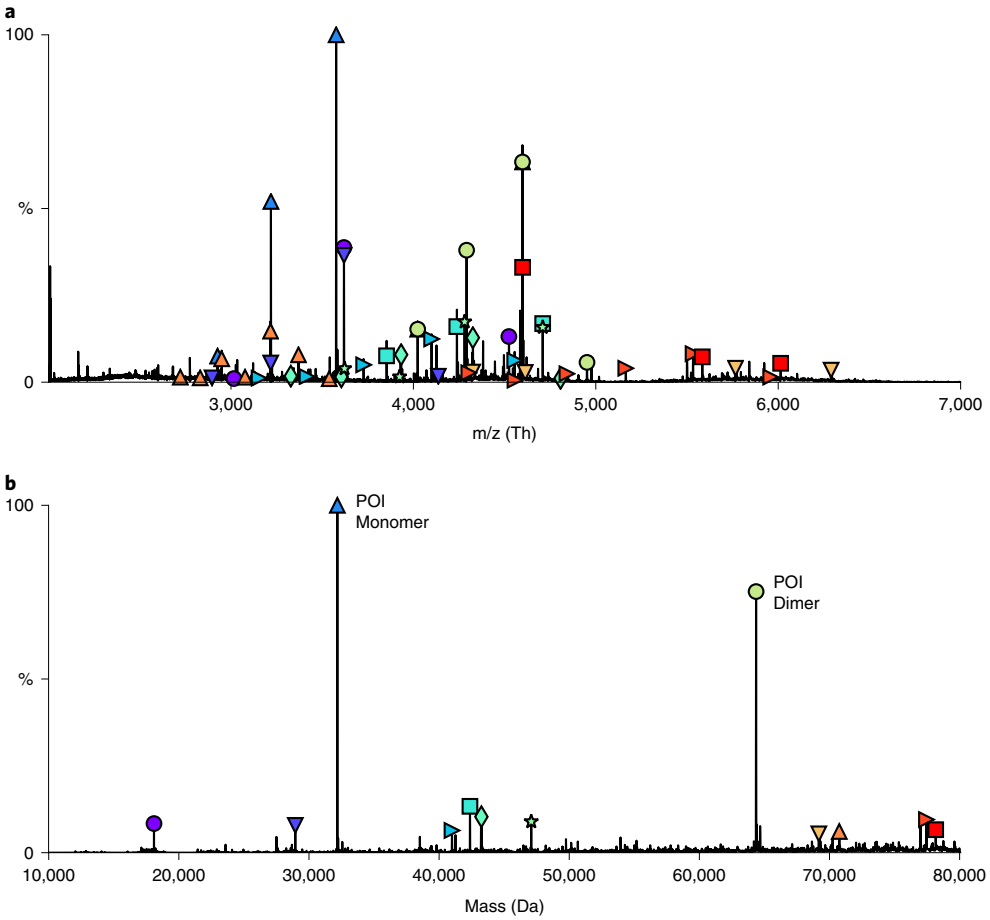

Fig. 5 | Detection of overexpressed proteins from a clarified cell lysate after OBE with a self-packed P6 column. a, Mass spectrum of a clarified cell lysate directly analyzed after online exchange to 200 mM ammonium acetate and recorded on an Exactive Plus EMR instrument. **b**, Deconvoluted (zero-charge) mass spectrum. The overexpressed protein of interest (POI) is labeled by a blue up triangle at 32 kDa (monomer) and a light-green circle at 64 kDa (dimer). All other symbols correspond to proteins in the overexpression system; no attempt was made to identify these proteins. The spectrum in **a** was deconvoluted using UniDec to produce the mass domain spectrum in **b**. Thomson (Th) is a unit of mass-to-charge ratio.

Coupling of OBE to MS

In an effort to establish the transferability of the OBE method to different MS platforms, we also analyzed three different proteins and protein complexes on a Bruker Solarix XR 15T Fourier-transform ion cyclotron resonance (FT-ICR) instrument and a Waters Synapt 'G1' HDMS quadrupole-ion mobility-time-of-flight (Q-IM-TOF) instrument. Streptavidin tetramer, CRP pentamer and NIST mAb prepared in PBS were buffer-exchanged online using a self-packed P6 column, and the results were compared to the experiments performed on the Thermo Exactive Plus EMR instrument (Fig. 6). All experiments that were recorded on the Q-IM-TOF and FT-ICR platforms produced spectra with good signal and easily resolvable charge states; however, the spectra obtained on these instruments resulted in protein ions with more adducting present than the spectra obtained on the Exactive Plus EMR platform. These results are consistent with the general trend observed when analyzing offline-desalted proteins and protein complexes by nanoESI on these instruments, which indicates that the lower amount of adducting present in spectra collected on the Exactive instrument is likely a result of more efficient desolvation and declustering of the ions in the source region of the Exactive instrument relative to the Solarix and Synapt instruments. To our knowledge, no systematic comparison of the in-source desolvation/declustering ability of different commercial MS platforms is currently available, but the use of source temperature and in-source collision voltage to clean up ions as they enter the mass spectrometer is well established in the literature[16–20].

We do note that the extra adducting present in the spectra obtained on the Synapt and Solarix instruments does not mean that OBE should not be implemented on these instruments. We

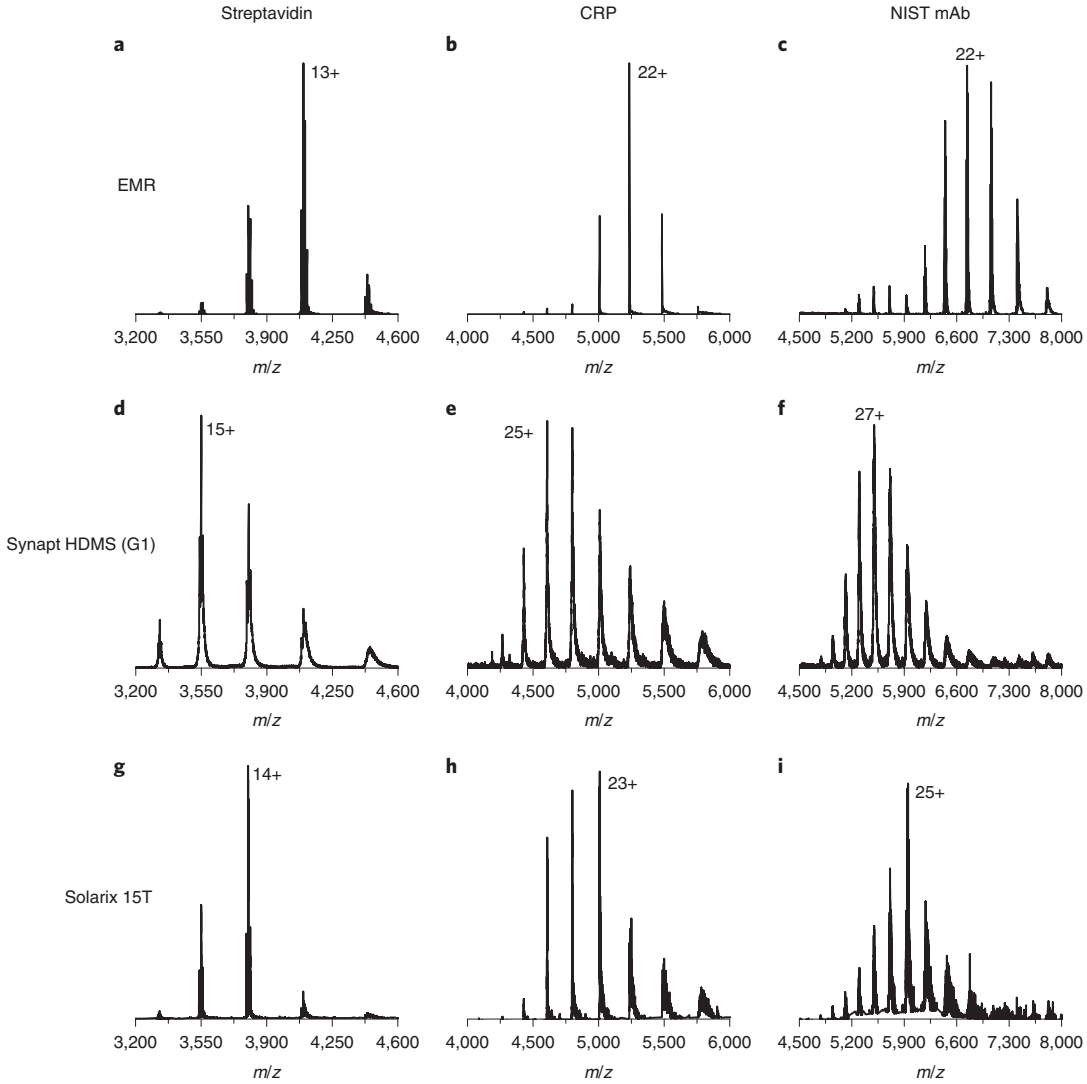

Fig. 6 | OBE coupled to different mass spectrometers. Mass spectra of streptavidin tetramer, CRP pentamer and NIST mAb were acquired on a Thermo Exactive Plus EMR mass spectrometer (**a**–**c**), a Waters Synapt 'G1' HDMS mass spectrometer (**d**–**f**) and a Bruker Solarix XR 15T FT-ICR mass spectrometer (**g**–**j**) after online exchange from PBS into ammonium acetate. Ion source temperature and collision voltage were tuned for optimal desolvation without causing dissociation or fragmentation. All proteins were present in 1× PBS before being buffer-exchanged online into 200 mM ammonium acetate with a self-packed P6 column. Differences in charge-state distributions likely result from differences in ESI probe positions and/or desolvation gas flow rates and are not indicative of structural changes of the analyte. The y-dimension of each spectrum represents relative intensity.

encourage the OBE method to be used on all three instrument platforms, especially with the high-resolution and ion mobility capabilities of the Solarix and Synapt instruments, respectively. Interestingly the charge-state distributions shifted slightly depending on which instrument was used. Although changes in charge-state distributions can indicate that conformational or structural changes to the ion have occurred (particularly in the case of increased charge)[21–23], we generally observed lower charge states by OBE compared to nanoESI and believe that the change in charge-state distributions between instruments is due to different ESI probe diameters, flow rates, probe positions and desolvation gas flow rates used on each instrument[24]. It should also be noted that the Exactive EMR instrument uses a higher source inlet temperature than the Synapt and Solarix instruments, which may explain some of the differences in charge-state distributions. Particularly, if ion formation was initially driven by charge carriers other than protons (i.e., sodium or ammonium ions), and then the ion is subsequently 'cleaned up' in source, it would explain the overall lower charge observed on the Exactive platform.

In addition to MS data (Fig. 6), it is also feasible to obtain MS/MS and mass spectrometry –ion mobility–mass spectrometry (MS/IM/MS) data using the OBE method on a Q Exactive UHMR instrument and the Synapt instrument, respectively. An example of a data-dependent acquisition

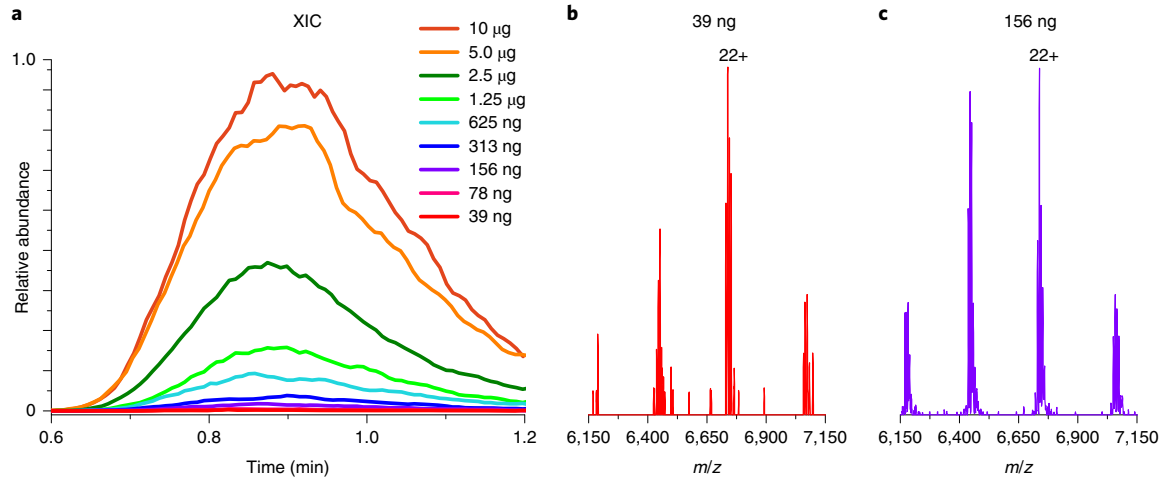

Fig. 7 | Limit of detection for OBE-MS on an EMR mass spectrometer. A dilution series from 10 µg to 39 ng NIST mAb in PBS were injected onto a self-packed P6 column and eluted with 200 mM ammonium acetate. **a**, Extracted ion chromatograms (XIC) (6,400–6,800 *m/z*) of NIST mAb. Mass spectra corresponding to 39 ng (**b**) and 156 ng (**c**) injected NIST mAb demonstrate acceptable signal to noise for OBE-MS even when low nanogram quantities are analyzed. The y-dimension of each spectrum represents relative intensity.

MS/MS experiment using CRP is shown in Supplementary Fig. 7. Within a single OBE run, a roughly 30-s protein elution window is available for precursor selection and dissociation, sufficient to perform MS/MS by preselecting the precursor *m/z* of interest (for example, in the case of routinely screening the same molecule) or by conducting a data-dependent experiment where the precursor is selected in real time by the software. However, it should be taken into account that MS/MS experiments require high precursor protein/protein complex ion signal and require the protein/protein complexes to readily dissociate by collision-induced dissociation. Although the data in Supplementary Fig. 7 were collected with a 5-µl, 4 µM injection, higher concentrations and/or chromatographic peak parking may be needed for other samples. Extending the OBE MS/MS method to use surface-induced dissociation to obtain connectivity and interface stability information of complexes, as well as using OBE MS/MS for top-down sequencing, is subject to ongoing and future work in our laboratory.

Limit of detection of the OBE method

Throughout the development of the OBE method, we have found that injecting samples of roughly 4 µM (5-µl injection) protein or protein complex results in favorable data regardless of a protein's ionization efficiency or which mass spectrometer is being used. However, we recognize that some samples are precious and difficult to obtain in such large quantities. Under these circumstances, it is often desirable to use the least amount of sample possible for screening purposes, as the remainder of the sample may be needed for additional experiments. In an effort to establish a reasonable lower concentration limit that can be analyzed using the OBE method, we conducted a set of dilution experiments with NIST mAb, online buffer-exchanged with a P6 column and acquired on an Exactive Plus EMR instrument. Figure 7a shows the extracted ion chromatogram of NIST mAb recorded at concentrations of 13 µM down to 53 nM (10 µg–39 ng loaded onto the column with a 5-µl injection volume). The charge states of NIST mAb are still well observable above the noise for the 39-ng injection (Fig. 7b), with a signal-to-noise (S/N) ratio of ~8. However, we feel that a more reasonable lower bound is ~156 ng, which results in an S/N of >50 (Fig. 7c).

Data analysis

With a routine data acquisition rate of <5 min, it is feasible to acquire mass spectra of >250 samples/24 h of instrument run time. Consequently, data analysis often becomes the rate-limiting step for OBE nMS. Many software options are available for deconvolution, analysis and reporting of data collected using the OBE method. We provide a summary below of the three most commonly used software packages in our laboratory. All three packages allow deconvolution and mass matching of detected species, making them a great option for reporting the protein identity, relative abundance, oligomeric state, heterogeneity, etc. of samples analyzed using the OBE method. A general guidance of their use is given in the Procedure.

- *Intact Mass, by Protein Metrics*. Intact Mass is a commercial software that is used for the spectral deconvolution and reporting of intact proteins as well as protein complexes, based on a parsimonious algorithm[25]. We find it particularly suitable for batch deconvolution and reporting of spectra produced by OBE screening. In addition, Intact Mass can be used with data collected on mass spectrometers from various vendors.
- *UniDec, by the laboratory of Michael Marty*. UniDec is a free and open source software suite based on a Bayesian deconvolution algorithm[26]. Deconvolution by UniDec is fast and easily implemented for mass and ion mobility spectra, with a focus on nMS data. A recently incorporated module 'MetaUniDec' also allows for high-throughput batch deconvolution of mass spectra[27]. UniDec is directly compatible with data collected on Thermo and Waters mass spectrometers and indirectly compatible with other mass spectrometer brands by first converting the raw data to mzML or .txt file format.
- *BioPharma Finder, by Thermo Scientific*. BioPharma Finder is a software used for the analysis of protein MS data for the characterization of proteins and biotherapeutics. When OBE data are acquired on a Thermo Scientific mass spectrometer, BioPharma Finder can be readily used for deconvolution and reporting of detected species.

Applications of the method

OBE nMS is particularly suitable for soluble protein and protein complex samples with masses ranging from roughly 10 to 800 kDa (we have not encountered an upper mass limit, but 800 kDa is the largest we have analyzed in our laboratory so far). The main purpose is to allow for rapid buffer exchange of sample aliquots and to obtain information on sample purity and quaternary structure, during or after the protein expression and purification process. MS/MS can be implemented for complex-down analysis (see Supplementary Fig. 7), and with appropriate amounts of sample and possible chromatographic peak parking, it may be feasible to conduct top-down protein analysis. Likewise, using an instrument with ion mobility, it may be feasible to utilize OBE nMS for automated collision cross-section determination. The rather short timescale for buffer exchange bears potential for measuring weak biomolecular interactions that would not be retained during size-exclusion chromatography[28]. Broader applications may include, but are not limited to, the analysis of RNA, DNA, (metal) cofactor-protein interactions, ligand-protein interactions, protein–nucleic acid interactions and protein-protein interactions. As protein samples in various buffers can be used for OBE nMS, this method is also useful for testing the effect of small molecules on protein and protein complex (long-term) stability.

Comparison with other methods

Information on oligomeric state and biomolecular interactions can to some extent be obtained by size-exclusion chromatography coupled with either UV detection or multiangle light scattering[29]. Whereas SEC coupled with UV detection provides only relative molecular weight information based on the apparent hydrodynamic radius, absolute molecular weights can be determined by SEC coupled with multigangle light scattering, albeit with relatively low accuracy and at low speed. Furthermore, a main disadvantage of this approach is the inability to determine distinct molecular weights of co-eluting species.

nMS is advantageous due to its ability to differentiate co-eluting species and resolve subtle mass differences such as post-translational modifications or small ligands[30]. Although several methods have been demonstrated that allow the nMS analysis of samples present in non-volatile buffers, we believe that OBE has advantages in speed, simplicity and robustness. Whereas proteins can be directly ionized from non-volatile buffers via nanoESI when small diameter tips are used[31–33], this procedure requires significant expertise and time to pull the proper tips, making it difficult to use as a routine method of analyzing dozens or even hundreds of samples. Additives[34], electrolytes[35,36] and supercharging reagents[37] can also help to counteract the effect of non-volatile buffer components on protein spectral quality, but their capability is generally limited to non-volatile concentrations lower than what would be used during protein purification, and the lack of non-volatile removal before ionization can increase the frequency of required instrument maintenance. Electrophoresis and dialysis can in principle also be used to remove small ions and small molecules, respectively[38–41]. Compared to OBE nMS, these methods have the clear advantage of a limited dilution of proteins during removal of small molecule non-volatiles. However, incomplete removal of non-volatiles

and/or a more challenging technical setup might hamper the widespread use of these methods for online salt/small molecule removal before MS.

Although the analysis of cell lysates using bottom-up[42] or top-down[43,44] MS and a combination of offline and/or online separation has become routine, it is perhaps more challenging to analyze non-covalent protein complexes directly from cell lysates. A common approach is to use extensive offline separation[45] or affinity purification[46,47] followed by proteolytic digestion and bottom-up MS; however, such a workflow does not provide a true picture of the sample at the protein complex level, as any complexes present are digested instead of being measured intact in the mass spectrometer. More recently, workflows have been developed for offline purification and nMS analysis of cell lysates to identify endogenous protein complexes[48–50]. An alternative approach to extensive offline purification that may be particularly useful for screening abundant protein complexes present in cell lysates is the introduction of the cell lysate directly into the mass spectrometer without any prior separation, under native conditions. Excellent work has recently demonstrated that intracellular and secreted proteins can be analyzed by nMS after overexpression via a so-called 'direct MS' method if non-volatile molecules are excluded in the resuspension solution and are first removed by washing the cell pellets[51–54]. The direct MS method is tailored for the analysis of cell lysates and supernatants, making it suitable for monitoring protein overexpression. In case additional purification steps are required due to low expression or weak ionization, this method typically cannot be used without a buffer exchange step due to the necessity of introducing non-volatiles (i.e., affinity chromatography requires the elution with a small, non-volatile competitor). The direct MS approach is thus complementary to OBE nMS. The focus of direct MS is on monitoring proteins during expression, whereas OBE nMS is mainly used for pre-purified proteins (albeit it is also feasible to analyze cell lysates as outlined above). In our laboratory, we often use OBE nMS for analysis of pre-purified samples, as after screening by OBE nMS, those samples can subsequently also be used in complementary biophysical characterization experiments as well as more extensive nMS measurements. As an example, we have recently shown that OBE nMS can be used to determine the quality of samples before their usage in mixing and subunit exchange experiments to determine the specificity of protein-protein interactions in complex mixtures by nMS[10].

Limitations

This protocol is specifically intended for the analysis of soluble proteins and protein complexes. Although they are areas of interest to us, we have not yet developed OBE nMS for the analysis of membrane proteins or nucleotide-protein complexes, which would require high amounts of non-volatile detergents and bivalent cations, respectively. It should be noted that the mass spectra obtained by OBE are comparable to those obtained by nanoESI after manual buffer exchange. In other words, OBE is specifically designed to be an automated, fast and efficient way of buffer-exchanging that will improve the spectral quality of samples, where heterogeneity is due to the presence of salt adducts. In contrast, OBE will not improve the spectral quality for samples where heterogeneity is due to the presence of an excess of proteoforms. However, OBE can help to readily identify protein heterogeneity and partial proteolysis and thus provide feedback to guide further optimization of protein expression and purification. In addition, because OBE does not typically provide separation between proteins present in the sample, ion suppression can become a problem with complex or heterogeneous samples. In such cases, use an SEC column that provides separation between proteins would be more beneficial.

Expertise needed to implement the protocol

Throughout this protocol, it will be assumed that the reader has a general understanding and expertise in MS as well as biological and chemical sample handling. Specifically, it is necessary to have expertise operating and tuning a mass spectrometer capable of performing nMS. Basic HPLC experience such as proper care, setup and troubleshooting is also assumed (i.e., degassing mobile phases, purging lines, flushing the system after use, etc.). Basic experience in solution preparation, sample handling, compressed gas cylinder handling and safety and interpretation of protein mass spectra is also assumed. In our experience, a knowledgeable undergraduate or graduate student can successfully and routinely perform this method. The robustness of the method makes it ideal for integration into core facilities as well as analytical divisions in pharmaceutical companies, given the availability of an HPLC and a mass spectrometer capable of transmitting and detecting high m/z ions. OBE can be easily connected and disconnected. We frequently change between OBE and direct infusion nanoESI, requiring only a few minutes for changing the source.

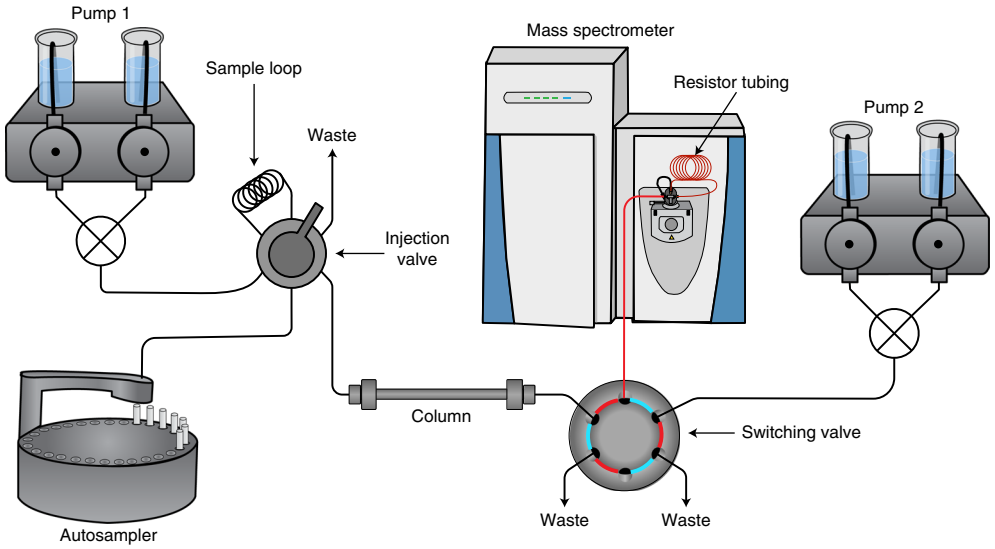

Fig. 8 | Experimental setup for OBE nMS. The sample is injected and separated from non-volatile salts by a size-exclusion column. The switching valve is used to divert salt to waste and to deliver the analyte toward the mass spectrometer via a second pump. Note that the initial position of the switching valve is designated by the red lines. The valve is switched to the second position (blue lines) for diversion of non-volatiles to waste.

Experimental design

OBE nMS can be used prior to or in parallel with additional protein characterization methods. For example, OBE can serve as a rapid method to assess protein identity, purity, oligomeric state, heterogeneity, etc. in parallel with techniques such as SDS-PAGE analysis and intact mass analysis, but before time-consuming techniques or techniques that require a large amount of sample, such as NMR spectroscopy or X-ray crystallography (see Anticipated results). OBE can be implemented at the protein expression level to monitor the production of the protein of interest, or it can be used after protein purification to assess protein quality.

In general, protein samples in common biological buffers are centrifuged to remove aggregates and are subsequently transferred into HPLC autosampler vials. Samples can be injected onto either a self-packed or a commercial short SEC column. Analytes are eluted with aqueous ammonium acetate solution. Proteins are directed to the MS, and subsequently eluting non-volatile small molecules are diverted to waste (Fig. 8).

Materials

Biological materials

- Pre-purified proteins or protein complexes (see Reagent setup). In the examples described in this protocol, we use BSA ≥96% (Sigma-Aldrich, cat. no. A2153), CRP human, recombinant (Sigma-Aldrich, cat. no. C1617), Cytochrome c from equine heart (Sigma-Aldrich, cat. no. C2506), NIST Monoclonal Antibody Reference Material (NIST, cat. no. 8671) and Chaperonin 60 (GroEL) from *E. coli* (Sigma-Aldrich, cat. no. C7688). The GroEL was refolded and prepared as described in the literature[55]. (See Applications of the method and Experimental design for general recommendations.)
- *E. coli* cell lysate (see Reagent setup). In this Protocol, we use Rosetta2 cells for protein expression, but other *E. coli* (DE3) derivatives are in principle also suitable

Reagents

- Ultrapure water (type 1) generated from a Sartorius Arium Pro water system (or suitable alternative); hereafter referred to as 'water'
- Ammonium acetate ≥99.99 trace metals basis (Sigma-Aldrich, cat. no. 431311) or ammonium acetate solution (7.5 M) (Sigma-Aldrich, cat. no. A2706)
- Methanol, LC-MS grade (Fisher Scientific, cat. no. A456) **! CAUTION** Methanol is a health hazard category 1, toxic hazard category 3 and flammability hazard category 2. Wear proper personal protective equipment (PPE) when handling and avoid contact with skin and eyes. Keep away from heat, sparks and open flame. Use per safety data sheet (SDS) recommendations.

- Bio-Rad P6 resin as spin columns or bulk resin (Bio-Rad, cat. no. 7326221 or 1504130)
- Cesium iodide (CsI) ≥99.999% trace metals basis (Sigma-Aldrich, cat. no. 203033)
- Isopropanol LC-MS grade (Fisher Scientific, cat. no. A461) **! CAUTION** Isopropanol is a flammability hazard category 2, eye irritant hazard category 2A and toxic hazard category 3. Wear proper PPE when handling and avoid contact with eyes. Keep away from heat, sparks and open flame. Use per SDS recommendations.
- Sodium phosphate, dibasic (Sigma-Aldrich, cat. no. 04272)
- Potassium phosphate, monobasic (Sigma-Aldrich, cat. no. P9791)
- Sodium chloride (Sigma-Aldrich, cat. no. S3014)
- Potassium chloride (Sigma-Aldrich, cat. no. 60130)
- Hydrochloric acid (Sigma-Aldrich, cat. no. H1758) **! CAUTION** Hydrochloric acid is a corrosive hazard category 1, eye irritant hazard category 1 and toxic hazard category 3. Wear proper PPE when handling, open only in a well ventilated area (such as a fume hood) and avoid contact with skin and eyes. Avoid unintentional reactions; hydrogen chloride can be produced as a decomposition product. Use per SDS recommendations.
- Perfluoroheptanoic acid (PFHA) (Sigma-Aldrich, cat. no. 342041) **! CAUTION** PFHA is a toxicity hazard category 4, skin irritant category 1B and eye irritant category 1. Wear appropriate PPE and handle according to SDS recommendations.
- Sodium bicarbonate (Sigma-Aldrich, cat. no. S6014)
- Acetonitrile LC-MS grade (Fisher Scientific, cat. no. A955) **! CAUTION** Acetonitrile is a flammability hazard category 2, toxic hazard category 4 and eye irritant category 2. Wear proper PPE when handling and avoid contact with eyes. Keep away from heat, sparks and open flame. Use per SDS recommendations.
- Bio-Rad Protein assay (Bradford reagent; Bio-Rad, cat. no. 5000001)
- Pierce bicinchoninic acid (BCA) Protein Assay Kit (Thermo Fisher, cat. no. 23225)
- Qubit Protein Assay Kit (Invitrogen, cat. no. Q33211)

Equipment

- Micropipettes (Eppendorf Research Plus, or similar) and appropriate tips
- Microcentrifuge tubes, 1.5 ml (Thermo Scientific, cat. no. 3448)
- Microcentrifuge capable of 21,000*g* (Thermo Scientific Sorvall Legend Micro 21 or similar alternative. Refrigerated models are recommended.)
- Assortment of volumetric flasks for solution preparation
- Glass bottles for buffers and mobile phases
- Nanodrop 2000c spectrophotometer (Thermo Scientific)
- Qubit fluorometer (Thermo Fisher)
- Glass funnel and filter flask for filtering of mobile phase
- Polytetrafluoroethylene (PTFE) membrane filters, 0.2 μm (Millipore, cat. no. JGWP04700)
- Ultrasonicator for degassing of mobile phases
- PEEK tubing, 0.005 inch i.d. (Sigma-Aldrich, cat. no. Z227307)
- PEEK tubing, 0.03 inch i.d. (Sigma-Aldrich, cat. no. Z226955)
- Tubing cutter (Sigma-Aldrich, cat. no. 57665-U)
- PEEK finger-tight fittings (Upchurch Scientific, cat. no. F-120x)
- Precolumn filters (Sigma-Aldrich, cat. no. 55215-U)
- Column packing station (Proxeon Biosystems, cat. no. SP036 or similar) **▲ CRITICAL** We use a Proxeon Biosystems packing station, however any packing station with a stirring function and ferrules to fit PEEK tubing can be used. An example of a possible alternative is cat. no. PC77-MAG from Next Advance (https://www.nextadvance.com/product/pressure-injection-cell/).
- Micro stir bar (Fisher Scientific cat. No. 14-513-63SIX)
- Compressed nitrogen cylinder with appropriate gas regulator capable of providing several hundred psi of pressure
- Dual pump HPLC system (Dionex/Thermo Scientific Ultimate 3000 RSLC series or similar) **▲ CRITICAL** We use an Ultimate 3000 liquid chromatograph; however, any liquid chromatograph capable of providing ~50–150 μl/min can be used.
- Short SEC columns (choose one of the following): Acclaim SEC-300 4.6 × 33 mm (Thermo Scientific, cat. no. 01425030), Acquity UPLC BEH125 4.6 × 30 mm (Waters, cat. no. 186006504) or Yarra SEC-3000 2.1 × 50 mm (Phenomenex prototype column) **▲ CRITICAL** Although we used the SEC columns

listed here, any short SEC column can be used for OBE. A column with an exclusion limit below the molecular weight of the analyte of interest is generally best for buffer exchange.
- Autosampler vials (Waters, cat. no. 186000384c or similar)
- Mass spectrometer capable of high *m/z*-range transmission and detection. We use an Exactive Plus EMR Orbitrap instrument equipped with an Ion Max ESI source and HESI-II probe fitted with a regular flow (100 µm i.d.) ESI needle (Thermo Scientific, Option 53010), a Synapt 'G1' HDMS Q-IM-TOF instrument equipped with a LockSpray ESI source and regular flow (90 µm i.d.) ESI needle (Waters, part no. 700000337) and a Solarix XR 15T FT-ICR instrument equipped with a standard ESI source and regular flow needle (150 µm i.d.) (Bruker) ▲ **CRITICAL** Although we use the three MS instruments listed here, any instrument that is capable of transmitting and analyzing the analyte of interest under native conditions can be used.
- Six-port switching valve (Idex, part no. MXT715)
- pH meter
- Analytical balance

Software
- Xcalibur version 3.0 or newer (Thermo Scientific): used to analyze data recorded on the Exactive Plus EMR instrument. https://www.thermofisher.com/order/catalog/product/OPTON-30965
- MassLynx version 4.1 or newer (Waters): used to analyze data recorded on the Synapt 'G1' HDMS instrument. https://www.waters.com/waters/en_US/MassLynx-MS-Software/nav.htm?locale=en_US&cid=513662
- Bruker Compass Data analysis version 5.0 or newer (Bruker Daltonics): used to analyze data recorded on the Solarix XR instrument. https://www.bruker.com/service/support-upgrades/software-downloads/mass-spectrometry.html
- MS deconvolution software: UniDec version 3.2.0 or newer (https://github.com/michaelmarty/UniDec/releases), Intact Mass version 3.1-19 or newer (Protein Metrics) and BioPharma Finder version 3.0 or newer (Thermo Scientific) ▲ **CRITICAL** We use the three software packages listed here, but not all three are necessary. One of these packages or a similar product can be used.

Reagent setup

Ammonium acetate mobile phase

To make 500 ml of a 200 mM ammonium acetate solution, add 7.7 g of ammonium acetate to ~300 ml of water, dissolve and then bring the final volume to 500 ml with water. Alternatively, if using a premade 7.5 M ammonium acetate stock solution, prepare 500 ml of 200 mM ammonium acetate by adding 13.3 ml to a volumetric flask and diluting with water to the calibration mark. Filter into a clean filter flask using a 0.2-µm PTFE membrane filter to remove any solids. Store at 4 °C in glass mobile phase bottles for ≤2 weeks. Degas the mobile phase solution by sonicating uncapped for 15 min before use. ! **CAUTION** Sonicating a capped bottle can cause the solution to heat up and the glass to explode. Ensure that any bottles are left uncapped. Always wear PPE such as hearing and eye protection per the manufacturer's recommendations.

PBS

To make 1 l of 1× PBS, combine 800 ml of water, 8.0 g of sodium chloride, 0.2 g of potassium chloride, 1.44 g of sodium phosphate dibasic and 0.24 g of potassium phosphate monobasic. Adjust to pH 7.4 at room temperature with hydrochloric acid. Adjust to a final volume of 1,000 ml. Store at 4 °C for ≤1 month.

CsI calibration solution

To make 5 ml of a 2-mg/mL CsI calibration solution, combine 2.5 ml of isopropanol with 2.5 ml of water. Dissolve 10 mg of CsI in the isopropanol:water solution. For best results, the calibration solution should be made fresh daily, or as needed for calibration.

PFHA calibration solution

To make a 10× stock solution, heat PFHA above its melting point of 54.3 °C and combine 1 µl of PFHA with 500 µl of isopropanol and 300 µl of 16.7 mM sodium bicarbonate. The stock solution can be stored at −20 °C for ≤1 year. To make the PFHA calibration solution, dilute the stock solution 10-fold in a 1:1 (vol/vol) isopropanol:acetonitrile solution. The calibration solution should be prepared fresh daily or as needed for calibration.

BSA stock solution

To prepare 1 ml of a 5-mg/mL BSA stock solution, combine 5 mg of BSA with 1 ml of PBS and dissolve. Divide into aliquots and store at −80 °C for ≤1 year. Before use, thaw an aliquot and centrifuge at high speed (~21,000*g*) at 4 °C for 15 min to pellet any solids to avoid column clogging.

***E. coli* cell lysate sample**

E. coli cell lysate samples can be prepared by mechanical cell lysis in a physiological buffer (we use PBS pH 7.4) after induction and protein (over)-expression. It is advantageous to perform all steps on ice to minimize proteolytic degradation of the proteins and/or protein complexes of interest. Protease inhibitors (i.e., Halt Protease and Phosphatase Inhibitor Cocktail, Thermo Fisher) can be added, but care must be taken that those do not lead to artifact formation due to protein binding or covalent protein modification. If the cell lysates are not measured immediately, it is advantageous to flash-freeze in liquid nitrogen and store at −80 °C. Cell debris can be removed by centrifugation (21,000*g*) at 4 °C for 15 min, and the clarified cell lysate can be directly used for OBE nMS.

Pre-purified protein or protein complex

To prepare a protein or protein complex sample for analysis by OBE nMS, the sample should be centrifuged at high speed to precipitate any solids, and the concentration of the sample should be measured. First, centrifuge the sample at high speed (21,000*g*) at 4 °C for 15 min, and, being careful not to disturb any pelleted precipitate, transfer the supernatant to a clean tube. Next, measure the protein concentration of the sample using a Nanodrop 2000 spectrophotometer or an assay such as Bradford, BCA or Qubit (see Estimation of protein concentration in a clarified cell lysate). The concentration of the sample should be adjusted to 1–20 μM protein or protein complex. Higher concentrations might result in partial retention of protein on the column, making it necessary to increase the regeneration time before applying the next sample. Store the sample on ice during preparation and before adding to the autosampler.

Equipment setup

Column-packing station setup

Fit a high-pressure helium or nitrogen gas cylinder with an appropriate regulator capable of delivering 100–200 p.s.i. Connect the gas regulator to the column-packing station, ensuring that all valves are safely closed. Clean the glass vial in the column-packing station that is used to hold the slurry. Fit the swage fitting on the column-packing station lid with an appropriately sized ferrule to fit the outer diameter of the PEEK tubing that will be used for column packing (usually 1/16 inch). **! CAUTION** This method uses high pressure gas; we recommend wearing safety glasses and performing the column-packing steps inside of a hood or behind an impact-resistant barrier.

HPLC setup

We use a Dionex Ultimate 3000 HPLC equipped with a 5-μl sample loop to deliver sample and mobile phase to the buffer exchange column. Use filtered and degassed 200 mM ammonium acetate as the mobile phase with a flow rate of 50–100 μl/min. The sample to be analyzed is loaded into the sample loop and injected using an autosampler by a full-loop method with an overfill factor of 1.2, or via a manual injection valve.

Coupling of the buffer exchange column, secondary pump and switching valve to the mass spectrometer

Connect the buffer exchange column to the switching valve so that flow from the column is directed to the mass spectrometer in position 1 and waste in position 2 (Fig. 8 and Supplementary Fig. 8a). Connect a secondary HPLC pump to the switching valve so that its flow of 200 mM ammonium acetate is directed to waste in position 1 and to the mass spectrometer in position 2. This configuration allows the protein of interest eluting from the column to be directed to the mass spectrometer in position 1, and the non-volatile salts eluting from the column to be sent to waste in position 2. Simultaneously, the secondary pump continues delivering the protein of interest through the switching valve to the mass spectrometer in position 2 while the non-volatile salts are being diverted to waste. Note that if a dual-pump HPLC is not available, a syringe pump with an appropriately large syringe can be used as the second pump because the pressure requirements are low.

Mass spectrometer

In this protocol, we demonstrate our approach using three different mass spectrometers: an Exactive Plus EMR Orbitrap instrument modified with a selection quadrupole and a surface-induced dissociation device[56], a Synapt 'G1' HDMS Q-IM-TOF instrument and a Solarix XR 15T FT-ICR instrument. We chose to use three instruments from different vendors to demonstrate that the OBE method is suitable for coupling with instruments from multiple vendors such as these or others. In each case, the instrument is tuned to maximize desolvation and transmission of the ions of interest. Tune settings for the Exactive Plus EMR instrument are provided in the table below:

Setting	Value
Scan range (*m/z*)	1,000–15,000
Resolution (at 200 *m/z*)	17,500
Microscans	2
AGC target	5.00×10^5
Max inject (ms)	100
Sheath gas (p.s.i.)	50
Aux gas (p.s.i.)	0
Sweep gas (p.s.i.)	0
Spray voltage (kV)	3.8
Capillary temperature (°C)	350
S-Lens RF level (V)	200
In-source dissociation (V)	10
HCD direct eV (V)	10
AGC mode	Prescan
Source DC offset (V)	40
Injection flatapole DC (V)	13
Inter flatapole lens (V)	13
Bent flatapole DC (V)	4
Trapping gas pressure setting	4

The tune settings for the Synapt and Solarix instrument can be found in Supplementary Table 2. The Synapt instrument is fitted with a Speedivalve to increase the backing and source pressures and assist in desolvation and transmission of large *m/z* ions as described by Sobott et al.[57]. The source regions of all three instruments are tuned to assist with desolvation by adjusting the source temperature, ESI gas and the in-source collision voltage.

Both the EMR and the Synapt instrument are fitted with a 10 ft. × 0.005 in. 'resistor' tube between the ESI probe and ground to reduce the electrospray current and make it possible to electrospray mobile phases with high ionic strength (Fig. 8 and Supplementary Fig. 8b). ▲ **CRITICAL** If resistor tubing is not used and ammonium acetate levels >20 mM are used as mobile phase, the electrospray current will likely exceed the maximum limit set in the instrument software, resulting in reduced sensitivity or loss of electrospray. The electrospray current as a function of mobile phase ionic strength recorded on an Exactive Plus EMR instrument is shown in Supplementary Table 3 for mobile phases up to 2 M ammonium acetate. ▲ **CRITICAL** It should be noted that a 10 ft. resistor tube is generally not necessary, and in most cases (mobile phase ionic strength <300 mM), a resistor tube of 2–3 ft. should be enough to keep the ESI current below the maximum limit while also reducing the post-column dead volume of the system; however, in the case where higher ionic strengths are necessary to retain the structural integrity of an analyte of interest, ammonium acetate concentrations in the molar range can readily be used by adjusting the length of the resistor tubing. The Solarix instrument does not require the resistor tubing, as the electrospray voltage is applied to the MS inlet rather than the ESI probe. ! **CAUTION** This method uses mass spectrometers with high-voltage electrospray sources. Ensure that the electrospray source is properly grounded. An improperly grounded electrospray source can result in high voltage being floated on the LC instrument, resulting in an electrical shock.

LC-MS method setup

The LC-MS method timing and acquisition parameters for an OBE experiment using a 12-cm long P6 column are given in the tables below. Note that these parameters and times may need to be optimized depending on the individual equipment setup and column used.

Time (min)	Steps
0	Start MS acquisition upon injection by LC
1.7	End acquisition (column flushes)
3	End method
Parameter	**Value**
Flow rate (pump 1 and 2)	100 µl/min
Injection volume	5 µl
Scan range	1,000–8,000 *m/z* or as appropriate for the analyte of interest

Procedure

Preparation of buffer exchange columns (optional) ● Timing ~60 min

▲ **CRITICAL** Preparation of buffer exchange columns is not a necessary step, as commercial options are available (see Suitable columns for OBE nMS section in the Introduction); however, self-packed columns are an economic option if you are working with unstable samples that may cause clogging. In addition, columns packed with P6 resin may perform better than commercial silica-based resins if the analyte of interest adsorbs to the silica.

1 Obtain a P6 spin column and mix well to obtain a uniform slurry. Alternatively, if using dry P6 resin, add a small amount (~250 mg) to 1.3 ml of water and mix into a uniform slurry.

2 Add 500 µl of the P6 slurry to 1.5 ml of water in the vial that came with the column-packing station (usually a standard HPLC vial).

3 Add a clean micro stir bar to the vial and place the vial in the chamber of the column-packing station. Set the stirrer to a medium speed.

4 Cut a piece of 0.03-in. i.d. PEEK tubing to ~14 cm and fit it with a finger-tight 1/16 in. male connector and pre-column filter on one end. Ensure that the filter is sufficiently tight that it will not move during the packing process.

5 Place the PEEK tubing (open end first) through the lid of the packing station. Assemble the lid onto the packing station and push the open end of the PEEK tubing down into the vial containing the slurry until it is ~3 mm from the bottom (making sure that the stir bar can move freely) (Supplementary Fig. 9a–c).

6 Tighten the lid to the column-packing station and tighten the swage nut to firmly hold the PEEK tubing in place.

! CAUTION Ensure that the packing station lid and swage nut are securely tightened before opening the gas valve. Failure to securely tighten either part could result in a dangerous release of pressure.

▲ **CRITICAL STEP** It is easy to crush thin-wall PEEK tubing if the nut is overtightened. Tighten the swage nut so that the tubing cannot be easily removed by hand, but not so tight that the tubing is crushed.

7 Set the pressure regulator to 100–200 p.s.i. and slowly open the valve on the column-packing station, being careful to keep your body and eyes clear of the packing station. Listen and visually inspect for leaks. Proper function will be indicated by a slow drip of solution (about one drop every 5 s) from the end of the column.

? TROUBLESHOOTING

8 After ~10 min, slowly relieve the pressure and inspect the column and slurry.

▲ **CRITICAL STEP** If the slurry in the packing station has gone dry, you can reform it by adding 1.5 ml of water. The packing process (Steps 5–8) can then be repeated one to two more times to ensure that the column is sufficiently packed. Narrower tubing may take longer packing times, higher packing pressure or multiple rounds of packing.

9 Trim the open end of the PEEK tubing to ~12 cm (length can be adjusted to your preference or application) and fit the open end with a finger-tight fitting and precolumn filter.

10 Attach the column to an HPLC and flush with 200 mM ammonium acetate at 50–100 µl/min for ≥30 min. Ensure that the HPLC pressure is stable (probably between 100 and 400 p.s.i. depending on the length of column) and not increasing over time.

? TROUBLESHOOTING

11 With a mobile phase flow rate of 50–100 µl/min, inject 5 µl of a 1-mg/mL BSA solution several times onto the column to passivate any sites that may adsorb protein. Flush the column with ammonium acetate for an additional 30 min.

PAUSE POINT When not in use, cap the column ends and store at 4 °C. We have found that the columns continue to perform well after >6 months when stored in this manner. If the column dries out because of a poor seal, it is advised that the column be repacked by repeating Steps 1–11.

Estimation of protein concentration

12 For estimating the protein concentration of a pre-purified protein or protein complex, follow Option A. For estimating the protein concentration of a clarified cell lysate, follow Option B. General guidance on how to prepare the protein samples or lysates can be found in the Reagent setup sections.

(A) **Estimation of protein concentration for a pre-purified protein or protein complex** ● **Timing <5 min per sample**

(i) Determine the molecular weight and estimated molar extinction coefficient from the sequence of the protein to be analyzed. Tools such as the ExPASy ProtParam (https://web.expasy.org/protparam/)[58] are useful for determining both of these values.

(ii) Select the 'Protein A280' option on the Nanodrop spectrophotometer and then select the 'Other protein (E & MW)' measurement type. Enter the molecular weight and molar extinction coefficient information on the screen.

(iii) Clean the sample pedestal by applying 2 µl of water to the lower pedestal and lowering the upper pedestal. Ensure that a liquid column forms between the pedestal and let sit for 2–3 min. Wipe with a lint-free laboratory wipe.

(iv) To blank the spectrophotometer, apply 1–2 µl of the sample buffer, lower the pedestal and select the blank option on the screen. After blanking is complete, wipe the pedestal and apply a fresh drop of buffer. Analyze the buffer drop as if it were a sample by choosing the measure button on the screen. If the resulting spectrum has minimal absorbance (<0.04 *A*), the blank was successful. If higher absorbance is observed, re-clean the pedestal and repeat the blanking procedure.

(v) To measure the protein concentration of your sample, pipette 1–2 µl of sample onto the lower pedestal and lower the upper pedestal. Ensure that a liquid column has formed successfully between the pedestals. Click the measure button and wait for the measurement to be completed. The resulting concentration value can be converted from milligrams per milliliter to molar concentration using the protein molecular weight information. Note: if the monomer molecular weight was used for the measurement of a protein complex, be sure to divide the resulting molar concentration by the oligomeric state of the protein complex to obtain the concentration of protein complex in the sample.

(B) **Estimation of protein concentration in a clarified cell lysate** ● **Timing ~45 min total**

▲ **CRITICAL** The protein concentration of a clarified cell lysate should be determined by a colorimetric assay to prevent interference from other biomolecules (DNA/RNA). Several microliters of clarified cell lysate will be consumed for the measurements.

(i) Choose the colorimetric assay based on requirements (sensitivity and compatibility). The following table shows three commonly used assays that are commercially available.

Assay	Quantitation limit	Main advantage
Bradford	20–2,000 µg/ml	Compatible with reducing agents
BCA	20–2,000 µg/ml	Compatible with detergents
Qubit	0.25–5 µg/ml	High sensitivity

(ii) Prepare a dilution series of BSA or other standard protein like IgG for a working range matching the quantification limit range.

(iii) Add the reagent according to the manufacturer's instructions; depending on the specific assay used, different incubation times are required for color development. The protein concentration can be determined based on the absorbance relative to that of a standard dilution series.

OBE ● Timing 20–30 min for determining the switching valve trigger time, ~5 min per sample for screening

13 *Determining switching valve trigger time (Steps 13–18).* Start by connecting the HPLC, column, switching valve and mass spectrometer as shown in Fig. 8.

14 Specify the switching valve method. The switching valve method in the table below will serve as a good starting point; however, the precise trigger time for the switching valve to divert non-volatile salts to waste will depend on the dead volume of the system, flow rate, the column length and specific retention times.

Time (min)	Steps
0	Pump 1: 100 µl/min, valve position 1–2 (column to MS), inject, start acquisition
0.85	Start pump 2: 100 µl/min
0.9	Switch valve to position 1–6 (column to waste)
1.7	End acquisition
1.8	Stop pump 2
3	Valve to position 1–2 (column to MS), end method

15 With the mass spectrometer set to start acquisition upon injection, inject 5 µl of a 5 µM BSA solution.

16 Observe as the BSA elutes into the mass spectrometer. Stop the acquisition and turn off the electrospray voltage as the salt starts to elute to avoid spraying non-volatile small molecules into the mass spectrometer.

? TROUBLESHOOTING

17 Set up a new LC-MS method with the switching valve set to trigger two-thirds of the way through the BSA peak from Step 16.

▲ CRITICAL STEP The precise timing of the switching valve relative to the detection of the BSA peak will depend on the dead volume in the system between the switching valve and the ESI source. With longer 'resistor' tubing, the switching valve will need to be triggered earlier relative to the detection time of the BSA peak.

18 Repeat Steps 15–17 until the switching time of the valve is optimized such that the BSA peak elutes without any non-volatile salt entering the mass spectrometer.

▲ CRITICAL STEP If a P6 column is used, the timing of the switching valve should not need to be further modified for different samples, as there is no significant separation between different-sized proteins (all proteins above 6 kDa are above the exclusion limit). However, if a different column is used, the timing of the switching valve may need to be slightly modified relative to this test with BSA due to differences in protein elution time.

▲ CRITICAL STEP It is important to divert all non-volatiles away from the mass spectrometer (to waste). If non-volatiles enter the mass spectrometer, it can lead to reduced sensitivity, spectral contamination and extensive downtime for instrument cleaning.

19 *Screening of proteins, protein complexes and clarified lysates (Steps 19–23).* Adjust all samples to 1–20 µM protein or protein complex using the mobile phase buffer or the buffer that the sample is already in. The lower the concentration that is used, the less carryover and the shorter the amount of time needed for flushing the column between runs.

20 Ensure that the mass range and tune parameters in Equipment setup and Supplementary Table 2 are amenable to the samples that will be injected, and, if not, adjust.

21 Load samples into LC vials and place in autosampler. If possible, cool the autosampler to 4–8 °C whenever samples are present.

22 Set up LC-MS method and switching valve method as in Equipment setup and Step 14, and add time for flushing of salt to waste between runs (adjust the total method time to be longer if samples are concentrated and more extensive flushing is needed between samples.)

23 Set up the sample sequence and vial position for each sample that needs to be analyzed and run the sequence. Observe the first couple of runs to ensure that the signal is appropriate, the switching valve is diverting salt to waste and the column is adequately flushed between runs.

? TROUBLESHOOTING

Table 1 | Deconvolution parameters for Intact Mass

Parameter	Value	Notes
'Basic' parameters		
Mass range	10,000–160,000	Adjust to mass range of your samples.
m/z range	600–9,000	Adjust to *m/z* range of your acquisition.
Minimum difference between mass peaks	15 (Da)	–
Maximum number of mass peaks	10	Increase if multiple species or proteoforms are present in one spectrum.
'Advanced' parameters		
Charge vector spacing	0.2	A larger value (1–2) may work better for nMS with broad *m/z* peaks.
Baseline radius (*m/z*)	15	Controls the stiffness of the baseline. Larger values (≥100) may be needed for nMS with broad *m/z* peaks.
Smoothing Sigma (*m/z*)	0.02	–
Spacing (*m/z*)	0.04	For nMS, higher values (0.05–0.1) can generally be used and will speed processing time.
Mass spacing (0.5)	0.5	Controls the spacing of points in the neutral mass spectrum. For spectra without isotopic resolution, a value of 0.2–1 is best for target molecules <200 kDa. Spacing of ≥10 is best for targets >300 kDa.
Iteration maximum	10	–
Charge range	3–35	Adjust to include the general charge range of species of interest.

Data analysis

24 Choose one or more data analysis software packages (options A–C) to deconvolute and process mass spectra. Intact Mass (Option A) supports data from all MS instrument vendors, provides the ability to batch process spectra and easily export them as reports. UniDec (Option B) is a freely available and open source software package for deconvolution of MS and IM-MS data. BioPharma Finder (Option C) is a software package sold by Thermo Scientific that can be used for the deconvolution and analysis of protein mass spectra acquired on a Thermo mass spectrometer.

(A) **Intact Mass by Protein Metrics ● Timing 10–90 minutes, depending on the number of data files and complexity of spectra**

(i) Open Intact Mass and select 'New Reference Project'.

(ii) Select and drag the acquired raw files into the sample input screen.

(iii) Add protein sequences under the 'protein input screen' by browsing for FAST-All (FASTA) files or by adding a row and importing each sequence manually. Alternatively, if the mass of each sample is known, import them as a comma-separated values (csv) file along with the protein name under the 'sample-protein input' screen (see csv template in Supplementary Table 4). Importing sequences or masses will allow for automatic mass matching and assignment of the deconvoluted signals.

(iv) Set deconvolution parameters under the 'Deconvolution' tab. Specific parameter values will depend on the types of samples being analyzed (mass, charge, resolution, etc.), but a good starting point for all parameters can be found in Table 1.

(v) If protein masses or sequences were included, check 'reference' under the 'Mass Matching' so that deconvoluted peaks will be matched to theoretical masses.

(vi) Check or uncheck common post-translational modifications (PTMs) if you would like them to be considered in the mass matching process.

(vii) Set the match tolerance to your preferred value. 6 Da is a good starting point for native spectra on a high-resolution instrument, but a larger value may need to be used for data collected on lower-resolution instruments.

(viii) If you wish to calculate the areas of each deconvoluted species, check 'compute areas of mass peaks' and set the integration width.

(ix) If a P6 column was used for the OBE, all data should have the same elution time. To speed up the deconvolution calculations, under the 'sample input' click the total ion chromatogram (TIC) button and under 'peak smoothing width' choose 'disable (single peak)'. This tells the software not to look for multiple peaks in the TIC. Next, under the 'Advanced' menu, type the following:
- [ElutionPeaks]
- ConstraintStartTime = X.X
- ConstraintEndTime = X.X

where X.X is replaced with the start time and end time of the elution peak in the TIC of all acquisitions. This tells the software to calculate the data only within the specific elution profile selected.

(x) Save the reference project by selecting 'save preset' and then start the deconvolution by selecting 'create'.

? TROUBLESHOOTING

(B) **UniDec ● Timing 10–20 minutes, depending on the number of data files and complexity of spectra**

(i) Unzip the downloaded UniDec release and open the folder. No installation is needed.

(ii) Open the UniDec launcher by clicking on GUI UniDec.exe and selecting the UniDec module.

(iii) Import individual spectra by selecting 'open' under the file dropdown menu (x y list, mzML or Thermo Raw format), by selecting 'open waters raw file' for Waters data or by selecting 'get spectrum from clipboard' if you have copied the spectrum list.

(iv) Select 'presets' from the file dropdown menu and choose the preset that best matches your collected data (low-resolution native, high-resolution native, isotopic resolution, etc.).

(v) Set the *m/z* range of interest and select 'process data'. Note that additional options (baseline subtraction and smoothing) are available under the data-processing tab but generally do not need to be adjusted if the appropriate preset option (step III) is used.

(vi) Set the appropriate charge range of all species present in the data (an estimate is okay; just make sure that all species fall within the range (i.e., make the range wider than you expect)).

(vii) Set the mass range to include the mass of all species possibly present in the data.

(viii) Select 'Run UniDec' to start the deconvolution process. After deconvolution has finished, a mass domain (zero-charge) spectrum is produced, as well as a charge versus *m/z* and charge versus mass plot. Ensure that the fitted data (shown as red in the original mass spectrum) align well with the original data. If they do not, the Peak Width under 'Additional Deconvolution Parameters' may need to be adjusted to better fit the data. Alternatively, the peak width tool under the 'tools' dropdown menu can be used. Inspect the mass domain spectrum and ensure that all species seem reasonable.

? TROUBLESHOOTING

(ix) Set an appropriate peak detection range (width between labeled peaks) and peak detection threshold (the threshold for labeling of peaks as a fraction of the most intense peak) and then select 'Peak Detection' to label the calculated species onto the original mass spectrum. Check that the assignments look appropriate. If the assignments do not look appropriate, it may be necessary to adjust some of the additional deconvolution parameters; however, in our experience this is often not necessary.

(x) Obtain additional information on peaks by clicking the 'plot peaks' button, by right clicking peaks in the list and through various tools in the Analysis menu.

(xi) Save the processed data figures by selecting 'save figure presets' from the file dropdown menu.

(xii) To batch-process spectra using UniDec, open the HDF5 Import Wizard on the UniDec Launcher page.

(xiii) Browse for a folder containing all your Raw files and select the files to convert by clicking the top file, holding shift and clicking the bottom file.

(xiv) Select 'add' to add all the files to the bottom screen.

(xv) Select 'Load All to HDF5' and write to an appropriate file location.

(xvi) Open MetaUniDec from the UniDec launcher screen.

(xvii) Select 'open' from the file dropdown and select the HDF5 file saved in step xv.

(xviii) Repeat steps iv–xi to process the data for all spectra.

(xix) Save the deconvoluted data as figures by selecting 'save figure presets' from the file dropdown menu.

(C) **BioPharma Finder ● Timing 10–90 min, depending on the number of data files and complexity of spectra**

(i) Open BioPharma Finder and select the Protein Sequence Manager.
(ii) Select 'New' to create a new protein sequence.
(iii) Provide a name and description for the sequence and select the experiment category it will be used for.
(iv) Import the protein sequence by manually copying and pasting into the 'Manual Input Protein Sequence' section, or import from a FASTA file by selecting 'Import Protein Sequence'.
(v) Set any variable modifications that may be present.
(vi) Save the protein sequence to the sequence manager.
(vii) Under the Home tab, select 'Intact Protein Analysis'.
(viii) Provide an experiment name and load one or more Thermo raw data files. If multiple files are loaded, check 'batch processing' as the result format.
(ix) Select the protein sequence(s) that should be considered for identification under the protein sequence menu.
(x) Under processing method, select the 'Default Native' method and select 'Edit Method'.
(xi) Under chromatogram parameters, set the time, scan range, *m/z* range and chromatogram type to be used for the deconvolution.
(xii) If a P6 column was used for OBE (no separation between proteins), select 'Average Over Selected Retention Time' under the Source Spectra Method window, and input the start and end time of the elution peak. If a different column was used that does result in separation between different proteins, the 'Sliding Windows' option should be used.
(xiii) Unless all peaks are isotopically resolved, select the ReSpect algorithm.
(xiv) Set the output mass range to an appropriate range for your data.
(xv) Check 'Show Advanced Parameters' and ensure that the 'Model Mass Range' and 'Charge State Range' are wide enough to contain all species in the data.
(xvi) Change the Rel. Abundance Threshold and Quality Threshold to a non-zero number to help clean up noisy data.
(xvii) Select the 'Identification' tab and set the sequence matching mass tolerance if you wish to match sequences to the deconvoluted results.
(xviii) Select the 'Report' tab and select the parameters that you wish to be included in the report. For example, figures of the deconvoluted data can be automatically saved in the reports.
(xix) Select the 'Save Method' and name the modified method. Select the Finish button.
(xx) Navigate back to the 'Intact Protein Analysis' tab, and with the newly saved method selected, select 'Add to Queue' to start the data analysis. Reports will be generated automatically as the data are processed.
(xxi) Load results by selecting the 'Load Results' tab. Each identified species can be viewed and evaluated for each raw file.
? TROUBLESHOOTING
(xxii) Save the results by selecting 'Save Result File As'.

Troubleshooting

Troubleshooting guidance can be found in Table 2.

Table 2 | Troubleshooting table

Step	Problem	Possible reason	Solution
7	Column drips too slow or too fast during packing.	Pressure used for column packing is inappropriate for the tubing size or the slurry viscosity.	Adjust pressure until the column drips about once every 5 s.
10	Pressure on the newly packed column increases over time or is unstable.	It is possible that the column bed has not settled, a frit is clogged or that the tubing was crushed during packing.	Reverse the column on the HPLC and pump at a low flow rate. Slowly increase the flow rate and observe if pressure is stable. Although uncommon, it may be necessary to repack the column.
		It is possible that the P6 resin has compressed and become unstable.	Repack a column using a lower gas pressure of ~100 p.s.i.

Table continued

Table 2 (continued)

Step	Problem	Possible reason	Solution
16,23	No protein signal on mass spectrometer	Electrospray current has exceeded maximum limit and stopped electrospray.	Most mass spectrometers have a maximum electrospray current programed into the system. If a high ionic strength mobile phase is used without proper resistor tubing, the ESI current may exceed the upper limit. Stop the experiment and add a longer piece of resistor tubing to reduce the ESI current.
	Low protein signal on mass spectrometer	High electrospray current leading to reduced sensitivity	Depending on the ionic strength of the mobile phase, the length of the resistor tubing may need to be optimized for best sensitivity. We have found that resistor tubing long enough to keep the ESI current below 50 μA seems to give the best sensitivity.
	High electrospray current	The resistor tube is too short or not connected properly.	Ensure that the resistor tube is connected between the ESI probe and ground. Use a longer resistor tube.
		Salt buildup on ESI probe tip	Sometimes if the ESI probe is not flushed sufficiently after use, salts can build up on the tip, causing high ESI currents and/or corona discharge. Ensure that the probe is flushed thoroughly with water and then methanol after every use.
23	High HPLC pressure	Precipitate from a sample has entered the column, or sample has aggregated on the column.	Sometimes it is possible to resolve this problem by disconnecting the column, reversing the direction and restarting flow at a low flow rate, letting the column exit drip into waste. After sufficient flushing time, reverse the column back to the original direction and continue the experiment. If the problem is not resolved, a new column may be necessary.
24Ax, Bviii,Cxxi	Deconvolution artifacts	Analyte does not fall within the set charge or mass range, or advanced parameters are not appropriate for experimental data.	Adjust the charge state range and mass range to include the analyte(s) present in the data. Although uncommon, if artifacts are still present, it may be necessary to change the advanced parameters.

Timing

Steps 1–11, (optional) column packing: 60 min
Step 12 Option A, estimation of protein concentration for purified proteins: <5 min per sample
Step 12 Option B, estimation of protein concentration in cell lysates: ~45 min
Steps 13–18, determination of switching valve trigger time: 20–30 min
Steps 19–23, screening of proteins: ~5 min per sample
Step 24, data analysis: 10–90 min

Anticipated results

OBE nMS can be readily set up and configured, given the availability of an analytical flow LC system and any mass spectrometer capable of transmitting and detecting high *m/z* ions as shown in Fig. 6 and Fig. 8. Columns for OBE are commercially available but can also be self-packed for a fraction of the cost as shown in the Procedure. High flow rates, relatively inert stationary phases and ESI provide sufficient robustness to allow for automated, routine measurement of hundreds of samples by nMS. Data generated by OBE nMS can be easily deconvoluted and annotated with software outlined above. OBE nMS can be used for analyzing cell lysates as well as purified proteins and protein complexes.

Cell lysates

Recent work has demonstrated the use of nMS to directly analyze cell lysates or supernatants to monitor protein expression and biomolecular interactions[51–53]. Generally, these methods require washing or buffer exchange steps prior to analysis of the sample by nanoESI. We therefore envision that these 'direct MS' methods are complementary to the OBE nMS method, as OBE will allow for automated buffer exchange of the cell lysate, bypassing the offline washing and/or buffer exchange steps. The LC ESI-MS–based OBE approach is generally less prone to clogging relative to static nanoESI due to the wide tubing and ESI needle diameter as well as the ability to increase the pressure to flush out minor aggregates. If the protein is overexpressed sufficiently, it is reasonable to determine protein identity, molecular weight, stoichiometry, modifications and possibly even top-down sequence information using the OBE nMS method. It should be noted, however, that OBE provides sufficient separation only between small molecules and proteins, whereas different proteins within a

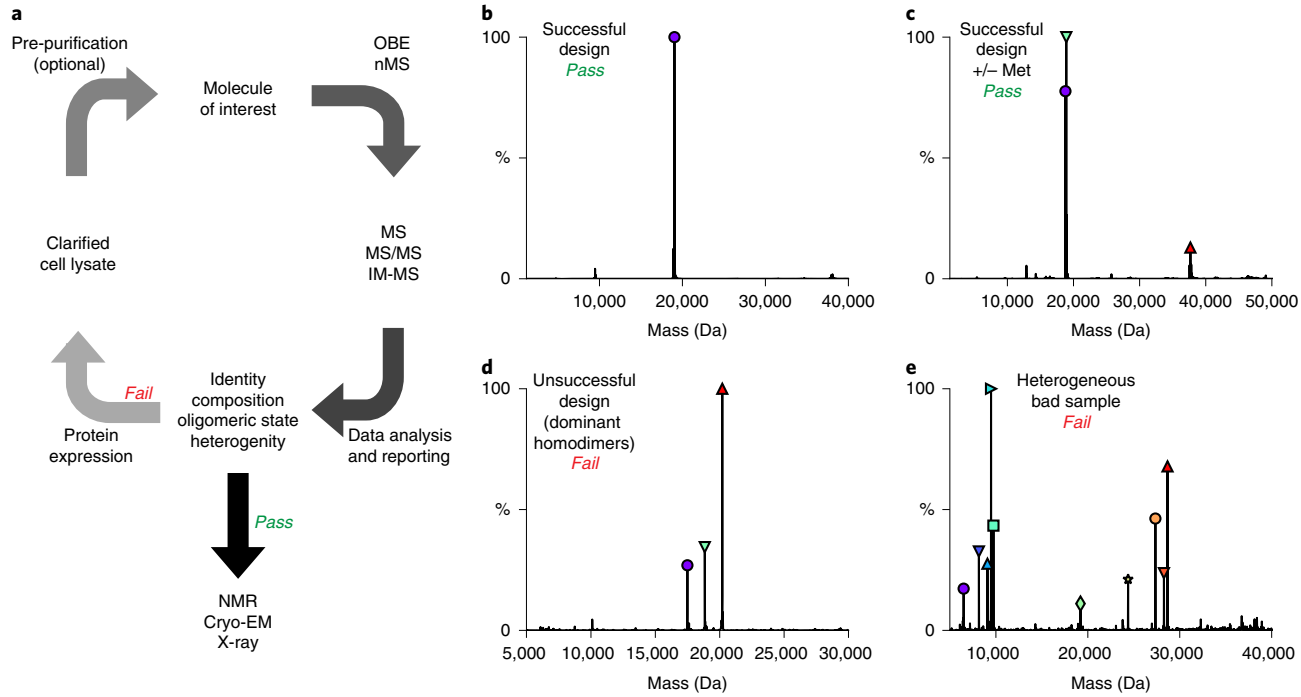

Fig. 9 | Implementation of OBE in structure-based protein screening. a, Flow chart showing the position of OBE nMS in a proposed workflow to accelerate the process from protein expression to structure determination. OBE nMS can be implemented to provide feedback on planning and execution of protein expression to optimize for more time-consuming structural biology characterization methods. **b–e**, Deconvoluted (zero-charge) mass spectra of computationally designed heterodimers screened using OBE. **b** and **c** are examples of successfully designed heterodimers that display the expected molecular weight except for partial N-terminal methionine cleavage for one of the subunits in **c**. **d**, An example of an unsuccessful design that forms homooligomers as the dominant species. **e**, An example of a heterogeneous sample where the expected heterodimer has low abundance relative to the contaminants in the sample. Spectra were deconvoluted using UniDec. The y-axis of each spectrum represents relative intensity.

sample are typically not separated. Consequently, host cell proteins can interfere with the detection of proteins of interest, if the proteins of interest are not sufficiently overexpressed or do not ionize well. Furthermore, routinely applying complex protein mixtures can decrease the column life due to some extent of protein aggregation and precipitation during exchange to MS-compatible solvent. We thus recommend using self-packed columns for this work, if budget is a concern, as they can be easily prepared in a larger quantity and changed at very low cost. It might also be necessary to increase the column regeneration time to remove smaller amounts of aggregated protein between runs.

Purified proteins and protein complexes

Commonly, proteins are expressed and purified in large quantities for subsequent usage, i.e., enzymatic and structural characterization. Frequently used buffers and additives are compatible with OBE nMS, making it possible to measure small sample aliquots without the need for prior buffer exchange. The acceleration in sample analysis can provide valuable feedback that can be used to set up a corresponding workflow (Fig. 9). In the illustrated case, we have used OBE nMS to determine the purity and oligomeric state of proteins that were designed to exclusively form heterodimers[10]. Guiding expression optimization, rapidly identifying complex formation and determining oligomeric state resulted in the identification of 94 out of 114 designs that successfully formed the anticipated heterodimer. Importantly, OBE nMS also helped to re-evaluate samples just immediately prior to further experiments to ensure that samples were not altered due to storage (i.e., by partial proteolysis). We also used OBE nMS for quality control purposes before mixing experiments to determine the specificity of the designed protein-protein interactions and now routinely use this method before more time-consuming experiments. We thus also consider OBE nMS to be a very useful tool to help rule out any artifact formation or degradation due to sample storage. In addition to full MS experiments, OBE can be used with MS/MS type experiments as well. In general, completing an MS/MS experiment will involve the same setup as a full MS experiment, with the MS method adjusted to perform MS/MS.

Reporting Summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

References

1. Structural Genomics Consortium. et al. Protein production and purification. *Nat. Methods* **5**, 135–146 (2008).
2. Arnau, J., Lauritzen, C. & Pedersen, J. Cloning strategy, production and purification of proteins with exopeptidase-cleavable His-tags. *Nat. Protoc.* **1**, 2326 (2006).
3. Nallamsetty, S. & Waugh, D. S. A generic protocol for the expression and purification of recombinant proteins in *Escherichia coli* using a combinatorial His_6-maltose binding protein fusion tag. *Nat. Protoc.* **2**, 383–391 (2007).
4. Bondos, S. E. & Bicknell, A. Detection and prevention of protein aggregation before, during, and after purification. *Anal. Biochem.* **316**, 223–231 (2003).
5. Vagenende, V., Yap, M. G. S. & Trout, B. L. Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochemistry* **48**, 11084–11096 (2009).
6. Kimple, M. E., Brill, A. L. & Pasker, R. L. Overview of affinity tags for protein purification. *Curr. Protoc. Protein Sci.* **73**, 9.9.1–9.9.23 (2013).
7. Cavanagh, J., Benson, L. M., Thompson, R. & Naylor, S. In-line desalting mass spectrometry for the study of noncovalent biological complexes. *Anal. Chem.* **75**, 3281–3286 (2003).
8. Waitt, G. M., Xu, R., Wisely, G. B. & Williams, J. D. Automated in-line gel filtration for native state mass spectrometry. *J. Am. Soc. Mass Spectrom.* **19**, 239–245 (2008).
9. Ehkirch, A. et al. An online four-dimensional HIC×SEC-IM×MS methodology for proof-of-concept characterization of antibody drug conjugates. *Anal. Chem.* **90**, 1578–1586 (2018).
10. Chen, Z. et al. Programmable design of orthogonal protein heterodimers. *Nature* **565**, 106 (2019).
11. Fekete, S., Ganzler, K. & Guillarme, D. Critical evaluation of fast size exclusion chromatographic separations of protein aggregates, applying sub-2 μm particles. *J. Pharm. Biomed. Anal.* **78–79**, 141–149 (2013).
12. Muddiman, D. C., Cheng, X., Udseth, H. R. & Smith, R. D. Charge-state reduction with improved signal intensity of oligonucleotides in electrospray ionization mass spectrometry. *J. Am. Soc. Mass Spectrom.* **7**, 697–706 (1996).
13. Mehmood, S. et al. Charge reduction stabilizes intact membrane protein complexes for mass spectrometry. *J. Am. Chem. Soc.* **136**, 17010–17012 (2014).
14. Marcoux, J. et al. Native mass spectrometry and ion mobility characterization of trastuzumab emtansine, a lysine-linked antibody drug conjugate. *Protein Sci.* **24**, 1210–1223 (2015).
15. Townsend, J. A., Keener, J. E., Miller, Z. M., Prell, J. S. & Marty, M. T. Imidazole derivatives improve charge reduction and stabilization for native mass spectrometry. *Anal. Chem.* (2019) https://doi.org/10.1021/acs.analchem.9b04263.
16. Benesch, J. L. P., Ruotolo, B. T., Simmons, D. A. & Robinson, C. V. Protein complexes in the gas phase: technology for structural genomics and proteomics. *Chem. Rev.* **107**, 3544–3567 (2007).
17. Sobott, F., McCammon, M. G., Hernandez, H. & Robinson, C. V. The flight of macromolecular complexes in a mass spectrometer. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **363**, 379–391 (2005).
18. Wen, J., Zhang, H., Gross, M. L. & Blankenship, R. E. Native electrospray mass spectrometry reveals the nature and stoichiometry of pigments in the FMO photosynthetic antenna protein. *Biochemistry* **50**, 3502–3511 (2011).
19. Rose, R. J., Damoc, E., Denisov, E., Makarov, A. & Heck, A. J. R. High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies. *Nat. Methods* **9**, 1084–1086 (2012).
20. Loo, J. A. Studying noncovalent protein complexes by electrospray ionization mass spectrometry. *Mass Spectrom. Rev.* **16**, 1–23 (1997).
21. Chowdhury, S. K., Katta, V. & Chait, B. T. Probing conformational changes in proteins by mass spectrometry. *J. Am. Chem. Soc.* **112**, 9012–9013 (1990).
22. Winston, R. L. & Fitzgerald, M. C. Mass spectrometry as a readout of protein structure and function. *Mass Spectrom. Rev.* **16**, 165–179 (1997).
23. Hall, Z. & Robinson, C. V. Do charge state signatures guarantee protein conformations? *J. Am. Soc. Mass Spectrom.* **23**, 1161–1168 (2012).
24. Li, Y. & Cole, R. B. Shifts in peptide and protein charge state distributions with varying spray tip orifice diameter in nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry. *Anal. Chem.* **75**, 5739–5746 (2003).
25. Bern, M. et al. Parsimonious charge deconvolution for native mass spectrometry. *J. Proteome Res.* **17**, 1216–1226 (2018).
26. Marty, M. T. et al. Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles. *Anal. Chem.* **87**, 4370–4376 (2015).

27. Reid, D. J. *et al.* MetaUniDec: high-throughput deconvolution of native mass spectra. *J. Am. Soc. Mass Spectrom.* (2018) https://doi.org/10.1007/s13361-018-1951-9.
28. Ren, C. et al. Quantitative determination of protein–ligand affinity by size exclusion chromatography directly coupled to high-resolution native mass spectrometry. *Anal. Chem.* **91**, 903–911 (2019).
29. Folta-Stogniew, E. Oligomeric states of proteins determined by size-exclusion chromatography coupled with light scattering, absorbance, and refractive index detectors. In *New and Emerging Proteomic Techniques* 97–112 (eds. Nedelkov, D. & Nelson, R. W.) (Humana Press, 2006). https://doi.org/10.1385/1-59745-026-X:97.
30. Lössl, P., van de Waterbeemd, M. & Heck, A. J. The diverse and expanding role of mass spectrometry in structural and molecular biology. *EMBO J.* **35**, 2634–2657, https://doi.org/10.15252/embj.201694818 (2016).
31. Susa, A. C., Xia, Z. & Williams, E. R. Native mass spectrometry from common buffers with salts that mimic the extracellular environment. *Angew. Chem. Int. Ed. Engl.* **56**, 7912–7915, https://doi.org/10.1002/anie.201702330 (2017).
32. Susa, A. C., Xia, Z. & Williams, E. R. Small emitter tips for native mass spectrometry of proteins and protein complexes from nonvolatile buffers that mimic the intracellular environment. *Anal. Chem.* **89**, 3116–3122 (2017).
33. Nguyen, G. T. H. et al. Nanoscale ion emitters in native mass spectrometry for measuring ligand–protein binding affinities. *ACS Cent. Sci.* **5**, 308–318, https://doi.org/10.1021/acscentsci.8b00787 (2019).
34. Clarke, D. J. & Campopiano, D. J. Desalting large protein complexes during native electrospray mass spectrometry by addition of amino acids to the working solution. *Analyst* **140**, 2679–2686 (2015).
35. Iavarone, A. T., Udekwu, O. A. & Williams, E. R. Buffer loading for counteracting metal salt-induced signal suppression in electrospray ionization. *Anal. Chem.* **76**, 3944–3950 (2004).
36. Flick, T. G., Cassou, C. A., Chang, T. M. & Williams, E. R. Solution additives that desalt protein ions in native mass spectrometry. *Anal. Chem.* **84**, 7511–7517 (2012).
37. Cassou, C. A. & Williams, E. R. Desalting protein ions in native mass spectrometry using supercharging reagents. *Analyst* **139**, 4810–4819 (2014).
38. Chen, Y., Mori, M., Pastusek, A. C., Schug, K. A. & Dasgupta, P. K. On-line electrodialytic salt removal in electrospray ionization mass spectrometry of proteins. *Anal. Chem.* **83**, 1015–1021 (2011).
39. Wilson, D. J. & Konermann, L. Ultrarapid desalting of protein solutions for electrospray mass spectrometry in a microchannel laminar flow device. *Anal. Chem.* **77**, 6887–6894 (2005).
40. Xu, N. et al. A microfabricated dialysis device for sample cleanup in electrospray ionization mass spectrometry. *Anal. Chem.* **70**, 3553–3556 (1998).
41. Xiang, F., Lin, Y., Wen, J., Matson, D. W. & Smith, R. D. An integrated microfabricated device for dual microdialysis and on-line ESI-ion trap mass spectrometry for analysis of complex biological samples. *Anal. Chem.* **71**, 1485–1490 (1999).
42. Zhang, Y., Fonslow, B. R., Shan, B., Baek, M.-C. & Yates, J. R. Protein analysis by shotgun/bottom-up proteomics. *Chem. Rev.* **113**, 2343–2394 (2013).
43. Meng, F. et al. Processing complex mixtures of intact proteins for direct analysis by mass spectrometry. *Anal. Chem.* **74**, 2923–2929 (2002).
44. Tian, Z. et al. Two-dimensional liquid chromatography system for online top-down mass spectrometry. *Proteomics* **10**, 3610–3620 (2010).
45. Camacho-Carvajal, M. M., Wollscheid, B., Aebersold, R., Steimle, V. & Schamel, W. W. A. Two-dimensional blue native/SDS gel electrophoresis of multi-protein complexes from whole cellular lysates: a proteomics approach. *Mol. Cell. Proteom.* **3**, 176–182 (2004).
46. Gingras, A.-C., Gstaiger, M., Raught, B. & Aebersold, R. Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* **8**, 645–654 (2007).
47. Bauer, A. & Kuster, B. Affinity purification-mass spectrometry. *Eur. J. Biochem* **270**, 570–578 (2003).
48. Skinner, O. S. et al. Native GELFrEE: a new separation technique for biomolecular assemblies. *Anal. Chem.* **87**, 3032–3038 (2015).
49. Skinner, O. S. et al. An informatic framework for decoding protein complexes by top-down mass spectrometry. *Nat. Methods* **13**, 237–240 (2016).
50. Skinner, O. S. et al. Top-down characterization of endogenous protein complexes with native proteomics. *Nat. Chem. Biol.* **14**, 36–41, https://doi.org/10.1038/nchembio.2515 (2017).
51. Catcott, K. C., Yan, J., Qu, W., Wysocki, V. H. & Zhou, Z. S. Identifying unknown enzyme–substrate pairs from the cellular milieu with native mass spectrometry. *Chembiochem* **18**, 613–617 (2017).
52. Gan, J. et al. Native mass spectrometry of recombinant proteins from crude cell lysates. *Anal. Chem.* **89**, 4398–4404 (2017).
53. Cveticanin, J. et al. Estimating interprotein pairwise interaction energies in cell lysates from a single native mass spectrum. *Anal. Chem.* **90**, 10090–10094 (2018).
54. Ben-Nissan, G. et al. Rapid characterization of secreted recombinant proteins by native mass spectrometry. *Commun. Biol.* **1**, 213 (2018).
55. Zhou, M., Jones, C. M. & Wysocki, V. H. Dissecting the large noncovalent protein complex GroEL with surface-induced dissociation and ion mobility–mass spectrometry. *Anal. Chem.* **85**, 8262–8267 (2013).
56. VanAernum, Z. L. et al. Surface-induced dissociation of noncovalent protein complexes in an extended mass range orbitrap mass spectrometer. *Anal. Chem.* **91**, 3611–3618 (2019).

57. Sobott, F., Hernández, H., McCammon, M. G., Tito, M. A. & Robinson, C. V. A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. *Anal. Chem.* **74**, 1402–1407 (2002).
58. Gasteiger, E. et al. ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res* **31**, 3784–3788 (2003).

Acknowledgements
The authors wish to thank R. Viner, A. Bailey and T. Zhang (Thermo Fisher Scientific) for assistance with the non-denaturing separations, M. Marty (University of Arizona) for assistance with UniDec, M. Bern and S.J. Skilton (Protein Metrics Inc.) for assistance with Intact Mass software, B. Rivera (Phenomenex) for helpful discussions and prototype Yarra columns, S. Thornton for assistance with figure production and S. Lai for careful proofing. The VP40 plasmid was a generous gift from the Ollmann Saphire research group (Scripps Research Institute). Work in the Wysocki laboratory was supported by National Institutes of Health Grant P41 GM128577, Ohio Eminent Scholar funds and a subaward from the University of Washington, Baker laboratory. The 15 T Bruker SolariXR FT-ICR instrument was supported by NIH Award Number Grant S10 OD018507. Design and preparation of proteins and protein complexes in the Baker laboratory was supported by the Howard Hughes Medical Institute and the generosity of Eric and Wendy Schmidt by recommendation of the Schmidt Futures program.

Author contributions
Z.L.V., F.B., M.J. and V.H.W. designed and technically developed the protocol. Z.C., S.E.B. and D.B. provided inspiration for the conception of the protocol and provided valuable ideas and feedback throughout the development and optimization of the protocol. Z.L.V., F.B., B.J.J., M.J. and A.S. performed experiments. Z.L.V. and F.B. wrote the manuscript with assistance from V.H.W. All authors discussed the results and commented on the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41596-019-0281-0.

Correspondence and requests for materials should be addressed to V.H.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 2 July 2019; Accepted: 6 December 2019;
Published online: 31 January 2020

Related links
Key references using this protocol
Chen, Z. et al. *Nature* **565**, 106–111 (2019): https://doi.org/10.1038/s41586-018-0802-y
Pyles, H. et al. *Nature* **571**, 251–256 (2019): https://doi.org/10.1038/s41586-019-1361-6
Schupfner, M. et al. *Chembiochem* **20**, 2747–2751 (2019): https://doi.org/10.1002/cbic.201900323

nature research

Corresponding author(s): Vicki H. Wysocki

Last updated by author(s): Nov 5, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☒	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☒	☐ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Xcalibur (Thermo Scientific) Version 3.0, MassLynx (Waters) Version 4.1, and Data Analysis (Bruker) Version 5.0 were used to record mass spectrometry data.
Data analysis	UniDec Version 3.2 and Intact Mass (Protein Metrics) version 3.1-19 were used for deconvolution of mass spectrometry data.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculation was performed. Due to the fact that this protocol describes online buffer exchange methods, sample sizes for discovery purposes are not relevant.
Data exclusions	No data was excluded
Replication	All replicate experiments are successful
Randomization	Due to the fact that this protocol describes online buffer exchange methods, randomization is not relevant.
Blinding	Due to the fact that this protocol describes online buffer exchange methods, blinding is not relevant.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	The NIST monoclonal antibody reference material 8671 (Lot # 14HB-D-002) was used in this study as a standard for MS analysis.
Validation	The NIST mAb has been extensively validated. All validation information can be found at: https://www.nist.gov/programs-projects/nist-monoclonal-antibody-reference-material-8671