PEPPI-MS: Strategies to enhance the extraction of electrophoretically separated proteins from polyacrylamide gels and their application to top-down/native mass spectrometry

Ayako Takemori[1*], Lissa C. Anderson[2*], Victoria M. Harman[3], Philip Brownridge[3], David Butcher[2], Keisuke Shima[4], Daisuke Higo[5], Jun Ishizaki[6], Hitoshi Hasegawa[6], Junpei Suzuki[7], Masakatsu Yamashita[7], Joseph A. Loo[8,9], Rachel R. Ogorzalek Loo[9], Robert J. Beynon[3], and Nobuaki Takemori[10,11]

1. Department of Bioresource Production Science, The United Graduate School of Agricultural Sciences, Ehime University, Matsuyama, Ehime, Japan
2. Ion Cyclotron Resonance Program, National High Magnetic Field Laboratory, Florida State University, Tallahassee, FL, USA.
3. Centre for Proteome Research, Institute of Integrative Biology, University of Liverpool, Liverpool, UK
4. Shimadzu Corporation, Nakagyo-ku, Kyoto, Japan.
5. Thermo Fisher Scientific K.K., Yokohama, Kanagawa, Japan.
6. Department of Hematology, Clinical Immunology and Infectious Disease, Graduate School of Medicine, Ehime University, Toon, Ehime, Japan.
7. Department of Immunology, Graduate School of Medicine, Ehime University, Toon, Ehime, Japan
8. Department of Chemistry and Biochemistry, University of California-Los Angeles, Los Angeles, CA, USA
9. Department of Biological Chemistry, UCLA/DOE Institute for Genomics and Proteomics, and UCLA Molecular Biology Institute, University of California-Los Angeles, Los Angeles, CA, USA
10. Division of Proteomics Research, Proteo-Science Center, Ehime University, Toon, Ehime, Japan
11. Division of Analytical Bio-Medicine, Advanced Research Support Center, Ehime University, Toon, Ehime, Japan

*A.T. and L.C.A. contributed equally to this work.
To whom correspondence should be addressed: Nobuaki Takemori (takemori@m.ehime-u.ac.jp)

Abstract

Polyacrylamide gel electrophoresis (PAGE) is a powerful technique for separating proteins from complex biological samples. However, the difficulty in recovering proteins with high yields from polyacrylamide matrices often precludes further analyses of intact proteins. Here, we propose a novel experimental workflow named Passively Eluting Proteins from Polyacrylamide gels as Intact species for MS ('PEPPI-MS'), which allows intact mass spectrometry (MS) of PAGE separated proteins. We discovered that staining proteins with certain Coomassie brilliant blue formulations immediately after PAGE improves the efficiency of extraction in a medium with pH 7–11. Post-staining, proteins spanning a broad range of molecular weights were recovered efficiently in a 10-minute procedure. High recovery yields were also obtained from dried and archived gels. This workflow is effective for top-down proteomics analysis of the target molecular region in the gel. An alternative procedure was developed for the extraction of protein complexes exceeding 400 kDa, which were separated using native PAGE, from unstained gels. Non-covalent hemoglobin tetramer, purified from cell lysate with two-dimensional native PAGE and extracted with the mild detergent octyl-β-D glucopyranoside, was amenable for native MS analysis. We anticipate that the established workflow will facilitate the purification, storage, and transport of proteins destined for detailed characterization by MS.

KEYWORDS: Gel electrophoresis, Passive protein extraction, Coomassie brilliant blue , Top-down proteomics, Native mass spectrometry

1. Introduction

Polyacrylamide gel electrophoresis (PAGE) is a standard protein separation technique vital to biochemistry, molecular biology, and proteomics research. Performing PAGE with sodium dodecyl sulfate (SDS), an anionic surfactant, is especially useful because it separates proteins at high resolution over a broad molecular weight range. The widely-used SDS-PAGE protocol established by Laemmli[1] separates linearized protein molecules denatured by SDS, based on their size as they migrate through the cross-linked polyacrylamide mesh under the influence of an applied electric field. Alternative protocols that are performed without strong ionic detergents (blue-native PAGE, clear-native PAGE, and/or non-denaturing PAGE), allow separated proteins to retain their native structures, which is desirable when purifying and analyzing non-covalent complexes.[2-5] SDS-PAGE is often used to prefractionate complex mixtures extracted from biological samples because it enables high-resolution separation of proteins at low cost.

In-gel digestion of proteins for analysis by mass spectrometry (MS) or tandem MS (MS/MS) has been a staple of proteome characterization for decades.[6-10] Proteins separated by 1D or 2D PAGE are fixed, visualized by staining with dyes (e.g., comassie or silver stain), and bands (1D) or spots (2D) containing proteins of interest are cut from the gel. These gel pieces are then saturated in solutions containing reagents to reduce, alkylate, and destain the proteins contained within. This is followed by in-gel protease digestion (usually with trypsin) and extraction of the resulting peptides. Tryptic peptides are generally small (< 3 kDa), are easily separated by reversed-phase liquid chromatography (LC), and fragment efficiently by collision-induced dissociation methods. Thus, this approach was widely adopted as it suits the resolution and sensitivity capabilities of most MS instruments.

Intact protein or “top-down” MS methods suffer several analytical limitations compared to their “bottom-up” (tryptic peptide-based) counterparts chiefly because signal-to-noise is inversely proportional to the molecular weight (MW) of the analyte.[11] As MW increases, ion current is distributed among wider charge state and isotope distributions, which ultimately translates to lower signal intensity. Additionally, specialized (and costly), high resolution mass analyzers are required for charge state and accurate mass determination.[12,13] These challenges, coupled with the ease and success of in-gel digestion for protein characterization, obsoleted efforts to recover intact proteins from traditional SDS-PAGE gels for subsequent MS analysis for a time.

Advances in MS instrumentation,[13-20] and fragmentation techniques[21-25] have ballooned interest in intact protein analysis, and subsequently sparked efforts to improve prefractionation of complex proteoform mixtures.[26-28] Size-based prefractionation methods are crucial to the success of top-down MS. Optimal online LC and MS data acquisition parameters that are appropriate for characterization of low-MW proteins differ greatly from those required for high-MW proteins (e.g., chromatography media, resolving power, activation type, m/z range, ion counts , transients summed/spectrum, etc.). Additionally, low-MW proteins often interfere with detection of coeluting proteins of higher MW as ion trapping instruments typically used for such work have limitations in charge capacity.[26,27]

In 2008, Tran & Doucette introduced Gel-eluted liquid fraction entrapment electrophoresis (GELFrEE), a method for partitioning mixtures of proteins into discrete MW ranges by SDS-PAGE.[29,30] A tube gel is housed between two chambers for sample loading and collection, and up to 500 µg of proteins are separated based on their electrophoretic mobilities. The MW-range of recovered proteins, elution times and the resolution of the separation are affected primarily by the percentage of the polyacrylamide medium. GELFrEE fractionation systems are commercially available. Additionally, Skinner et al. have combined GELFrEE with blue native and clear native electrophoresis to perform separations without SDS to preserve native structures and separate intact protein complexes.[31] More recently, Cai et al. developed a serial-size exclusion chromatography (sSEC) method in which multiple columns with size exclusion media of different pore sizes are connected serially for prefractionation of complex protein mixtures in MS-compatible buffers.[32] This approach enables a higher resolution separation compared to conventional SEC, and when combined with online reversed-phase LC, significantly improved the detection of higher MW proteins (up to 223 kDa).

While not routinely implemented, applications involving global intact proteome and targeted top-down proteoform characterization have improved our understanding of numerous biological systems, and cemented a place for intact protein analysis by MS in the pantheon of proteomic approaches.[26,27] However, traditional SDS-PAGE is not routinely preferred for the separation of proteins prior to MS analysis of intact species, primarily because procedures for efficient intact protein recovery are lacking. Proteins separated by PAGE are tightly trapped in the gel matrix and thus, require additional treatments for their recovery, such as “active” extraction by electroelution or “passive” extraction by diffusion.[33-37] Such extraction procedures usually employ organic solvents or high concentrations of surfactants to overcome the affinities of the proteins toward the gel matrix.[38-41] However, it often

takes a long time to extract a sufficient amount of protein, and the passive extraction of high-MW proteins, such as those over 60 kDa, is usually difficult.[42] Furthermore, detergents must be thoroughly removed from the recovered solution prior to analysis.

Establishing an efficient method to draw whole proteins from the traditional gel would facilitate the preparation of samples for intact analysis by MS in almost any laboratory, as no expensive, specialized equipment would be required (LCs or GELFrEE station). In this study, we aimed to establish a novel workflow and optimize conditions for passive extraction of proteins to improve yields markedly, while ensuring compatibility with intact MS. To improve protein recovery, Sheer and Ryan developed the syringe maceration extraction method (SME), which grinds the gel by repeatedly extruding it through the syringe orifice.[43] Thorough disruption of the gel by SME was effective in improving protein recovery, and more than 90% of the 160-kDa standard protein could be recovered without the use of surfactants. However, SME was only effective in recovery from unstained gels; it involved an extreme reduction in protein recovery from gels stained with Coomassie brilliant blue (CBB), a very popular staining reagent for PAGE-separated proteins. In conventional CBB formulations, the CBB dye is dissolved in an acidic solution. In the acidic environment, proteins form electrostatic and hydrophobic bonds with CBB dyes[44] and are strongly immobilized to the surrounding gel matrix. Although such protein fixation method prevents its diffusion in the gel matrix after PAGE, it majorly contributes to the decrease of protein recovery in passive extraction. Besides, the use of organic solvent (e.g., methanol or ethanol) for preparing conventional CBB exposes the hydrophobic parts of the protein, thereby enhancing its affinity to the surrounding gel matrix. Currently, aqueous formulations that avoid organic solvents and acetic acid have become more popular.[45] Rapid and sensitive aqueous CBB reagents are commercially available (e.g. Atto's EzStain AQua), although no report of their use in passive extraction has been reported yet. In this study, we investigated passive extraction conditions for overcoming the excessive fixation of CBB-protein complexes in the gel matrix using aqueous CBB. Subsequently, we established an experimental workflow for introducing proteins, obtained by rapid passive extraction with aqueous CBB, into the mass spectrometer, and evaluated protein recovery by quantitative MS analysis. Efficacy of this workflow for pre-fractionation of complex mixtures of intact proteoforms for top-down analysis with high-resolution MS was also evaluated.

We also explored the use of octyl-β-D-glucoside ('octylglucoside'), a non-ionic mild surfactant[46], in a passive extraction protocol for the structural analysis of proteins extracted from unstained clear native

gels. CBB-based passive extraction protocols require CBB removal after protein recovery, prior to further MS analysis. Typically, organic solvents are employed to remove CBB. However, these would alter the native structure of the recovered proteins. Meanwhile, low concentrations of octylglucoside have been reported to have no discernible influence on electrospray ionization (ESI) MS of protein samples.[47] Thus, an appropriate concentration of octylglucoside for the recovery of proteins separated by native-PAGE could be expected to enable recovery with low impact on the native structure of the recovered proteins. In this study, we established a procedure to extract protein complexes from unstained clear native gels using octylglucoside and evaluated their application for native MS analysis.[48]

2. Experimental Procedures

2.1. Materials

Acrylamide, chloroform, acetonitrile, methanol, trifluoroacetic acid (TFA), formic acid, glycerol, CBB-G250 (CAS #6104-58-1), acetic acid, boric acid, phosphoric acid, pure water, and bovine serum albumin (BSA) were obtained from Wako (Osaka, Japan). *N,N'*-methylene-bis-acrylamide (Bis) was purchased from GE Healthcare (Pittsburg, PA, USA). *N,N'*-bis(acryloyl)cystamine (BAC) was purchased from PolySciences (Warrington, PA, USA). n-Octyl-β-D-glucoside ('octyl glucoside') was purchased from Dojindo (Kumamoto, Japan). Sequence grade trypsin was purchased from Promega (Madison, WI, USA). Human hemoglobin (Hb) was purchased from Sigma Aldrich (St. Louis, MO, USA). Recombinant green fluorescent protein (GFP) was purchased from Vector Laboratories (Burlingame, CA, USA). Ammonium acetate, dithiothreitol (DTT), and b-phycoerythrin were purchased from Thermo Fisher Scientific (Waltham, MA, USA). Human serum (pooled, 100 mL) was purchased from Cosmo Bio (Tokyo, Japan). iTRAQ Reagents 8-plex Kit was purchased from SCIEX (Framingham, MA, USA). The Britton-Robinson buffers (pH 3-11) were prepared from an equal mixture of 120 mM acetic acid, 120 mM boric acid and 120 mM phosphoric acid. The pH value was adjusted by addition of 1 M NaOH.

2.2. SDS-PAGE

SDS-PAGE separations used NuPAGE Bis-Tris polyacrylamide gels (4-12%) and NuPAGE MOPS running buffer (Thermo Fisher Scientific). NuPAGE LDS Sample Buffer (Thermo Fisher Scientific) was used to prepare protein samples for loading onto SDS-PAGE gels. Electrophoresis was performed at 23°C with a constant 180 V and monitored with wide-view prestained protein size marker III (Wako). After

electrophoresis, the gel (removed from its plastic cassette) was placed on a disposable plastic tray, gently shaken with 40 mL aqueous CBB solution (EzStain Aqua CBB solution, ATTO, Tokyo, Japan) for 8 minutes, and washed by incubating in 200 mL of distilled water for 30 minutes. For conventional CBB staining, the gel was stained with CBB G-250 staining solution (0.2% (w/v) CBB G-250, 40% (v/v) methanol, and 10% (v/v) acetic acid) for 30 min, destained with 10% (v/v) methanol/10% (v/v) acetic acid for 30 min, and washed for 30 min in water. Stained gels were scanned using a GELSCAN transmission scanner, followed by quantification using ImageJ software.[49] The drying of CBB-stained gels was performed using a GelAir drying system (BioRad, Hercules, CA, USA) as described in the manufacturer's instructions. Immediately before drying, the gels were immersed in a 5% (v/v) glycerol solution for 10 minutes. The gels were sandwiched between two wet cellophane sheets, fixed on a plastic frame, and heat treated for 1 hour on the gel drier. The dried gels, shielded from light in a dark box, were stored at 20°C until use.

2.3. Protein Extraction from SDS-PAGE Gels

Protein bands of interest were then excised from wet gels with a razor blade and transferred to a BioMasher disposable homogenizer tube (Nippi, Tokyo, Japan). The excised gel segments were uniformly ground for 30 seconds using a plastic pestle. To extract proteins, 500 µL of protein extraction solution A (100 mM ammonium bicarbonate, pH 8) or B (0.1% (w/v) SDS/100 mM ammonium bicarbonate, pH 8) were added to the macerated gels in the homogenizer tube and shaken vigorously (1500 rpm) at room temperature for 10 minutes on a MixMate desktop tube shaker (Eppendorf, Hamburg, Germany). After filtration through a 0.45-µm cellulose acetate membrane within a Spin-X centrifuge tube filter (Corning, Corning, NY, USA), the protein filtrate was concentrated using an Amicon centrifugal 3-kDa ultrafiltration device (Merck Millipore, Darmstadt, Germany).

2.4. Protein Precipitation

Purification of recovered proteins by methanol/chloroform/water precipitation was performed employing Wessel and Flügge's method[50] with slight modifications. Briefly, the protein sample (200–300 µL) was mixed with 600 µL methanol, 150 µL chloroform, and 400 µL water in a 1.5 mL microfuge tube, and then centrifuged at 18,000 × *g* for 3 min. Without disturbing the interface, most of the upper layer was withdrawn and discarded. Next, 400 µL methanol was added to the lower phase, which was vortexed and centrifuged at 18,000 × *g* for 3 minutes. The precipitated protein pellets were air-dried at room temperature for 30 minutes and used for subsequent analysis.

2.5. Native Gel Electrophoresis

Native PAGE- Native PAGE employed BAC cross-linked tris-acetate polyacrylamide gels (8% (w/v) acrylamide) with Novex tris-glycine native running buffer (Thermo Fisher Scientific). NativePAGE Sample Buffer (Thermo Fisher Scientific) was mixed with protein samples for loading native PAGE gels. At 4°C in a laboratory refrigerator, the gels were pre-electrophoresed at a constant 200 V for 10 minutes, after which the protein/buffer mixtures were deposited into sample wells and separated at 200 V for approximately 2 hours.

Native 2D-PAGE (Figure 7e)- A mouse T-cell pellet was prepared as described previously.[51] The cell pellet was homogenized with Novex isoelectric focusing (IEF) sample buffer (pH 3–10; Thermo Fisher Scientific) and the supernatant obtained after centrifugation was subjected to two-dimensional gel electrophoresis. Just prior to IEF, 30 µL of the T-cell protein sample (1 µg total protein/µL) was mixed with 10 µL GFP (1 µg/µL) and 40 µL of 2X Novex IEF buffer (40 mM lysine, 40 mM arginine, 30% (v/v) glycerol). Prepared samples were loaded into a Novex pH 3-10 IEF gels (5% (w/v) polyacrylamide, 2% (v/v) ampholytes) and subjected to native IEF as described in the manufacturer's instructions. After electrophoresis, the sample lane was excised from the IEF gel. The excised lane was gently incubated in Novex tris-glycine native running buffer, and transferred to the top of a BAC-crosslinked 8% (w/v) acrylamide gel for the second dimension separation by native PAGE. A shark tooth-shaped comb was used to hold the lane on the gel surface during electrophoresis. Fluorescence images of separated GFP were obtained using a FluoroPhoreStar (Anatech, Tokyo, Japan).

2.6. Protein Extraction from Native PAGE Gels

Proteins of interest were excised from unstained native PAGE gels immediately following electrophoresis. Gel segments were then transferred to disposable homogenizer tubes and macerated. Novex tris-glycine native running buffer (Thermo Fisher Scientific) containing 0.1% (w/v) octylglucoside was added to the homogenizer tube, which was gently shaken (500 rpm) for 10 minutes at room temperature. (Alternatively, 0.1% octyl glucoside in H_2O replaced the running buffer.) After filtration through a 0.45-µm Spin-X centrifuge tube filter, the recovered protein was concentrated using an Amicon centrifugal 3-kDa ultrafiltration device (Merck Millipore).

2.7. Quantitative Analysis of SDS-PAGE Gel-Recovered Proteins using iTRAQ

The compound eyes from wild-type *Drosophila melanogaster* (Canton-S) were isolated as described previously[52] and used for the isobaric tags for relative and absolute quantitation (iTRAQ) analysis. The compound eye tissues (n = 150) were homogenized in 300 µL NuPAGE LDS sample buffer (Thermo Fisher Scientific) containing 50 mM DTT. After centrifugation at 18,000 × *g* for 5 min, the supernatant was incubated at 70 °C for 10 min, followed by alkylation with 75 µL 1 M iodoacetamide for 30 minutes at room temperature. Protein concentrations were determined using a Qubit protein assay kit (Thermo Fisher Scientific) according to the manufacturer's instructions. After SDS-PAGE of the tissue extract (15 µg of total protein), the separated proteins were recovered by passive extraction using 0.1% (w/v) SDS/100 mM ammonium bicarbonate (with shaking and filtration as described in 'Protein Extraction from SDS-PAGE Gels'). Recovered proteins ('recovered sample') were purified by methanol/chloroform/water precipitation as described above, and the precipitated protein pellet was dissolved in 20 µL iTRAQ buffer containing 0.1% (w/v) of RapiGest surfactant (Waters, Milford, MA, USA). Another 15 µg aliquot of the extract (one that had not been subjected to SDS-PAGE) was dissolved in 20 µL iTRAQ buffer containing 0.1% (w/v) of RapiGest following methanol/chloroform/water precipitation and used as a control for quantitative analysis ('original sample'). After in-solution digestion of the recovered and original sample with 0.2 µg trypsin for 16 hours at 37 °C, the digested peptides were labelled with iTRAQ reagents (iTRAQ 113, 114, and 115 for original sample; iTRAQ 116, 117, and 118 for recovered sample) for 2 hours at room temperature. For labelling, each iTRAQ reagent was dissolved in 50 µl of isopropanol and added to the respective protein digests. The labelled samples were mixed with 25 µL 2% (v/v) TFA, incubated for 30 minutes at 37 °C, and combined in a 1.5 mL microfuge tube. After centrifugation at 18,000 × *g* for 10 min, the digested peptides in the supernatant were fractionated using a Pierce high pH reversed-phase peptide fractionation kit (Thermo Fisher Scientific) and purified using self-made C18 STAGE tips. The obtained peptides were dried using vacuum centrifugation, resuspended in 0.1% (v/v) TFA, and subjected to LC-MS/MS analysis.

The data-dependent iTRAQ analysis was performed using an Ultimate 3000 RSLCTM nano system (Thermo Fisher Scientific) coupled to a Fusion mass spectrometer (Thermo Fisher Scientific). The sample (10 µL corresponding to approximately 500 ng of protein) was loaded onto the trapping column (Thermo Fisher Scientific, PepMap100, C18, 75 µm X 20 mm), using partial loop injection, for 7 minutes at a flow rate of 9 µL/min with 0.1% (v/v) TFA. The sample was resolved on the analytical column (Easy-Spray C18 75 µm x 500 mm 2 µm bead) using a gradient of 96.2% A (0.1% (v/v) formic acid) 3.8% B

(79.9% (v/v) acetonitrile, 20% (v/v) water, 0.1% (v/v) formic acid) to 50% A 50% B over 30 min at a flow rate of 300 nL/min. The MS acquisition program followed the multi-notch MS3 method.[53] The program consisted of a 60,000 resolution full-scan MS scan (AGC set to 4e5 ions with a maximum fill time of 50 ms) with MS/MS using quadrupole ion selection with a 0.7 m/z window, HCD fragmentation with a normalized collision energy of 35 and LTQ analysis using the turbo scan setting and a maximum fill time of 50 msec. The 10 most intense MS2 peaks were simultaneously selected for MS3 analysis using SPS isolation, HCD fragmentation at collision energy 65 and analysis in the orbitrap using 50,000 resolution over a mass range of 100-500 m/z. The machine was set to perform as many MSMS scans as to maintain a cycle time of 3 sec. To avoid repeated selection of peptides for MS/MS the program used a 60 s dynamic exclusion window.

The data was processed using Proteome Discoverer V2.2 (Thermo Fisher Scientific) in combination with Mascot V2.6 (Matrix Science, Boston, MA, USA) search engine for protein identification. Peptide identification was performed against the reviewed entries from the UniProt *Drosophila melanogaster* reference proteome (23298 entries) using a precursor mass tolerance of 10ppm, a product mass tolerance of 0.6 Da, fixed cysteine carbamidomethyl modification and variable iTRAQ 8 plex modification of lysine and the N-terminus. Peptide and protein identifies were accepted using 1% FDR for strict filtering and 5% for relaxed filtering. Unique peptides were only accepted for quantification if all reporter channels were present with an average S:N of 10 and below a co-isolation threshold of 50%. As a systematic quantification shift was expected, no sample normalization was performed.

2.8. Protein Quantification using SILAC

Tris-glycine polyacrylamide gels for PAGE were prepared with 12% (w/v) acrylamide concentration. Electrophoresis was carried out using Bio-Rad Mini-PROTEAN Tetra Vertical Electrophoresis Cell with a PowerPac Power Supply (Bio-Rad). SDS-PAGE gels were run in duplicate with identical lanes containing 15 µg of *Saccharomyces cerevisiae* whole cell lysate in 50 mM ammonium bicarbonate. Yeast had been grown in 200 ml batch culture in minimal media (24 mM ammonium sulphate, 15 mM potassium dihydrophosphate, 2.2 mM magnesium sulphate, 1.7 mM sodium chloride, 0.6 mM calcium chloride, 0.2 mM uracil, 0.1 mM histidine, 0.8 mM leucine, 0.2 µM zinc sulphate, 0.06 µM copper sulphate, 0.2 µM boric acid, 0.06 µM potassium iodine, 0.2 µM iron III chloride, 0.3 mM inositol, 41.5 µM thiamine hydrochloride, 23.6 µM pyridoxine, 8.4 µM calcium pantothenate, 1.2 µM biotin, 1% (w/v) glucose) supplemented with unlabeled arginine (0.5 mM) and lysine (1 mM). Lysis was achieved using a Minilys

bead beater (Bertin Corp., Rockville, USA) bead beating with glass beads for 15 rounds of 30 seconds, 1 minute rest between on ice. All lysed material was recovered including cell debris. Just prior to SDS-PAGE analysis, samples were combined with 2X sample buffer (60 mM Tris buffer pH 6.8, 10% (v/v) glycerol, 2% (w/v) SDS, 0.005% (w/v) bromophenol blue, 0.1 M dithiothreitol) and heated to 95 °C for 5 minutes. SDS-PAGE was completed at room temperature with a constant voltage of 200 V for 30 minutes. One gel was stained with aqueous CBB (ATTO EzStain) for 8 minutes, and washed for 30 minutes in deionized water. The other was stained using conventional CBB (40% (v/v) methanol, 10% (v/v) acetic acid, PhastGel tablet (GE Healthcare) one per 2 L) for 60 minutes, and destained for 30 minutes (10% (v/v) methanol, 10% (v/v) acetic acid) before washing for 30 minutes in deionized water.

After washing, individual lanes were excised using a razor blade and split into six sections. Each section was transferred to a BioMasher disposable homogenizer tube (Nippi). The excised gel sections were uniformly ground for 30 seconds using a plastic pestle. Five hundred microliters of protein extraction solution (0.1% (w/v) SDS, 100 mM ammonium bicarbonate) was added to the crushed gels in the homogenizer and grinding resumed using the plastic pestle until all fragments of gel were visibly homogenized. Each sample was then shaken vigorously (400 rpm) at room temperature for 10 minutes using a Vibrax benchtop orbital shaker (IKA England LTD, Oxford, UK). The fractions for each lane were recombined, resulting in pooled extracted protein for each lane.

For stable isotope labeling with amino acids in cell culture (SILAC) experiment, three pools from each CBB stain were mixed with 15 µg of stable isotope labelled ([$^{13}C_6$] Arg, [$^{13}C_6$] Lys) yeast lysate. Yeast had been grown in 200 ml batch culture in minimal media (see above) supplemented with stable isotope labelled ([$^{13}C_6$] arginine (0.5 mM) and ([$^{13}C_6$] lysine (1 mM). Lysis was achieved using a Minilys bead beater (as above). Three control samples were created using 15 µg unlabeled (‘light’) and 15 µg stable isotope labelled (‘heavy’) yeast. One pool from aqueous CBB extraction and one from conventional CBB extraction were retained as ‘light-only’ for repeat SDS-PAGE analysis. All samples were precipitated using methanol/chloroform/water precipitation as above. Light-only samples were resuspended in 10 µl 1X sample buffer and incubated at 95 °C for 5 minutes before SDS-PAGE analysis.

All heavy/light samples were resuspended in 50 mM ammonium bicarbonate, 0.05% (w/v) RapiGest SF Surfactant (Waters) and heated 80 °C for 10 minutes. Samples were then incubated with 3 mM DTT at 60 °C for 10 minutes, followed by alkylation with 9 mM iodoacetamide for 30 minutes at room

temperature. After in-solution digestion with 0.5 µg trypsin for 16 hours at 37 °C, RapiGest was precipitated from the digest with the addition of 1 µl TFA and incubation at 37 °C for 45 minutes. After centrifugation at 13,500 × g, 4 °C for 30 minutes the cleared digest was analyzed using LC-MS. Samples were analyzed using a Q Exactive HF Quadrupole-Orbitrap Mass Spectrometer coupled to a Dionex U300 RSLCnano system (Thermo Fisher Scientific). 1.5 µl of digest (1 µg peptides on column assuming full recovery) was injected on column and analyzed using a two hour shallow gradient. SILAC data was processed using Progenesis QI for proteomics (Nonlinear Dynamics, Newcastle upon Tyne, UK) with the Proteolabels module (Omic Analytics Ltd., UK) for SILAC data and database searching using Mascot V2.6 (Matrix Science) against the Saccharomyces Cerevisiae proteome. Raw ratio light/heavy data were exported for both proteins and individual peptides. The DDA data files (SILAC and iTRAC) have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD010592.

2.9. Protein Identification by MALDI-ISD MS

A human serum sample was passed through MARS Human 14 spin column (Agilent Technologies, Santa Clara, CA, USA) according to the manufacturer's instructions and the 14 bound major protein components were used for the analysis. The bound proteins were eluted with 2 mL of MARS elution buffer (Agilent Technology) and purified by methanol/chloroform/water precipitation. After separating the 14 proteins (15 µg total protein/lane) by SDS-PAGE, the selected protein band was excised from the gel and the protein component was recovered by passive extraction with 0.1% (w/v) SDS/100 mM ammonium bicarbonate (as described in ‘2.6. Protein Extraction from Native PAGE Gels’). The recovered protein was desalted using methanol/chloroform/water precipitation and reconstituted in 5 µL of 0.1% (v/v) TFA.

A 10 mg/mL solution of 1,5- diaminonaphthalene in 50% (v/v) acetonitrile/0.1% (v/v) TFA was used as the matrix-assisted laser desorption/ionization (MALDI) matrix. The sample was spotted onto a stainless steel MALDI sample plate and the matrix solution was overlaid on the sample spot. MALDI in-source decay (ISD) analysis was performed using a SHIMADZU MALDI-8020 mass spectrometer (Shimadzu, Kyoto, Japan). The obtained MALDI spectra were processed by the MALDI Solutions Data Acquisition software V2.5.1 and searched against the SwissProt proteome database (2017_02) using the online Mascot server (Matrix Science). In the MASCOT database search, the ion peak at m/z 3689 in the obtained MALDI-ISD spectrum was used as the "virtual precursor ion" and ISD fragments lower

than m/z 3689 were considered as product ions derived from the virtual precursor ion. The following parameters were used for the database search: instruments type, MALDI ISD; taxonomy, mammals; Enzyme, none; mass values, average; protein mass, unrestricted; peptide mass tolerance, 2 Da; fragment mass tolerance, 2 Da; max missed cleavages, 0.

2.10. Top-Down Proteomics

Sample Preparation- Escherichia coli, Strain K12, lyophilized cells (Sigma Aldrich) were used for the top-down MS analysis. Prior to PAGE separation, the *E. coli* extract was mixed with 50 µL 4 X NuPAGE LDS sample buffer (Thermo Fisher Scientific), 50 µL 200 mM DTT, and 100 µL water in a 1.5 mL microfuge tube. The protein sample (6 µg of total protein/lane) was then subjected to SDS-PAGE as described in Experimental Section '2.2. SDS-PAGE'. After electrophoresis, the separated proteins were recovered by passive extraction using 0.1% (w/v) SDS/100 mM ammonium bicarbonate as described in Experimental Section '2.3. Protein Extraction from SDS-PAGE Gels'. The bands collected from total nine lanes were used for a single PEPPI experiment.

Liquid Chromatography- Following methanol/chloroform/water precipitation, *E. coli* PEPPI fractions were reconstituted in 50 µL of solvent A. These were diluted by a factor of two immediately prior to injection with solvent A and analyzed by reversed-phase LC-MS/MS. For each injection, 4 µL was loaded onto an in-house-fabricated 360 µm O.D. x 150 µm I.D. fused-silica microcapillary trap column packed 2.5 cm with PLRP-S resin (5 µm particle, 1000 Å pore, Agilent Technologies, Palo Alto, CA, USA). The LC system (Acquity M-Class, Waters, Milford, MA, USA) was operated at a flow rate of 2.5 µL/min for loading onto the trap column and washed with 95% solvent A for 10 min. Separation was achieved on an in-house-fabricated analytical column packed 17.5 cm with PLRP-S resin. Samples were gradient eluted (gradients used for each fraction given in Table S1) at a flow rate of 0.3 µL/min over 95 min. The gradients utilized solvent A: 0.3% formic acid and 5% acetonitrile in water, and solvent B: 47.5% acetonitrile, 47.5% 2-propanol, 4.7% water and 0.3% formic acid (% all expressed as v/v). Following separation, proteins were directly ionized by microelectrospray ionization (2.75 kV source voltage) using a 15 µm fused-silica PicoTip emitter (New Objective, Woburn, MA) packed with 3 mm PLRP-S resin.

Mass Spectrometry- The instrument was operated in data-dependent mode with Xcalibur software (Thermo Fisher Scientific). All spectra were collected in the ion cyclotron resonance (ICR) mass analyzer

at 21 tesla (T). Data acquisition parameters were varied based upon the expected molecular weight range of the proteins contained in each fraction. For MS1 spectra – resolving power (RP) was set to 300,000 at m/z 400; 1E6 automatic gain control (AGC) target; 3-6 microscans per spectrum; 600-2000 m/z range; source induced dissociation 20 V. For MS2 spectra – RP was set to 150,000 or 300,000 at m/z 400; 5E5 AGC target; 1-2 microscans per spectrum; 300-2000 m/z range. CID activation was performed with 10 m/z isolation width, 35% normalized collision energy, 10 ms activation period, 0.25 q, and 3-5 fragment ion fills of the multipole storage device were performed such that cumulative fragment ion targets were 1.5-2.5E6 prior to detection in the ICR cell. Dynamic exclusion was enabled with a repeat count of 1, repeat duration of 60 sec, and exclusion duration of 90 sec. Data-dependent selection was allowed over m/z 700-1400. Charge state rejection was enabled for $[M+1H]^{+}$ and $[M+2H]^{2+}$.

Data Analysis- The data (.raw files) were uploaded to the National Resource for Translational and Developmental Proteomics Galaxy [54] web portal (http://nrtdp.northwestern.edu/tdportal-request) for performing top-down proteomics database searches. This platform (TDPortal [55]) utilizes three search modes: a narrow absolute mass search (with intact mass measurement tolerance of 2.2 Da and 10 ppm fragment mass tolerance), a biomarker search (similar to traditional "no-enzyme" search) with biomarker and fragment mass tolerances set to 10 ppm, and a "find unexpected modifications" search that utilized an intact mass tolerance of 200 Da ("delta m" mode enabled) and a 10 ppm fragment mass tolerance.

2.11. Native MS Analysis

Native MS analysis of Hb standard sample was performed on a ZipChip capillary electrophoresis (CE) electrospray ionization (ESI) interface (908 Devices, Boston, MA, USA) coupled to a mass spectrometer. After native PAGE separation, separated Hb was recovered from a native PAGE gel with 0.1% (w/v) octylglucoside. The buffer in this Hb solvent was subsequently replaced with 20 mM ammonium acetate (pH 6.8) using an Amicon 30-kDa centrifugal ultrafiltration filter device (Merck Millipore). Gel-recovered Hb solution (10 µL) was injected onto a ZipChip HR chip with a capillary length of 22 cm (908 Devices). For CE separations (field strength: 500 V/cm), 10% (v/v) methanol/20 mM ammonium acetate adjusted to selected pH values (4.6, 5.5, 6.0, and 6.8) was used as the background electrolyte. For native MS analysis, the Q-Exactive HF-X orbitrap mass spectrometer (Thermo Fisher Scientific) SLens RF value was set at 150, the source CID was 75 eV, the transfer tube temperature was 125 °C,

and the HCD collision energy was 20 eV. The obtained MS data were processed using the Thermo Scientific Xcalibur software V4.1.31.9 (Thermo Fisher Scientific) and interpreted manually.

For native MS analysis of the Hb tetramer in human red blood cells (RBCs), human peripheral blood was collected from a healthy volunteer in accordance with the procedures approved by the human ethics committee of Ehime University. The peripheral blood sample was suspended in an equal volume phosphate-buffered saline, and gently layered onto 25 mL of Ficoll-Paque Plus (GE Healthcare). After centrifugation at 1500 rpm for 30 min, the RBC layer was carefully transferred to a 50-mL centrifuge tube by pipetting and was washed twice in PBS with centrifugation at 1500 rpm for 5 min. After removing the supernatant, the cell pellet (~3 x 10^9 cells) was homogenized with 500 µL of Novex IEF sample buffer (Thermo Fisher Scientific) and the supernatant obtained after centrifugation at 15000 rpm for 15 min at 4 °C was concentrated until 50 µL by ultrafiltration. After resuspension in 500 µL of Novex IEF sample buffer, the sample was again concentrated again to 50 µL by ultrafiltration. Prepared samples were loaded onto Novex pH 3-10 IEF gels (5% (w/v) polyacrylamide, 2% (v/v) ampholytes) and subjected to native IEF as described above. After IEF separation, the gel pieces containing Hb bands were cut out and Hb was recovered from the gel pieces uaing the PEPPI workflow with 0.1% (w/v) octylglucoside. The recovered Hb was further separated using Novex NativePAGE 4-16% Bis-Tris Protein gel according to the manufacturer's instructions. The gel-separated Hb was recovered again by the PEPPI workflow with 0.1% (w/v) octylglucoside. The recovered Hb was cleaned using a micro Bio-Spin size exclusion column (BioRad). For native MS analysis, the purified Hb solution was directly infused into the nano ESI source using a syringe pump at a flow rate of 1 µL/min.

3. Results and Discussion

3.1. Highly Efficient Passive Extraction after CBB Staining

We first evaluated the use of aqueous CBB to develop a novel extraction workflow that would overcome the challenges associated with in-gel protein recovery. The commercially available aqueous CBB used in this study (ATTO's EzStain AQua) consists of CBB and a stabilizing agent in water acidified with citric acid. The aqueous CBB was capable of staining the gel with low background, and 16 ng BSA could be detected after washing the gel for 30 min (Figure S1). Aqueous CBB is characterized by very short staining and de-staining times, which causes a significant reduction in overall protocol time. Introduction of the efficient gel-grinding technique is also indispensable for improving recovery. Here,

we used a disposable plastic masher to crush the gel. The inner surface of the masher and the pestle were dimpled to allow the gel to be quickly pulverized into fine pieces. To reduce sample loss by protein adsorption, all the steps from gel crushing to protein extraction were performed sequentially in the same masher tube. The recovered solution was concentrated by ultrafiltration.

The results evaluating the aqueous CBB effect in overcoming the affinities of the proteins toward the gel matrix are shown in Figures 1a and 1b. After SDS-PAGE of BSA, the polyacrylamide gel was stained with commercial aqueous CBB solution and the recovery of in-gel BSA by passive extraction with three different conditions (0.1% (w/v) SDS/100 mM ammonium bicarbonate, 100 mM ammonium bicarbonate, and pure water) was examined (Figure 1a). When passively extracting with pure water, we observed that the gel pieces remained stained after shaking for 10 minutes and no released BSA was recovered. On the other hand, extraction with 0.1% (w/v) SDS/100 mM ammonium bicarbonate released CBB dye from crushed gel pieces rapidly, and the shredded gel was destained after 10 minutes. No CBB stain was visible in the polyacrylamide residue after filtering the extract. SDS-PAGE analysis of the extraction solution demonstrated that BSA was extracted by shaking for 10 minutes (Figure 1b). We have previously reported the recovery workflow of PAGE-separated proteins, employing a dissolvable polyacrylamide gel made with BAC-crosslinker with a disulfide bond.[56] That method recovers proteins in 30 minutes by completely dissolving the BAC gel. However, passive extraction post-CBB staining recovers BSA from gels in only 10 minutes, clearly faster than gel dissolution. We evaluated extractions employing a 100 mM ammonium bicarbonate solution lacking SDS. Gels destained rapidly in the solution, and the BSA recovery was the same as that obtained with 0.1% (w/v) SDS/100 mM ammonium bicarbonate. In contrast, proteins stained with conventional CBB (containing methanol and acetic acid), were immobilized, apparently locked in place by methanol-acetic acid fixation. No BSA was extracted from such gels, despite shaking with 100 mM ammonium bicarbonate for 10 minutes (Figure 1b).

Passive extraction without SDS was also performed at different pH conditions (Figure 1c). Different pH Britton-Robinson buffers (3 to 11) were used for extraction. At pH 3 and 4, the blue color of the stain did not fade and BSA was not recovered. CBB removal increased with elevated pH, as did the amount of BSA recovered. Maximum recovery was achieved at pH 7 to 11. These results indicate that the recovery of aqueous CBB-stained proteins is pH-dependent. Next, we quantified passive extraction recoveries after aqueous CBB staining (Figures 1d and 1e), comparing recoveries immediately following

electrophoresis and from dried gels. Extraction with 100 mM ammonium bicarbonate yielded a recovery efficiency of 76 ± 5%. This recovery was superior to the previously reported recovery of BSA (59.7%) by dissolution of BAC gel. [56] Even when gels were dehydrated for 1 hour with a gel dryer and stored at 20 °C for 1 week, a recovery efficiency of 44 ± 6% was achieved. Finally, extraction using 0.1% SDS/100 mM ammonium bicarbonate yielded almost the same recovery (75± 2%) as that obtained before gel drying (72 ± 3%). CBB binds electrostatically to lysine and arginine residues of proteins in an acidic environment; however, loss of positive charge from these amino acid residues in an alkaline environment causes the dye to dissociate from the protein. Jin and Manabe had previously reported that CBB release in a strong alkaline environment (0.1% NaOH, pH 13) reduces affinity of the protein to its surrounding gel, leading to improved efficiency of passive extraction.[57] In this study, we demonstrated that CBB dye could dissociate from the aqueous CBB-stained gel even in an environment with pH 7–11. While a strong alkaline environment causes β-elimination of proteins, this study revealed that aqueous CBB could extract proteins across pH 7–11. Aqueous CBB may also be used for surfactant-free protein extraction, which is a significant advantage for direct MS analysis of recovered proteins. In case of surfactant-free extraction, it would be desirable to select the pH of the recovery buffer based on the pI of the target protein; in case of retrieval of membrane proteins, decreased recovery rate may be a concern. Therefore, use of a low concentration of surfactant (e.g., 0.1% SDS) for passive extraction of samples is desirable for proteomics analysis.

3.2. Passive Extraction Workflow for MS-based Proteomics

With aqueous CBB acting as a recovery enhancer, we developed a high-speed passive extraction workflow to overcome the challenges associated with in-gel whole protein recovery (Figure 2a). We refer to this workflow by the acronym 'PEPPI-MS' (Passively Eluting Proteins from Polyacrylamide gels as Intact species for MS) for brevity. The PEPPI-MS workflow includes 1) PAGE separation, 2) CBB staining and gel band cutting, 3) gel crushing, 4) passive extraction, and 5) purification of the recovered protein. Following passive extraction, additional sample handling steps are required to further purify the proteins and to remove CBB and other contaminants. However, losses in these pretreatment processes result in reduced final recovery rates. In this workflow, methanol/chloroform/water precipitation[50] which is effective for purification of trace protein was adopted. Figure 2b illustrates the SDS-PAGE fractionation of *Drosophila melanogaster* proteins (15 µg) and their recovery by PEPPI using 0.1% (w/v) SDS/100 mM ammonium bicarbonate. SDS-PAGE analysis of the recovered fractions revealed efficient extraction of proteins from the gel over a ~25-250 kDa molecular weight range. To

quantify PEPPI performance, we evaluated protein recovery with iTRAQ and performed tandem MS (Figures 2c and 2d). Initial and extracted protein solutions were digested (after passive extraction from the gel), labeled, mixed, and analyzed. For proteins below 100 kDa, the median protein recovery rate was 68%, and for high-MW proteins greater than 100 kDa, the median recovery rate was 57%. Among the 560 proteins identified by MS, 91% (508 proteins) were recovered at levels exceeding 50%. However, six proteins delivered recovery rates exceeding 100% (maximum 156%), suggesting that a change in protein structure due to SDS-PAGE may have increased the digestion efficiency relative to the unresolved proteins.

We further verified the recovery efficiency by proteome-wide quantification using SILAC (Figure 3). In this experiment, effectiveness of the PEPPI method for gels stained with conventional CBB solution (containing methanol and acetic acid) was also verified. After separation on 12% (w/v) SDS-PAGE, *Saccharomyces cerevisiae* protein extracts were separately stained using different CBB dyes and subjected to PEPPI (Figures 3a-3c). To quantify the recovery efficiency following PEPPI, stable isotope-labelled yeast extract was added as an internal standard to the recovered protein prior to tryptic digestion. The recovery rates following PEPPI using aqueous CBB were similar to those observed in the iTRAQ labeled *Drosophila* samples (Figure 3d). Although conventional CBB staining was expected to decrease the recovery from PEPPI because of methanol fixation, its recovery was actually somewhat better than that from aqueous CBB (Figure 3e). The addition of 0.1% (w/v) SDS improved the recovery rate from conventional CBB stained (i.e., methanol-fixed) gel samples. The release of CBB was also confirmed using gels stained with conventional CBB, in 0.1% SDS/bicarbonate ammonium solution. SDS is considered to contribute to the improvement of protein extraction by dissociating CBB from the protein and lowering affinity between the two. For the recovery of low-MW proteins, fixed gels were superior to aqueous CBB stained gels (Figure 3f), suggesting that some in-gel proteins, especially small proteins, partially diffuse away during aqueous CBB staining, as it does not employ fixation. In contrast to protocols including fixation steps, aqueous CBB is fast and requires only a single step, thus greatly shortening the time required for recovering proteins. In the following top-down proteomic analysis, we used the PEPPI workflow with aqueous CBB.

3.3. Top-Down Mass Spectrometry of Gel Recovered Proteins

Although trypsin and/or Lys-C are normally used for in-gel digestions in bottom-up proteomics, the release of intact proteins from gel matrices enables a broader range of enzymes and digestion

conditions to be employed. Moreover, sequence-analyzing intact proteins directly using MS without protease treatment can further reduce the time needed to identify proteins. In particular, highly sensitive MALDI-ISD MS is an attractive approach to apply to gel-separated proteins.[58] Figure 4 shows an example of protein identification by MALDI-ISD MS following PEPPI. The PEPPI-MS workflow enabled us to excise the proteins of interest from a band-dense area, which was effective in reducing contamination from other proteins (Figures 4a,b). In particular, selective extraction from a 1 mm-wide band was possible when using a dried gel. If the protein recovered from a single lane was insufficient for subsequent analysis, the extraction of identical bands from multiple lanes resulted in increased recovery (Figure 4c).

The effective recovery of intact proteins following high-resolution fractionation will be particularly useful for large-scale analysis by top-down MS (top-down proteomics). Most top-down proteomics approaches involve extensive offline fractionation (e.g., GELFrEE[59], high-pH/low-pH[60], and sSEC[32]). The PEPPI-MS workflow would be a powerful addition to these approaches, because of the rapid high-resolution separation performance as well as the widespread availability of traditional SDS-PAGE equipment. As proof of principle that PEPPI can be used to fractionate complex mixtures for top-down proteomic analysis, *E. coli* whole cell lysate was separated with a 4-12% Nu-PAGE gel and bands 1-4 were subjected to PEPPI (Figure 5a). Recovered proteins were sent to the National High Magnetic Field Laboratory (Tallahassee, FL, USA) for analysis by LC-MS/MS on the 21 tesla (T) FT-ICR mass spectrometer.[18,61] Each fraction was run in triplicate, for a total of 12 LC-MS/MS experiments (Figure 5b). The data (.raw files) were uploaded to the National Resource for Translational and Developmental Proteomics (NRTDP; based at Northwestern University) Galaxy web portal[54] for performing top-down proteomics searches (TDPortal)[55] and searched concurrently against forward and decoy databases (UniProt May 2016). This search resulted in identification of 323 proteoforms at a 1% false discovery rate (FDR) (Table S2). Representative data obtained from a single injection of fraction 1 are shown in Figure 6. The base peak chromatogram is shown in panel a along with single-scan MS1 spectra, which show the charge-state distributions of proteins eluting over 9 chromatographic peaks. The protein identities (given by accession number) were manually validated. Global *q*-values and sequence coverage (% Cleavages) of each protein are also indicated. Panel b shows a zoom inset of the $[M+12H]^{12+}$ charge state of DNA-binding protein HU-Alpha. These peaks are assigned with 0.7 ppm RMS error when compared to theoretical isotope distributions based upon the known elemental composition of this HU-Alpha proteoform (proteoform repository number 244781;

http://atlas.topdownproteomics.org/). A single-transient CID MS/MS spectrum of HU-Alpha is given in panel c along with a fragment map of the 65% sequence coverage derived from the spectrum (bottom).

Extensive offline fractionation is a hallmark of top-down proteomics as it enables observation of less abundant proteoforms, and reduces the complexity of the sample introduced to the mass spectrometer, thereby improving many facets of the analysis (e.g. observed S:N, required resolving power, co-isolation of precursors prior to MS/MS, etc.). However, a drawback of these approaches is that the same proteoforms are often observed across multiple fractions. This complicates label free quantitation strategies, and necessitates highly reproducible sample handling processes across multiple biological and technical replicates to ensure that observed changes in proteoform abundance are biologically affected.[28,62] A very promising aspect of the PEPPI-MS workflow for top-down proteomics observed here is the limited amount of overlap in proteoforms observed across multiple "fractions". A Venn diagram of proteoforms identified in each PEPPI fraction is shown in Figure 5c (left). While each fraction is derived from the pooling of proteins recovered from several individual PAGE gels and gel lanes, 263 of the 323 of the proteoforms identified (81%) were observed in just one fraction.

The observed molecular weight distribution of unique proteoforms identified at 1% FDR is shown in Figure 5c (right) and spans 2.3-35.2 kDa. As expected, the number of proteoforms identified decreased as a function of molecular weight. This is due in part to lower observed signal-to-noise (S:N) ratios inherent to mass spectrometric analysis of large proteins electrosprayed under denaturing conditions, as well as the need for better separation of larger proteins.[11,32] However, manual inspection of the data revealed that many of the proteoforms exhibited several iterations of mass shifts of 57 Da, corresponding to incomplete/over-complete carbamidomethylation of cysteine residues (Figure 5d). Reduction and alkylation with dithiothreitol and iodoacetamide (respectively) were performed to curb the creation of acrylamide adducts formed during the PAGE process, which we anticipated would cause a similar increase in heterogeneity. Such chemical artifacts have devastating effects in top-down proteomics, especially as molecular weight increases. Signal is divided among several m/z isotopic envelopes, decreasing S:N as well as dynamic range for data-dependent MS/MS precursor selection. Additionally, allowing variable modification database searching of top-down data dramatically increases search time and FDR. We are currently working to further optimize the PEPPI workflow for

offline fractionation of complex samples, to improve efficiency or eliminate the need for carbamidomethylation of cysteine residues, and to demonstrate its utility in quantitative top-down proteomics. These efforts will be detailed in a separate manuscript.

3.4. Native Mass Spectrometry of Gel Recovered Protein Complexes

Next, the PEPPI-MS workflow established for SDS-PAGE gels was applied to recover proteins and protein complexes separated by native PAGE. We explored the use of octylglucoside, which can be removed more easily than CBB, in a passive extraction protocol with native PAGE for protein structural analysis. In this study, a colored standard protein complex (b-phycoerythrin) was used to optimize the recovery conditions. After native PAGE separation, the gel band containing b-phycoerythrin was excised and subjected to passive extraction with 0.1% (w/v) octylglucoside (Figure 7a). The gel was shaken at a low speed (500 rpm for 10 min) to reduce the risk of denaturation. The use of octylglucoside improved b-phycoerythrin extraction, and the recovered b-phycoerythrin complex showed the same mobility as the original sample when re-separated by native PAGE (Figure 7b). We also succeeded in selectively recovering a higher MW protein complex (24-subunit complex Apoferritin, 480 kDa) (Figure 7c). The combination of native PAGE and PEPPI was also applicable to the purification of recombinant red fluorescent protein from a cell-free translation system (Figure 7d). To improve the resolution of protein separation, we combined native IEF with native PAGE and selectively recovered the target proteins from biological samples (Figures 7e and 7f). Using a similar extraction method, firefly luciferase was recovered with retention of its enzyme activity (Figure S2). Passive extraction with 0.1% (w/v) octylglucoside is effective for rapid protein recovery following native PAGE separation, and the native structure and enzymatic activity of the recovered proteins are maintained.

To demonstrate that protein complex structures were preserved after extraction with octylglucoside, we performed native MS on recovered human Hb. After buffer exchange of the recovered hemoglobin sample with 20 mM ammonium acetate (pH 6.8) by 30-kDa ultrafiltration (Figure 8a), we performed native MS analysis on-line with CE (Figure 8b). Although some adduction of the acrylamide monomer was observed in the recovered Hb, the 64 kDa Hb tetramer was detected by CE-MS in neutral pH buffer (Figures 8c and 8d). As demonstrated above (Figure 7e), two-dimensional native PAGE (native IEF + native PAGE) allows high-resolution separation of target proteins. We further evaluated whether PEPPI with octylglucoside after two-dimensional PAGE separation can be effective in the pretreatment for native MS analysis of protein complexes in crude biological samples (Figure 9a).

The Hb tetramer contained in human RBCs was selected as a test case. Two-dimensional separation demonstrated effectiveness for selective Hb recovery (Figure 9b), and no additional separation was necessary. After simple gel filtration, the recovered Hb was subjected to native MS analysis by direct infusion ESI, and the Hb tetramer could be detected (Figure 9c). Although we show a limited example of targeting an abundant protein complex, our study demonstrated that an octylglucoside-based PEPPI workflow can provide a solution for combining native PAGE gel with native MS analysis. In addition to octylglucoside, other nonionic saccharide surfactants, which are compatible with MS, such as *n*-dodecyl glucoside and *n*-hexyl glucoside can also be effective in the extraction workflow. The advantage of PEPPI is that target protein complexes can be selectively recovered by a simple process, and our proposed workflow enables acquisition of high-quality native MS information on target proteins with reduced contamination even from crude purified samples. While we demonstrated a workflow using the easily detectable colored protein Hb in this report, establishment of a workflow for more common non-colored protein samples will require additional steps to rapidly detect the proteins separated by native PAGE without staining.

4. Concluding Remarks

We established a method to efficiently recover intact proteins from acrylamide gels after PAGE separation for subsequent measurement using MS methodologies. PEPPI after SDS-PAGE enabled selectively recovery of proteins visualized with CBB. Additionally, PEPPI using octyl glucoside instead of CBB was effective in recovering the protein complex separated by native PAGE, and the use of octylglucoside ensured that the proteins remained intact for native MS analysis. Conventional passive extraction generally requires shaking the gel for several hours, whereas PEPPI allows rapid protein extraction within 10 minutes and provides highly efficient recovery from a wide molecular weight range of the PAGE gel. PEPPI offers an efficient link between the separation of complex biological samples by PAGE and MS for intact protein analysis and top-down proteomics.

Author Information

Author Contributions

A.T., L.C.A., V.M.H., P.B., H.H., M.Y., J.A.L., R.R.O.L., R.J.B., and N.T. designed the research; A.T., L.C.A., V.M.H., P.B., D.B., D.H., K.S., J.I., J.S., and N.T. performed the research; A.T., L.C.A., V.M.H., P.B., D.B., D.H., K.S., R.J.B., and N.T. analyzed data; and A.T., L.C.A., V.M.H., P.B., J.A.L., R.R.O.L., R.J.B., and N.T. wrote the paper.

Notes

The authors declare no competing financial interest.

Acknowledgments

This study was supported from JSPS KAKENHI Grants (16K08937, 18H04559, and 19K05526 to N.T.). The work was also supported by grants from the Biotechnology and Biological Sciences Research Council (BB/M025756/1 and BB/R005311/1 to R.J.B.) and the US National Institutes of Health (R01GM104610 and R01GM103479 to R.R.O.L. and J.A.L.). A portion of this work was performed at the National High Magnetic Field Laboratory, which is supported by National Science Foundation Cooperative Agreement No. DMR-1644779* and the State of Florida* (* L.C.A. and D.B. support). We gratefully acknowledge Joseph B. Greer of the National Resource for Translational and Developmental Proteomics for facilitating data analysis via TDPortal (Grant P41 GM108569 to Neil L. Kelleher from the National Institute of General Medical Sciences, National Institutes of Health, and supported by the Sherman Fairchild Foundation).

References

1. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **1970**, *227* (5259), 680-685.
2. Schagger, H.; von Jagow, G. Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* **1991**, *199* (2), 223-231.
3. Schagger, H.; Cramer, W. A.; von Jagow, G. Analysis of molecular masses and oligomeric states of protein complexes by blue native electrophoresis and isolation of membrane protein complexes by two-dimensional native electrophoresis. *Anal. Biochem.* **1994**, *217* (2), 220-230.
4. Camm, E. L.; Green, B. R. Fractionation of Thylakoid Membranes with the Nonionic Detergent Octyl-beta-d-glucopyranoside: RESOLUTION OF CHLOROPHYLL-PROTEIN COMPLEX II INTO TWO CHLOROPHYLL-PROTEIN COMPLEXES. *Plant Physiol.* **1980**, 66 (3), 428-432.
5. Wittig, I.; Schagger, H., Features and applications of blue-native and clear-native electrophoresis. *Proteomics* **2008,** *8* (19), 3974-3990.
6. Rosenfeld, J.; Capdevielle, J.; Guillemot, J. C.; Ferrara, P., In-gel digestion of proteins for internal sequence analysis after one- or two-dimensional gel electrophoresis. *Anal Biochem* **1992,** *203* (1), 173-179.
7. Hellman, U.; Wernstedt, C.; Gonez, J.; Heldin, C. H., Improvement of an "In-Gel" digestion procedure for the micropreparation of internal protein fragments for amino acid sequencing. *Anal Biochem* **1995,** *224* (1), 451-455.
8. Borchers, C.; Peter, J. F.; Hall, M. C.; Kunkel, T. A.; Tomer, K. B., Identification of in-gel digested proteins by complementary peptide mass fingerprinting and tandem mass spectrometry data obtained on an electrospray ionization quadrupole time-of-flight mass spectrometer. *Anal Chem* **2000,** *72* (6), 1163-1168.
9. Gygi, S. P.; Corthals, G. L.; Zhang, Y.; Rochon, Y.; Aebersold, R., Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc Natl Acad Sci U S A* **2000,** *97* (17), 9390-9395.
10. Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M., In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat Protoc* **2006,** *1* (6), 2856-2860.
11. Compton, P. D.; Zamdborg, L.; Thomas, P. M.; Kelleher, N. L., On the scalability and requirements of whole protein mass spectrometry. *Anal Chem* **2011,** *83* (17), 6868-6874.
12. Tolmachev, A. V.; Robinson, E. W.; Wu, S.; Pasa-Tolic, L.; Smith, R. D., FT-ICR MS optimization for the analysis of intact proteins. *Int J Mass Spectrom* **2009,** *281* (1-3), 32-38.

13. Ahlf, D. R.; Compton, P. D.; Tran, J. C.; Early, B. P.; Thomas, P. M.; Kelleher, N. L., Evaluation of the compact high-field orbitrap for top-down proteomics of human cells. *J Proteome Res* **2012,** *11* (8), 4308-4314.
14. Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R., The Orbitrap: a new mass spectrometer. *J Mass Spectrom* **2005,** *40* (4), 430-443.
15. Armirotti, A.; Benatti, U.; Damonte, G., Top-down proteomics with a quadrupole time-of-flight mass spectrometer and collision-induced dissociation. *Rapid Commun Mass Spectrom* **2009,** *23* (5), 661-666.
16. Second, T. P.; Blethrow, J. D.; Schwartz, J. C.; Merrihew, G. E.; MacCoss, M. J.; Swaney, D. L.; Russell, J. D.; Coon, J. J.; Zabrouskov, V., Dual-pressure linear ion trap mass spectrometer improving the analysis of complex protein mixtures. *Anal Chem* **2009,** *81* (18), 7757-7765.
17. Senko, M. W.; Remes, P. M.; Canterbury, J. D.; Mathur, R.; Song, Q.; Eliuk, S. M.; Mullen, C.; Earley, L.; Hardman, M.; Blethrow, J. D.; Bui, H.; Specht, A.; Lange, O.; Denisov, E.; Makarov, A.; Horning, S.; Zabrouskov, V., Novel parallelized quadrupole/linear ion trap/Orbitrap tribrid mass spectrometer improving proteome coverage and peptide identification rates. *Anal Chem* **2013,** *85* (24), 11710-11714.
18. Hendrickson, C. L.; Quinn, J. P.; Kaiser, N. K.; Smith, D. F.; Blakney, G. T.; Chen, T.; Marshall, A. G.; Weisbrod, C. R.; Beu, S. C., 21 Tesla Fourier Transform Ion Cyclotron Resonance Mass Spectrometer: A National Resource for Ultrahigh Resolution Mass Analysis. *J Am Soc Mass Spectrom* **2015,** *26* (9), 1626-1632.
19. Riley, N. M.; Mullen, C.; Weisbrod, C. R.; Sharma, S.; Senko, M. W.; Zabrouskov, V.; Westphall, M. S.; Syka, J. E.; Coon, J. J., Enhanced Dissociation of Intact Proteins with High Capacity Electron Transfer Dissociation. *J Am Soc Mass Spectrom* **2016,** *27* (3), 520-531.
20. Earley, L.; Anderson, L. C.; Bai, D. L.; Mullen, C.; Syka, J. E.; English, A. M.; Dunyach, J. J.; Stafford, G. C., Jr.; Shabanowitz, J.; Hunt, D. F.; Compton, P. D., Front-end electron transfer dissociation: a new ionization source. *Anal Chem* **2013,** *85* (17), 8385-8390.
21. Syka, J. E.; Coon, J. J.; Schroeder, M. J.; Shabanowitz, J.; Hunt, D. F., Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc Natl Acad Sci U S A* **2004,** *101* (26), 9528-9533.
22. Shaw, J. B.; Li, W.; Holden, D. D.; Zhang, Y.; Griep-Raming, J.; Fellers, R. T.; Early, B. P.; Thomas, P. M.; Kelleher, N. L.; Brodbelt, J. S., Complete protein characterization using top-down mass spectrometry and ultraviolet photodissociation. *J Am Chem Soc* **2013,** 135 (34), 12646-12651.

23. Anderson, L. C.; English, A. M.; Wang, W.; Bai, D. L.; Shabanowitz, J.; Hunt, D. F., Protein derivatization and sequential ion/ion reactions to enhance sequence coverage produced by electron transfer dissociation mass spectrometry. *Int J Mass Spectrom* **2015,** *377*, 617-624.
24. Riley, N. M.; Westphall, M. S.; Coon, J. J., Activated Ion-Electron Transfer Dissociation Enables Comprehensive Top-Down Protein Fragmentation. *J Proteome Res* **2017,** *16* (7), 2653-2659.
25. Shaw, J. B.; Malhan, N.; Vasil'ev, Y. V.; Lopez, N. I.; Makarov, A.; Beckman, J. S.; Voinov, V. G., Sequencing Grade Tandem Mass Spectrometry for Top-Down Proteomics Using Hybrid Electron Capture Dissociation Methods in a Benchtop Orbitrap Mass Spectrometer. *Anal Chem* **2018,** *90* (18), 10819-10827.
26. Toby, T. K.; Fornelli, L.; Kelleher, N. L., Progress in Top-Down Proteomics and the Analysis of Proteoforms. *Annu Rev Anal Chem (Palo Alto Calif)* **2016,** *9* (1), 499-519.
27. Chen, B.; Brown, K. A.; Lin, Z.; Ge, Y., Top-Down Proteomics: Ready for Prime Time? *Anal Chem* **2018,** *90* (1), 110-127.
28. Schaffer, L. V.; Millikin, R. J.; Miller, R. M.; Anderson, L. C.; Fellers, R. T.; Ge, Y.; Kelleher, N. L.; LeDuc, R. D.; Liu, X.; Payne, S. H.; Sun, L.; Thomas, P. M.; Tucholski, T.; Wang, Z.; Wu, S.; Wu, Z.; Yu, D.; Shortreed, M. R.; Smith, L. M., Identification and Quantification of Proteoforms by Mass Spectrometry. *Proteomics* **2019,** *19* (10), e1800361.
29. Tran, J. C.; Doucette, A. A., Gel-eluted liquid fraction entrapment electrophoresis: an electrophoretic method for broad molecular weight range proteome separation. *Anal Chem* **2008,** *80* (5), 1568-1573.
30. Tran, J. C.; Doucette, A. A., Multiplexed size separation of intact proteins in solution phase for mass spectrometry. *Anal Chem* **2009,** *81* (15), 6201-6209.
31. Skinner, O. S.; Do Vale, L. H.; Catherman, A. D.; Havugimana, P. C.; de Sousa, M. V.; Compton, P. D.; Kelleher, N. L., Native GELFrEE: a new separation technique for biomolecular assemblies. *Anal Chem* **2015,** *87* (5), 3032-3038.
32. Cai, W.; Tucholski, T.; Chen, B.; Alpert, A. J.; McIlwain, S.; Kohmoto, T.; Jin, S.; Ge, Y., Top-Down Proteomics of Large Proteins up to 223 kDa Enabled by Serial Size Exclusion Chromatography Strategy. *Anal Chem* **2017,** *89* (10), 5467-5475.
33. Vanfleteren, J. R., Sequential sodium dodecyl sulfate-polyacrylamide gel electrophoresis and reversed-phase chromatography of unfolded proteins. *Anal Biochem* **1989,** *178* (2), 385-390.
34. Jeno, P., Horst, M.: Electroelution of Proteins from Polyacrylamide Gels. In: Walker, J.M. (ed.) *The Protein Protocols Handbook*. pp. 207-214. Humana Press, Totowa, NJ (1996)

35. Cohen, S. L.; Chait, B. T., Mass spectrometry of whole proteins eluted from sodium dodecyl sulfate-polyacrylamide gel electrophoresis gels. *Anal Biochem* **1997,** *247* (2), 257-267.
36. Loo, J. A.; Brown, J.; Critchley, G.; Mitchell, C.; Andrews, P. C.; Ogorzalek Loo, R. R., High sensitivity mass spectrometric methods for obtaining intact molecular weights from gel-separated proteins. *Electrophoresis* **1999,** *20* (4-5), 743-748.
37. Galvani, M.; Bordini, E.; Piubelli, C.; Hamdan, M., Effect of experimental conditions on the analysis of sodium dodecyl sulphate polyacrylamide gel electrophoresis separated proteins by matrix-assisted laser desorption/ ionisation mass spectrometry. *Rapid Commun Mass Spectrom* **2000,** *14* (1), 18-25.
38. Asquith, T.N., Keough, T.W., Takigiku, R., Lacey, M.P., Purdon, M.P., Gauggel, D.L.: Accurate Determination of Molecular Mass for Proteins Isolated from Polyacrylamide Gels. In: Angeletti, R.H. (ed.) Techniques in Protein Chemistry IV. pp. 99-106. Academic Press, San Diego (1993)
39. Galvani, M.; Hamdan, M.; Righetti, P. G., Two-dimensional gel electrophoresis/matrix-assisted laser desorption/ionisation mass spectrometry of a milk powder. *Rapid Commun Mass Spectrom* **2000,** *14* (20), 1889-1897.
40. Galvani, M.; Hamdan, M., Electroelution and passive elution of gamma-globulins from sodium dodecyl sulphate polyacrylamide gel electrophoresis gels for matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* **2000,** *14* (8), 721-723.
41. Jorgensen, C. S.; Jagd, M.; Sorensen, B. K.; McGuire, J.; Barkholt, V.; Hojrup, P.; Houen, G., Efficacy and compatibility with mass spectrometry of methods for elution of proteins from sodium dodecyl sulfate-polyacrylamide gels and polyvinyldifluoride membranes. *Anal Biochem* **2004,** *330* (1), 87-97.
42. Seelert, H.; Krause, F., Preparative isolation of protein complexes and other bioparticles by elution from polyacrylamide gels. *Electrophoresis* **2008,** *29* (12), 2617-2636.
43. Scheer, J. M.; Ryan, C. A., A method for the quantitative recovery of proteins from polyacrylamide gels. *Anal Biochem.* **2001,** 298 (1), 130-132.
44. Georgiou, C. D.; Grintzalis, K.; Zervoudakis, G.; Papapostolou, I., Mechanism of Coomassie brilliant blue G-250 binding to proteins: a hydrophobic assay for nanogram quantities of proteins. *Anal Bioanal Chem* **2008,** *391* (1), 391-403.
45. Lawrence, A. M.; Besir, H. U., Staining of proteins in gels with Coomassie G-250 without organic solvent and acetic acid. *J Vis Exp* **2009,** (30).

46. Saito, S.; Tsuchiya, T., Characteristics of n-octyl beta-D-thioglucopyranoside, a new non-ionic detergent useful for membrane biochemistry. *Biochem J* **1984,** *222* (3), 829-832.

47. Loo, R. R.; Dales, N.; Andrews, P. C., Surfactant effects on protein structure examined by electrospray ionization mass spectrometry. *Protein Sci* **1994,** *3* (11), 1975-1983.

48. Li, H.; Sheng, Y.; McGee, W.; Cammarata, M.; Holden, D.; Loo, J. A., Structural Characterization of Native Proteins and Protein Complexes by Electron Ionization Dissociation-Mass Spectrometry. *Anal Chem* **2017,** *89* (5), 2731-2738.

49. Schneider, C. A.; Rasband, W. S.; Eliceiri, K. W., NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* **2012,** *9* (7), 671-5.

50. Wessel, D.; Flugge, U. I., A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids. *Anal Biochem* **1984,** *138* (1), 141-143.

51. Kuwahara, M.; Suzuki, J.; Tofukuji, S.; Yamada, T.; Kanoh, M.; Matsumoto, A.; Maruyama, S.; Kometani, K.; Kurosaki, T.; Ohara, O.; Nakayama, T.; Yamashita, M. The Menin-Bach2 axis is critical for regulating CD4 T-cell senescence and cytokine homeostasis. *Nat. Commun.* **2014**, *5*, 3555.

52. Takemori, N.; Komori, N.; Thompson, J. N., Jr.; Yamamoto, M. T.; Matsumoto, H. Novel eye-specific calmodulin methylation characterized by protein mapping in *Drosophila melanogaster*. *Proteomics* **2007**, *7* (15), 2651-2658.

53. McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wuhr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal. Chem.* **2014**, *86* (14), 7150-7158.

54. Afgan, E.; Baker, D.; Batut, B.; van den Beek, M.; Bouvier, D.; Cech, M.; Chilton, J.; Clements, D.; Coraor, N.; Grüning, B.A.; Guerler, A.; Hillman-Jackson, J.; Hiltemann, S.; Jalili, V.; Rasche, H.; Soranzo, N.; Goecks, J.; Taylor, J.; Nekrutenko, A.; Blankenberg, D. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res*. **2018**, 46(W1), W537-W544.

55. Toby, TK.; Fornelli, L.; Srzentić, K.; DeHart, C. J.; Levitsky, J.; Friedewald, J.; Kelleher, N. L. A comprehensive pipeline for translational top-down proteomics from a single blood draw. *Nat. Protoc.* **2019**, 14(1), 119-152.

56. Takemori, N.; Takemori, A.; Wongkongkathep, P.; Nshanian, M.; Loo, R. R. O.; Lermyte, F.; Loo, J. A., Top-down/Bottom-up Mass Spectrometry Workflow Using Dissolvable Polyacrylamide Gels. *Anal Chem* **2017,** *89* (16), 8244-8250.

57. Jin, Y.; Manabe, T., High-efficiency protein extraction from polyacrylamide gels for molecular mass measurement by matrix-assisted laser desorption/ionization-time of flight-mass spectrometry. *Electrophoresis* **2005,** *26* (6), 1019-1028.

58. Loo, R. R.; Cavalcoli, J. D.; VanBogelen, R. A.; Mitchell, C.; Loo, J. A.; Moldover, B.; Andrews, P. C., Virtual 2-D gel electrophoresis: visualization and analysis of the E. coli proteome by mass spectrometry. *Anal Chem* **2001,** *73* (17), 4063-4070.

59. Tran, J.C.; Zamdborg, L.; Ahlf, D.R.; Lee, J.E.; Catherman, A.D.; Durbin, K.R.; Tipton, J.D.; Vellaichamy, A.; Kellie, J.F.; Li, M.; Wu, C.; Sweet, S.M.; Early, B.P.; Siuti, N.; LeDuc, R.D.; Compton, P.D.; Thomas, P.M.; Kelleher, N.L., Mapping intact protein isoforms in discovery mode using top-down proteomics. *Nature* **2011,** 480(7376), 254-258.

60. Wang, Z.; Ma, H.; Smith, K.; Wu, S., Two-Dimensional Separation Using High-pH and Low-pH Reversed Phase Liquid Chromatography for Top-down Proteomics. *Int J Mass Spectrom.* **2018,** 427, 43-51.

61. Anderson, L. C.; DeHart, C. J.; Kaiser, N. K.; Fellers, R. T.; Smith, D. F.; Greer, J. B.; LeDuc, R. D.; Blakney, G. T.; Thomas, P. M.; Kelleher, N. L.; Hendrickson, C. L., Identification and Characterization of Human Proteoforms by Top-Down LC-21 Tesla FT-ICR Mass Spectrometry. *J Proteome Res* **2017,** *16* (2), 1087-1096.

62. Ntai, I.; Toby, T.K.; LeDuc, R.D.; Kelleher, N.L., A Method for Label-Free, Differential Top-Down Proteomics. *Methods Mol Biol.* **2016**, 14 (10), 121-133.

Figures

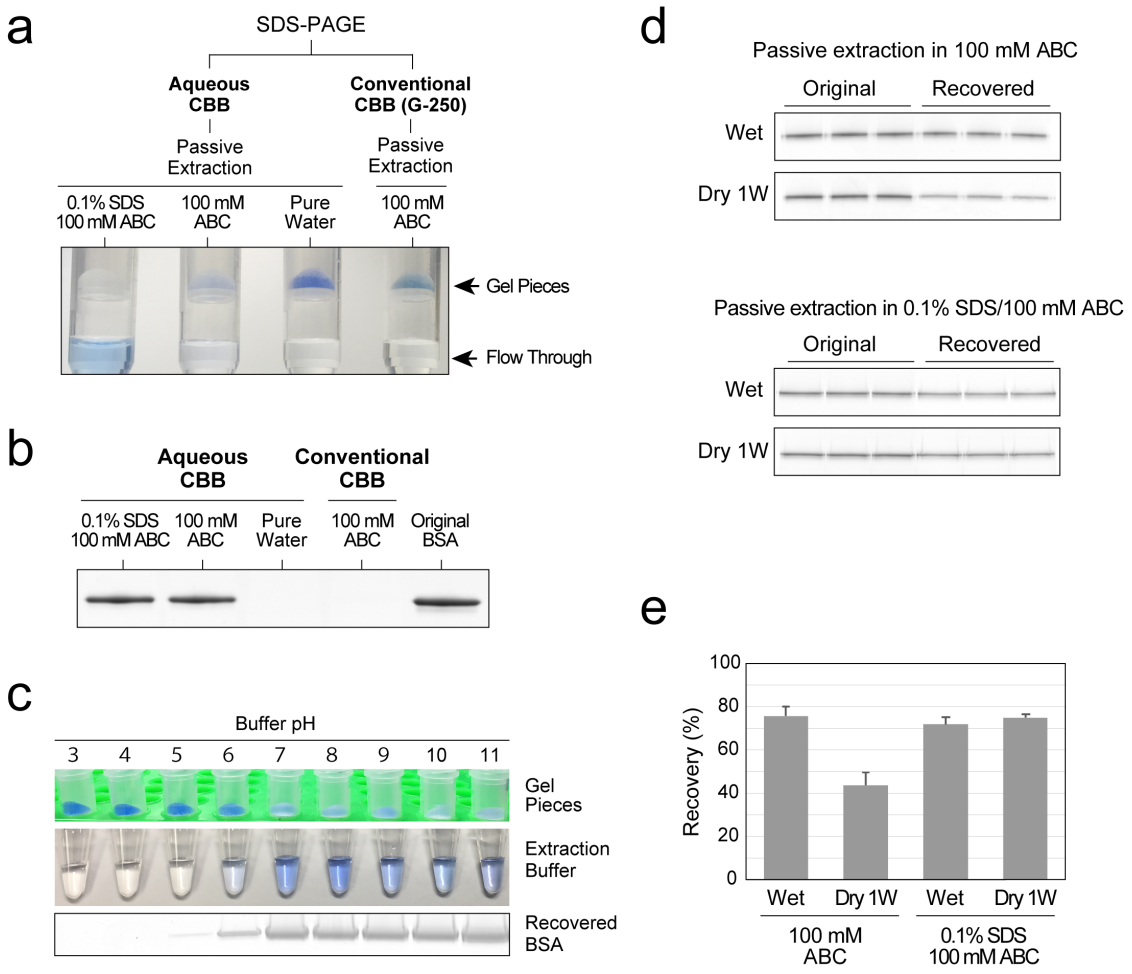

Figure 1. Passive extraction of in-gel proteins separated by SDS-PAGE

(a) Experimental workflow for passive extraction of CBB-stained BSA. For aqueous CBB staining, the gel was stained with EzStain Aqua CBB (ATTO) for 8 min, and washed for 30 min in water. Inset photo: gel pieces after protein extraction. Most of the CBB dye released into the solution upon passive extraction with 100 mM ABC was adsorbed on the membrane of the spin column. (b) SDS-PAGE images of BSA recovered from polyacrylamide gels stained with aqueous CBB and conventional CBB G-250. (c) Passive protein extraction under different pH conditions. (d) SDS-PAGE images of gel-recovered BSA. Separated BSA bands were visualized using BioSafe CBB (Bio-Rad). CBB-stained BSA was extracted from polyacrylamide gels treated two different ways: extraction performed immediately after PAGE (Wet) and extraction performed after gel dehydration and storage for 1 week (Dry 1W). (e) Recoveries for passive extraction. The recovery efficiency was estimated based on the intensity of each BSA band (Figure 1d).

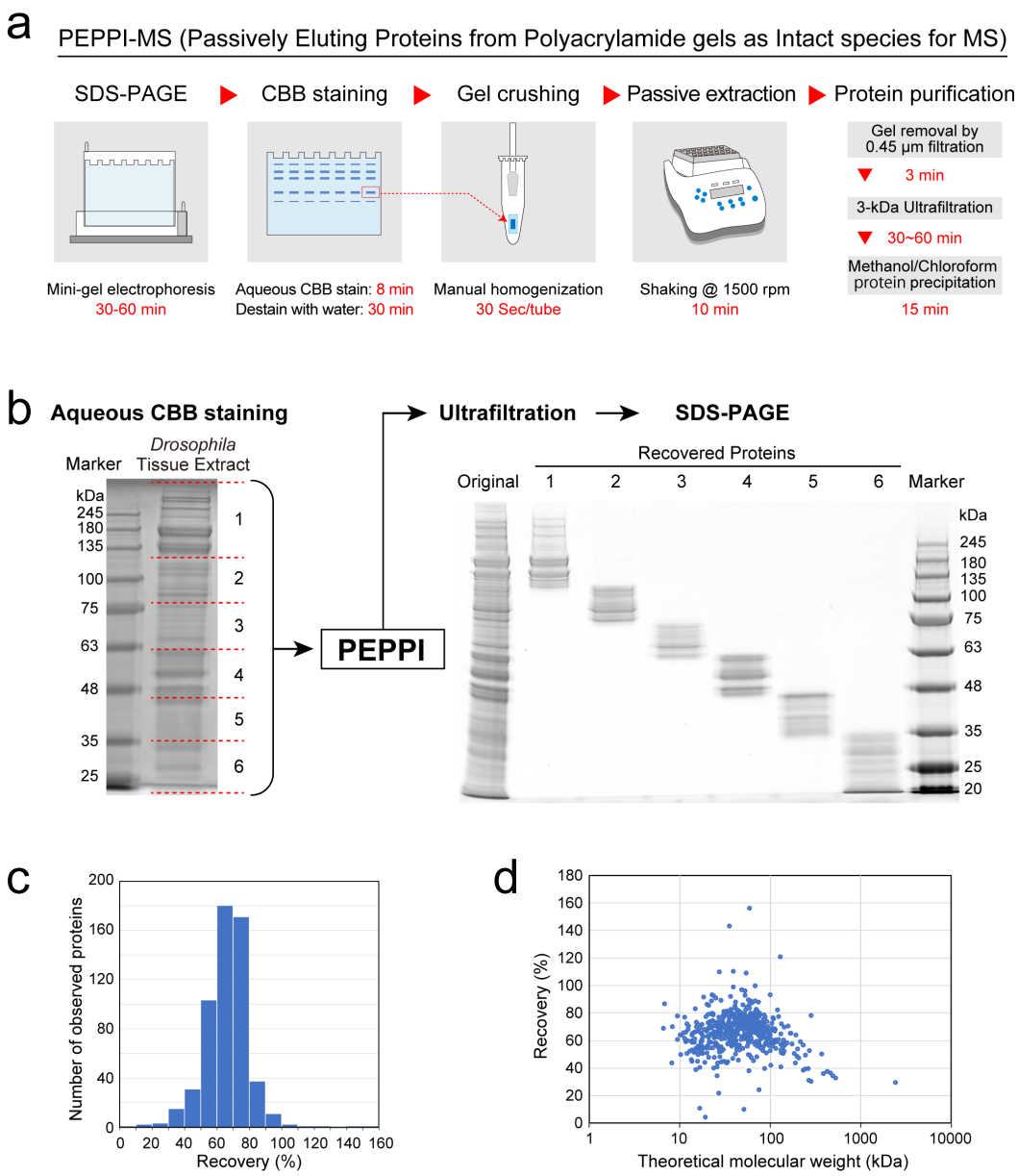

Figure 2. Large-scale verification of PEPPI-MS performance

(a) Schematic illustration of the PEPPI-MS workflow. Red characters indicate the required time for each experiment. (b) Protein fractionation using the PEPPI workflow. Protein components extracted from *Drosophila* compound eyes were separated by SDS-PAGE and stained with aqueous CBB. After dividing the sample lane into six portions, proteins were extracted from each split gel using PEPPI. Extracted proteins were displayed on SDS-PAGE gels. (c) Histogram showing the protein recovery efficiency. Recovery rate of gel-separated *Drosophila* proteins was estimated by iTRAQ-based quantitative analysis. (d) Distribution of recovery efficiency with the theoretical molecular weight of recovered proteins.

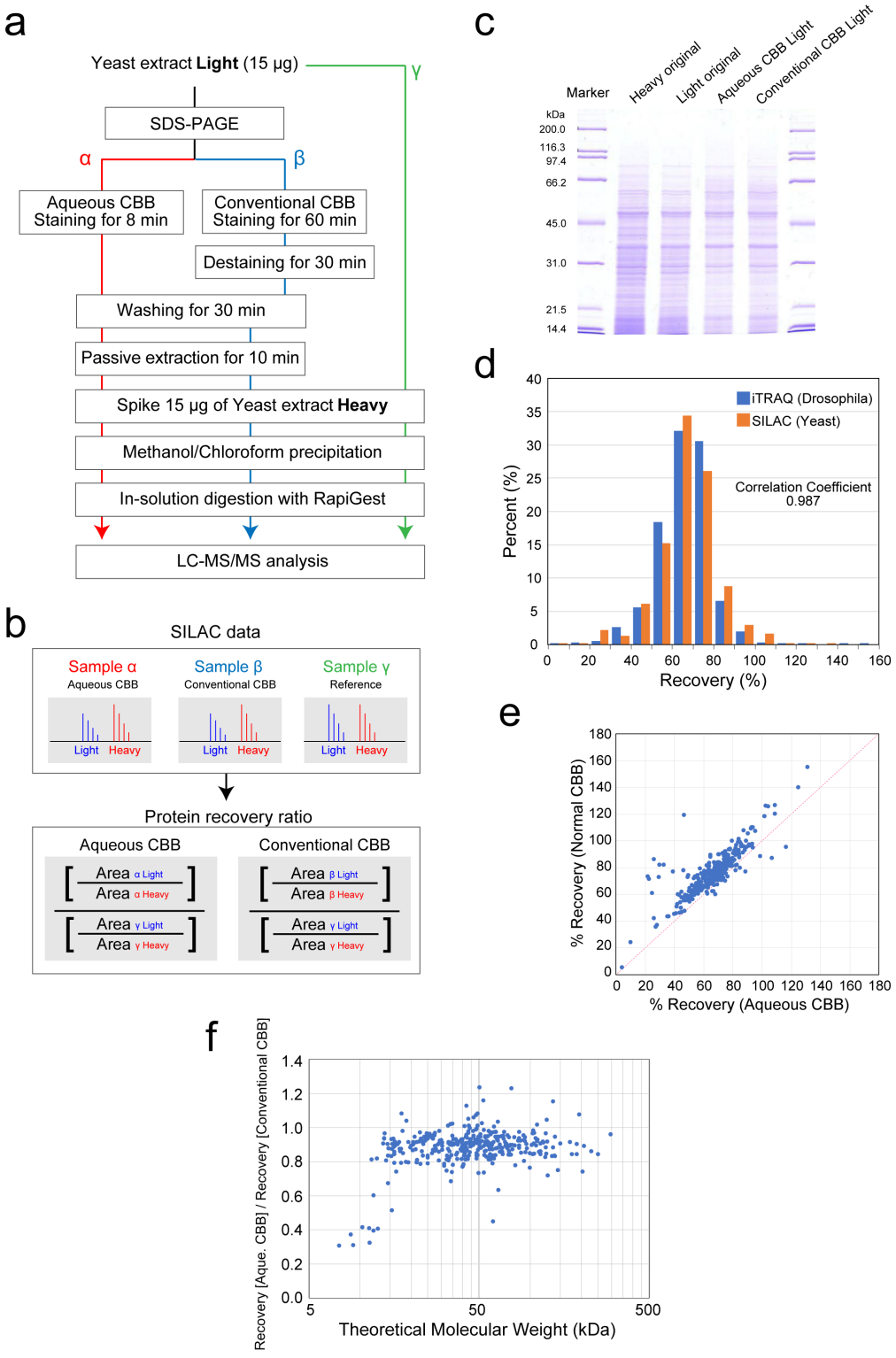

Figure 3. Quantitative evaluation of protein recovery by SILAC

(a) Experimental workflow to estimate the recovery efficiency of the gel-separated yeast proteome by SILAC. Yeast protein extract (Light) separated using SDS-PAGE was recovered using PEPPI and mixed

with 13C/15N-labeled (Heavy) yeast extract. The tryptic digests of the recovered proteins were analyzed by LC-MS/MS. (b) The recovery rate of the protein in the gel was calculated using the L/H ratio determined from each LC-MS/MS data set (α, β, and γ). (c) A representative SDS-PAGE image of recovered proteins from the aqueous CBB-stained gel and conventional CBB-stained gel. (d) Histogram of estimated protein recovery. (e) Comparison of protein recovery performance under different staining conditions. After SDS-PAGE of a yeast protein extract, separated proteins were independently stained with two formulations of CBB (aqueous CBB and conventional CBB containing methanol and acetic acid) and were extracted using PEPPI. Protein recovery rates between the two different CBB-staining conditions were plotted in a scatter plot. (f) Distribution of the protein recovery rate. Molecular weight of each protein and its recovery rate were plotted.

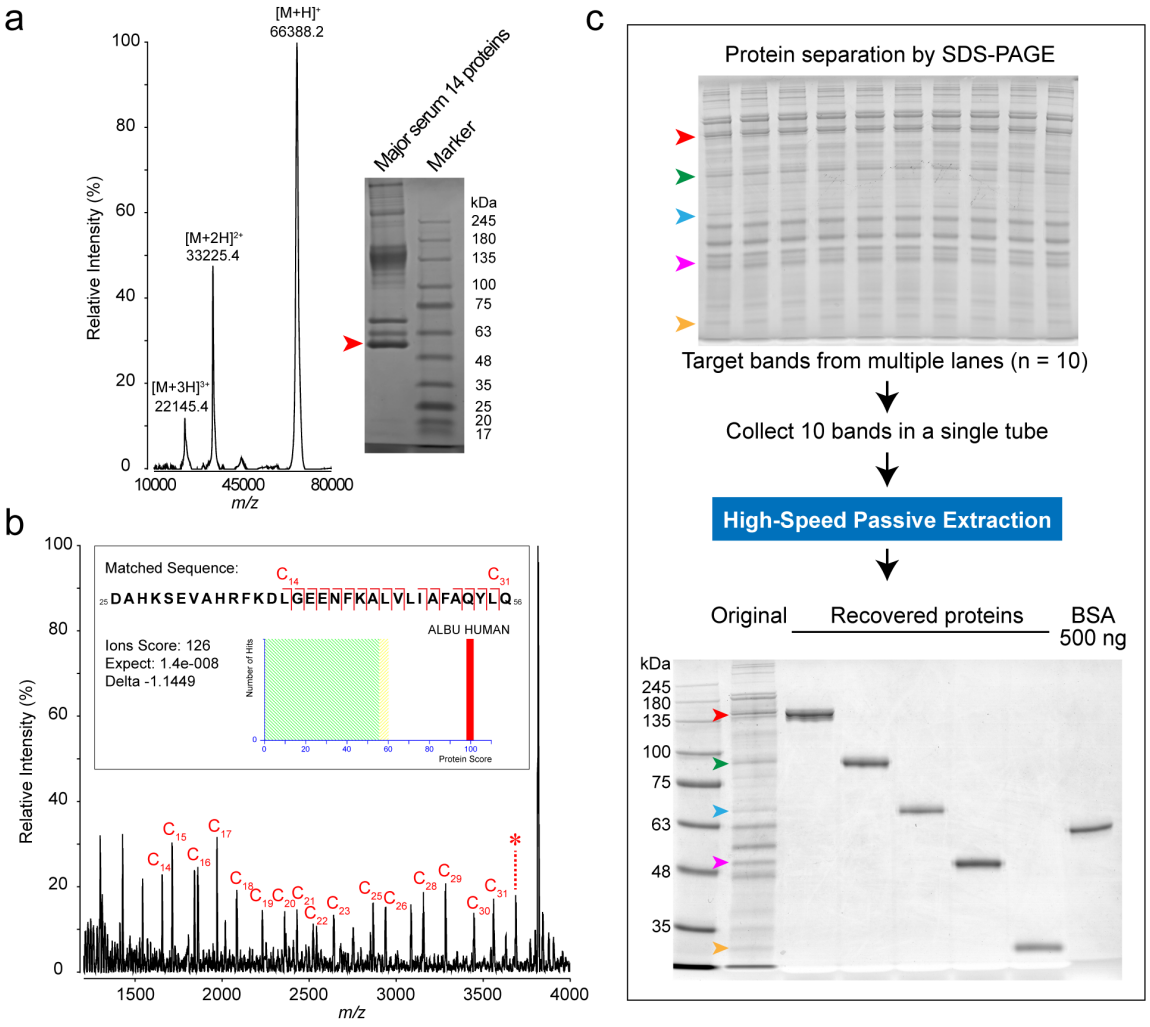

Figure 4. Selective protein recovery from crude biological samples

(a) MALDI-TOF MS analysis of a gel-recovered serum protein. After gel separation of 14 major human serum proteins, the selected band (red arrow) was excised, and the extracted protein was subjected to MALDI-TOF MS analysis. (b) MALDI-ISD of the gel-recovered protein. The N-terminal amino acid sequence of the recovered protein (Human serum albumin) was determined using product ions generated by MALDI-ISD. The ion peak at m/z 3689 (asterisk) was used as "virtual precursor ion" in the MASCOT database search. Inset: MASCOT database search results. (c) Selective enrichment of target proteins from crude biological samples. The crude protein extract from *Drosophila* compound eyes was separated by SDS-PAGE and the target protein bands (arrows) were excised. The recovered proteins were separated again by SDS-PAGE and stained with CBB.

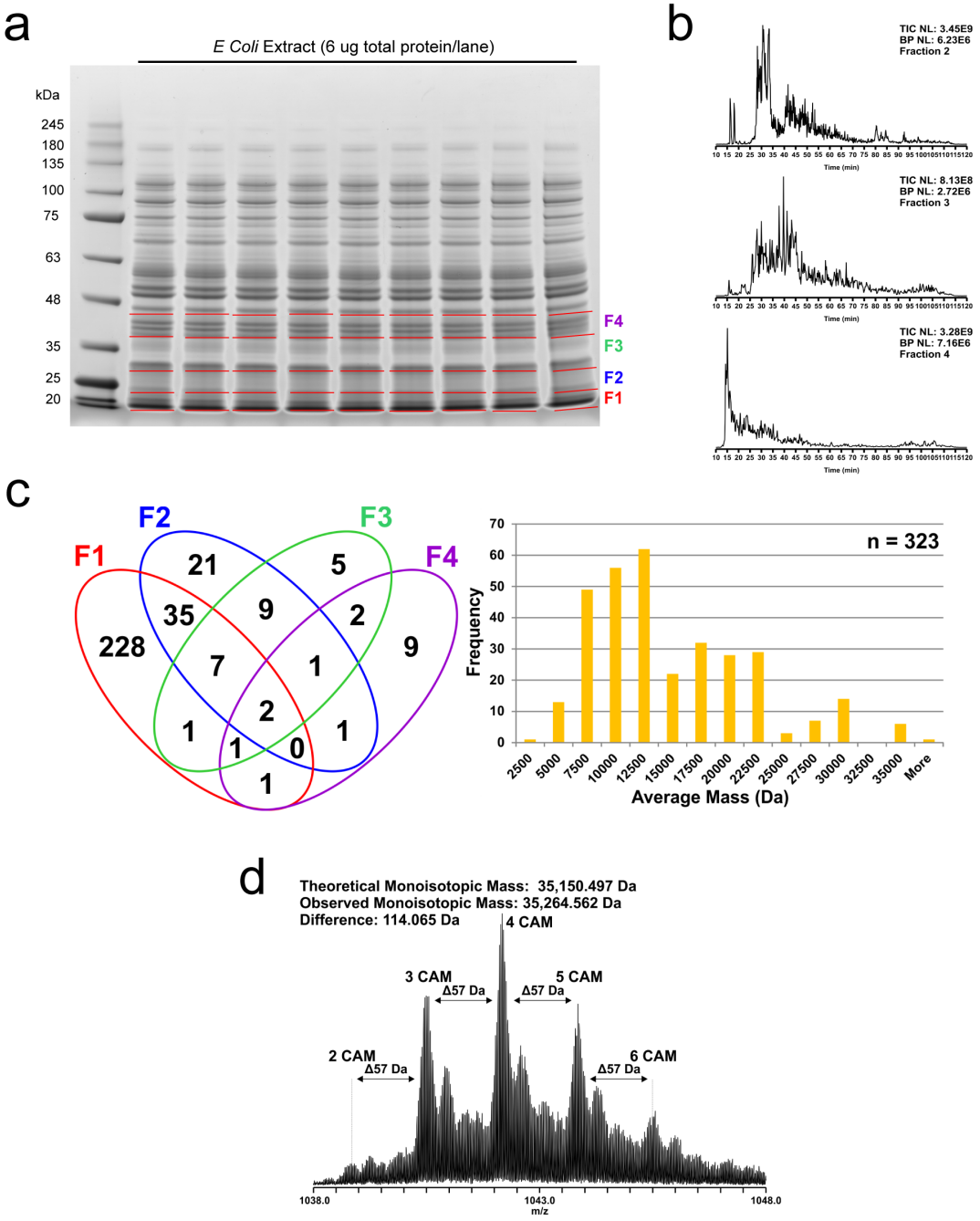

Figure 5. Gel-based top-down proteomics of *E. coli* proteome

(a) Representative SDS-PAGE image of *E. coli* extract. *E. coli* proteins were separated with a 4-14% Nu-PAGE gel and visualized with aqueous CBB. After PAGE separation, the selected bands were subjected to PEPPI. (b) Base peak chromatograms of PEPPI fractions 2-4 (*E. coli* whole cell lysate). (c) (left) Venn diagram showing proteoforms identified within each PEPPI fraction. (right) Histogram depicting the molecular weight distributions of *E. coli* proteoforms identified at 1% FDR in PEPPI fractions 1-4 (323

proteoforms total). (d) Post-FT spectral average (~50 scans acquired over ~5 min) of proteoforms derived from outer membrane protein A (ompA; proteoform repository identifier number 242792), the largest protein identified in the TDPortal search of PEPPI-fractionated (observed in fractions 2 and 3) *E. coli* whole cell lysate. This proteoform sequence contains 2 cys residues. Theoretical and observed monoisotopic masses differed by 114 Da, the result of carbamidomethylation of cys residues (2 X 57 Da). This proteoform is observed along with additions of up to 6 total carbamidomethyl groups (CAM). The TDPortal search identified proteoforms containing 2 and 3 CAM groups (max 3 possible; intact mass tolerance 200 Da). These assignments were manually inspected and observed mass measurement accuracies of the most abundant isotopologues of each CAM form agree to theoretical m/z values within 1.5-2 ppm (external calibration only).

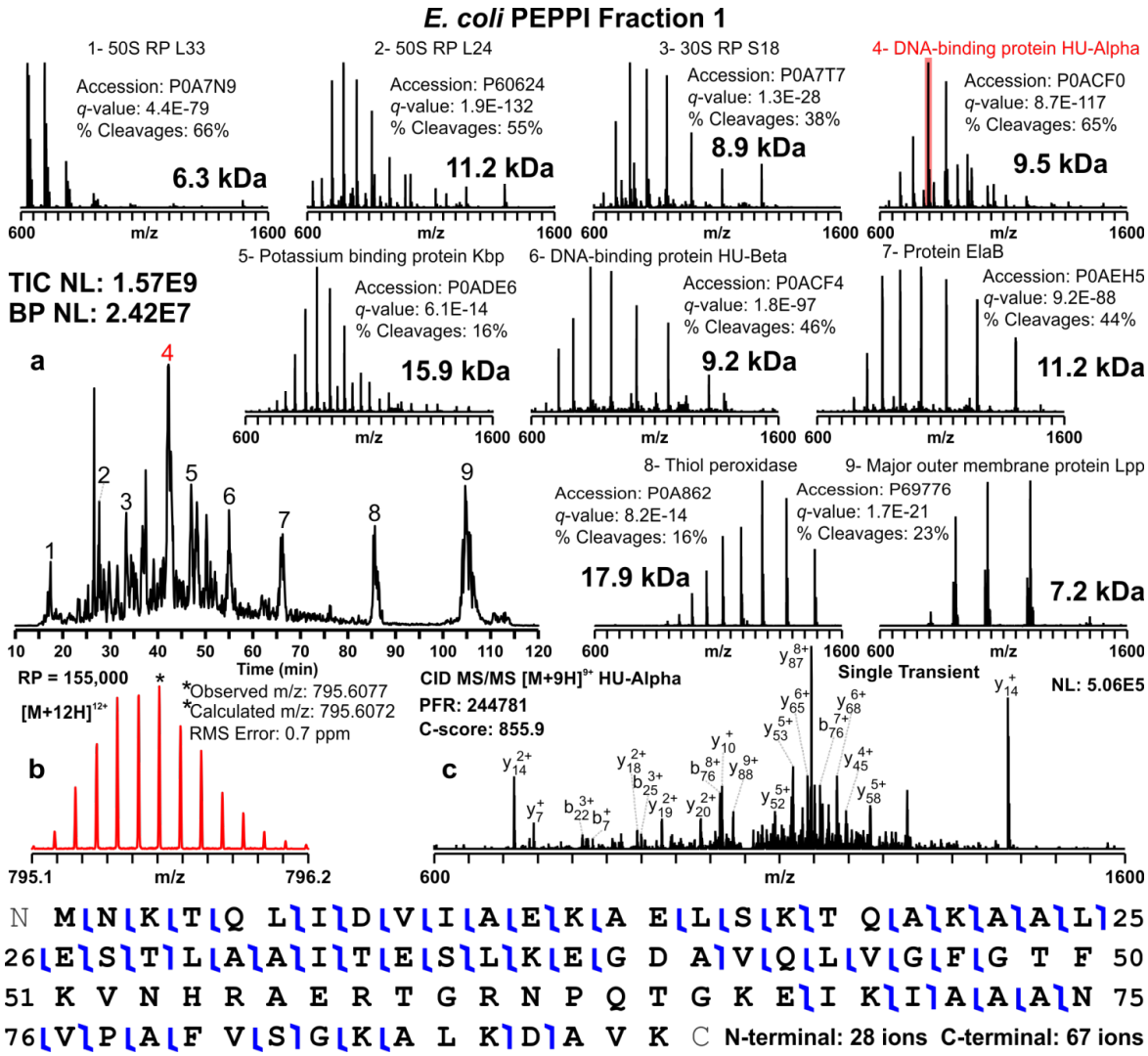

Figure 6. LC-21 T FT-ICR MS/MS of a single injection of PEPPI fraction 1

(a) Base peak chromatogram. Nine example identifications are indicated (UniProt accession numbers included) along with single-scan MS1 spectra, global q-values (from TDPortal search), molecular weight, and sequence coverage (% Cleavages). (b) m/z-scale expanded segment of the single-scan MS1 spectrum (a; red highlight) depicting the isotopic distribution of the $[M+12H]^{12+}$ charge state of DNA-binding protein HU-Alpha. These peaks matched the theoretical isotope distribution simulated from the known elemental composition of this HU-Alpha proteoform (proteoform repository identifier number 244781) with an RMS error of 0.7 ppm. (c) Single-transient CID MS2 spectrum of HU-Alpha. Ninety-five fragments were matched to the sequence with a 10 ppm mass tolerance, yielding 65% sequence coverage.

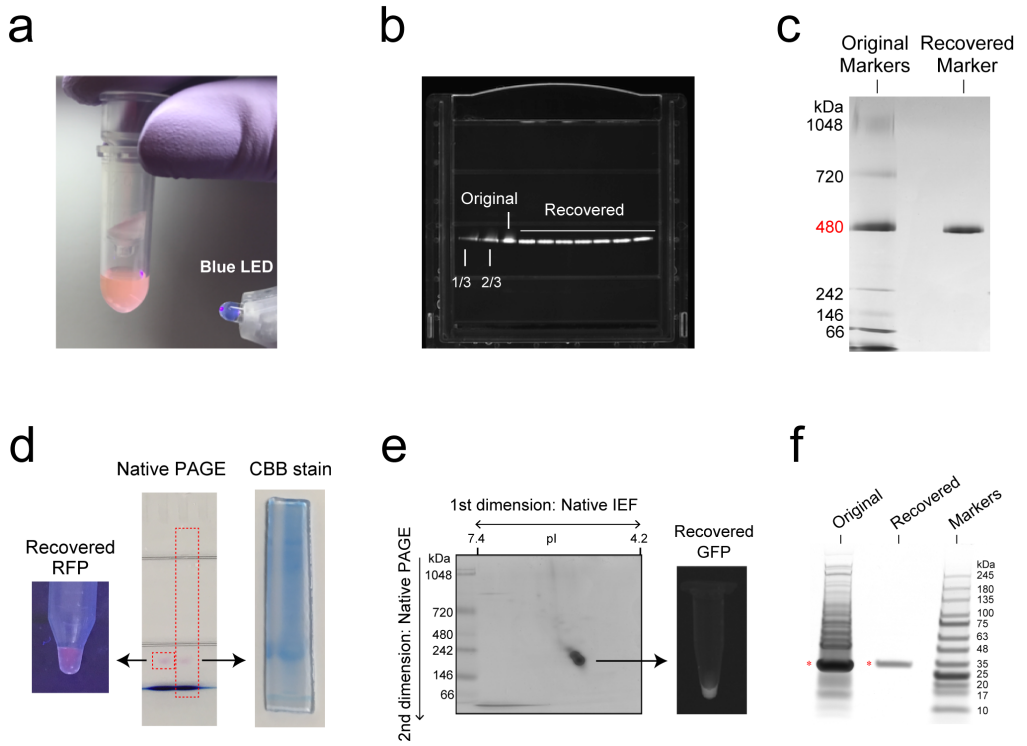

Figure 7. Protein recovery after native PAGE separation

(a) B-phycoerythrin separated by native PAGE was extracted with Novex tris-glycine native running buffer (Thermo Fisher Scientific) containing 0.1% (w/v) octylglucoside. Recovery of b-phycoerythrin after filtration was confirmed by blue light emitting diode (LED) irradiation. (b) Recovered b-phycoerythrin was separated again using native PAGE and the fluorescence image after electrophoresis was acquired. (c) Selective recovery of protein markers after native PAGE separation. After separating the NativeMark unstained protein standard (Thermo Fisher Scientific) using native PAGE, a band of the 480-kDa marker (Apoferritin) was cut out and extracted. The recovered marker was electrophoresed again using native PAGE and stained with CBB. (d) Recombinant red fluorescent protein (RFP) synthesized in a wheat cell-free system was selectively recovered from a native PAGE gel. After separating proteins from the crude synthesis solution, a region containing the target protein was excised using fluorescence as a marker and extracted with 0.1% (w/v) octylglucoside. Fluorescence was observed from the extracted RFP sample by UV irradiation. Wheat cell-free synthesized RFP was a gift from Prof. Yaeta Endo (Ehime University). (e) Recombinant green fluorescent protein (GFP)-spiked in mouse T cell lysates was separated by two-dimensional native gel electrophoresis. After the first separation by native isoelectric focusing, the gel region containing GFP was excised and subjected to the further separation by native PAGE. (f) SDS-PAGE image of recovered GFP. *, GFP band; Original, GFP-spiked mouse T cell extract.

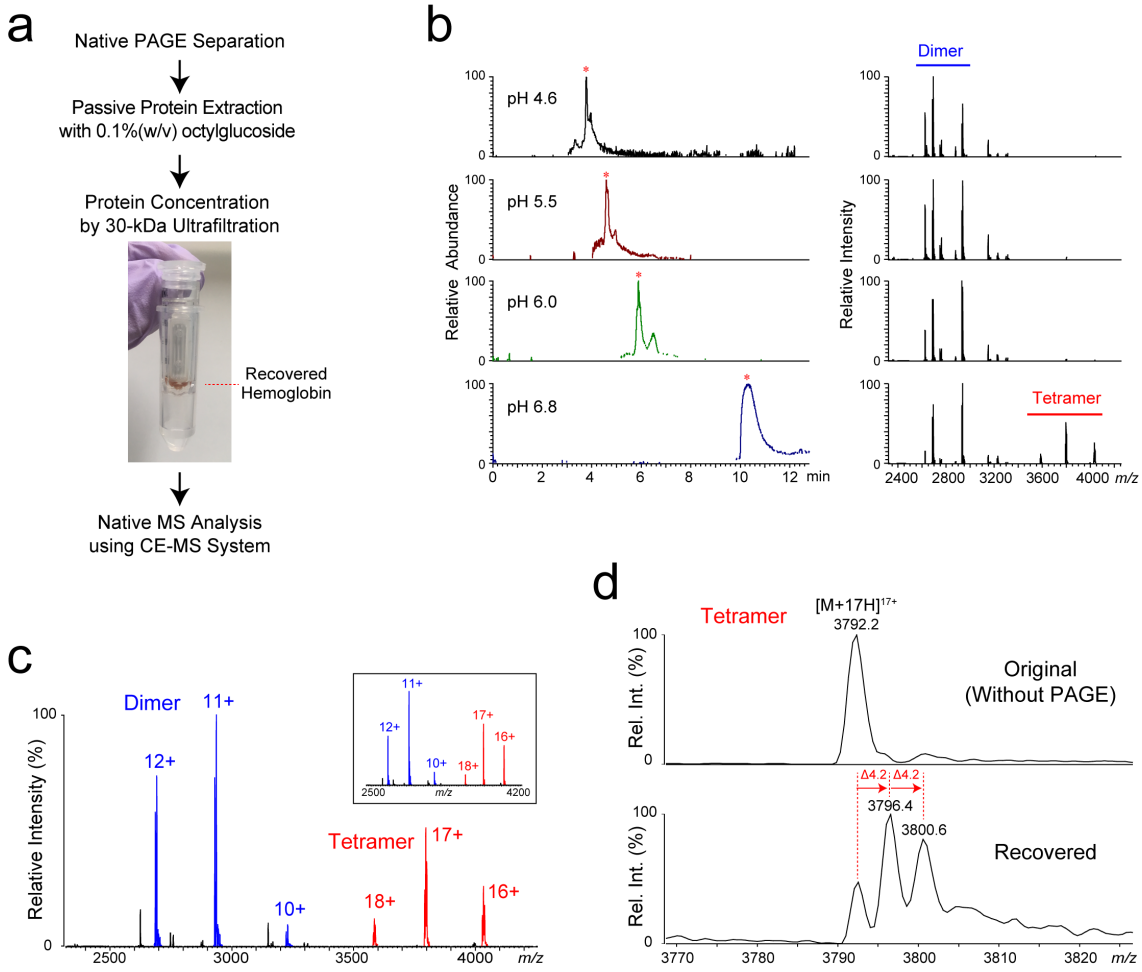

Figure 8. Native MS analysis of gel-extracted human hemoglobin

(a) Human Hb separated by native PAGE was recovered with 0.1% (w/v) octylglucoside and solution replacement with 20 mM ammonium acetate (pH 6.8) was carried out using a 30-kDa centrifugal ultrafiltration device. Native MS analysis of human Hb was performed using a ZipChip CE-MS system. (b) Native MS analyses using a ZipChip CE orbitrap mass spectrometer. Hb samples were separated in a ZipChip CE system using a background electrolyte with different pH conditions and were directly ionized using online nanoflow ESI. (c) Native MS spectrum of gel recovered hemoglobin (Hb). Multiply charged ions of Hb dimer and tetramer separated by ZipChip CE were measured using a Q-Exactive HF-X orbitrap MS. Inset: native MS spectrum of original Hb sample without PAGE treatment, showing little change in dimer-tetramer distribution between the original sample and following PEPPI from the native gel. (d) Mass shift caused by acrylamide modification to Hb molecules during electrophoresis. Ion peaks at m/z 3792.2, 3796.4 and 3800.6 correspond to unmodified Hb, modified Hb with one acrylamide monomer, and modified Hb with two acrylamide monomers, respectively.

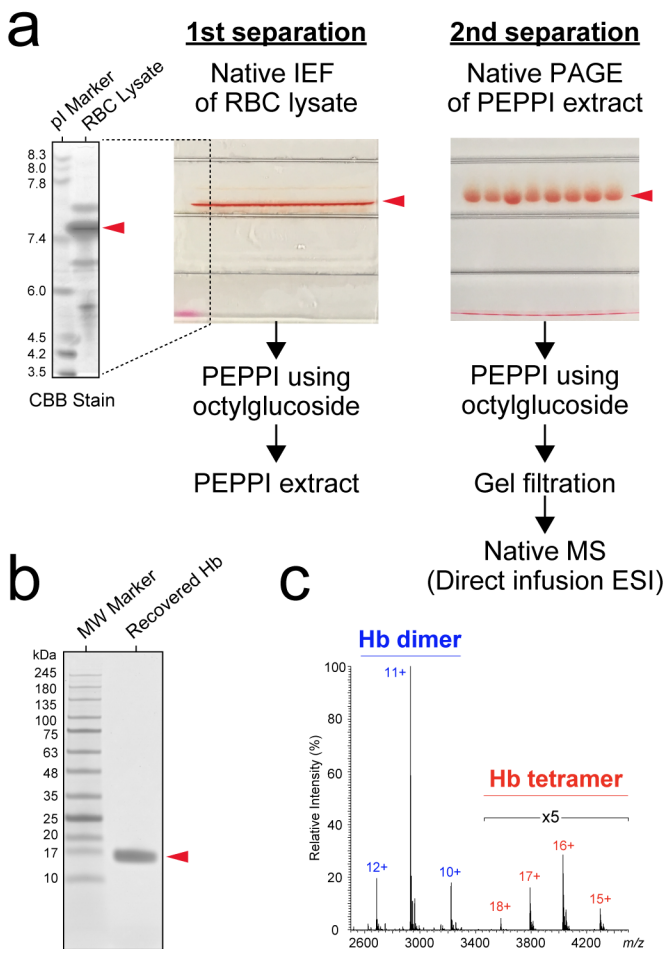

Figure 9. MS-based characterization of recovered human hemoglobin (Hb) from a native PAGE gel

(a) Scheme for the recovery of Hb from human red blood cell (RBC) samples by two-dimensional native PAGE and PEPPI. In the 1st dimension, RBC lysates were separated by native IEF (pI 3-10) and the gel band of Hb (red arrow) was subjected to PEPPI using 0.1% octyl glucoside. A portion of the IEF gel was stained with CBB. In the second dimension, the recovered Hb was further separated by native PAGE and again recovered by PEPPI. The recovered sample was purified on a gel filtration spin column and subjected to native MS analysis by direct infusion ESI. (b) A representative SDS-PAGE gel image of Hb recovered with PEPPI (red arrow). The Hb sample purified by 2D native PAGE was separated with a 4-12% NuPAGE gel and visualized with BioRad's bio-safe CBB. (c) Native MS spectra of the recovered Hb.

Supplementary Information

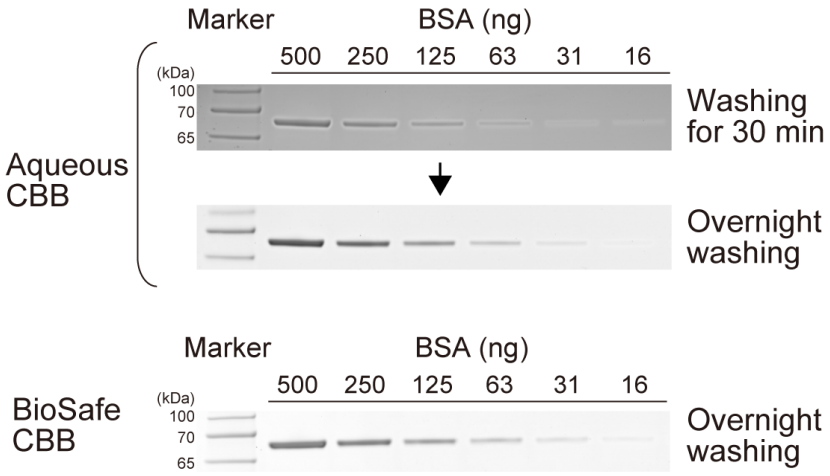

Figure S1. Comparison of two different CBB staining methods.
Different concentrations of BSA standard (Wako) were separated by NuPAGE Precast 4-12% Bis-Tris gradient gels (1-mm-thick, 10 sample wells, Thermo Fisher Scientific) with NuPAGE MES SDS Running Buffer (Thermo Fisher Scientific). After separation, the gels were stained with ATTO's EzStain Aqua CBB ("Aqueous CBB") or BioRad's BioSafe CBB.

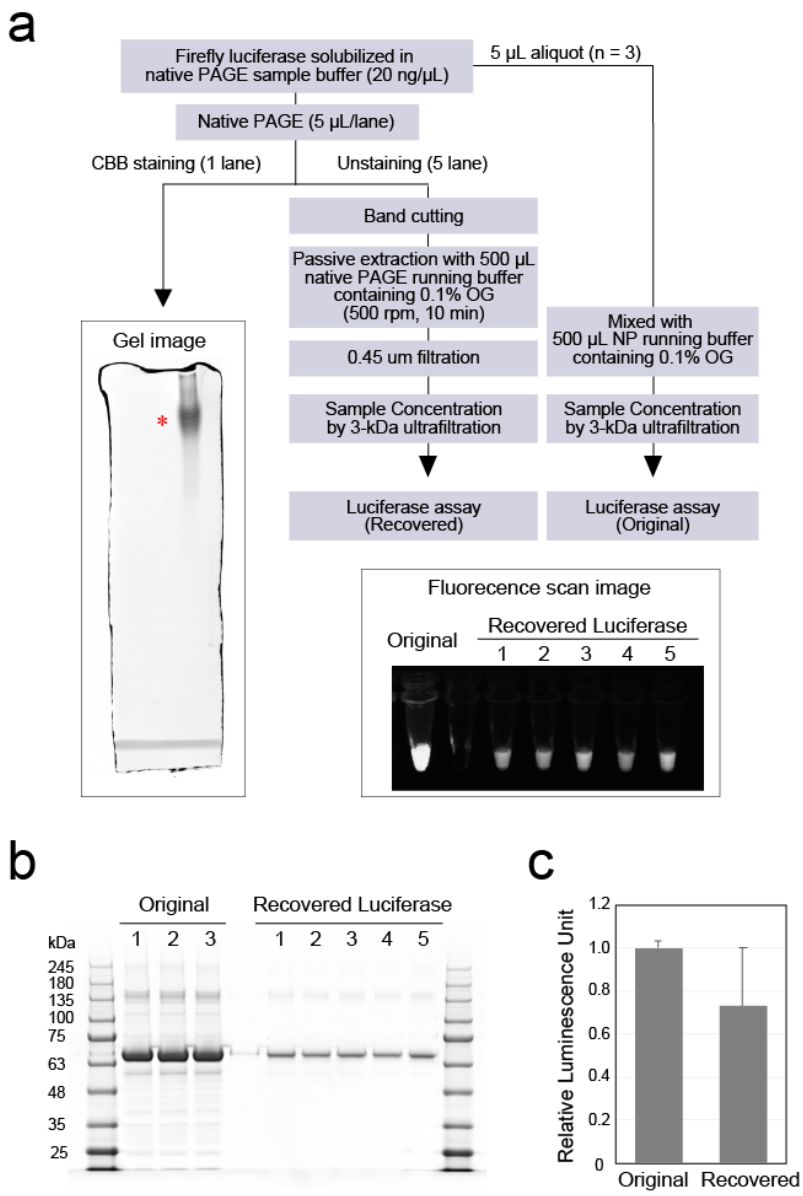

Figure S2. Analysis of gel-extracted luciferase activity following native-PAGE separation

(a) Experimental workflow. After native-PAGE separation, the gel band containing firefly luciferase (Photalight, Kikkoman Biochemifa, Tokyo, Japan) was excised and transferred to a homogenizer tube. Part of the gel was stained with CBB to determine the position of the separated luciferase in the gel (the CBB staining image is shown in the inset). In-gel luciferase was recovered in native PAGE running buffer (Novex tris-glycine native running buffer) containing 0.1% (w/v) octylglucoside by gentle shaking of the crushed gel band. Recovered luciferase samples (n = 5) and original luciferase samples (n = 3) were subjected to a luciferase activity assay. (b) Representative SDS-PAGE image of recovered luciferase. Proteins were stained with aqueous CBB. (c) Luciferase activity assay. The luciferase assay was performed with Luciferase Assay System (Promega, Madison, WI, USA) according to the manufacturer's instruction. Luminescence intensity was measured using a GloMax 20/20 luminometer (Promega). Luminescence values were normalized to the concentration of luciferase in each sample.

Table S1. Gradients for separation of PEPPI *E. coli* fractions 1-4.

Fraction	Time (min)	%A	%B
1	Initial	95	5
	5	85	15
	85	50	50
	95	30	70
	100	25	75
	102	25	75
	110	95	5
	125	95	5
2	Initial	95	5
	5	77.5	22.5
	80	40	60
	95	25	75
	98	25	75
	105	95	5
	120	95	5
3	Initial	95	5
	5	75	25
	80	50	50
	90	30	70
	95	25	75
	97	25	75
	107	95	5
	120	95	5
4	Initial	95	5
	5	72.5	27.5
	80	52.5	47.5
	90	30	70
	95	25	75
	97	25	75
	107	95	5
	120	95	5

Table S2. *E. coli* proteoforms derived from PEPPI fractions 1-4 identified by TDPortal (1% FDR).

(MS Excel file)