In-Line Separation by Capillary Electrophoresis Prior to Analysis by Top-Down Mass Spectrometry Enables Sensitive Characterization of Protein Complexes

Xuemei Han,[†] Yueju Wang,[†] Aaron Aslanian,[†,‡] Bryan Fonslow,[†,§] Beth Graczyk,[⊥] Trisha N. Davis,[⊥] and John R. Yates, III*,[†]

[†]Department of Chemical Physiology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States

[‡]Molecular and Cell Biology Laboratory, Salk Institute for Biological Studies, 10010 North Torrey Pines Road, La Jolla, California 92037, United States

[§]Sciex Separations, 250 South Kraemer Boulevard, Brea, California 92821, United States

[⊥]Department of Biochemistry, University of Washington, Seattle, Washington 98195, United States

Ⓢ Supporting Information

ABSTRACT: Intact protein analysis via top-down mass spectrometry (MS) provides a bird's eye view over the protein complexes and complex protein mixtures with the unique capability of characterizing protein variants, splice isoforms, and combinatorial post-translational modifications (PTMs). Here we applied capillary electrophoresis (CE) through a sheathless CE−electrospray ionization interface coupled to an LTQ Velos Orbitrap Elite mass spectrometer to analyze the Dam1 complex from Saccharomyces cerevisiae. We achieved a 100-fold increase in sensitivity compared to a reversed-phase liquid chromatography coupled MS analysis of recombinant Dam1 complex with a total loading of 2.5 ng (12 amol). N-terminal processing forms of individual subunits of the Dam1 complex were observed as well as their phosphorylation stoichiometry upon Mps1p kinase treatment.

KEYWORDS: capillary electrophoresis, top-down mass spectrometry, protein complexes, phosphorylation site mapping, phosphorylation stoichiometry, post-translational modification
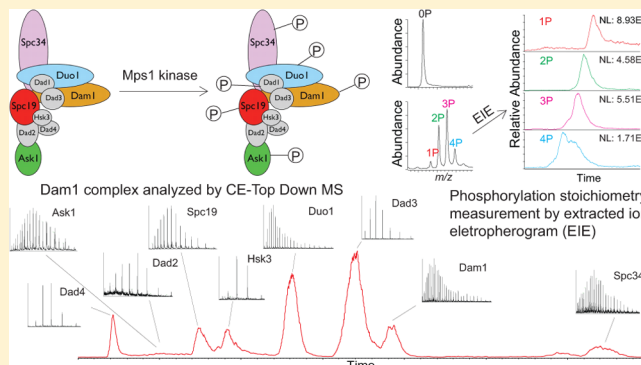
INTRODUCTION

Cellular processes are driven by molecular interactions. These include the stable association of proteins as part of a functional complex as well as the transient association proteins, such as a kinase with its substrates, leading to their phosphorylation. Disruption of these molecular interactions can lead to disease; therefore, understanding how proteins associate with one another and the consequence of these associations is of great interest. Mass spectrometry (MS)-based proteomics has grown to become a major analytical tool for the study of biologically relevant protein−protein interactions and protein PTMs.

The prevailing MS approach in the proteomics field employs proteolytic digestion of samples into peptides that are then separated, analyzed, and used to infer protein identity. Although this bottom-up MS approach is high-throughput, proteoform information pertaining to variants, splice isoforms, and combinatorial PTMs are often lost after proteolysis.[1,2] Additionally, protein processing events and unexpected PTMs are often missed during standard bottom-up MS analysis. These important protein features are better addressed by the top-down MS approach, in which the mass spectrometer directly measures and fragments intact proteins as a whole. Since its initial introduction, this technique has predominantly been utilized for targeted interrogation of individual proteins or simple protein mixtures.[3−5]

Within the past few years, the development of high-resolution MS instrumentation,[6,7] front-end separations,[8−10] and top-down software,[11−13] top-down analysis of various organisms has been reported on a proteomic scale.[14−18] However, these studies still lag behind the capabilities of bottom-up analyses by an order of magnitude in regard to protein identification and require significantly greater sample consumption. Additionally, the multidimensional front-end separations employed by these studies make it difficult to determine the biological context in which these proteoforms function.

An unexplored application that is more in line with the current top-down MS capabilities is the analysis of protein complexes. They are not as complicated as a whole proteome but still present important biological questions about which isoforms and PTMs participate in the complex, particularly in light of regulation by processes such as phosphorylation. These protein complexes are usually obtained by affinity purification, and the low yields as well as the poor purity associated with this kind of sample preparation typically compromises intact protein analysis by the top-down MS approach. The commonly used reversed-phase liquid chromatography (RPLC)-based protein separation often suffers from irreversible protein adsorption to the stationary phase and a relatively high diffusion coefficient leading to high sample consumption and poor separation efficiency and resolution for intact protein separation. These can be better addressed by capillary electrophoresis (CE) separation. With recent developments in interfacing CE to electrospray ionization (ESI)−MS,[19−23] CE has shown its application in bottom-up proteomics with its superior sensitivity and separation efficiency.[24−27] CE has also been explored for top-down analysis of intact model proteins[28,29] and simple protein mixtures.[30,31]

The Dam1 complex consists of 10 subunits as a stable heterodecamer. It is an essential component of kinetochores in yeast, making a direct contribution to tip attachment and microtubule (MT)-driven movement during mitosis.[32−34] Progression through mitosis is regulated by the activity of protein kinases, although precise kinase−substrate relationships are not completely understood. In vitro kinase assays are one of the primary means of identifying potential phosphorylation sites on substrates by a specific kinase. Previously, in vitro kinase assays have demonstrated that Dam1 can be phosphorylated by multiple mitotic kinases including Mps1p, Ipl1p, and Plo1p.[35−37] Phosphorylation of the protein Dam1 within the complex by Mps1p kinase is necessary for coupling of the kinetochore to the plus-ends of MTs. Mps1p phosphorylation sites on other components of the Dam1 complex have yet to be identified, although a CIEF analysis of the Dam1 complex following in vitro phosphorylation by Mps1p suggested that additional subunits are also substrates of the kinase.[38] Furthermore, analysis of human mitotic protein complexes has demonstrated that most protein complexes are phosphorylated on multiple subunits, and often a few subunits exhibit hyper-phosphorylation.[39] Therefore, it is reasonable to expect that additional subunits of the Dam1 complex may also be substrates of Mps1p.

In this report, we have coupled CE with an LTQ Velos Orbitrap Elite mass spectrometer through the prototype sheathless capillary electrophoresis−electrospray ionization (CESI) interface for the characterization of subunits of the *Saccharomyces cerevisiae* Dam1 complex as well as their Mps1p phosphorylated forms.

■ EXPERIMENTAL SECTION

Reagents and Chemicals

Unless otherwise stated, chemicals were purchased from Thermo Fisher Scientific (Waltham, MA). Deionized (DI) water (18.2 MΩ, Barnstead, Dubuque, IA), hydrochloric acid (Mallinckrodt Chemicals), sodium hydroxide (Sigma-Aldrich), ammonium acetate (Spectrum Chemical MFG Corp), ammonium hydroxide (J.T. Baker), methanol (J.T. Baker), acetonitrile (J.T. Baker), and isopropanol (J.T. Baker) were used for preparations.

Dam1 Complex Preparation and Mps1p Kinase Assay

Saccharomyces cerevisiae Dam1 complex was expressed in *Escherichia coli* and purified by affinity chromatography as described previously.[40] *Saccharomyces cerevisiae* Mps1p kinase was prepared as described previously.[38] The Dam1 complex was phosphorylated in 50 mM HEPES buffer, pH 7.2, containing 25 mM $MgCl_2$, 10 mM ATP, 150 mM NaCl, 10.9 mM $NaPO_4$, 120 mM imidzaole, 4 μM Dam1 complex, and 80 nM Mps1p kinase in a 25 μL reaction volume. The reaction was incubated at 30 °C for 90 min.

CE Separation of Intact Proteins

Prototype fused-silica CE capillaries (90 cm length, 30 μm ID, 150 μm OD) with a 3 cm porous tip for electrospray were obtained from Beckman Coulter, Inc. (Brea, CA). The CE capillary was positively coated using polyethylenimine (PEI) as described in U.S. Patent 6923895. CE separations were performed on a PA 800 Plus Pharmaceutical Analysis System (Beckman Coulter, Inc.). The CE capillary with a 3 cm long porous segment at one end was inserted into a sprayer interface and enabled the electric contact through a conductive liquid delivered by a second capillary. For the Dam1 complex experiments, 5% acetic acid (HAc) with 5% isopropanol (IPA) was used as the background electrolyte (BGE) buffer as well as the conductive liquid. The protein complex samples (400 ng/ μL) were first loaded into the CE capillary using 5 psi pressure for 10 s, followed by BGE injection for 10 s at 5 psi. Under the current experimental conditions, 5 psi loading pressure corresponds to a loading rate of ∼38 nL/min. The separation was then conducted by applying −30 kV to the injection end and 1.5 kV at the electrospray tip. DI water and methanol were used to rinse the capillary between runs.

Reversed-Phase Liquid Chromatography Separation of Intact Proteins

Protein samples were pressure-loaded onto a fused silica capillary tubing (75 μm i.d., with a 5 μm pulled tip) packed with 12 cm of PSDVB resins (PolyRP, 5 μm, 1000 Å, Sepax Technologies). The Agilent 1200 HPLC pump (Palo Alto, CA) was operated with a flow split to elute proteins at a constant flow rate of 300 nL/min using the flowing gradient: 5% buffer B (80% acetonitrile, 20% water, 0.1% formic acid) at start, 20% buffer B at 5 min, 65% buffer B at 50 min, 100% buffer B at 55 min, and 5% buffer B at 60 min.

Top-Down Mass Spectrometry and Data Analysis

An LTQ Velos Orbitrap Elite (Thermo Scientific) was used for top-down MS analysis. Orbitrap automatic gain control (AGC) target values for MS1, MSn, and SIM scans were 1E6. For the molecular weight profiling experiment for the Dam1 complex, four SIM scans (200 m/z window, 20 microscans, 120 000 resolving power at m/z = 400) were used to cover the mass range of 675−1450 m/z. For MS/MS experiments, fragmentation was performed in a data-dependent fashion (top 1−3 precursors from MS1 or SIM scans) with higher-energy collisional dissociation (HCD) and collision-induced dissociation (CID, 5 m/z isolation window, 10−50 microscans, 60 000 or 120 000 resolving power at m/z = 400). Dynamic exclusion was enabled with a repeat count of two, an exclusion duration of 180 s, and a repeat duration of 60 s. For HCD, the normalized collision energy setting was 15−30%, while CID used 41%.
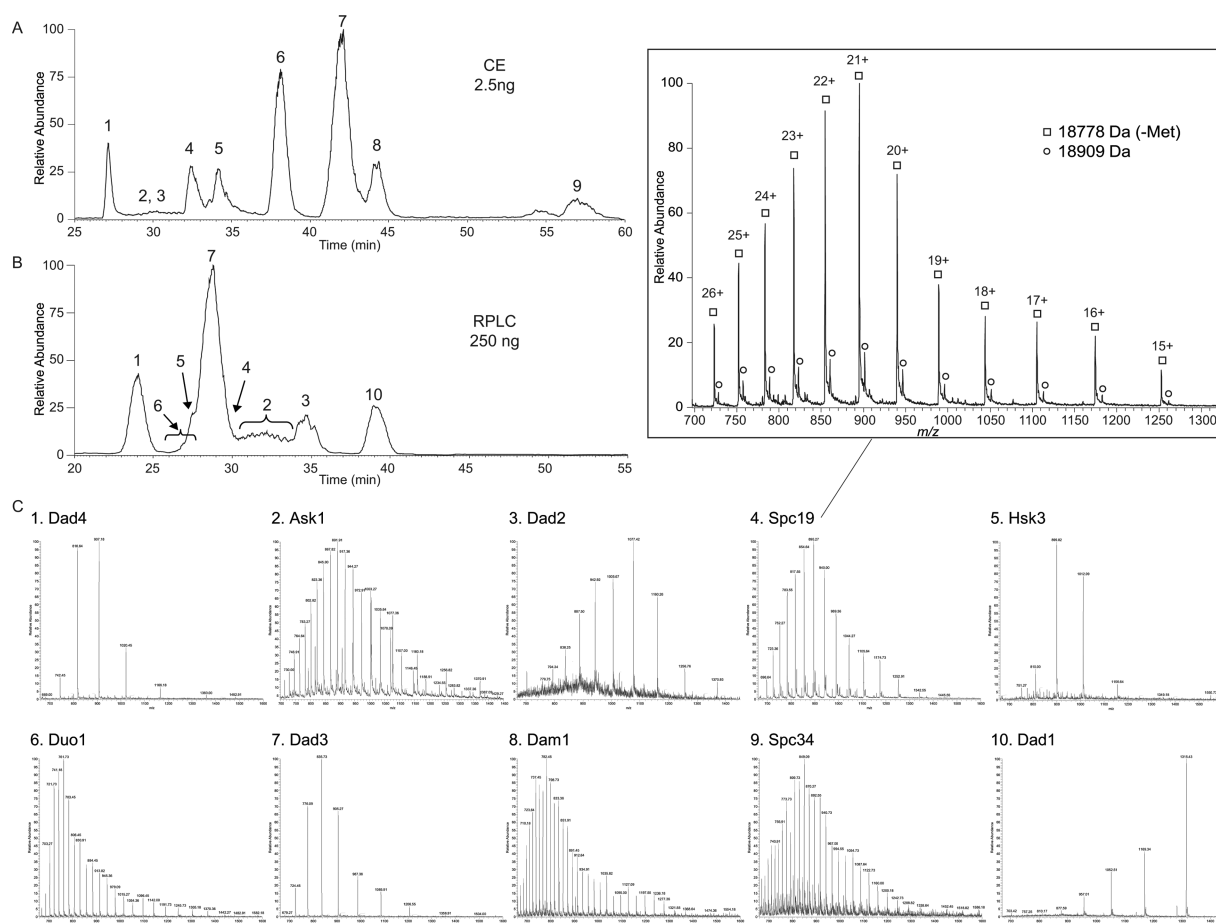
Figure 1. (A) Electropherogram of Dam1 complex subunit separation by sheathless CESI−MS (2.5 ng loaded). CE separation conditions: PEI coated capillary; voltage −30 kV; BGE 5% acetic acid, 5% IPA. (B) Chromatogram of Dam1 complex subunit separation by RPLC−MS (250 ng loaded). RPLC separation conditions: 75 um i.d. capillary with 5 μm 1000 Å PSDVB resin; flow rate of 300 nL/min. For more details see Methods and Materials section. (C) Representative MS precursor scans of the 10 protein subunits from Dam1 complex. Subunits 1−9 are from (A) the CESI−MS experiment; subunit 10 is from (B) the RPLC−MS experiment. The insert is the zoomed-in spectrum for Spc19p, showing two different protein envelopes that represent the sequence predicted form (indicated by ○, 18 909 Da, average mass) and the N-terminal methionine excision form (indicated by □, 18 778 Da, average mass).

The top-down MS/MS database searching was performed with ProSightPC 3.0 (Thermo Scientific). CID and HCD spectra were converted to monoisotopic masses using Xtract and searched against a database made via shotgun annotation[41] from the *Saccharomyces cerevisiae* Uniprot release 2012_10. The database contained 9703 basic protein sequences with 32 343 protein forms considering N-terminal acetylation and initial methionine cleavage as well as annotated PTMs. Searches were performed using a tiered, iterative absolute mass search logic that used a 20 ppm window followed by a 200 Da window of precursor mass values for protein-spectrum matches with *E*-value scores less than 10^{-4} in the 20 ppm search. Fragment ion tolerance was 15 ppm using monoisotopic masses.

Identification of Phosphorylation Sites by Bottom-Up and Middle-Down Mass Spectrometry

For bottom-up analysis, the Mps1p treated Dam1 complex proteins were either digested with trypsin (Promega) followed by a modified 10-step MudPIT analysis[42] or went through a double digestion with trypsin (Promega) and chymotrypsin (Promega) followed by a TiO$_2$-based phosphopeptide enrichment[43] and analyzed by a 4 h gradient RPLC−MS/MS. In both cases, the peptide mixtures were separated using an Agilent 1200 quaternary high-performance liquid chromatography

(HPLC) pump (Agilent Technologies) and analyzed by an LTQ Orbitrap XL mass spectrometer (Thermo Scientific). Briefly, a cycle consisted of one full scan mass spectrum (300−1600 *m/z*, 1 microscan, AGC = 1E6, 120 000 resolving power at *m/z* = 400) followed by 10 data-dependent CID MS/MS spectra (one microscan, AGC = 1E4 in the ion trap, 2 *m/z* isolation window, normalized collision energy at 35%). Application of mass spectrometer scan functions and HPLC solvent gradients were controlled by the Xcalibur data system (Thermo Scientific).

Protein identification was done with Integrated Proteomics Pipeline−IP2 (Integrated Proteomics Applications, Inc., San Diego, CA. http://www.integratedproteomics.com/) using ProLuCID and DTASelect2. MS/MS spectra were extracted using RawXtract (version 1.9.9).[44] MS/MS spectra were searched with the ProLuCID algorithm[45] against a *Saccharomyces cerevisiae* database concatenated to a decoy database in which the sequence for each entry in the original database was reversed.[46] The ProLuCID search was performed using differential modification of serine and threonine due to phosphorylation (79.9663). No enzymatic cleavage conditions were imposed on the database search. ProLuCID search results were assembled and filtered using the DTASelect (version 2.0)

Table 1. Intact Molecular Weight Measurements of Dam1 Complex Subunits Using Orbitrap Elite and Calculated Phosphorylation Stoichiometry upon in Vitro Mps1p Treatment. All Mass Values in This Table Are Monoisotopic Masses

| protein | accession | observed mass (Da) | | theoretical mass (Da) | Δmass (ppm) | phosphorylation stoichiometry | | average phosphorylation stoichiometry |
		untreated sample	Mps1p treated sample			distribution of phosphorylation states		
Hsk3	P69852	8082.998	8082.998	8083.006	−1.0			
		7951.970 (-Met)	7951.970 (-Met)	7951.966	0.6			
Dad4	P69851	8150.194	8150.194	8150.200	−0.7			
Dad1	Q12248	10509.18	10509.18	10509.19	−1.0	0P	90%	0.10
		10551.18 (Ac)	10551.18 (Ac)	10551.20	−2.0			
			10589.14 (P)	10589.16	−1.5	1P	10%	
			10631.14 (Ac, P)	10631.17	−2.5			
Dad3	P69850	10841.62	10841.62	10841.63	−0.9			
Dad2	P36162	15062.26	15062.26	15062.29	−2.0			
Spc19	Q03954	18766.59 (-Met)	18766.59 (-Met)	18766.61	−1.0	0P	100%	≪ 0.01
		18897.57	18897.57	18897.65	−4.2			
			18846.60 (-Met)	18846.58	1.3	1P[a]		
Duo1	P53168	27367.08 (-Met, Ac)		27367.17	−3.3			
			27447.14 (-Met, AC, P)	27447.14	0.1	1P	5%	2.85
			27527.10 (-Met, AC, 2P)	27527.10	−0.1	2P	27%	
			27607.07(-Met, AC, 3P)	27607.07	0.0	3P	47%	
			27687.03 (-Met, AC, 4P)	27687.04	−0.2	4P	20%	
			27767.08 (-Met, AC, 5P)	27767.00	2.8	5P	1%	
			27847.08 (-Met, AC, 6P)	27846.97	4.0	6P[a]		4.71
Ask1	P35734	32052.23		32052.20	1.0			
			32212.14 (2P)	32212.13	0.2	2P[a]		
			32292.04 (3P)	32292.10	−1.8	3P	12%	
			32372.20 (4P)	32372.07	4.2	4P	33%	
			32452.14 (5P)	32452.03	3.3	5P	34%	
			32531.87 (6P)	32532.00	−3.9	6P	16%	
			32612.05 (7P)	32611.96	2.6	7P	6%	
			32691.81 (8P)	32691.93	−3.7	8P[a]		
Spc34−6XHis	P36131	34747.85 (-Met)	34747.85 (-Met)	34747.89	−1.2	0P	54%	0.31
			34827.90 (-Met, P)	34827.86	1.2	1P	35%	
			34907.88 (-Met, 2P)	34907.82	1.6	2P	11%	
Dam1	P53267	38267.38 (-Met)		38267.41	−0.8			5.85
			38507.24 (-Met, 3P)	38507.31	−1.8	3P[a]		
			38587.41 (-Met, 4P)	38587.27	3.5	4P	15%	
			38667.19 (-Met, 5P)	38667.24	−1.3	5P	23%	
			38747.25 (-Met, 6P)	38747.21	1.1	6P	35%	
			38826.98 (-Met, 7P)	38827.17	−5.0	7P	17%	
			38907.09 (-Met, 8P)	38907.14	−1.3	8P	10%	
			38987.16 (-Met, 9P)	38987.11	1.4	9P[a]		

[a]MS1 spectra detected but EIE peak cannot be extracted due to the low abundance, resulting in too few data points for constructing an EIE peak accurately.

algorithm.[47] The peptide identification false positive rate was kept below one percent, and all peptide-spectra matches had less than 10 ppm mass error. After database searching and DTASelect2 filtering, we used DeBunker[48] and Ascore[49] to further validate phosphorylation identifications and localizations. A DeBunker score ≥ 0.5 suggests that the phosphorylation identification is valid. An Ascore value ≥ 20 indicates the phosphorylation site is localized with 99% certainty, and ≥ 13 for 95% certainty. Ascore and DeBunker analyses were done through the IP2 software.

For middle-down analysis, the Mps1p treated Dam1 complex proteins were digested with OmpT as described previously.[50] The resulting peptide mixtures were separated by RPLC as described for intact protein separation. Samples eluted from RPLC were analyzed with an LTQ Orbitrap Velos mass spectrometer (Thermo Scientific) using a data-dependent top-three method. CID or HCD was applied with 4 m/z isolation window and normalized collision energy of 41% and 35%, respectively. For MS1, four microscans at 60 000 resolving power at 400 m/z were used with a target value of 1E6 and scan range of m/z 500−2000 in the Orbitrap. For MS2, four microscans at 30 000 resolving power were used with a target value of 5E5 in the Orbitrap. The data processing procedure for middle-down proteomics analysis was done the same way as for bottom-up data analysis described previously, extracted with RawXtract, searched with ProLuCID, and filtered with DTASelect2 on IP2. The ProLuCID search was performed using differential modification of serine and threonine due to phosphorylation (79.9663) under fully tryptic cleavage condition. The PSMs were filtered with 10 ppm peptide mass
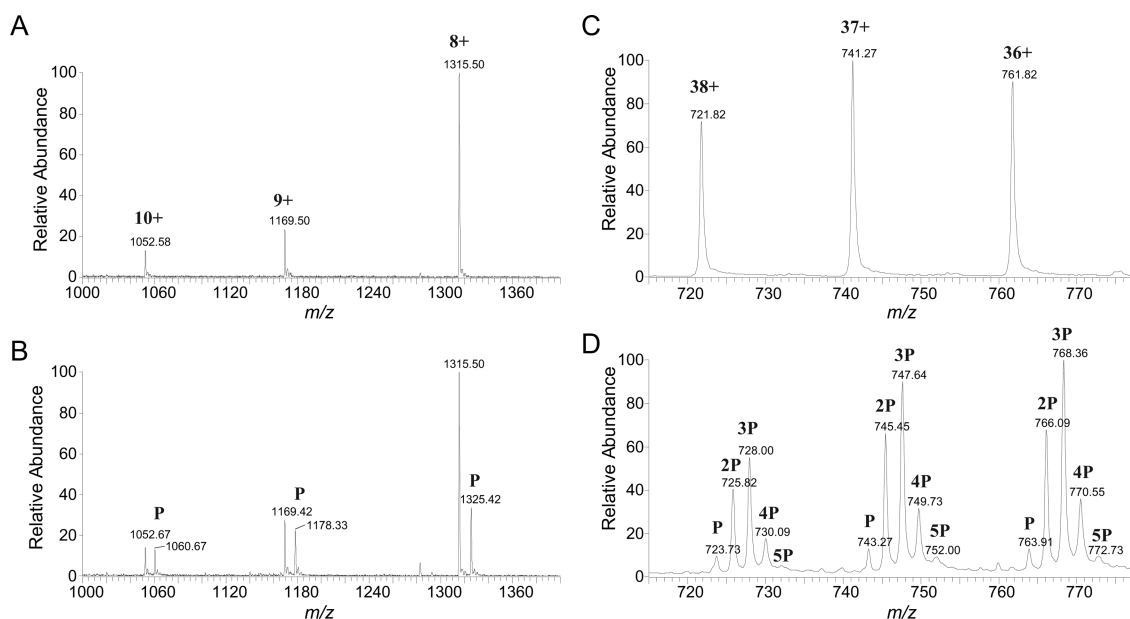
Figure 2. Comparison of the phosphorylation states of untreated Dam1 complex subunits (A) Dad1p and (C) Duo1p, and Mps1p kinase-treated Dam1 complex subunits (B) Dad1p and (D) Duo1p on selected charge states. The phosphorylated proteoforms were labeled with #P, where # is the number of phosphorylation states.

tolerance, and the peptide identification false positive rate was kept below one percent.

■ RESULTS AND DISCUSSION

CE Is Effective and Sensitive for Intact Protein Separations

Capillary electrophoresis is a powerful separation technique that has advantages such as high separation efficiency and short analysis time as well as low sample consumption. However, the use of CE−MS in proteomics is not as popular as LC-based techniques. One of the main reasons is the difficulty and lack of robustness associated with the CE−MS interface. The development of different CE−MS interfaces has begun to facilitate the successful application of CE−MS in proteomics.[20−28,51,52] In this study, we have applied the prototype CESI interface developed by Beckman Coulter Inc. to intact protein complex separation and top-down MS analysis. To prevent protein−capillary wall interactions, the CESI capillary inner surface was positively coated with PEI. The BGE buffer, including the buffer ionic strength as well as the organic modifier, was optimized for effective separation of the individual protein subunits of the Dam1 complex (Figure S1, Supporting Information). Among the BGEs tested, 5% acetic acid with 5% isopropanol gave the best protein separation as well as sensitive MS detection and therefore was chosen for the rest of the study. As shown in Figure 1, panel A, CE enabled effective separation of the individual proteins with a total loading of 2.5 ng (12 amol) of the Dam1 complex. The signal-to-noise ratio (S/N) of the protein peaks calculated by the XCalibur software (Table S1, Supporting Information) ranged from 11 (peaks 2,3) to 629 (peak 7). Also shown in Figure 1, panel C are the individual protein spectra detected by the LTQ mass spectrometer indicating the charge state distribution of each protein subunit and demonstrating the separation efficiency of CE. We also conducted LC separations of the Dam1 complex using an in-house packed reversed-phase column (Figure 1B) and were able to detect eight of the 10 protein subunits at similar S/N ratios (Table S1, Supporting Information).

However, the total sample required was 250 ng, which is 100-fold more than that required for CE-based separation. CE separation combined with LTQ detection was able to reproducibly detect nine of the 10 protein subunits, missing Dad1, a small and acidic protein (MW = 10512 Da, pI = 4.3), whereas the LC-based separation had difficulty detecting larger proteins (Dam1p and Spc34p), probably due to irreversible protein adsorption onto the stationary phase. Moreover, LC-based separation had difficulty baseline resolving some of the protein subunits, which led to protein coelution and made the downstream MS characterization of the individual proteins challenging. Overall, CE showed better separation efficiency and resolution with 100-fold less sample consumption compared to an LC-based intact protein separation.

Intact Molecular Weight Measurements and N-Terminal Processing Determination of the Dam1 Complex Subunits

By utilizing the optimized separation conditions obtained by CE, we further conducted CE−MS analysis using the LTQ Velos Orbitrap Elite. The Orbitrap mass analyzer was used at a mass resolution of 120 000 (at m/z = 400) for intact protein mass detection. Table 1 lists the detected masses of the 10 protein Dam1 complex subunits. Compared to the database reported masses, the average mass accuracy is 1.84 ppm. With the advantage of this accurate mass detection, we were able to identify different proteoforms associated with each of the protein subunits (Table 1). For example, N-terminal methionine excision (NME) forms were detected for Hsk3p and Spc19p as well as their nonprocessed forms; Spc34p and Dam1p were detected only in their NME form; and both the N-terminal acetylated form as well as the nonprocessed form were detected for Dad1p, while Duo1p shows a single NME form with acetylation. Figure 1, panel C (inset) shows two distinct protein envelopes for Spc19p, the deconvoluted mass of the minor protein envelopes matches well with the sequence predicted mass of Spc19p, while the major protein envelopes correlate with the NME form.

Phosphorylation Comparison of the Mps1p-Treated versus Untreated Dam1 Complex

Phosphorylation of the protein Dam1p within the complex by the Mps1p kinase has been studied previously and shown to play an important role in regulating the molecular switch for coupling kinetochores to the plus-ends of MTs.[35] However, the phosphorylation of the other subunits within the complex by Mps1p kinase is not well-studied, although it is known that another mitotic kinase, Ipl1p, phosphorylates multiple sites on four of the subunits of the Dam1 complex (Ask1p, Dam1p, Spc34p, and Spc19p).[36] Previously, a capillary isoelectric focusing (CIEF) analysis of the Dam1 complex was able to separate two different phosphorylation states from an in vitro kinase assay with Mps1p kinase,[38] although the nature of these two states was not investigated. Further analysis with MS is necessary to elucidate the phosphorylation stoichiometry within the entire Dam1 complex. Under similar reaction conditions to the previous study, the Mps1p kinase (90 min incubation) was capable of phosphorylating six subunits of the Dam1 complex. The CE electropherogram and RPLC chromatogram for the separation of the Mps1p kinase-treated Dam1 complex subunits compared to the untreated sample are shown in the Figure S2 of the Supporting Information. The data indicate that this charge-influencing PTM (especially hyper-phosphorylation on some subunits) affected the CE separation in terms of retention time and peak shape/intensity as phosphorylated and non-phosphorylated proteins carry different charges and masses that render different electrophoretic mobilities. To examine specific subunits in greater detail, the comparison of mass spectra of the Mps1p-treated and untreated samples for the subunits Dad1p and Duo1p is shown in Figure 2. Upon Mps1p treatment, Dad1p showed a single phosphorylated form, although the residual presence of the unphosphorylated form indicated that the in vitro reaction was not complete (Figure 2A,B). In contrast, the kinase treatment generated multiple phosphory-lated forms (mainly 1P to 5P) of Duo1p with no detectable unphosphorylated form (Figure 2C,D). This suggests that Duo1p is a better Mps1p substrate than Dad1p. For further phosphorylation analysis, we also plotted the extracted ion electropherogram (EIE) for the different phosphorylation states of each of the subunits. Figure 3 is an example EIE for the Duo1p major phosphorylated forms. The elution time clearly shows a partial separation of Duo1p forms with different numbers of phosphorylation events, with the more phosphory-lated form (4P for Duo1p in Figure 3) eluting earlier than the less phosphorylated ones. The reason for this is that CE separates analytes according to their mass and charge differences, and phosphorylation increases both the total protein mass and more importantly the overall negative charge, which results in a change in electrophoretic mobility and thus the CE separation of protein forms with different numbers of phosphorylation events. Using the EIEs constructed from protein forms with different numbers of phosphorylation events, the peak area was calculated and used as an estimation of the overall phosphorylation stoichiometry upon treatment with Mps1p (Table 1).

In summary, this intact protein molecular weight profiling provided information on the proteoforms for each of the Dam1 complex subunits and showed diverse phosphorylation stoichiometry of six proteins within the complex (Table 1).
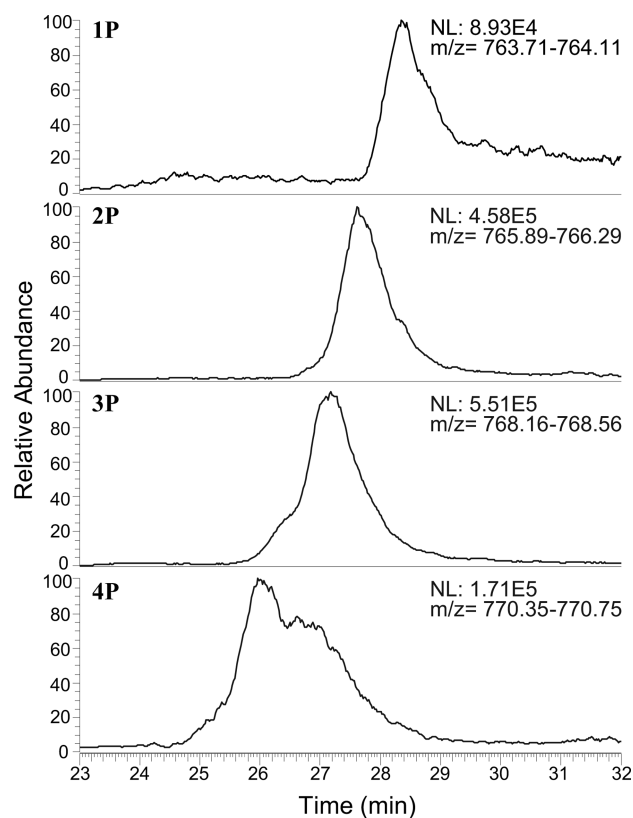

Figure 3. EIEs of the four major phosphorylation forms (1P, 2P, 3P, 4P at charge state 36+) for Duo1p following in vitro Mps1p kinase treatment. The MS detection was done with LTQ full scan mode, and the window of ±0.2 amu was used for peak extraction. The integrated peak areas were used for phosphorylation stoichiometry calculations in Table 1.

Identification of Phosphorylation Sites of the Dam1 Complex

Initially, top-down tandem MS was applied to localize the phosphorylation sites at the intact protein level. This efficiently identified a phosphorylation site (T87) on the singly phosphorylated protein Dad1p (10 kDa) (Figure 4) with 100% sequence coverage and 44% cleavage percentage (proportion of the number of observed interresidue cleavages considering all ion series among the total number of interresidue bonds in the protein). This also identified a phosphorylation site near the N-terminus of the protein Duo1p (27 kDa), although no site-determining cleavage event was observed to pinpoint the exact phosphorylation site, either S10 or T11. For some other proteins, top-down MS/MS was only able to localize the sites of phosphorylation down to certain regions rather than individual amino acids. This was partially due to the relatively low phosphorylation stoichiometry of some proteins (Spc19p and Spc34p) for which online precursor isolation failed to isolate enough precursor ions to generate good quality MS/MS spectra. The relatively large protein size (>27 kDa) poses an additional challenge for intact protein MS/MS efficiency in an electrophoregraphic time-scale. Both the CID and HCD fragmentation techniques available on our instrument were used in this study. They showed similar fragmentation efficiency for unmodified smaller proteins of the Dam1 complex, although HCD provided better fragmentation efficiency for large proteins with varying collision energies (Table S2, Supporting Information). These fragmentation
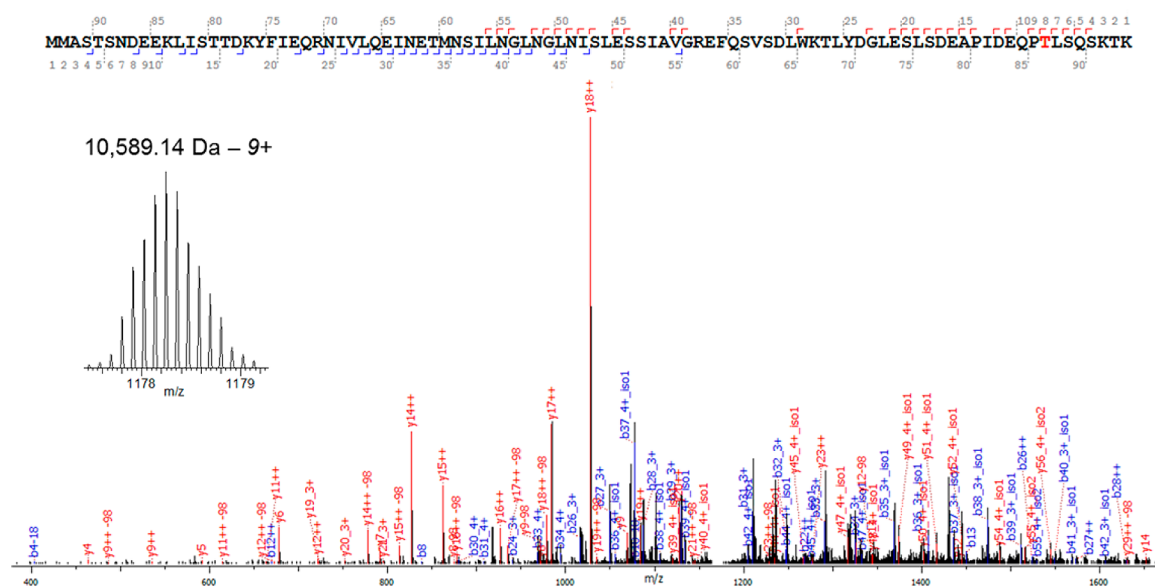
Figure 4. CID fragmentation of selected precursor ion (1178.13^{9+}) and protein identification along with the fragmentation map for Dad1p (measured MW = 10 589.14 Da). The MS/MS spectrum corresponded to the identification of Dad1p with one phosphorylation site (T87, in red) localized. The b- and y-ions were indicated as blue and red in the fragmentation map and the MS/MS spectrum.

Table 2. Identified Phosphorylation Sites on Dam1 Complex Subunits Following 90 min in Vitro Treatment With Mps1p. Combined Results from Top-Down, Bottom-Up, Middle-Down Experiments

protein name	number of phosphorylated states indicated by intact MW detection	number of phosphorylation sites identified	phosphorylation sites identified
Dad1	1	1 (0)	**T87**
Spc19	1	1 (0)	**T15**
Duo1	6	3 (6)	(S10/T11)^a; (T30/T31); **S150**; **T151**; (S221)^b; (S225)^b; **T231**^{b}
Ask1	8	11 (7)	(S56/T57); **T67**; **S134**; **T140**; **T144**; **S145**; **T146**; **T160**; **S167**; **T174**; **T189**; (S244/S249/S250)^b; (S275/T276); **T284**
Spc34	2	2 (0)	**T145**; **T199**
Dam1	9	10 (8)	**S13**^{b}; **S27**^{b}; (S36/S37); **T48**; **S49**; (S50); **T216**; **S217**; (S218); (S221); **T227**^{b}; **S232**^{b}; (S234/S235)^b; **T261**; **S265**^{b}; (S292)

^{a}Identified by top-down MS/MS experiments only. ^{b}Identified by middle-down experiments only. Numbers in the third column are localized phosphorylation sites with additional number of poorly localized sites in parentheses. Identified phosphorylation sites in bold are localized phosphorylation sites; poorly localized sites are in parentheses.

techniques tend to cleave at the protein termini, which leaves the center of the protein intact.[5] The nonergodic electron capture dissociation and electron transfer dissociation (ECD/ETD) technique and in-source fragmentation (prefolding dissociation) method[5] may be able to overcome the intractability of large proteins with conventional slow-heating ion-dissociation methods. Although the latter technique might not be helpful in the case of phosphorylation site mapping of coeluting multiply phosphorylated protein species.

Given the limitations of top-down MS/MS in mapping phosphorylation sites on the larger subunits of the Dam1 complex, bottom-up mass spectrometry was applied with either MudPIT analysis or TiO$_2$-based phosphorylation enrichment analysis. This identified and confidently localized 22 phosphorylation sites (Table S3, Supporting Information) on the same six subunits of the Dam1 complex shown to be phosphorylated using the top-down MS approach. There were an additional 11 potential phosphorylation sites identified on these proteins with poor localization scores. For the extensively phosphorylated proteins (Duo1p, Ask1p, and Dam1p), the bottom-up MS method revealed many phosphorylation sites, some of which appeared on multiply phosphorylated peptides. Bottom-up MS

experiments provided high sequence coverage of all Dam1 complex subunits. However, for Duo1p, the number of phosphorylation sites identified was less than that inferred by top-down MW measurements (Table 2). The "missed" sites may be a result of low spectra count coverage of specific sequences within a protein.

To provide better coverage of the missing sequences from the top-down and bottom-up MS approaches, a middle-down strategy was applied. We chose OmpT to produce large peptides because of its robust and restricted proteolysis as demonstrated by the Kelleher lab.[50] This method revealed new phosphorylation sites on several proteins (Table 2; Table S4, Supporting Information). Overall, the combination of top-down, bottom-up, and middle-down MS approaches has resulted in extensive mapping of in vitro phosphorylation sites on six subunits of the Dam1 complex. Among these, for the Dam1 protein, six sites (S13, S49, S217, S218, S221, and S232) were previously identified. Phosphorylation of two of the sites (S218 and S221) is required to target kinetochores to the plus-ends of MTs.[35] For Duo1p, Ask1p, and Dam1p, the total numbers of phosphorylation sites identified (9, 18, and 18, respectively) were more than inferred by top-down MW

measurements (6, 8, and 9, respectively). The reason for this discrepancy may be that not all possible phosphorylation sites are phosphorylated on any individual protein molecule. Alternatively, the heavily modified proteoforms may be of very low abundance and below the detection limit. And since we are studying the protein complex under denaturing conditions, the identified phosphorylation sites only reflect the overall effect of phosphorylation on each subunit instead of the whole complex; a previous native CIEF study[38] suggested that there are at least two different phosphorylation states of the Dam1 complex.

■ CONCLUSIONS

The in-line separation of intact proteins by CE prior to top-down mass spectrometric analysis is very suitable to analyze protein complexes that are moderately complex with very low sample consumption. This provides evidence for the biologically relevant proteoforms present within the protein complex. Top-down MS is well-suited to characterize protein variants, splice isoforms, and PTM stoichiometry. Bottom-up and middle-down MS provide complementary information regarding the site localization of PTMs.

■ ASSOCIATED CONTENT

ⓢ Supporting Information

Figure S1: Optimization of CE separation of Dam1 complex subunits. Figure S2: Electropherograms of CE separation and chromatograms of RPLC separation of Dam1 complex under no kinase treatment and Mps1 kinase treatment. Table S1: The signal-to-noise ratio (S/N) of each Dam1 complex subunit detected in CE–MS and RPLC–MS experiments. Table S2: Number of fragments in CID and HCD spectra for each subunit of the Dam1 complex. Table S3: Identified phosphorylation sites of Dam1 complex by bottom-up MS approach. Table S4: Identified phosphorylation sites of Dam1 complex by middle-down MS approach. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: jyates@scripps.edu. Phone: 858-784-8862. Fax: 858-784-8883.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Chait, B. T. Chemistry. Mass spectrometry: Bottom-up or top-down? *Science* **2006**, *314* (5796), 65–66.

(2) Kelleher, N. L. Top-down proteomics. *Anal. Chem.* **2004**, *76* (11), 197A–203A.

(3) Kelleher, N. L.; Lin, H. Y.; Valaskovic, G. A.; Aaserud, D. J.; Fridriksson, E. K.; McLafferty, F. W. Top-down versus bottom-up protein characterization by tandem high-resolution mass spectrometry. *J. Am. Chem. Soc.* **1999**, *121* (4), 806–812.

(4) Ge, Y.; Lawhorn, B. G.; ElNaggar, M.; Strauss, E.; Park, J.-H.; Begley, T. P.; McLafferty, F. W. Top down characterization of larger proteins (45 kDa) by electron capture dissociation mass spectrometry. *J. Am. Chem. Soc.* **2002**, *124* (4), 672–678.

(5) Han, X.; Jin, M.; Breuker, K.; McLafferty, F. W. Extending top-down mass spectrometry to proteins with masses greater than 200 kDas. *Science* **2006**, *314* (5796), 109–12.

(6) 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.

(7) 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.

(8) Vellaichamy, A.; Tran, J. C.; Catherman, A. D.; Lee, J. E.; Kellie, J. F.; Sweet, S. M.; Zamdborg, L.; Thomas, P. M.; Ahlf, D. R.; Durbin, K. R.; Valaskovic, G. A.; Kelleher, N. L. Size-sorting combined with improved nanocapillary liquid chromatography–mass spectrometry for identification of intact proteins up to 80 kDa. *Anal. Chem.* **2010**, *82* (4), 1234–1244.

(9) Lee, J. E.; Kellie, J. F.; Tran, J. C.; Tipton, J. D.; Catherman, A. D.; Thomas, H. M.; Ahlf, D. R.; Durbin, K. R.; Vellaichamy, A.; Ntai, I.; Marshall, A. G.; Kelleher, N. L. A robust two-dimensional separation for top-down tandem mass spectrometry of the low-mass proteome. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (12), 2183–2191.

(10) Meng, F.; Cargile, B. J.; Patrie, S. M.; Johnson, J. R.; McLoughlin, S. M.; Kelleher, N. L. Processing complex mixtures of intact proteins for direct analysis by mass spectrometry. *Anal. Chem.* **2002**, *74* (13), 2923–2929.

(11) LeDuc, R. D.; Taylor, G. K.; Kim, Y.-B.; Januszyk, T. E.; Bynum, L. H.; Sola, J. V.; Garavelli, J. S.; Kelleher, N. L. ProSight PTM: An integrated environment for protein identification and characterization by top-down mass spectrometry. *Nucleic Acids Res.* **2004**, *32* (Suppl. 2), W340–W345.

(12) Frank, A. M.; Pesavento, J. J.; Mizzen, C. A.; Kelleher, N. L.; Pevzner, P. A. Interpreting top-down mass spectra using spectral alignment. *Anal. Chem.* **2008**, *80* (7), 2499–2505.

(13) Tsai, Y.; Scherl, A.; Shaw, J.; MacKay, C. L.; Shaffer, S.; Langridge-Smith, P. R.; Goodlett, D. Precursor ion independent algorithm for top-down shotgun proteomics. *J. Am. Soc. Mass Spectrom.* **2009**, *20* (11), 2154–2166.

(14) Bunger, M. K.; Cargile, B. J.; Ngunjiri, A.; Bundy, J. L.; Stephenson, J. L., Jr. Automated proteomics of *E. coli* via top-down electron-transfer dissociation mass spectrometry. *Anal. Chem.* **2008**, *80* (5), 1459–1467.

(15) Kellie, J. F.; Catherman, A. D.; Durbin, K. R.; Tran, J. C.; Tipton, J. D.; Norris, J. L.; Witkowski, C. E., 2nd; Thomas, P. M.; Kelleher, N. L. Robust analysis of the yeast proteome under 50 kDa by molecular-mass-based fractionation and top-down mass spectrometry. *Anal. Chem.* **2012**, *84* (1), 209–215.

(16) Roth, M. J.; Parks, B. A.; Ferguson, J. T.; Boyne, M. T., 2nd; Kelleher, N. L. Proteotyping: Population proteomics of human leukocytes using top-down mass spectrometry. *Anal. Chem.* **2008**, *80* (8), 2857–2866.

(17) 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.

(18) Ansong, C.; Wu, S.; Meng, D.; Liu, X.; Brewer, H. M.; Deatherage Kaiser, B. L.; Nakayasu, E. S.; Cort, J. R.; Pevzner, P.; Smith, R. D.; Heffron, F.; Adkins, J. N.; Pasa-Tolic, L. Top-down proteomics reveals a unique protein S-thiolation switch in *Salmonella typhimurium* in response to infection-like conditions. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110* (25), 10153–10158.

(19) Smith, R. D.; Barinaga, C. J.; Udseth, H. R. Improved electrospray ionization interface for capillary zone electrophoresis–mass spectrometry. *Anal. Chem.* **1988**, *60* (18), 1948–1952.

(20) Moini, M.; Huang, H. Application of capillary electrophoresis/electrospray ionization—mass spectrometry to subcellular proteomics of *Escherichia coli* ribosomal proteins. *Electrophoresis* **2004**, *25* (13), 1981−1987.

(21) Moini, M. Simplifying CE—MS operation. 2. Interfacing low-flow separation techniques to mass spectrometry using a porous tip. *Anal. Chem.* **2007**, *79* (11), 4241−4246.

(22) Wojcik, R.; Dada, O. O.; Sadilek, M.; Dovichi, N. J. Simplified capillary electrophoresis nanospray sheath-flow interface for high efficiency and sensitive peptide analysis. *Rapid Commun. Mass Spectrom.* **2010**, *24* (17), 2554−2560.

(23) Zhong, X.; Maxwell, E. J.; Chen, D. D. Y. Mass transport in a micro flow-through vial of a junction-at-the-tip capillary electrophoresis−mass spectrometry interface. *Anal. Chem.* **2011**, *83* (12), 4916−4923.

(24) Busnel, J. M.; Schoenmaker, B.; Ramautar, R.; Carrasco-Pancorbo, A.; Ratnayake, C.; Feitelson, J. S.; Chapman, J. D.; Deelder, A. M.; Mayboroda, O. A. High capacity capillary electrophoresis−electrospray ionization mass spectrometry: Coupling a porous sheathless interface with transient-isotachophoresis. *Anal. Chem.* **2010**, *82* (22), 9476−9483.

(25) Ramautar, R.; Busnel, J. M.; Deelder, A. M.; Mayboroda, O. A. Enhancing the coverage of the urinary metabolome by sheathless capillary electrophoresis−mass spectrometry. *Anal. Chem.* **2012**, *84* (2), 885−892.

(26) Faserl, K.; Sarg, B.; Kremser, L.; Lindner, H. Optimization and evaluation of a sheathless capillary electrophoresis−electrospray ionization mass spectrometry platform for peptide analysis: Comparison to liquid chromatography−electrospray ionization mass spectrometry. *Anal. Chem.* **2011**, *83* (19), 7297−7305.

(27) Wang, Y.; Fonslow, B. R.; Wong, C. C.; Nakorchevsky, A.; Yates, J. R., 3rd Improving the comprehensiveness and sensitivity of sheathless capillary electrophoresis−tandem mass spectrometry for proteomic analysis. *Anal. Chem.* **2012**, *84* (20), 8505−8513.

(28) Haselberg, R.; de Jong, G. J.; Somsen, G. W. Capillary electrophoresis−mass spectrometry of intact basic proteins using polybrene-dextran sulfate-polybrene-coated capillaries: System optimization and performance. *Anal. Chim. Acta* **2010**, *678* (1), 128−134.

(29) Sun, L.; Knierman, M. D.; Zhu, G.; Dovichi, N. J. Fast top-down intact protein characterization with capillary zone electrophoresis−electrospray ionization tandem mass spectrometry. *Anal. Chem.* **2013**, *85* (12), 5989−5995.

(30) Zhao, Y.; Sun, L.; Champion, M. M.; Knierman, M. D.; Dovichi, N. J. Capillary zone electrophoresis−electrospray ionization-tandem mass spectrometry for top-down characterization of the *Mycobacterium marinum* secretome. *Anal. Chem.* **2014**, *86* (10), 4873−4878.

(31) Li, Y.; Compton, P. D.; Tran, J. C.; Ntai, I.; Kelleher, N. L. Optimizing capillary electrophoresis for top-down proteomics of 30−80 kDa proteins. *Proteomics* **2014**, *14* (10), 1158−1164.

(32) Cheeseman, I. M.; Enquist-Newman, M.; Muller-Reichert, T.; Drubin, D. G.; Barnes, G. Mitotic spindle integrity and kinetochore function linked by the Duo1p/Dam1p complex. *J. Cell Biol.* **2001**, *152* (1), 197−212.

(33) Jones, M. H.; He, X.; Giddings, T. H.; Winey, M. Yeast Dam1p has a role at the kinetochore in assembly of the mitotic spindle. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (24), 13675−13680.

(34) Asbury, C. L.; Gestaut, D. R.; Powers, A. F.; Franck, A. D.; Davis, T. N. The Dam1 kinetochore complex harnesses microtubule dynamics to produce force and movement. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103* (26), 9873−9878.

(35) Shimogawa, M. M.; Graczyk, B.; Gardner, M. K.; Francis, S. E.; White, E. A.; Ess, M.; Molk, J. N.; Ruse, C.; Niessen, S.; Yates, J. R., 3rd; Muller, E. G.; Bloom, K.; Odde, D. J.; Davis, T. N. Mps1 phosphorylation of Dam1 couples kinetochores to microtubule plus ends at metaphase. *Curr. Biol.* **2006**, *16* (15), 1489−1501.

(36) Cheeseman, I. M.; Anderson, S.; Jwa, M.; Green, E. M.; Kang, J.; Yates, J. R., 3rd; Chan, C. S.; Drubin, D. G.; Barnes, G. Phospho-regulation of kinetochore−microtubule attachments by the Aurora kinase Ipl1p. *Cell* **2002**, *111* (2), 163−172.

(37) Buttrick, G. J.; Lancaster, T. C.; Meadows, J. C.; Millar, J. B. Plo1 phosphorylates Dam1 to promote chromosome bi-orientation in fission yeast. *J. Cell Sci.* **2012**, *125* (Pt 7), 1645−1651.

(38) Fonslow, B. R.; Kang, S. A.; Gestaut, D. R.; Graczyk, B.; Davis, T. N.; Sabatini, D. M.; Yates, J. R., 3rd Native capillary isoelectric focusing for the separation of protein complex isoforms and subcomplexes. *Anal. Chem.* **2010**, *82* (15), 6643−6651.

(39) Hegemann, B.; Hutchins, J. R.; Hudecz, O.; Novatchkova, M.; Rameseder, J.; Sykora, M. M.; Liu, S.; Mazanek, M.; Lenart, P.; Heriche, J. K.; Poser, I.; Kraut, N.; Hyman, A. A.; Yaffe, M. B.; Mechtler, K.; Peters, J. M. Systematic phosphorylation analysis of human mitotic protein complexes. *Sci. Signaling* **2011**, *4* (198), rs12.

(40) Gestaut, D. R.; Graczyk, B.; Cooper, J.; Widlund, P. O.; Zelter, A.; Wordeman, L.; Asbury, C. L.; Davis, T. N. Phosphoregulation and depolymerization-driven movement of the Dam1 complex do not require ring formation. *Nat. Cell Biol.* **2008**, *10* (4), 407−414.

(41) Roth, M. J.; Forbes, A. J.; Boyne, M. T., 2nd; Kim, Y. B.; Robinson, D. E.; Kelleher, N. L. Precise and parallel characterization of coding polymorphisms, alternative splicing, and modifications in human proteins by mass spectrometry. *Mol. Cell. Proteomics* **2005**, *4* (7), 1002−1008.

(42) Washburn, M. P.; Wolters, D.; Yates, J. R., 3rd Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* **2001**, *19* (3), 242−247.

(43) Cantin, G. T.; Shock, T. R.; Park, S. K.; Madhani, H. D.; Yates, J. R., 3rd Optimizing TiO_2-based phosphopeptide enrichment for automated multidimensional liquid chromatography coupled to tandem mass spectrometry. *Anal. Chem.* **2007**, *79* (12), 4666−4673.

(44) McDonald, W. H.; Tabb, D. L.; Sadygov, R. G.; MacCoss, M. J.; Venable, J.; Graumann, J.; Johnson, J. R.; Cociorva, D.; Yates, J. R., 3rd MS1, MS2, and SQT-three unified, compact, and easily parsed file formats for the storage of shotgun proteomic spectra and identifications. *Rapid Commun. Mass Spectrom.* **2004**, *18* (18), 2162−2168.

(45) Xu, T.; Venable, J. D.; Park, S. K.; Cociorva, D.; Lu, B.; Liao, L.; Wohlschlegel, J.; Hewel, J.; Yates, J. R. ProLuCID, a fast and sensitive tandem mass spectra-based protein identification program. *Mol. Cell. Proteomics* **2006**, *5* (10), S174−S174.

(46) Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC−MS/MS) for large-scale protein analysis: The yeast proteome. *J. Proteome Res.* **2003**, *2* (1), 43−50.

(47) Tabb, D. L.; McDonald, W. H.; Yates, J. R., 3rd DTASelect and Contrast: Tools for assembling and comparing protein identifications from shotgun proteomics. *J. Proteome Res.* **2002**, *1* (1), 21−26.

(48) Lu, B.; Ruse, C.; Xu, T.; Park, S. K.; Yates, J., 3rd Automatic validation of phosphopeptide identifications from tandem mass spectra. *Anal. Chem.* **2007**, *79* (4), 1301−1310.

(49) Beausoleil, S. A.; Villen, J.; Gerber, S. A.; Rush, J.; Gygi, S. P. A probability-based approach for high-throughput protein phosphorylation analysis and site localization. *Nat. Biotechnol.* **2006**, *24* (10), 1285−1292.

(50) Wu, C.; Tran, J. C.; Zamdborg, L.; Durbin, K. R.; Li, M.; Ahlf, D. R.; Early, B. P.; Thomas, P. M.; Sweedler, J. V.; Kelleher, N. L. A protease for 'middle-down' proteomics. *Nat. Methods* **2012**, *9* (8), 822−824.

(51) Maxwell, E. J.; Zhong, X.; Chen, D. D. Y. Asymmetrical emitter geometries for increased range of stable electrospray flow rates. *Anal. Chem.* **2010**, *82* (20), 8377−8381.

(52) Heemskerk, A. A.; Wuhrer, M.; Busnel, J. M.; Koeleman, C. A.; Selman, M. H.; Vidarsson, G.; Kapur, R.; Schoenmaker, B.; Derks, R. J.; Deelder, A. M.; Mayboroda, O. A. Coupling porous sheathless interface MS with transient-ITP in neutral capillaries for improved sensitivity in glycopeptide analysis. *Electrophoresis* **2013**, *34* (3), 383−387.