Mass Spectrometry Analysis of Proteome-Wide Proteolytic Post-Translational Degradation of Proteins

7 AUTHORS, INCLUDING:

Yufeng Shen
Pacific Northwest National Laboratory
113 PUBLICATIONS **8,785** CITATIONS

SEE PROFILE

Kim Hixson
Pacific Northwest National Laboratory
41 PUBLICATIONS **2,027** CITATIONS

SEE PROFILE

Ronald J Moore
Pacific Northwest National Laboratory
125 PUBLICATIONS **5,436** CITATIONS

SEE PROFILE

Richard D Smith
Pacific Northwest National Laboratory
1,132 PUBLICATIONS **46,141** CITATIONS

SEE PROFILE

Mass Spectrometry Analysis of Proteome-Wide Proteolytic Post-Translational Degradation of Proteins

Yufeng Shen,*,† Kim K. Hixson,‡ Nikola Tolić,‡ David G. Camp,† Samuel O. Purvine,‡ Ronald J. Moore,† and Richard D. Smith*,†,‡

Biological Science Division, and Environmental Molecular Sciences Laboratory, Pacific Northwest National Laboratory, Richland, Washington 99352

Protein proteolytic degradation is an essential component to proper cell function and its life cycle. Here, we study the protein degradation in yeast *Saccharomyces cerevisiae* cells on a proteome-wide scale by detection of the intermediate peptides produced from the intracellular degradation of proteins using sequencing-based tandem mass spectrometry. By tracing the detected ∼1100 peptides and their ∼200 protein−substrate origins we obtain evidence for new insights into the proteome-wide protein-selective degradation in yeast cells. This evidence shows that the yeast cytoplasm is the largest pool for the degradation of proteins with both biochemical and geometric specificities, whereas the yeast nucleus seems to be a proteolysis-inert organelle under the condition studied. Yeast V-ATPase subunits appear to be degraded during their disassembly, and yeast mitochondrial proteins functioning as precursors, transport carriers, and gates are preferentially degraded. Ubiquitylation may be unnecessary for the proteasomal degradation of yeast cytoplasmic regulatory and enzyme proteins according to our observations. This study shows that the intracellular peptides are informational targets for directly probing the protein degradation-involved molecular mechanisms and cell biology processes.

Protein degradation is a post-translational modification necessary for protein quality control and cell survival. In eukaryotic cells degradation of proteins is primarily carried out in either the proteasomes or lysosomes. Proteasomes[1,2] are responsible for protein quality control by breaking down misfolded proteins,[3] and they have a well-defined geometric proteolysis catalytic center[4,5] that breaks down the proteins to produce peptides with specific lengths.[6,7] Ubiquitylation[7,8] of proteins is the typical initial step for protein degradation through the proteasomes (note: ubiquitylation also functions for nonproteolytic purposes[10]). Resultant peptides from proteasomal degradation can then be further cleaved into amino acids by cytosolic aminopeptidases.[11] Lysosomes, on the other hand, are membrane-bound organelle-like structures containing multiple proteases that degrade protein aggregates and membrane proteins.[3,12] Proteins to be degraded are transported to the lysosome through, e.g., the ATGs/ signaling/regulating systems[13,14] and the following proteolyses are catalyzed by endoproteinases[15,16] that are matured and activated within the lysosome. The resultant peptides continue to break down into amino acids by the lysosomal aminopeptidases and carboxypeptidases,[16,17] and the products (e.g., amino acids and others such as sugars, cholesterols) are then transported back out into the cytosol by the solute transporters in the lysosomal membrane and can then be recycled as nutrients making this degradation process important for cell survival when cells are subjected to stresses such as nutrient limitation.[13,14]

Protein degradation has been primarily studied using the methods based on radioactivity,[3] which have proven to be informative on an overall level but provided little information regarding the degradation steps occurring to individual proteins. Such information currently has been gained by performing in vitro experiments[6,7,13–17] with individual model proteins/peptides and proteases, but evidence provided from these in vitro experiments do not represent the true in vivo scenarios due to the inability of a few single proteins to truly mimic the regulatory processes and complexity of degradation occurring to multiple proteins at the same time within a cell. Additionally, current ideas on protein degradation concede that in vivo monitoring of overall cellular protein degradation is less possible because it is believed that

* To whom correspondence should be addressed. E-mail: yufeng.shen@pnl.gov (Y.S.), dick.smith@pnl.gov (R.D.S.).
† Biological Science Division.
‡ Environmental Molecular Sciences Laboratory.

(1) Goldberg, A. L. *Nature* **2003**, *426*, 895–899.
(2) Demartino, G. N.; Gillette, T. G. *Cell* **2007**, *129*, 659–662.
(3) Lee, D. H.; Sherman, M. Y.; Goldberg, A. L. *Mol. Cell. Biol.* **1996**, *16*, 4773–4781.
(4) Goldberg, A. L.; Rock, K. L. *Nature* **1992**, *357*, 375–379.
(5) Baumeister, W.; Walz, J.; Zühl, F.; Seemüller, E. *Cell* **1998**, *92*, 367–380.
(6) Wenzel, T.; Eckerskorn, C.; Lottspeich, F.; Baumeister, W. *FEBS Lett.* **1994**, *349*, 205–209.
(7) Kisselev, A. F.; Akopian, T. N.; Woo, K. M.; Goldberg, A. L. *J. Biol. Chem.* **1999**, *274*, 3363–3371.
(8) Hershko, A.; Ciechanover, A. *Annu. Rev. Biochem.* **1998**, *67*, 425–479.
(9) Weissman, A. M. *Nat. Rev. Mol. Cell Biol.* **2001**, *2*, 169–178.
(10) Welchman, R. L.; Cordon, C.; Mayer, R. J. *Nat. Rev. Mol. Cell Biol.* **2005**, *6*, 599–609.
(11) Scaric, T.; Graef, C. I.; Goldberg, A. L. *J. Biol. Chem.* **2004**, *279*, 46723–46732.
(12) Rubinsztein, D. C. *Nature* **2006**, *443*, 780–786.
(13) Klionsky, D. J.; Emr, S. D. *Science* **2000**, *290*, 1717–1721.
(14) Yoshimori, T. *Cell* **2007**, *128*, 833–836.
(15) Knop, M.; Schiffer, H. H.; Pupp, S.; Wolf, D. H. *Curr. Opin. Cell Biol.* **1993**, *5*, 990–996.
(16) Jones, E. W. *Methods Enzymol.* **2002**, *351*, 127–150.
(17) Rendueles, P. S.; Wolf, D. H. *FEMS Microbiol. Rev.* **1988**, *4*, 17–45.

intermediate peptides being produced by the proteasomes or lysosomes are broken down so quickly into amino acids by peptidases that the sequence of protein degradation is undetectable except for special cases such as with the major histocompatibility complexes.[18,19] The proteolytic degradation of proteins in cells on the proteomewide level basically is still not understood fully.

In this work, we introduce a mass spectrometry (MS)-based method to study proteome-wide protein degradation. Yeast *Saccharomyces cerevisiae* was used as a model organism due to its simplicity and mechanistic homology to higher eukaryotic cells. Various aspects of information for yeast are now available including protein sequences (ftp://genome-ftp.stanford.edu/pub/yeast/data_download/sequence/GenBank/), localization,[20] expression levels,[21] lifetimes,[22] and proteolytic activities of proteasomal and lysosomal (vacuolar) proteinases and peptidases (summarized in the Supporting Information). We demonstrate that protein degradation can be probed on a proteome-wide scale by sequencing the protein degradation intermediate peptides with high-precision tandem MS (i.e., MS/MS). The obtained information provides evidence for insights and better understanding of the molecular mechanisms and processes involved in protein proteolytic degradation.

RESULTS

Identification of Intracellular Peptides from Yeast Lysate Using Sequencing-Based Mass Spectrometry. Figure 1 shows an LC−MS display of molecular species having masses <10 000 u from a yeast cell lysate. In this display, ∼8000 different molecular species exist. Identification of these displayed molecules was accomplished with accurate MS/MS data that were concurrently acquired during the LC−MS measurements (see Methods). Differing from the conventional probability-based scoring approaches, we developed a sequencing-based UStag method (see Methods) to achieve unambiguous identification of various lengths and termini of peptides with expected/unexpected post-translational modifications. In total, we identified 1111 peptides (Supporting Information Table 1) having lengths of 6−100 amino acids (those having ∼20 amino acids were most frequently observed, see Supporting Information Figure 1). Among these peptides, 59 were found as modified with mutations, acetylations, amidations, etc. The correctness of the identifications was evidenced with spectral similarity, correct molecular masses and isotopic envelope patterns, correct fragmental masses, correct residues, and the database uniqueness of the sequences constructed from the consecutive fragments and the residues measured even with consideration of various possible modifications (see Methods). The random false positives were tested by searching all acquired MS/MS spectra (30 760 in total) against a reverse-ordered yeast protein sequence database using the same method as that used for peptide identification, and no false positives were found except for one with a symmetric sequence (i.e., EKAKE, from YAL038W).
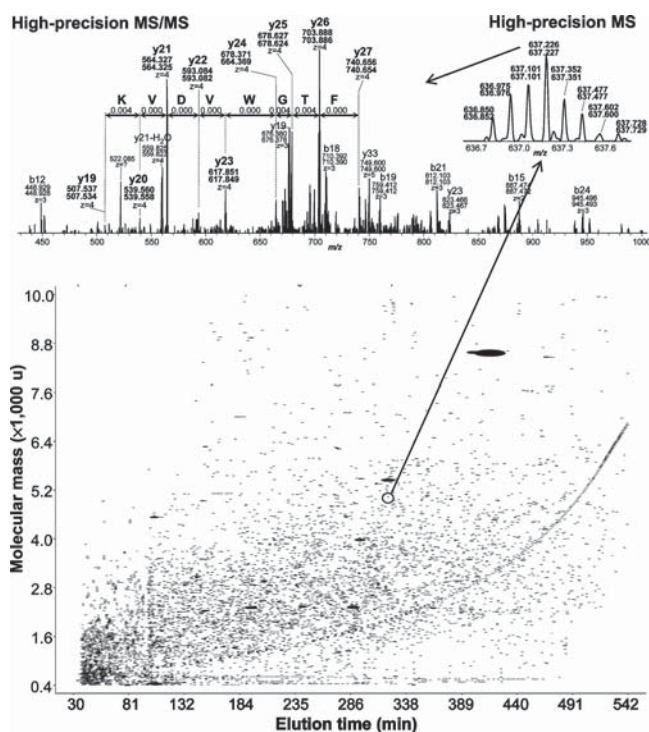
(18) Falk, K.; Rötzschke, O.; Rammensee, H. G. *Nature* **1990**, *348*, 248–251.
(19) Reits, E. *Immunity* **2003**, *18*, 97–108.
(20) Huh, W.-K.; Falvo, J. V.; Gerke, L. C.; Carroll, A. S.; Howson, R. W.; Weissman, J. S.; O'Shea, E. K. *Nature* **2003**, *425*, 686–691.
(21) Ghaemmaghami, S.; Huh, W.-K.; Bower, K.; Howson, R. W.; Belle, A.; Dephoure, N.; O'Shea, E. K.; Weissman, J. S. *Nature* **2003**, *425*, 737–741.
(22) Belle, A.; Tanay, A.; Bitincka, L.; Shamir, R.; O'Shea, E. K. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 13004–13009.

Figure 1. LC−MS display of molecular species having molecular masses of 400−10 000 u in a yeast *Saccharomyces cerevisiae* lysate. Species with >3 MS scans were counted. The mass accuracies for both amino acid sequencing and individual mass peaks are illustrated. The sequencing-based interpretation of high-precision MS/MS and experimental conditions are described in detail in the Methods section.

Peptide Origins Found from the Yeast Cell Lysate. Each peptide identified was examined to determine its possible biochemical/biological origins. These origins include a protein maturation process (protein cleavage that transforms inactive proteins into active conformations), a partially assembled protein (a protein in the middle of being synthesized by a ribosome at the time of cell harvest), a product of a protein localization process, a part of a protein signaling process, or a product of a protein degradation process. The identified peptides that were a component of the first 70 amino acids of the full predicted sequence of each protein were analyzed using SignalP[23] to examine the signal peptides, and only one signal peptide was found from the yeast protoplast secreted protein 2 precursor. The possible products from proprotein convertases (Kex2 in yeast),[24] which cleaves polypeptide sequences at dibasic amino acid sites (i.e., RR, RK, KR, and KK) to convert proproteins into mature protein forms and which leaves propeptides behind, were examined; and 4 of the identified peptides were potential candidates due to the dibasic characteristic cleavage sites. However, these four peptides were more likely the products of degradation or partially synthesized sequences, as they were found also to be the components of peptide ladders which were typically seen from protein degradation shown in this study. We did not observe any evidence of intramembrane proteolysis under the growth condition studied, and it is unknown whether or not yeast *S. cerevisiae* can generate

(23) Bendtsen, J. D.; Nielsen, H.; von Heijne, G.; Brunak, S. *J. Mol. Biol.* **2004**, *340*, 783–795.
(24) Rockwell, N. C.; Krysan, D. J.; Komiyama, T.; Fuller, R. S. *Chem. Rev.* **2002**, *102*, 4525–4548.

Figure 2. Peptide ladders showing yeast phosphoglycerate kinase sequences occurring in the cells at the time of cell harvest. The peptides observed had lengths of 12−70 amino acid residues; the curved sequence unnecessarily represents the protein real folding.

<10 000 u protein fragments from protease cascade processing for these membrane proteins. The protein synthesis, starting from the protein N-termini, may also generate N-terminal peptides due to the interruption of synthesis at the time of cell harvest, and we found that peptides corresponding to 15 proteins were potential products of the protein partial synthesis as they were not accompanied with other peptide ladders.

Figure 2 is an example of the peptide ladder pattern observed for a yeast protein (phosphoglycerate kinase) in the cell lysate. Observed is a sequential cutting of the protein into smaller and smaller pieces from the long peptides (the longest peptide detected having 70 amino acids for this example). The peptides were verified to be present in the lysate before LC−MS analysis as these ladder sequences eluted at different LC elution time profiles (not shown). Nondegradation-related intracellular proteolysis processes (e.g., N-terminal signal peptide, propeptides, and partial protein synthesis etc.) do not explain these ladders as many specific cleavages were located within the middle of the sequence or near the C-terminus of the protein. These ladders were similar to the proteasome peptide map observed in previous in vitro experiments with model proteins,[25] and the peptides detected from our experiment have a large range in size (e.g., up to 75 AA residues for the shown example) and display cleavage preferences (e.g., various peptides in the ladders starting at specific points).

(25) Nussbaum, A. K.; Dick, T. P.; Keilholz, W.; Schirle, M.; Dietz, K.; Heinemeyer, W.; Groll, M.; Wolf, D. H.; Huber, R.; Rammensee, H.-G.; Schild, H. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12504–12509.

Organellar Selectivity of the Detected Peptides' Protein Origins. We examined the protein origins for the peptides identified and found that 205 yeast proteins were the source of the 1111 peptides identified (Supporting Information Table 1). These proteins have expressions ranging from low (e.g., ~150 molecules/cell for Hsp 10 kDa) to high (e.g., ~1×10^6 molecules/ cell for fructose-bisphosphate aldolase) levels[21] and half-lives ranging from short (e.g., 10 min for vacuolar ATPase subunit E) to long (e.g., >1500 min for viral protein U-binding protein) ones.[22] The number of peptides identified did not necessarily correlate with the protein expression levels; e.g., only two peptides were detected for both Hsp SSB2 (1×10^5 molecules/cell documented) and the 60S ribosomal protein L3 (450 molecules/cell documented).

We next looked for the subcellular location of these proteins using information from http://db.yeastgenome.org/cgi-bin/search/ featureSearch. Of the 205 proteins found, 87 colocalize in multiple organelles and 118 belong to a single organelle (Supporting Information Table 2). Figure 3 shows the organellar distribution of the 118 single organelle-specific proteins. It clearly shows that these single organelle-specific proteins showed a high selectively from a specific yeast organelle such as cytoplasm and mitochondria, whereas other organelles including the yeast nucleus, cell wall, and cytoplasmic membrane did not release any single peptide (either hydrophilic or hydrophobic one). This finding excludes the possible contribution of artifact proteolysis (see the Discussion section below) during and after the cell lysis process, which was completed in ~200 s in the presence of protease inhibitors (see
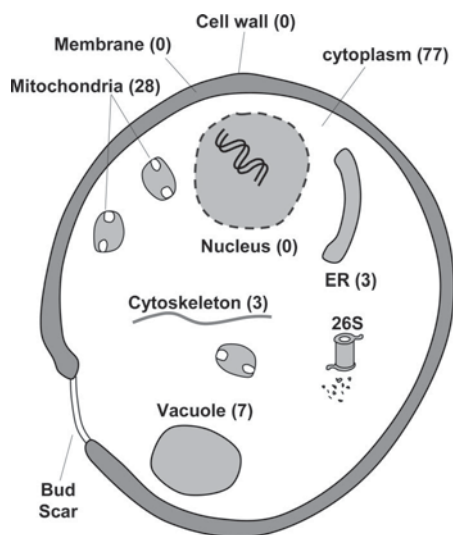
Figure 3. Localization of 118 yeast proteins observed with the existence of their intracellular peptides. Proteins observed in each organelle are described in the following sections.

Methods), and suggests that the identified peptides were products from the breakdown of specific proteins in specific locations. Excluding those from potential nonproteolysis processes described above, we ascribe the identified peptides to the products of the protein intracellular in vivo degradation. Additionally, the organelle selectivity observed for protein breakdown also excluded lysosome-involved proteolyses as these types of proteolyzes are preferential in the breakdown of membrane proteins,[3,13,14] which were not found. This claim is consistent with fact that there was a lack of inducers to start the lysosome-involved apoptosis and autophagy from the nutrient-rich, rapid growth environments that the yeast cells were exposed to at the time of harvest in this study.

Proteolytic Degradation of Yeast Cytoplasmic Proteins. Yeast cytoplasm was found to be the largest pool where the breakdown products of 77 of 118 single organelle-specific proteins were found. Ribosomal, stress-response, regulatory, and enzymatic proteins contributed to the large portion of peptides found from cytoplasmic proteins (Table 1). Most ribosomal and stress-response protein products found were ubiquitylation proteins (http://ubiprot.org.ru/), and their proteolysis may be explained by the ubiquitylation–proteasome pathway. In contrast, most proteins annotated as regulatory and enzymatic in function were not ubiquitylation proteins. Proteasomes are the major proteolytic machine in yeast cytoplasm, which leads us to hypothesize that the degradation of these regulatory and enzymatic proteins found (or some of those) should be carried out through proteasomes without ubiquitylation. Next we examined the proteolytic specificity of the peptides detected (Figure 4). At the amino acid cleavage site P1, trypsin-like and chymotrypsin-like proteolyses, i.e., the activities of yeast proteasome $\beta2$/PUP1 and $\beta5$/PRE2, were observed. The cleavage preferences for Leu and Ala were also observed, which agrees with the specificity of yeast proteasomal $\beta1$/PRE3, $\beta2$/PUP1, and $\beta5$/PRE2.[25] Removal of Met from the protein N-terminus[26] was responsible for the high-frequency observation of P1 Met (of the 57 P1 Met found, 47 were located at the N-termini). At P1′ we found the cleavage preference for

(26) Li, X.; Chang, Y. H. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 12357–12361.

Table 1. Properties of the Yeast Cytoplasm-Specific Proteins Observed from Intermediate Peptides Detected in This Study[a]

Ribosomal Proteins 60S:		
P0	NR	Ubi
L10	NR	Ubi
L11-B	300 (71)	Ubi
L16-B	NR	Ubi
L19	48 (985)	Ubi
L2	−257	Ubi
L20-A	137 (300)	Ubi
L28	NR	Ubi
L3	120	Ubi
L32	NR	Ubi
L34-B	69	Ubi
L38	NR	NR
L4-A	NR	Ubi
L4-B	300	Ubi
L5	NR	NR
L7-A	113	NR
40S:		
S12	NR	Ubi
S15	NR	NR
S1-B	NR	Ubi
S2	NR	NR
S20	NR	Ubi
S27-A	NR	Ubi
S3	NR	Ubi
S4	53	Ubi
S5	NR	Ubi
S6	100	Ubi
S7-B	300	Ubi
Enzyme Proteins:		
hexokinase-1	49	NR
NADP cytoplasmic	69	NR
phosphoglycerate mutase 1	300	Ubi
phosphoserine aminotransferase	NR	NR
mannose-1-phosphate guanyltransferase	NR	NR
phosphoenolpyruvate carboxykinase	NR	NR
argininosuccinate lyase	NR	NR
fructose-bisphosphate aldolase	104	NR
alcohol dehydrogenase 1	NR	Ubi
γ-cystathionase	NR	NR
glutamate decarboxylase	62	NR
LeuRS	81	NR
TSL1	39	NR
aspartokinase	78	NR
6-phosphofructokinase subunit α	NR	NR
6-phosphofructokinase subunit β	96	NR
Regulatory Proteins:		
ran-specific GTPase-activating protein 1	NR	NR
pyruvate kinase 1	NR	NR
translation machinery-associated protein 22	NR	Ubi
translationally controlled tumor protein homologue	166	NR
eukaryotic peptide chain release factor GTP-binding subunit	NR	NR
elongation factor 2	155 (300)	Ubi
elongation factor 1-β	NR	NR
elongation factor 3A	191	NR
eIF-4E	NR	NR
eIF-4B	67	NR
Stress-Response Proteins:		
Hsp26	300	NR
Hsp104	61	Ubi
HspSSA3	59	Ubi
HspSSA4	300	Ubi
HspSSB1	149	NR
HspSSB2	712	Ubi
HspSSE1	101	NR
HspSTI1	89	Ubi
enolase 1	NR	Ubi
enolase 2	27	Ubi

[a] The number after each protein is its half-life time (min) from ref 22. Ubi: ubiquitylation proteins listed in according to http://ubiprot.org.ru/. NR: not reported.

amino acids having short side chains (e.g., Gly, Ala, Ser, and Thr), i.e., sites with small steric hindrance. This steric driven preference is likely imposed by proteasomes where a partially folded protein
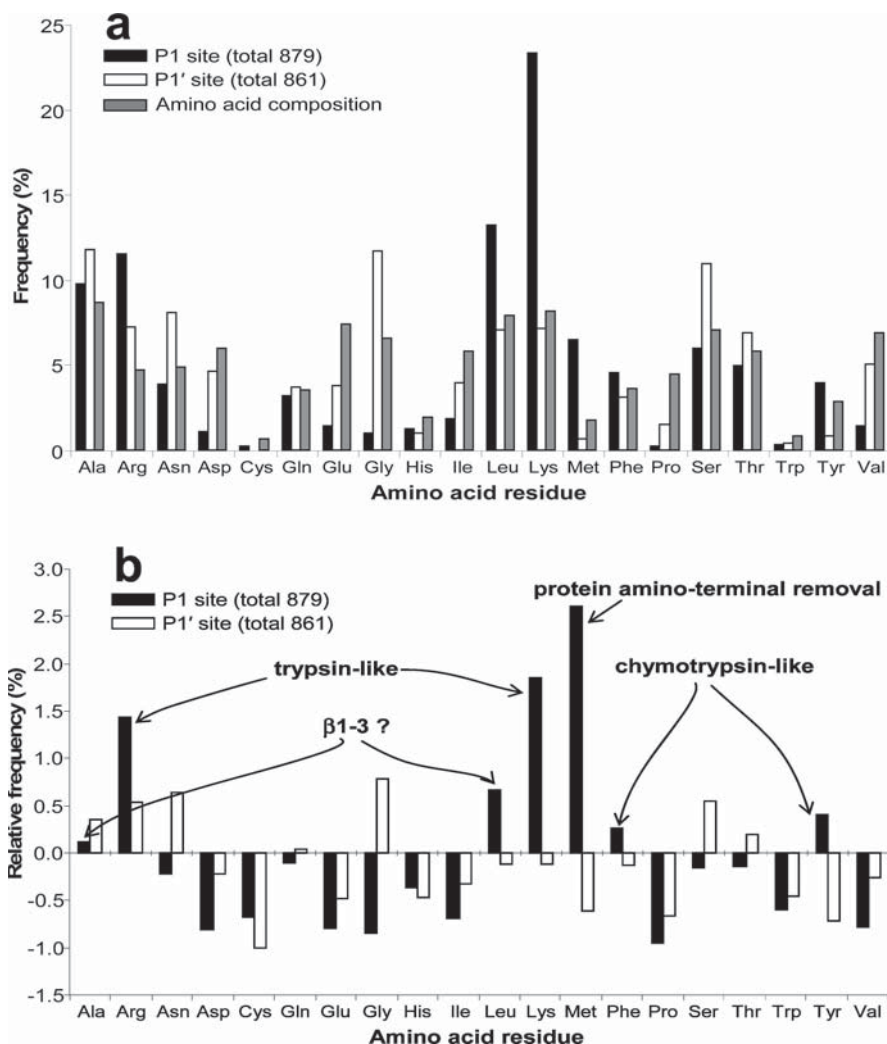
Figure 4. Proteolytic cleavage preferences of yeast *S. cerevisiae* cytoplasmic proteins. A total of 462 peptides observed for the 77 cytoplasm-specific proteins were analyzed. The cleavage sites were defined as follows: $NH_2-... -P1 \downarrow P1'-...$ COOH, where the arrow represents the proteolytic cleavage site: (a) the percentage (or frequency) of amino acids at the two cleavage sites; (b) the relative frequency (%) = [(percentage of the amino acid in all detected P1 or P1' termini − percentage of the amino acid in detected proteins)/percentage of the amino acid in detected proteins] × 100.

"loop"[27] at the location of the small amino acid enters the proteasome more easily, and these protein portions thus can access the tiny catalytic cavity[28] that allows for the protein hydrolysis. The protein "loop" model can also elucidate our observations of peptides significantly longer than those predicted by the "molecular ruler" (i.e., generation of ~eight-residue peptides) theory of proteasomal degradation[6,7] with additional consideration of kinetic factors[29] for proteolysis of the cellular proteins.

Yeast Nucleus Is Not an Active Site for Protein Proteolytic Degradation. In contrast to observations of a significant number of degraded cytoplasm-specific proteins, no single peptide from the breakdown of any nucleus-specific protein was found. In comparison to ~1600 cytoplasm-specific proteins, yeast *S. cerevisiae* is predicted to have ~900 nucleus-specific proteins that have

a variety of documented lifetimes (e.g., 2−13 000 min) and expression levels (e.g., 100−10 000 molecules/cell). The absence here in observation of nucleus-specific protein degradation products suggests that the yeast nucleus itself is a proteolysis-inert site in the cells studied. The 26S proteasomes seem not to be located in the yeast nucleus (http://www.nature.com/nrm/poster/ubiquitin/poster.pdf), whereas 20S proteasomal subunit complexes have been observed to be imported into the yeast nucleus.[30] Our results here reveal that the 20S complexes imported (if it happened) were not matured or functional in the nucleus to warrant proteolysis.

Nuclear proteins of yeast that colocalized in other proteolysis-active organelles or locations were observed to be degraded in this study. For example, 55 of the 58 yeast nucleus-related proteins (as defined by yeastgenome.org, see Supporting Information Table 1) were colocalized within the highly active cytoplasm; the other three nucleus-related proteins found included two short-lived

(27) Piwko, W.; Jentsch, S. *Nat. Struct. Mol. Biol.* **2006**, *13*, 691–697.

(28) Löwe, L.; Stock, D.; Jap, B.; Zwickl, P.; Baumeister, W.; Huber, R. *Science* **1995**, *268*, 533–539.

(29) Niedermann, G.; King, G.; Butz, S.; Birsner, U.; Grimm, R.; Shabanowitz, J.; Hunt, D. F.; Eichmann, K. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 8572–8577.

(30) Lehmann, A.; Janek, K.; Braun, B.; Kloetzel, P.-K.; Enenkel, C. *J. Mol. Biol.* **2002**, *317*, 401–413.
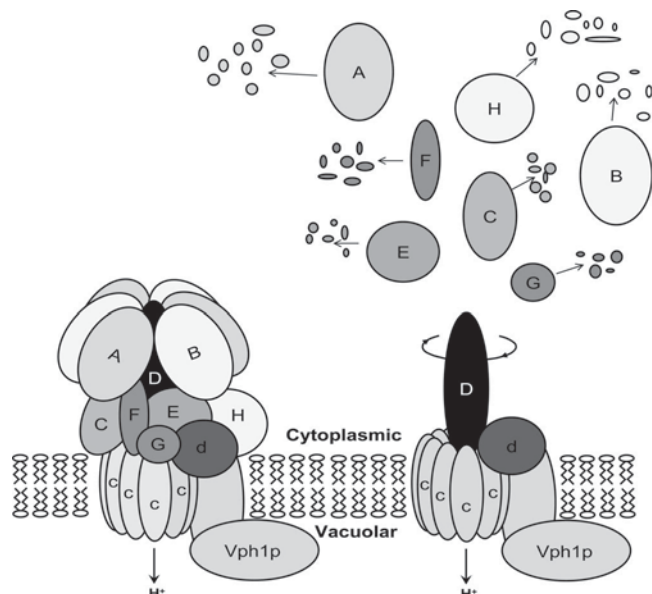
Figure 5. Observed degradation of the yeast V-ATPase subunits suggesting that the process happens after disassembly of subunits from subunit D. The V-ATPase structure diagram was redrawn based on ref 33.

nucleus pore proteins NSP1[31] (colocalized with mitochondrion) and NUP2[32] (colocalized with the endoplasmic reticulum and cytoskeleton) and a yeast protein transporter histone (colocalized within the vacuole). Proteolytic degradation of these nucleus-related yeast proteins are thus hypothesized to take place by the proteolysis-active enzymes found in the cytoplasm (containing 26S proteasomes) or other organelles.

Yeast V-ATPase Degradation. V-ATPase is composed of a membrane-associated V_0 domain and a peripheral membrane protein complex that makes up the V_1 domain.[33] Yeast V-ATPase V_0 domain contains six subunits (a, d, c, c′, c″, and e), and the V_1 domain is an assembly of eight subunits A–H. We detected intermediate peptides from the breakdown of all seven V_1 subunits including A, B, C, E, F, G, and H (H was identified only from its N-terminal peptides, see Supporting Information Table 1), only missing the stalk subunit D that connects seven other subunits to the V_0 domain[33] although the subunit D protein has a higher expression level (8500 molecules/cell) than some other subunit proteins (4000 and 2300 molecules/cell for F and G, respectively).[21] (The lifespan of subunit D is not available, while the other seven subunit proteins have half-lives of 10−300 min.[22]) Peptide products from the V_0 membrane domain subunit proteins were not found. From these observations, we deduce that degradation of the V_1 subunit proteins is a process where subunits A−C and E−H first uncouple (dissemble) from the D, become a free floating state in cytoplasm, and then are degraded there by cytoplasmic proteasomes (Figure 5). Subunit D remains connected to the V_0 domain, whereas the rest of the V_1 subunits disassemble, which protects subunit D from the degradation. It is unlikely that the degradation of the V-ATPase subunit proteins happened during their transport, localization, and assembly processes; otherwise,

all subunits including D should be equally degraded and detectable. The observed degradation of the V_1 subunit proteins also reveals that the assembly-disassembly of the V-ATPase may actually be irreversible and require new protein synthesis to make up a deficiency of the degraded subunit proteins which maintain normal function of the vacuolar proton pump. In other words, it appears as though the subunits are not simply reused from disassembled subunits to reform the complex and make it active again.[34]

Selectivity/Specificity for Proteolysis of Yeast Mitochondrial Proteins. Yeast mitochondria with their own proteolytic activities[35] are the second largest pool for protein degradation observed in this study. In total, we found 28 mitochondrion-specific proteins (Supporting Information Table 1) with the detection of their degradation-produced peptides, and Figure 6 shows the compartmental locations of these proteins. Of the 28 proteins, 17 were located in the mitochondrial matrix (MM), proteolytic degradation of which should be catalyzed by proteasome-like ClpP/ClpX and Yta12p.[35] The transport carrier and gate proteins in the mitochondrial inner membrane (MIM, containing proteolytic protease *Yta12p*[35]) were selectively degraded, including four transport-related proteins, two ADP/ATP carrier proteins, two CcO proteins, one acryl carrier protein, one TIM 23 subunit protein (TIM 23 was found with only N-terminal peptide ladders), and the tricalbin-3 (MIM-attached, facing the MM side). For the mitochondrial outer membrane (MOM) compartment, only transport gate proteins (OM45 and TOM70[36,37]) appeared to be degraded. Degradation of the mitochondrial intermembrane space (MIMS) proteins was not found, which agrees with the lack in observation of the proteolytic proteases in yeast MIMS.[35] The yeast mitochondrial proteases' preference is determined to be in the following order: Lys > Leu > Ser > Phe > Arg = Ala > Val > Gln > . . . at P1 and Gly > Arg > Ser > Lys > Glu > Thr = Ala > Asp > . . . at P1′ according to the peptides detected.

Proteolytic Degradation of Other Organellar Proteins Observed. In addition to the major proteolytic activities observed above, proteolytic degradation of a few cytoskeleton- and endoplasmic reticulum (ER)-specific proteins was also found. The three cytoskeleton-specific proteins found included actin, actin-capping protein (CAP1), and protein cofilin promoting actin filament remodeling and turnover,[38,39] all of which play key roles in endocytosis. The three ER-specific proteins found include the secretory-regulating protein BFR1, regulator protein GRP 78 (a member of the Hsp70 family) involved in incorrectly folded protein regulation,[40,41] and a sterol-acyltransferase 2.

(31) Fahrenkrog, B.; Aris, J. P.; Hurt, E. C.; Panté, N.; Aebi, U. *J. Struct. Biol.* **2000**, *129*, 295–305.

(32) Matsuura, Y.; Lange, A.; Harreman, M. T.; Corbett, A. H.; Stewart, M. *EMBO J.* **2003**, *22*, 5358–5369.

(33) Finbow, M. E.; Harrison, M. A. *Biochem. J.* **1997**, *324*, 697–712.

(34) Holliday, L. S.; Lu, M.; Lee, B. S.; Nelson, R. D.; Solivan, S.; Zhang, L.; Gluck, S. L. *J. Biol. Chem.* **2000**, *275*, 32331–32337.

(35) Bota, D. A.; Davies, K. J. A. *Mitochondrion* **2001**, *1*, 33–49.

(36) Ohba, M.; Schatz, G. *EMBO J.* **1987**, *6*, 2109–2115.

(37) Van der Laan, M.; Rissler, M.; Rehling, P. *FEMS Yeast Res.* **2006**, *6*, 849–861.

(38) Theriot, J. A. *J. Cell Biol.* **1997**, *136*, 1165–1168.

(39) Hubberstey, A. V.; Mottillo, E. P. *FASEB J.* **2002**, *16*, 487–499.

(40) Kassenbrock, C. K.; Garcia, P. D.; Walter, P.; Kelly, R. B. *Nature* **1988**, *333*, 90–93.

(41) Jackson, C. J.; Képès, F. *Genetics* **1994**, *137*, 423–437.

(42) Washburn, M. P.; Wolters, D.; Yates, J. R. *Nat. Biotechnol.* **2001**, *19*, 242–247.
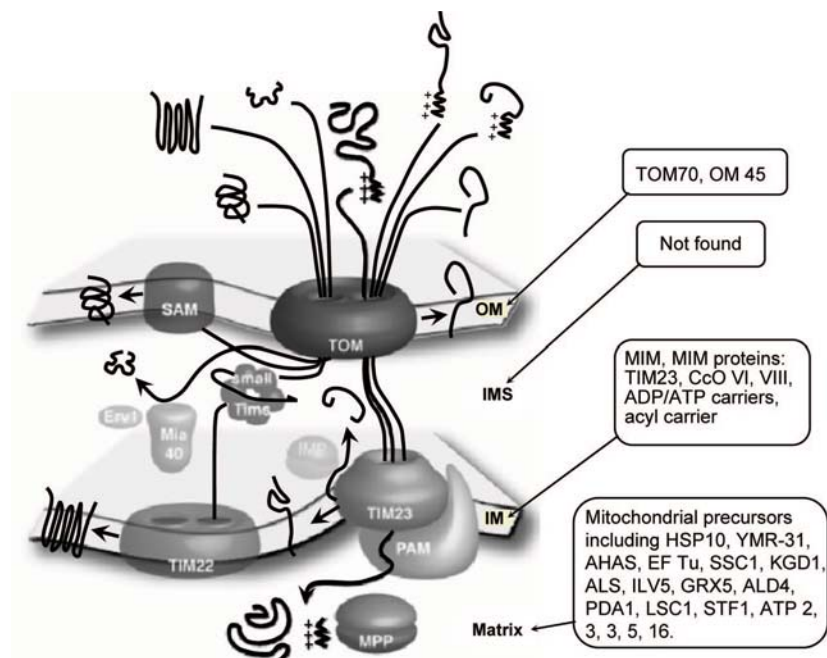
Figure 6. Distribution of the degraded yeast mitochondrial proteins. The translocation diagram was modified based on ref 35.

DISCUSSION

With the use of yeast *S. cerevisiae* as a model organism, we demonstrated that thousands of molecular species with masses of <10 000 u were detectable for cells at the time of harvest (shown in Figure 1). Identification of peptides from this mixture challenges the method[42] for identification accuracy without bias against the peptide's terminus and size. The accurate MS/MS sequencing-based unique sequence tags method used in this study, yielding extremely few random false positives, resolves this challenge. This sequencing-based method is desirable to probe the protein proteolysis processes and the activities and specificities of the related proteolytic proteases on a proteome-wide scale (as demonstrated in the Results section), which also covers the peptide scope of peptodomics[43] (study of intercellular peptides) and degradomics[44] (study of enzyme activity). With the use of the reliable identification method, we identified ~1100 peptides among the relatively small molecules observed from the cell lysate. These peptides have a range of sequence lengths from 6 to 100 amino acid residues and show up with specific cleavage sites and show up overall as an array of protein pieces that produce sequence ladders (see Figure 2).

Both in vivo protein proteolysis (i.e., intracellular degradation) and protein breakdown during/after the breaking of the cells and the releasing of proteins and their degradation products from their specific cellular locations (i.e., extracellular proteolysis or artifacts) can contribute peptides in the lysate of cells. In this study, we used pressure cycling technique (PCT[45]) for the cell lysis and peptide/protein extraction. There was no evidence that the mechanical processes applied to the cells (including the cell freeze−thaw and pressurized

lysis) attributed to any sort of protein cleavage and, furthermore, the cleavage selectivity pattern seen in this study (Figure 4b). (In fact, the high pressure has been shown to only alter protein higher structures, e.g., protein folding/unfolding and stability of protein aggregates.[46]) The extracellular proteolysis of proteins during sample processing are expected to favor detection of high-abundance proteins located in various locations, but the proteins assigned from the peptides detected in this study had little correlation with the proteins' abundances (Supporting Information Figure 2). For example, the highly abundant proteins located in the yeast nucleus, e.g., yeast histone H4 (GNL030W) with a documented abundance of >600 000 molecules/cell and a GRAVY value of −0.5, were not observed. The lack in detecting such high-abundance proteins is less possible from the incomplete disruption of cells by PCT, as PCT has been demonstrated as being able to release proteins that are difficult to be released by traditional bead beating, pulverization under liquid nitrogen, and sonication methods widely used for sample preparation in 2D gel.[45] Additionally, many highly abundant cytoplasmic proteins (e.g., >350 proteins having abundances of >10 000 molecules/cell), including the most abundant cytoplasmic protein CWP2 (YKL096W-a, ~1 600 000 molecules/cell), were not in our protein identification list (Supporting Information Table 1) where the abundance of proteins identified can be as low as ~450 molecules/cell. This evidence reveals that postlysis artifact proteolysis by the released proteases during the step-by-step PCT cell breaking and protein lysis contributed little to the peptides detected, and the proteolysis that occurred within the cells (or intracellular degradation of proteins) at the time of cell harvest should be responsible for the peptide observations.

(43) Svensson, M.; Sköld, K.; Nilsson, A.; Fälth, M.; Svenningsson, P.; Andrén, P. E. *Biochem. Soc. Trans.* **2007**, *35*, 588–593.

(44) López-Otín, C.; Overall, C. M. *Nat. Rev. Mol. Cell Biol.* **2002**, *3*, 509–519.

(45) Smejkal, G. B.; Witzmann, F. A.; Ringham, H.; Small, D.; Chase, S. F.; Behnke, J.; Ting, E. *Anal. Biochem.* **2007**, *363*, 309–311.

(46) Katrusiak, A., McMillan, P., Eds. *High-Pressure Crystallography. Proceedings of the NATO Advanced Research Workshop on High-Pressure Crystallography*; Erice, Italy, 2003; pp 543−556.

Through detection of intracellular peptides, we now can "see" the proteome-wide protein post-translational proteolytic degradation. It was observed that only proteins located in some special organelles and compartments contributed the peptides that were detected (see Figures 3, 5, and 6). The potential influence of use of a hydrophilic PCT solvent for extraction of peptides analyzed (see the Methods section) on such observations were investigated by examination of the hydrophobicity of peptides detected. The GRAVY values for the peptides detected were distributed over a range from −1.9 to 1.7 (Supporting Information Figure 3). Both soluble and insoluble (membrane) proteins can generate hydrophilic and moderately hydrophobic peptides located in this GRAVY range.[47] The peptides detected for selective organelles and compartments should be mainly derived from the well-controlled selective proteolysis processes, instead of the lysis solvent and less-selective lysosome-involved proteolyses. We plan to explore the use of hydrophobic solvents for extension of the method to study peptides containing hydrophobic sequences, e.g., membrane domain peptides and signal peptides.

One might predict that the yeast cytoplasm would be the most active location for proteolytic breakdown of proteins for cells that experienced a nutrient-rich, rapid growth environment at the time of harvest, as proteasomes are mainly distributed in the yeast cytoplasm. We did observe the proteasomal trypsin-like and chymotrypsin-like activities for the breakdown of yeast cytoplasmic proteins, but found less evidence for the caspase-like activity. From the terminal amino acids of the peptides detected, we observed the steric preference for proteolytic degradation of yeast cytoplasmic proteins. This steric preference could be related to a strict requirement for protein proteolysis to occur in proteasomes, but is less necessary for protein breakdown by lysosomal proteinases (see the Supporting Information). Further investigation (e.g., using an inhibitor to control specific degradation pathways) is needed to better understand the mechanism of protein proteasomal degradation. Our data also indicates that the cytoplasmic proteasomes could also be responsible for the degradation of the ATPase V_1 domain proteins after their disassembly and most of the multiple organelle colocalized proteins observed in this study (e.g., 83 of the 87 multiple organelle colocalized proteins assigned are associated as being localized to the cytoplasm).

Whether the yeast nucleus is proteolysis-active or not has not been clearly elucidated yet, and few experimental data have reported the proteolysis activities happening in the yeast nucleus. Proteasomes (>100 Å in size) cannot be directly imported into the yeast nucleus through the smaller nucleus pores (<90 Å for free diffusion) in the envelope, but the proteasome subunit complexes have been reported to be imported into yeast nucleus as precursor complexes.[30] Existence of proteasome precursor complexes in the yeast nucleus, however, cannot necessarily conclude the existance of proteolytic activity since after importation, these complexes would have to be assembled to form the matured, proteolysis-active proteasomes. Our results provide evidence to the contrary, at least for the growth condition we explored. Further studies involving the isolation of the nucleus could help to determine if there is indeed proteolytic activity occurring in the nucleus of yeast.

Ubiquitylation is a well-established pathway for protein proteasomal degradation, and almost all ribosomal and stress-response proteins identified in this study by the detection of the intermediate peptides were ubiquitylation proteins (http://ubi-prot.org.ru/). However, an overwhelming majority of cytoplasmic regulatory and enzyme proteins that produced the degradation intermediate peptides identified are not associated as being ubiquitylation proteins (see Table 1). The documented ubiquitylation protein list may be incomplete, but this is unlikely the reason for our lack in comprehensive observation of these proteins as the possibility for the measurements of protein ubiquitylation should be the same for ribosomal and stress-response proteins and regulatory and enzyme proteins with consideration that these different proteins from different functional groups have a similar range of lifetime span and expression levels. From our results of the proteome-wide measurements of intracellular peptides, we speculate that some nonubiquitylation−proteasome pathways[48,49] or even undefined proteolytic processes that need further study were also possibly functioning in the yeast cytoplasm which was responsible for the observation of degraded regulatory and enzyme proteins.

Mass spectrometry-based methods have been widely used for various proteomics applications. Information of protein post-translational degradation, however, has been generally understudied or totally ignored in most analyses attempting to elucidate protein expressions and functions in the cell biological processes involved. Since degradation events are ever present in cells and are as biologically important as protein creation, more emphasis needs to take place in the field of proteomics on the understanding of the simultaneous protein breakdown occurring. Developments in liquid phase separations, mass spectrometers, and informatics now allows for reliable assessment of proteome-wide protein degradation and can provide proteomics with additional information of protein degradation/turnover for more precise elucidation of protein expressions and cellular function. Finally, in combination with the control of some cellular process (e.g., with use of some inhibitors) the specific pathways of protein degradation and the related diseases[50,51] could be more directly, in a high-throughput fashion, studied with the method described here for better understanding of what is taking place within the cell than with the traditional approaches currently used in proteomics and protein degradation studies.

METHODS

Preparation of Soluble Yeast Lysate. *S. cerevisiae* (ATCC 90845) was grown in a batch shaker flask at 30 °C on YPD broth. Cells were harvested by centrifugation at 4000 rpm for 10 min at midlogarithmic phase. Cells were washed twice with 4 °C phosphate-buffered saline, pH 7.4 containing a protease inhibitor cocktail (complete, mini tablet) at the concentration recommended by the manufacturer (Roche Applied Science,

(47) Shen, Y.; Moore, R. J.; Zhao, R.; Blonder, J.; Auberry, D. L.; Masselon, C.; Pasa-Tolic, L.; Hixson, K. K.; Auberry, K. J.; Smith, R. D. *Anal. Chem.* **2003**, *75*, 3596–3605.

(48) Murakami, Y.; Matsufuji, S.; Kameji, T.; Hayashi, S.-i.; Igarashi, K.; Tamura, T.; Tanaka, K.; Ichihara, A. *Nature* **1992**, *360*, 597–599.
(49) Jariel-Encontre, I. *J. Biol. Chem.* **1996**, *270*, 11623–11627.
(50) Reinstein, E.; Ciechanover, A. *Ann. Intern. Med.* **2006**, *145*, 676–684.
(51) Meijer, A. J.; Codogno, P. *Mol. Aspects Med.* **2006**, *27*, 411–425.

Mannheim Germany). The cells were washed by adding 2× the volume of buffer described above to the cell pellet followed with a gentle vortex and a repeat of the centrifugation followed by a decantation of the supernatant. After the second the supernatant was decanted from the cell pellet and immediately the yeast cells were flash frozen with liquid N_2 and stored at -80 °C until ready for further processing. To a 200 μL volume sized cell pellet of yeast S. cerevisiae, 1.3 mL of 18.3 MΩ water containing a protease inhibitor cocktail (complete, mini tablet) at the concentration recommended by the manufacturer (Roche Applied Science) was added and the suspension was vortexed. Lysis was achieved using pressure cycling technology capable with the Barocycler (Pressure BioSciences, West Bridgewater, MA) for 10 cycles going between ambient pressure for 20 s and 35 000 psi for 20 s at 4 °C. The lysate was ultracentrifuged at 100 000 rpm (355 040 rcf) with a Beckman (Fullerton, CA) Optima TL ultracentrifuge at 4 °C for 20 min. The supernatant was collected and placed immediately on ice. A BCA protein assay (Pierce, Rockford IL) was performed to determine the protein concentration using a bovine serum albumin standard. The sample was flash frozen with liquid N_2.

High-Resolution LC−High-Precision MS/MS Experiments. The prepared lysate was analyzed using high-resolution LC−high-precision MS/MS as follows. The high-resolution LC was performed on an LC instrument as previously reported.[52] A 120 cm × 100 μm i.d. fused-silica capillary column containing 3 μm (120 Å size) C4−silica particles (Phenomenex, Terrence, CA) was used as the LC column. The sample (1 μg/μL protein content) was loaded into a 10 μL loop, switched onto the LC column, and separated with a gradient from mobile phase A acetonitrile/H_2O/acetic acid (25:75:0.2, v/v/v) to mobile phase B acetonitrile/H_2O/acetic acid/trifluroacetic acid (90:10:0.2: 0.1, v/v/v/v) at room temperature. (The use of 25% ACN in mobile phase A was to prevent precipitation of proteins existing in the lysate, and the use of TFA in the mobile phase B was to improve elution of various sizes and properties of peptides/ proteins.) The separation was completed in 600 min at an operation pressure of 15K psi; the reproducibility of such a long high-resolution LC coupled with the mass spectrometers has been systematically evaluated,[52] and no repeated runs were made herein. The high-precision MS/MS was performed on an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific., San Jose, CA). FT-MS and FT-MS/MS data were collected with AGC targets of 1×10^6 and 2×10^5, respectively. Spectra were acquired at a 30K resolution, using a survey scan with $400 \leq m/z \leq 2000$ followed by FT-MS/MS of the five most intense ions from the survey scan (monoisotopic precursor selection was not enabled). FT-MS/MS employed an isolation window of 3 m/z units and 35% normalized collision energy. Dynamic exclusion was enabled with no repeat counts using a mass window of ±1.5 m/z units and a duration cycle of 25 s. Mass calibration was performed according to the method provided by the instrument manufacturer.

Analysis of High-Precision MS/MS Data. The previously described method[53] was used to process the high-precision MS/MS data. Briefly, the FT-MS/MS data were initially processed using SEQUEST (Sequest 27_master, rev 12, Thermo Fisher Scientific) with molecular mass tolerances of ±5 u and ±210 u in two database searches, respectively. The amino acid sequences available from the high-precision MS/MS spectra were recalculated for the SEQUEST-recommended top 10 peptide candidates using ICR2LS software developed in-house for analysis of mass spectrometry data (http://ncrr.pnl.gov/ software/ICR2LS.stm; the ICR2LS parameter settings used in this study are detailed in the Supporting Information). The theoretical isotopic envelopes for the fragments of each peptide candidate and its precursor were generated by use of ICR2LS. For the peptide candidate fragments, we matched the b and y fragments predicted from the candidates against the deisotoped monoisotopic masses (i.e., deconvolved isotopic envelopes) of the spectrum using ICR2LS. If the match was achieved with a mass error tolerance of 10 ppm, the b and y fragments were accepted as assigned. The consecutive b or y fragments were used to construct the amino acid sequence(s). During construction of amino acid sequence(s), we required that each sequence must contain at least three fragments that had the isotopic envelopes (≥2 isotopic peaks) matched with theoretical ones within the 10 ppm mass tolerance. The constructed sequences from all MS/MS spectra were then searched against the yeast protein database to examine the sequence uniqueness with consideration of influence from various possible modifications,[53] and only the unambiguous unique sequences (referred as unique sequence tags, UStags) were used for identification of the peptides. For each UStag-identified peptide, we matched its theoretical isotopic envelope to that of the corresponding precursor that we required to have at least three most abundant isotopic peaks for matching. If the match was within a mass error tolerance of 10 ppm, the peptide was considered to be identified without modification. If the mass error was larger than 10 ppm, we assumed existence of some modifications. According to the mass difference between the peptide identified and its precursor, we examined the spectra (including with manual inspection by use of XCalibur 2.0.5 and ICR2LS) to determinate the modifications. At this step, we simultaneously corrected the charge states wrongly assigned by SEQUEST (Supporting Information Table 3). The mass errors for various types of matching, including molecular mass and averaged values for fragmental mass and amino acid residues, are given in the identification list (Supporting Information Table 1). The data set search results for the identified peptides are given in Supporting Information Table 3 for additional information, and 20 spectra randomly selected for various charge states were given in Supporting Information Table 4.

In this study, the hydrophobity GRAVY values were calculated according to the method reported in ref 54.

(52) Shen, Y.; Zhang, R.; Moore, R. J.; Kim, J.; Metz, T. O.; Hixson, K. K.; Zhao, R.; Livesay, E. A.; Udseth, H. R.; Smith, R. D. Anal. Chem. **2005**, 77, 3090–3100.

(53) Shen, Y.; Tolic, N.; Hixson, K. K.; Purvine, S. O.; Pasa-Tolic, L.; Qian, W.-J.; Adkins, J. N.; Moore, R. J.; Smith, R. D. Anal. Chem. **2008**, 80, 1871–1882.

(54) Kyte, J.; Doolittle, R. F. J. Mol. Biol. **1982**, 157, 105–132.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.