Mass Spectrometric Characterization of the Affinity-Purified Human 26S Proteasome Complex†

Xiaorong Wang,‡ Chi-Fen Chen,§ Peter R. Baker,|| Phang-lang Chen,§ Peter Kaiser,§ and Lan Huang*,‡

Departments of Physiology and Biophysics and of Developmental and Cell Biology, University of California, Irvine, California 92697, Department of Biological Chemistry, University of California, Irvine, California 92697, and Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94143

Received September 26, 2006; Revised Manuscript Received December 18, 2006

ABSTRACT: The 26S proteasome is a multisubunit complex responsible for degradation of ubiquitinated substrates, which plays a critical role in regulating various biological processes. To fully understand the function and regulation of the proteasome complex, an important step is to elucidate its subunit composition and posttranslational modifications. Toward this goal, a new affinity purification strategy has been developed using a derivative of the HB tag for rapid isolation of the human 26S proteasome complex for subsequent proteomic analysis. The purification of the complex is achieved from stable 293 cell lines expressing a HB-tagged proteasome subunit and by high-affinity streptavidin binding with TEV cleavage elution. The complete composition of the 26S proteasome complex, including recently assigned new subunits, is identified by LC−MS/MS. In addition, all known proteasome activator proteins and components involved in the ubiquitin−proteasome degradation pathway are identified. Aside from the subunit composition, the N-terminal modification and phosphorylation of the proteasome subunits have been characterized. Twelve novel phosphorylation sites from eight subunits have been identified, and N-terminal modifications are determined for 25 subunits, 12 of which have not been previously reported in mammals. We also observe different N-terminal processing of subunit Rpn2, which results in identification of two different N-termini of the protein. This work presents the first comprehensive characterization of the human 26S proteasome complex by affinity purification and tandem mass spectrometry. The detailed proteomic profiling obtained here is significant to future studies aiming at a complete understanding of the structure−function relationship of the human 26S proteasome complex.

The 26S proteasome is the macromolecular machine of the ubiquitin proteasome-dependent degradation pathway that is responsible for most of the nonlysosomal protein degradation in both the nucleus and cytosol (*1*, *2*). It is involved in many important biological processes such as cell cycle progression, apoptosis, and DNA repair. With a molecular mass of approximately 2.5 MDa, the 26S proteasome consists of the 20S core particle (CP) capped by the 19S regulatory complex at each end (RP, also called CAP or PA700). The eukaryotic 20S core particle is composed of at least 14 subunits with molecular masses of 21−34 kDa and varied charges (pI 3−10). On the basis of sequence similarities, these subunits can be divided into α and *β* subfamilies, which assemble into four stacked heptameric rings in α*ββ*α order (*3*, *4*). Throughout evolution, the cylindrical structure of the 20S proteasome has been conserved, whereas its subunit composition has changed. In the eukaryotic 20S proteasome, the α subunits are catalytically inactive but are responsible for the assembly of the complex and its interactions with the regulatory complex, therefore controlling the access of the substrates to the catalytic chamber for degradation. Three of the seven *β* subunits, *β*1(Y), *β*2(Z), and *β*5(X), are catalytically active and responsible for various proteolytic activities (*1*, *4*). Compared to the 20S core, which degrades small peptides and fully unfolded proteins in an ATP-independent manner, protein degradation by the 26S proteasome generally requires not only ATP but also the presence of a polyubiquitin chain conjugated to the substrate protein (*5*).

The 19S regulatory complex is composed of at least 18 different subunits, which are assembled into two main subcomplexes: a base that contains six ATPases and two non-ATPase subunits and abuts the proteasome α ring and a lid subcomplex containing at least 10 non-ATPase subunits that sit on top of the base (*1*, *6*). In contrast to the 20S core, the structure and function of the 19S regulatory complex are less understood. Various biochemical and genetic studies reveal that the regulatory complex carries out a number of biochemical functions, including recognition of polyubiq-

† This work was supported by National Institutes of Health Grants GM-74830 (to L.H.) and GM-66164 (to P.K.), the U.S. Army (Grant PC-041126 to L.H.), the California Breast Cancer Research Program (Grant 11NB-0177 to P.K.), and a UCI cancer center postdoctoral fellowship to X.W.

* To whom correspondence should be addressed: Medical Science I, D233, Department of Physiology and Biophysics and Department of Developmental and Cell Biology, University of California, Irvine, CA 92697-4560. Phone: (949) 824-8548. Fax: (949) 824-8540. E-mail: lanhuang@uci.edu.

‡ Departments of Physiology and Biophysics and of Developmental and Cell Biology, University of California, Irvine.

§ Department of Biological Chemistry, University of California, Irvine.

|| University of California, San Francisco.

10.1021/bi061994u CCC: $37.00
Published on Web 02/27/2007

uitinated substrates (*7*, *8*), cleavage of polyubiquitin chains to recycle ubiquitin (*9*), unfolding of substrates, assisting in opening the gate of the 20S chamber, and subsequently translocating unfolded substrates into the catalytic chamber (*2*). The activities of the 19S regulatory complex and its assembly with the 20S proteasome have been shown to be strictly ATP-dependent (*2*, *5*).

In addition to binding to the 19S regulatory complex, the 20S core can be capped by two other evolutionarily conserved activator protein complexes, PA28 (*10*) and PA200 (*11*), which can activate the proteolytic activity of the 20S proteasome against model peptide substrates. Different from the 26S proteasome, the resulting activated proteasome complexes do not recognize ubiquitinated protein substrates, and their activities and assembly are ATP-independent. There are three PA28 homologues, α, *β*, and *γ*. The PA28α and -*β* subunits form a heteroheptameric complex to activate the 20S proteasome and are mainly cytoplasmic. This complex is highly abundant in immune tissues and can be induced by interferon-*γ*, suggesting that PA28α*β* is involved in the production of class I antigens for presentation (*12*). In comparison, the PA28*γ* subunit forms a homoheptameric activator complex and is localized in the nucleus. Although the function of PA28*γ* is not completely clear, genetic studies with mice and flies indicate that PA28*γ* may play a role in the cell cycle and apoptosis (*13*). Recent studies have suggested that either PA28 activators or PA200 can bind to one end of the 20S proteasome with the other end capped by a 19S complex to form a hybrid proteasome complex, possibly for immune response as well as in general intracellular proteolysis (*6*, *14*).

Intracellular degradation of ubiquitinated substrates by the ubiquitin−26S proteasome pathway is tightly regulated so normal cell growth and survival can be maintained. Therefore, aberrations in this pathway can contribute to human diseases, including cancer (*15*). Due to the central role of the 26S proteasome, it has become an attractive therapeutic target for cancer treatment. Velcade (bortezomib, PS-341), a potent and selective proteasome inhibitor, has been approved by the Federal Drug Administration and used for the treatment of relapsed and refractory multiple myeloma. However, the mechanism by which proteasome inhibitors such as Velcade function in cancer treatment remains elusive since blockage of the ubiquitin−proteasome-dependent degradation affects many regulatory pathways in the cell. In addition, previous studies have implicated that posttranslational modifications of the subunits may play a role in regulating the proteasome assembly and functions (*16*−*20*), but the details of these modifications are unknown. To fully understand the function and regulation of the human 26S proteasome complex, it is essential to first investigate its composition and posttranslational modifications. Although proteomic analyses of proteasome complexes from different species have previously been reported, the majority of them were focused on the 20S proteasome complex (*21*−*26*), and only a few extended the analyses to the 26S proteasome complex (*5*, *27*−*29*). Many of these studies have been mainly performed on yeast proteasomes and only recently on proteasomes purified from the murine heart. However, no systematic proteomic analysis of the human 26S proteasome complex has been carried out to date.

Here we present the development of a new affinity purification strategy for rapid and effective isolation of the human 26S proteasome. Liquid chromatography−tandem mass spectrometry (LC−MS/MS)[1] is employed in elucidating the proteomic profile of the purified human 26S proteasome to further the understanding of this degradation machinery.

EXPERIMENTAL PROCEDURES

Chemicals and Reagents. ImmunoPure streptavidin, HRP-conjugated antibody, and Super Signal West Pico chemiluminescent substrate were from Pierce Biotechnology (Rockford, IL). Sequencing grade trypsin was purchased from Promega Corp. (Madison, WI), which has been modified by methylation, rendering it resistant to proteolytic digestion. Endoproteinase Lys-C was from WAKO chemicals (Osaka, Japan). Chymotrypsin was obtained from Roche Diagnostics (Palo Alto, CA). All other general chemicals for buffers and culture media were purchased from Fisher Scientific and/or VWR International.

Plasmid and Cloning. The HTBH tag was constructed by PCR using the HBTH and HBH tags as templates along with designed primers. Those for the HT fragment were as follows: forward, 5′ CCTTAATTAACCATCATCACCAC-CATCATGACTACGATATACCCACAACC 3′; and reverse, 5′ ACCGGCCTTTCCAGCTGAACCCCTCATCCC-TAACTC 3′. Those for the TH fragment were as follows: forward, 5′ GGGATGAGGGGTTCAGCTGGAAAGGCC-GGTGAAGGT 3′; and reverse, 5′ GGAATTCTCATT-AATGATGGTGGTGATGATGAACGCCGATCTTGATTAGAC 3′. HTBH was constructed by PCR using the forward primer of HT and the reverse primer of TH. hRpn11 was PCR amplified using an ATCC clone (ATCC 7031739) as the template with the following primers: forward, 5′ ATGCG-GCCGCATGGACAGACTTCTTAGACTTGGAGGA 3′; and reverse, 5′ CCTTAATTAATTTAAATACGACAGTATC-CAACATAGC 3′. A hRpn11 DNA fragment containing a NotI site at the 5′ end and a PacI site at the 3′ end and a HTBH DNA fragment containing a PacI site at the 5′ end and an EcoRI site at the 3′ end were cloned into pQCXIP (Clontech, Mountain View, CA) by 3 fragment ligation. The schematic representation of hRpn11-HTBH is shown in Figure 1A, whereas the detailed plasmid maps of pQCXIP-hRpn11-HTBH and pQCXIP-HTBH are illustrated in panels B and C of Figure 1, respectively.

Generation of Retrovirus for Protein Expression and 293 Stable Cell Lines. The procedure for making the retrovirus and the 293 stable cell lines containing hRpn11-HTBH and the HTBH tag alone was similar to that reported previously (*30*). Briefly, a 293 GP2 cell line was cotransfected with pQCXIP-hRpn11-HTBH and a plasmid expressing VSV-G. The retrovirus was produced and released to the medium between 36 and 96 h after transfection. The medium containing the retrovirus was used to transduce 293 cells, which were subsequently selected with puromycin to estab-

[1] Abbreviations: HB tag, histidine and biotin tag; LC−MS/MS, liquid chromatography−tandem mass spectrometry; HRP, horseradish peroxidase; SCX, strong cation exchange; TAP, tandem affinity purification; TEV, tobacco etch virus; TEB, TEV elution buffer; TOFMS, time-of-flight mass spectrometry; UBA, ubiquitin-associated domain; UBL, ubiquitin-like domain; PIP, proteasome interacting partner; VSV-G, vesicular stomatitis virus envelope glycoprotein; AA, amino acids.

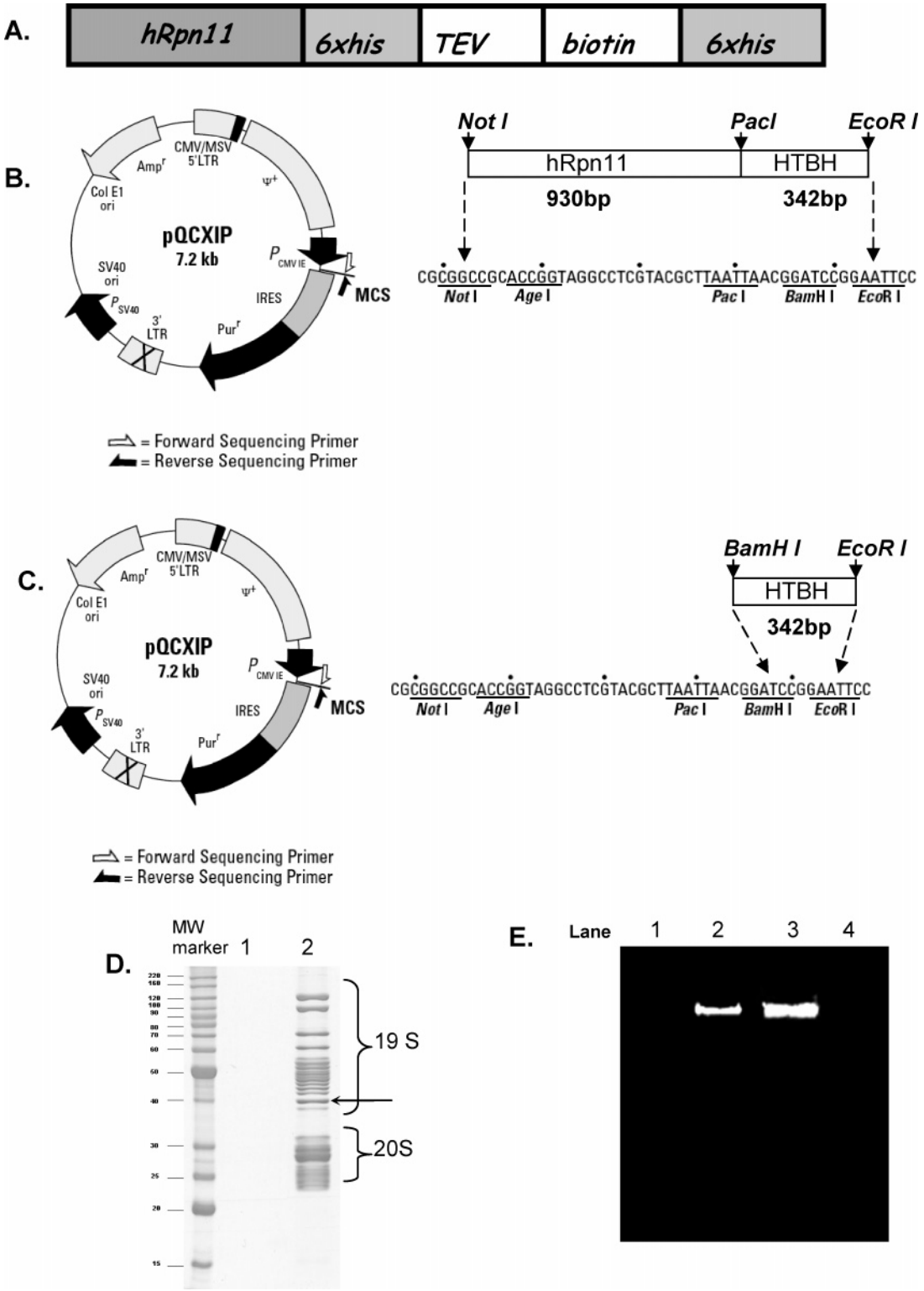

FIGURE 1: (A) Schematic representation of hRpn11 fused to the HTBH tag at its C-terminus. The HTBH tag consists of two hexahistidine tags, a TEV cleavage site, and a bacterially derived peptide that induces biotinylation in vivo. (B) The hRpn11 cDNA fused to the HTBH tag was cloned into pQCXIP as indicated. (C) The HTBH tag was cloned into pQCXIP as a negative control. (D) 1D SDS−PAGE patterns of the purified samples from stable cell lines expressing the HTBH tag (control, lane 1) or hRpn11-HTBH (lane 2). The arrow indicates the band containing the tagged hRpn11 as identified by in-gel digestion and LC−MS/MS. (E) Detection of the proteolytic activity of the proteasome by a gel overlay assay using the fluorogenic peptide substrate SUC-LLVY-AMC: lane 1, purified proteins from stable 293 cells expressing the HTBH tag (negative control); lane 2, cell lysate from wild-type 293 cells (positive control); lane 3, purified proteasome complex from stable 293 cells expressing hRpn11-HTBH in the presence of ATP; and lane 4, same as lane 3 but without ATP.

lish the stable cell lines expressing hRpn11-HTBH. Similarly, the stable cell lines expressing the HTBH tag were also generated.

Tandem Affinity Purification of the Human 26S Proteasome. Stable 293 cell lines expressing either hRpn11-HTBH or the HTBH tag were grown to confluence in DMEM

containing 10% FBS and 1% Pen/strep, then trypsinized, and washed three times with PBS buffer. The cell pellets were collected and lysed in buffer A [100 mM sodium chloride, 50 mM sodium phosphate, 10% glycerol, 5 mM ATP, 1 mM DTT, 5 mM $MgCl_2$, 1× protease inhibiter (Roche), 1× phosphatase inhibitor, and 0.5% NP-40 (pH 7.5)]. The lysates were centrifuged at 13 000 rpm for 15 min to remove cell debris, and the supernatant was incubated with streptavidin resin overnight at 4 °C. The streptavidin beads were then washed with 20 bed volumes of the lysis buffer, followed by a final wash with 10 bed volumes of TEB buffer [50 mM Tris-HCl (pH 7.5)] containing 10% glycerol. Beads were incubated in 2 bed volumes of TEB buffer with 1% TEV at 30 °C for 1 h. The 26S proteasome complex was eluted from the beads with TEB and stored in 10% glycerol at −80 °C. For the phosphorylation study, cells were treated with 50 nM calyculin for 30 min before the cells were harvested (*16*).

Assay of Proteasome Proteolytic Activity. Gel overlay assays were carried out to monitor the proteolytic activity of the proteasome using the fluorogenic peptide substrate SUC-LLVY-AMC as described previously (*31*).

Gel Electrophoresis and Immunoblotting. Cell lysates and flow through, wash, and elution fractions were separated by one-dimensional (1D) SDS−PAGE. Proteins were transferred to a PVDF membrane and analyzed by immunoblotting. Biotinylated proteins were detected using a streptavidin−HRP conjugate (1:5000).

Mass Spectrometric Analysis via Liquid Chromatography and Tandem Mass Spectrometry (LC−MS/MS). The purification of the human 26S proteasome complex was evaluated by 1D SDS−PAGE (12.5%), and the protein bands were visualized by Coomassie Brilliant Blue staining. The procedure for in-gel digestion was as described previously (*26*). For shotgun analysis, the 26S proteasome was precipitated with 25% TCA and the pellet was resolubilized in 25 μL of 50 mM NH_4HCO_3 with 8 M urea. Proteins were then reduced with 2 mM TCEP at 37 °C for 15 min, alkylated with 25 mM iodoacetamide in the dark at room temperature for 30 min, and digested with endopeptidase Lys-C at 37 °C for 4 h. For some experiments, cysteine was not reduced and alkylated prior to enzymatic digestion for the sake of simplicity. Finally, the urea concentration was decreased to 1.5 M for subsequent trypsin digestion at 37 °C overnight. The two-dimensional (2D) LC−MS/MS analysis of the digested mixture was similar to that described previously using a quadrupole orthogonal time-of-flight tandem mass spectrometer (QSTAR XL, Applied Biosystems/MDS Sciex) (*28*). Briefly, the digests were first separated by strong cation exchange (SCX) chromatography. Solvent A (5 mM KH_2PO_4 and 30% acetonitrile, pH 3 adjusted with formic acid) and solvent B (solvent A with 350 mM KCl) were used to develop a salt gradient. The digests were separated using a 2.1 mm × 10 cm polysulfoethyl A column (Nest group) at a flow rate of 200 μL/min. Peptide elution was monitored by UV detection at 215 and 280 nm. A typical separation employed 0% B for 10 min to allow for sample loading and removal of nonpeptide species, followed by a gradient from 0 to 100% B over 30 min. Fractions were manually collected on the basis of UV absorbance. All of the SCX fractions were desalted off-line using Vivapure C18 micro spin columns (Vivascience) prior to LC−MS/MS. LC−MS/MS was carried out by nanoflow reverse phase liquid chromatography (RPLC) (Ultimate LC packings, Dionex) coupled on-line to QSTAR XL MS. RPLC was performed using a capillary column [75 μm (inside diameter) × 150 mm (length)] packed with Polaris C18-A resin (Varian Inc.), and the peptides were eluted using a linear gradient from 0 to 35% B over 80 min at a flow rate of 250 nL/min (solvent A, 98% H_2O/2% acetonitrile/0.1% formic acid; solvent B, 98% acetonitrile/2% H_2O/0.1% formic acid). The MS acquisition was performed in an information-dependent mode in which each full MS scan was followed by three MS/MS scans where the three most abundant peptide molecular ions were dynamically selected for collision-induced dissociation (CID) to generate tandem mass spectra. To increase the number of MS/MS spectra acquired from the sample and improve the dynamic range of mass spectrometric analysis, two or three LC−MS/MS runs were performed on the same sample with exclusion of the ions being sequenced from the previous LC−MS/MS runs. To obtain complementary information, the purified proteasome complex was also digested with chymotrypsin in 50 mM NH_4HCO_3 at 37 °C overnight.

Database Searching for Protein Identification and Characterization. Monoisotopic masses (*m*/*z*) of both parent ions and corresponding fragment ions, parent ion charge states (*z*), and ion intensities from the tandem mass spectra (MS/MS) were obtained by using an automated version of the Mascot script from Analyst QS (version 1.1 QS, MDS Sciex/Applied Biosystems) within the development version of Protein Prospector (version 4.25.0, University of California, San Francisco, CA). The mass accuracy for parent ions and fragment ions was set as ±100 and ±200 ppm, respectively. Chemical modifications such as protein N-terminal acetylation, methionine oxidation, N-terminal pyroglutamine, carbamylation of the N-terminus, and lysine were selected as variable modifications during the search using Batch-tag in Protein Prospector. Carbamylation was introduced during the purification through the use of urea-containing buffers. When cysteine was reduced and alkylated, carbamidomethylation was chosen as a fixed modification. Otherwise, no cysteine modification was selected. Deamidation of asparagine was also selected due to its occurrence during sample preparation. When phosphorylation sites were mapped, phosphorylations of serine, threonine, and tyrosine were chosen as variable modifications. Both the Uniprot and NCBInr public databases were queried to identify the purified proteins since each database contains unique protein entries. The Search Compare program in Protein Prospector was used for summarization, validation, and comparison of results. Search Compare gives a peptide score (*32*) and an expectation value (*33*) for each peptide. To validate the expectation values, the data were searched against a reverse decoy database (*34*) to obtain a plot of false positive rate against expectation value threshold (see Figure 1 of the Supporting Information). A false positive rate of <0.1% was chosen, which corresponds to an expectation value cutoff of 10^{-5}. All reported modified peptides were confirmed by manual inspection of the MS/MS spectra.

RESULTS

Affinity Purification of the Human 26S Proteasome Complex. To purify the human 26S proteasome complex, the non-ATPase subunit hRpn11 was tagged at its C-terminus

with HTBH (Figure 1A), a derivative of the HB tag (*28*, *35*). Rpn11 was chosen because it is a successful affinity bait for purifying yeast proteasome without affecting proteasome function (*5*, *27*, *28*). The HTBH tag consists of two hexahistidine tags, a TEV cleavage site, and a signal sequence for in vivo biotinylation, which potentially allows for a three-step purification involving binding to Ni^{2+} chelate resins (step 1), elution and binding to streptavidin resins (step 2), and specific elution by cleavage with TEV protease (step 3). On the basis of experiments in yeast, initial purification of proteins tagged with the HB tags by Ni^{2+} chelate chromatography was considered important for removal of endogenous biotinylated proteins as they efficiently bind to streptavidin resins (*28*, *35*). However, endogenous biotinylated proteins are less abundant in mammalian cells than in yeast, and elution from streptavidin by TEV cleavage was sufficiently specific to prevent copurification of these proteins. Therefore, the first step of Ni^{2+} chelate chromatography was omitted, and a single-step purification strategy was employed in the direct binding to streptavidin resins followed by TEV-based elution under native conditions.

The purified samples from stable 293 cell lines expressing either hRpn11-HTBH or the HTBH tag alone (control) were separated by 1D SDS−PAGE (Figure 1D). The single-step purification of the complex (lane 2) using tagged hRpn11 appears to be highly specific. The overall gel pattern of the purified 26S proteasome complex was very similar to that previously reported (*36*). Since ATP stabilizes interactions between the 19S and 20S proteasome, the purification in the absence and presence of ATP was performed and compared. As expected, the full composition of the 26S proteasome complex was captured only in the presence of ATP, whereas in the absence of ATP, only the 19S proteasome was efficiently isolated while the 20S proteasome complex was partially copurified (data not shown).

Transient transfection often causes overexpression of a selected gene, leading to nonphysiological results. To minimize overexpression of hRpn11-HTBH, a retrovirus containing hRpn11-HTBH was produced and transduced to 293 cells to obtain a stable cell line. Previous studies have shown that most of the selected clones derived from this type of retrovirus transduction integrated only one copy of the DNA (*30*). From Figure 1D, hRpn11-HTBH appears to have an abundance similar to those of other proteasome subunits, implying that little free hRpn11-HTBH is present.

Activity of the Purified 26S Proteasome Complex. To determine whether the HTBH tag on hRpn11 affects the proteolytic function of purified proteasomes, purified complexes were subjected to native gel electrophoresis, followed by a gel overlay assay using a fluorogenic peptide substrate to measure the chymotryptic activity of the proteasome (*37*) (Figure 1E). The 26S proteasome complex purified in the presence of ATP appears to have strong peptidase activity and migrates at the same position (lane 3) as untagged wild-type proteasomes (lane 2), indicating the tag does not significantly interfere with the proteasome function. We cannot exclude, however, the possibility that HTBH-tagged proteasomes may have subtle functional difference as compared to untagged proteasomes, though the tag has no apparent effect on function and stable cell lines expressing the tagged proteasome subunit grow normally. Note that the complex purified in the absence of ATP (lane 4) did not exhibit any peptidase activity along with the negative control (lane 1), consistent with the requirement of ATP for the purification of the fully functional proteasome complex (*2*, *5*, *27*).

Table 1: Summary of Proteasome Subunits and Other Components Identified from the Purified Complex by LC−MS/MS

	subunit name	gene name	accession number	no. of unique peptides	sequence coverage (%)
19S	Rpt1	PSMC2	P35998	55	80.6
	Rpt2	PSMC1	P62191	59	71.4
	Rpt3	PSMC4	P43686	58	81.8
	Rpt4	PSMC6	P62333	41	58.6
	Rpt5	PSMC3	P17980	70	73.1
	Rpt6	PSMC5	P62195	60	73.2
	Rpn1	PSMD2	Q13200	139	77.8
	Rpn2	PSMD1	Q99460	101	71.1
	Rpn3	PSMD3	O43242	74	79.2
	Rpn5	PSMD12	O00232	56	59.6
	Rpn6	PSMD11	O00231	65	72
	Rpn7	PSMD6	Q15008	55	76.1
	Rpn8	PSMD7	P51665	40	70.1
	Rpn9	PSMD13	Q9UNM6	53	75.3
	Rpn10	PSMD4	P55036	31	32.4
	Rpn11	PSMD14	O00487	36	64.8
	Rpn12	PSMD8	P48556	24	63.0
	Gankyrin	PSMD10	O75832	7	33.2
	Rpn13/ADRM1	ADRM1	Q16186	13	21.4
	Rpn14/ PAAF1	WDR71	Q9BRP4	2	5.6
	Rpn15/DSS1	SHFM1	P60896	3	48.6
20S	α1	PSMA6	P60900	16	59.8
	α2	PSMA2	P25787	20	54.5
	α3	PSMA4	P25789	24	61.3
	α4	PSMA7	O14818	26	67.3
	α5	PSMA5	P28066	27	46.1
	α6	PSMA1	P25786	43	74.9
	α7	PSMA3	P25788	22	55.9
	β1	PSMB6	P28072	14	60.7
	β2	PSMB7	Q99436	24	61
	β3	PSMB3	P49720	33	62.4
	β4	PSMB2	P49721	31	69.7
	β5	PSMB5	P28074	21	62.5
	β6	PSMB1	P20618	29	66.4
	β7	PSMB4	P28070	21	47.0
other	PA28α	PSME1	Q06323	4	18.9
	PA28 γ	PSME3	P61289	8	31.9
	PA28β	PSME2	Q9UL46	7	41.2
	PA200i isofrom	PSME4	Q14997	38	26.6
	PA200i	PSME4	62467430	24	27.4
	UCH37	UCHL5	Q9Y5K5	30	66.9
	ubiquitin ligase E3C	UBE3C	Q15386	11	15.4
	ubiquitin ligase E3A	UBE3A	Q05086	23	23.5
	ubiquitin	UBC	P62988	9	61.8
	hHR23B	RAD23B	P54727	7	17.6
	hPLIC-1	UBQLN1	Q9UMX0	4	12.4

Protein Identification of the Purified Human 26S Proteasome Complex. A comprehensive mass spectrometric analysis of purified protein complexes was carried out in an attempt to characterize subunit composition and posttranslational modifications. A shotgun proteomics approach (*38*) was used to directly analyze in-solution protein tryptic digests by 2D LC−MS/MS. The results are summarized in Table 1. The complete subunit composition of the 26S proteasome complex is captured and identified, which includes six ATPase and 15 non-ATPase subunits of the 19S regulatory complex and seven α subunits and seven β subunits of the 20S core. Among the identified non-ATPase subunits, three are newly assigned: Rpn13/ADRM1, Rpn14/PAAF1, and Rpn15/DSS1 (*36*, *39*−*43*). All of the proteasome subunits were identified from analysis of tryptic digests except for the smallest

subunit, Rpn15/DSS1, which was identified from chymotryptic digests. This result indicates that alternative enzymatic digestion can be very useful for the identification of complete components of a given protein complex by shotgun proteomics. The results were reproducible and corresponded well with the proteins identified by a 1D SDS−PAGE/in-gel digestion/mass spectrometry approach (data not shown). Detailed peptide reports for protein identification are given in Table 1 of the Supporting Information.

In addition to components of the 26S proteasome complex, subunits of the 11S activator complex (PA28α, -β, and -γ) and recently identified nuclear proteasome activator, PA200 (*10*, *11*), were also identified. This result is interesting since these alternative activators were found in proteasome complexes that were purified on the basis of the affinity tag fused to a 19S component, further supporting the previous notion that hybrid proteasomes containing one 19S and one of the alternative activators exist in vivo (*6*, *14*). In addition, we found two homologous PA200 sequences, one in UniProt (accession number Q14997, 1798 amino acids) and the other in NCBInr (accession number 62467430, 1210 amino acids), between which 1160 amino acids are identical. However, Q14997 has 633 additional amino acids at the C-terminus, whereas 62467430 has a 40-amino acid extension at its N-terminus. Mass spectrometric analysis identified multiple peptides in both the N-terminal and C-terminal extension regions of these two sequences. As a result, the correct sequence of PA200 is the combination of the two sequences with a total of 1843 amino acids as reported by Ustrell et al. (*11*). Surprisingly, this correct sequence is not found in any public databases. Coincidentally, the mouse PA200 (accession number 117956381) also contains 1843 amino acids and has a sequence 95% identical to that of human PA200.

Several proteins involved in the ubiquitin−proteasome degradation pathway, including polyubiquitin, deubiquitinating enzyme UCH37, and ubiquitin ligases E3A and E3C, were identified in the purified proteasome fraction. Ubiquitin ligase E3A is the human ortholog of yeast Hul5, which has been shown to interact with the yeast proteasome (*27*). Previously, polyubiquitination at Lys48 of ubiquitin has been detected during purification of the yeast 26S proteasome complex after in vivo cross-linking (*28*). In this study, a tryptic peptide of ubiquitin was identified as LIFAGK*QLEDGR, where K* was ubiquitinated (Table 1 of the Supporting Information). This sequence is conserved from yeast to human and is indicative of Lys48-linked ubiquitin chains. Although other lysine residues of ubiquitin can be ubiquitinated (*35*, *44*), none were detected in this experiment. The result further supports the idea that Lys48-linked polyubiquitin chains are the major signal for protein degradation (*2*).

Two known nonproteasome ubiquitin receptor proteins, hHR23B and hPLIC-1, were copurified and identified with the human 26S proteasome complex under native conditions (Table 1). In contrast, the yeast orthologs, i.e., Rad23 and Dsk2, were not copurified with yeast 26S proteasomes under native conditions (*5*), but only after in vivo cross-linking (*28*). Native purification usually captures stable interactions, whereas in vivo cross-linking stabilizes weak and transient interactions. Our results show that hHR23B and hPLIC-1 interact with the 26S proteasome as expected and that our purification strategy is rapid and efficient for capturing even dynamically interacting proteins. Although hHR23 and hPLIC exist in mammalian cells as two different isoforms, we were not able to identify hHR23A and hPLIC-2 in this work. Recent studies indicate that hHR23A has a weaker interaction with the proteasome than hHR23B (*45*), a possibility for the missed capture.

Identification of the N-Terminal Peptides of the 26S Proteasome Subunits. It has been demonstrated for yeast and murine heart proteasomes that the N-termini of subunits can be chemically modified (*20*, *29*, *46*). However, the detailed characterization of the modifications in human 26S proteasome subunits is lacking. In this study, among the 25 subunits with the identified N-termini, 21 were from tryptic digests and four from chymotrypsin digests. As an example, Figure 2 displays the ESI-MS/MS spectrum of a chymotryptic peptide with MH_3^{3+} 1037.12, and the sequence was unambiguously determined to be acetyl-MFRNQYDNDVTV-WSPQGRIHQIEYA on the basis of the observed fragment ions (details of ion assignments in Table 2 of the Supporting Information). This sequence was identified as the N-terminus of α6. In total, 15 subunits were found to be acetylated at their N-termini. Eight subunits, Rpt3, Rpt5, Rpn1, Rpn2, Gankyrin (a non-ATPase subunit), α5, α6, and β4, are acetylated at the first methionine residue, whereas seven subunits, Rpt4, Rpt6, Rpn6, Rpn13/ADRM1, α4, α7, and β3, are acetylated at the second amino acid after the removal of the first methionine residue. Other subunits such as Rpt1, Rpn7, Rpn8, and Rpn10 have free N-termini without the first methionine.

Interestingly, subunit Rpn2 was found to exist in two different N-terminal forms, as shown in the MS/MS spectra of the two corresponding peptides with MH_2^{2+} 1093.6 in Figure 3A and MH_2^{2+} 1080.5 in Figure 3B. The overall fragmentation patterns of the two peptides are very similar, with identical y ions, whereas the b ions show a 26 Da shift, suggesting that the difference between these two peptides was at the N-terminus. The presence of the b_2 ion (287.16) in Figure 3A indicates acetylation of the N-terminal methionine, whereas the characteristic loss of 64 Da ($-SOCH_4$) from the b ion series in Figure 3B indicates the presence of oxidized methionine at the N-terminus. On the basis of these results, the peptide sequences in panels A and B of Figure 3 were determined to be acetyl-MITSAAGIISLLDEDE-PQLK and M(ox)ITSAAGIISLLDEDEPQLK, respectively. These sequences unambiguously match the N-terminal peptide of subunit Rpn2. The different Rpn2 N-termini have been reproducibly observed in multiple purifications, indicating that isoforms of Rpn2 exist in proteasome complexes in vivo: one with an acetylated N-terminus and the other with a free N-terminus.

In addition to N-acetylation, the 19S subunit Rpt2 subunit has been shown to be myristoylated (*20*, *47*). Our LC−MS/MS analysis revealed that the human Rpt2 subunit is indeed modified by myristoylation with a mass addition of 210.2 at the N-terminus after removal of the first methionine. As shown in the MS/MS spectrum (MH_2^{2+} 710.89, Figure 4), a unique b_1 ion was observed due to the myristoylation of glycine at the N-terminus. The peptide sequence was unambiguously determined to be myristoyl-GQSQSGGH-GPGGGK, the N-terminus of subunit Rpt2. Most recently, Gomes and co-workers also identified N-myristoylation of murine heart Rpt2 (*29*). Although the function of N-

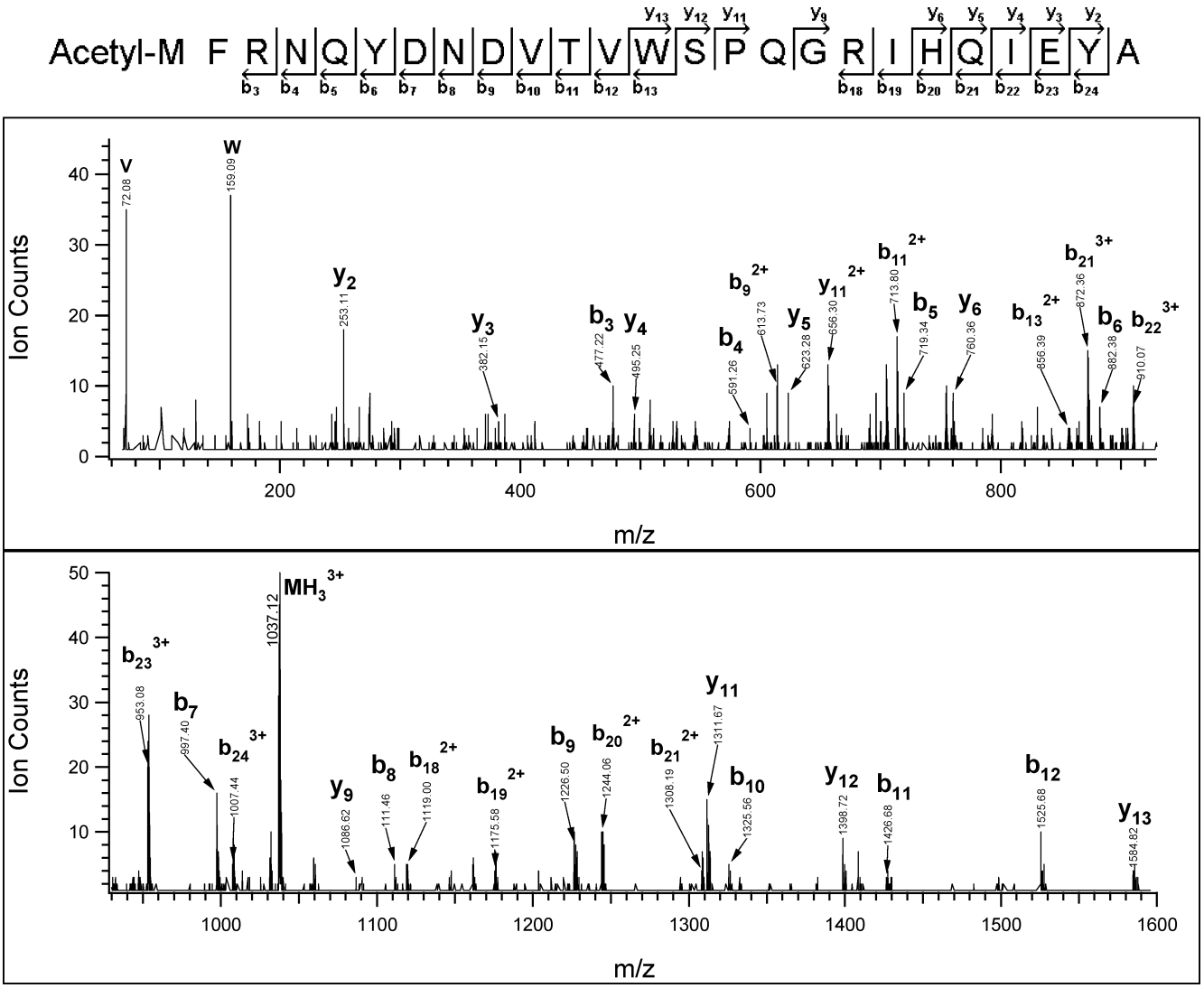

FIGURE 2: ESI-MS/MS spectrum of a peptide with MH_3^{3+} 1037.12 from a chymotryptic digest. The determined sequence, acetyl-MFRNQYDNDVTVWSPQGRIHQIEYA, matches the N-terminus of proteasome subunit α6. The detailed assignments of fragment ions are listed in Table 2 of the Supporting Information.

myristoylation in Rpt2 is not clear, the conservation of this modification on Rpt2 from yeast to human suggests an important function for Rpt2 myristoylation. Since this modification promotes protein−protein and protein−membrane interactions, we speculate that Rpt2 may play a role in the interaction of the 26S proteasome with membranes or other proteins.

In comparison to other proteasome subunits, some of the *β* subunits, especially the catalytic *β* subunits, are required to be processed at N-termini from their prosequences prior to assembly with other subunits to form a functional proteasome complex (*1*, *4*). Indeed, the N-termini of the three catalytic *β* subunits, *β*1, *β*2, and *β*5, were identified as the processed forms with the N-terminal sequences starting at double threonine residues for their catalytic activities (*1*, *2*, *4*). In addition, the *β*6 subunit appeared to be processed with a free N-terminus. Table 2 summarizes the results on N-terminal peptides of human 26S proteasome subunits identified in this work and compares them to the N-termini of yeast proteasome subunits. Clearly, the N-termini can be differently processed among subunit orthologs, although N-terminal posttranslational modifications are conserved among different eukaryotic proteasomes.

Identification of Phosphorylation Sites of the 26S Proteasome Subunits. Phosphorylation of proteasome subunits may play an important role in modulating the assembly of 26S proteasome complexes, their stability, and their functions (*17*, *48*). So far, no detailed characterization of proteasome phosphorylation has been carried out for any organisms. In this work, a total of 16 phosphorylation sites were identified from 10 proteasome subunits, and 12 of these sites have not been reported previously. As an example, Figure 5A illustrates the MS/MS spectrum of a tryptic peptide (MH_2^{2+} 906.93) from the digest of the complex. A fragment ion labeled as (MH_2^{2+})* (i.e., MH_2^{2+} 857.92) was obtained due to the loss of H_3PO_4 (−98 Da) from the parent ion (MH_2^{2+} 906.93), indicating that the peptide is phosphorylated. On the basis of the series of b and y ions, the peptide was identified as the phosphorylated form of the peptide ^{303}-TSSAFVGKTPEASPEPK^{319}, matching the sequence of subunit Rpn2. In this sequence, there are five potential phosphorylation sites. From the nonphosphorylated forms of b_3−b_8 and y_1−y_8 ions, the phosphorylation most likely occurs at Thr311. The detection of phosphorylated forms of b_9, y_9−y_{13}, and their neutral loss ions confirms this assignment unambiguously. In addition to single phosphorylation, a

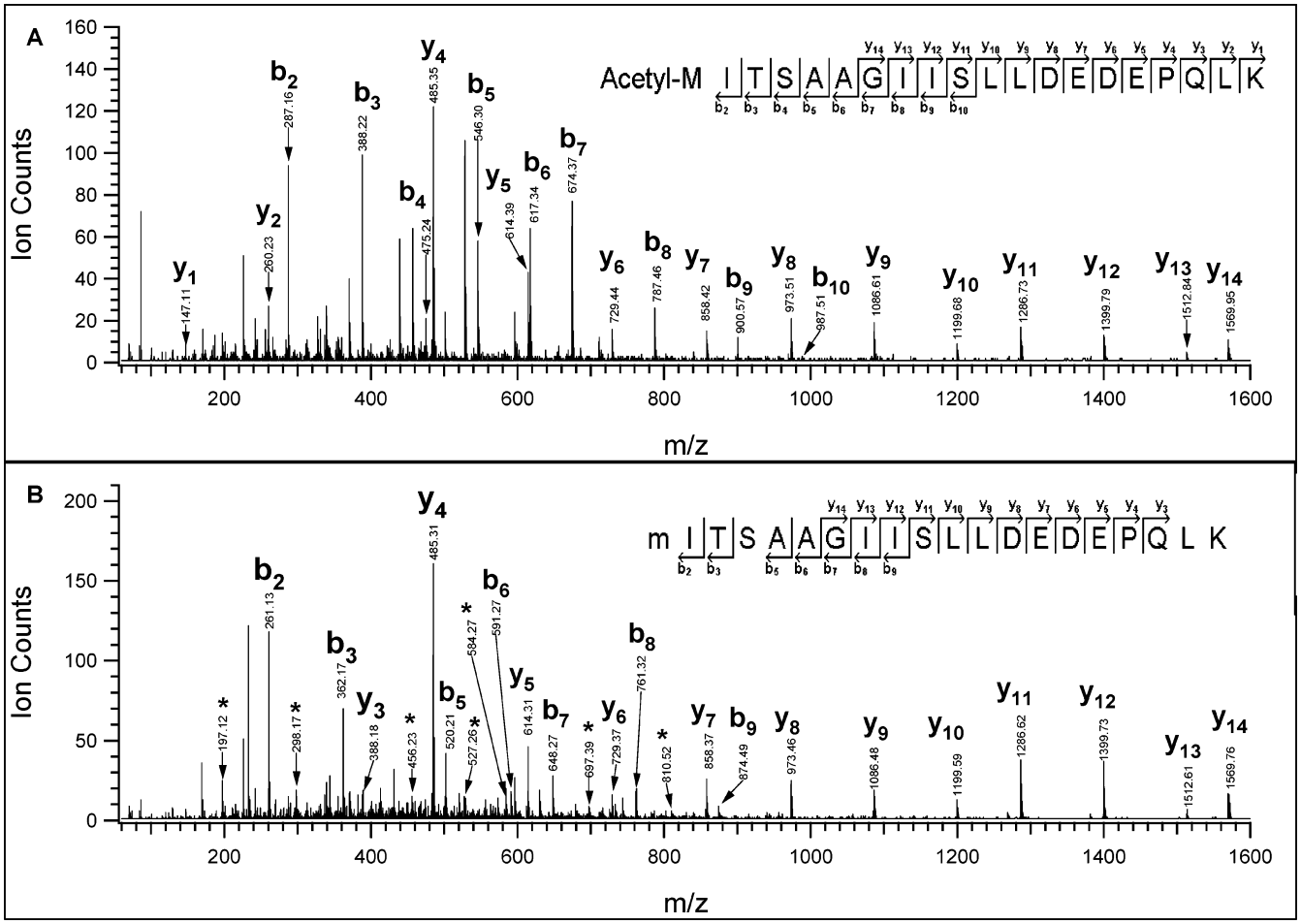

FIGURE 3: (A) ESI-MS/MS spectrum of a tryptic peptide with MH_2^{2+} 1093.6. The determined sequence, acetyl-MITSAAGIISLLDEDEPQLK, matches the N-terminus of proteasome subunit Rpn2. (B) ESI-MS/MS spectrum of a tryptic peptide with MH_2^{2+} 1080.5. Its sequence is determined to be M(ox)ITSAAGIISLLDEDEPQLK, which also matches the N-terminus of subunit Rpn2. The first methionine in panel B is oxidized instead of acetylated in comparison with the peptide in panel A.

doubly phosphorylated peptide (MH_2^{2+} 946.92) was identified as shown in Figure 5B. The detection of two fragment ions (MH_2^{2+} 897.89 and 848.94) due to loss of one and two H_3PO_4 groups, respectively (−98 and −196 Da, respectively), indicated that this peptide was modified by two phosphate groups. The second phosphorylation was determined at Ser315. More detailed results are given in Table 3 of the Supporting Information. The mapping of a total of 16 phosphorylation sites in 10 different proteasome subunits is summarized in Table 3. The corresponding MS/MS spectra of the identified phosphorylated peptides are included in Figure 2A−N of the Supporting Information.

It is worth noting that proteasome-bound hHR23B was also identified as phosphorylated on the basis of detection of both singly (MH_3^{3+} 1009.14) and doubly phosphorylated (MH_3^{3+} 1035.78) forms of a peptide from hHR23B (^{145}-QEKPAEKPAETPVATSPTATDSTSGDSSR^{173}). From the fragment ions, Ser160 was identified as being phosphorylated for both peptides; hence, Thr155 appeared most likely to be the additional phosphorylation site in the doubly phosphorylated peptide (Figure 2O,P of the Supporting Information). Although the biological functions of the phosphorylation sites identified in this work are yet to be determined, the results are useful for future functional studies aimed at elucidating the role of phosphorylation in proteasome function and regulation.

DISCUSSION

We report in this work a new affinity purification strategy using a derivative of the HB tag, which was originally designed for protein purification under fully denaturing conditions in yeast (*28*, *35*). The single-step purification strategy with streptavidin resin followed by TEV elution proves to be sufficient for rapid and effective purification of human 26S proteasomes from stable cell lines under native conditions. Previously, different affinity tags have been applied to the purification of yeast proteasomes (*5*, *27*, *28*); however, purification of mammalian proteasome complexes for proteomic analysis still largely relies on conventional chromatographic methods (*23*−*26*, *29*) and immunoaffinity purification strategies using specific proteasome antibodies (*21*, *22*). Although effective, the conventional procedure requires a large amount of starting materials and is time-consuming, and the antibody-based immunoaffinity purification prohibits downstream applications due to the presence of antibodies and the use of harsh elution conditions. Our purification strategy requires less starting material and is rapid, comparable to methods described for yeast proteasomes using other affinity tags (*5*, *27*).

Mass spectrometric analyses of the purified 26S proteasome complex have determined that the complex consists of 14 20S subunits and 21 19S subunits. Of 21 19S subunits, 18 are commonly known and three are recent new assign-

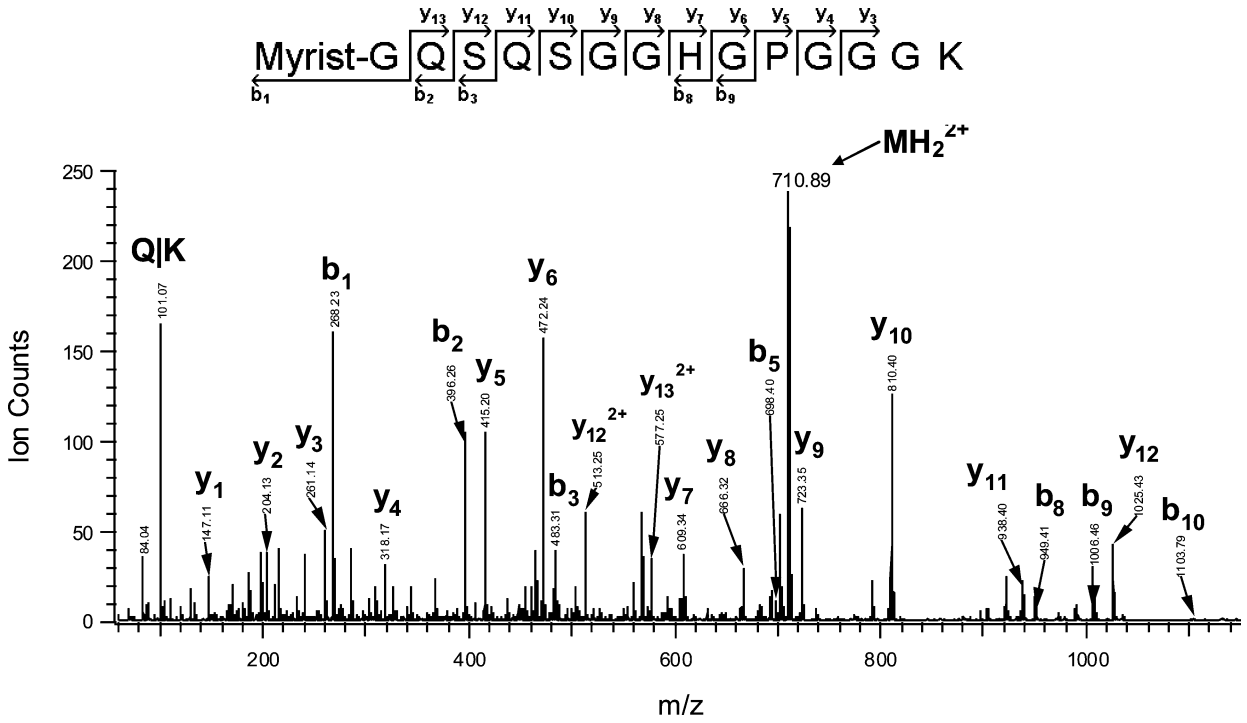

FIGURE 4: ESI-MS/MS spectrum of a tryptic peptide with MH_2^{2+} 710.89. Its sequence is determined to be myristoyl-GQSQSGGHGPGGGK, which matches subunit Rpt2.

ments (Rpn13/ADRM1, Rpn14/PAAF1, and Rpn15/DSS1) (*36*, *39*−*43*). To the best of our knowledge, this information has not been reported before from a single purification. Compared to the yeast 26S proteasome, all of the subunits identified here are conserved, and all but one (Rpn14/PAAF1) were identified in the yeast complex (*5*, *27*, *28*). Rpn13/ADRM1 is the human ortholog of the yeast Rpn13 subunit and is responsible for recruiting the deubiquitinating enzyme UCH37 to the 26S proteasome (*39*−*42*). For the two other new assignments, Rpn14/PAAF1 functions as a negative regulator of the human proteasome by controlling the assembly and disassembly of the proteasome (*36*), and Rpn15/DSS1 is the functional ortholog of yeast subunit Sem1, which appears to play a role in proteasome assembly and function (*43*). Although Rpn13/ADRM1 (42 kDa) and Rpn14/PAAF1 (42 kDa) are considered as new proteasome subunits (*39*, *40*), the number of peptides and protein sequence coverage obtained for these two subunits are much lower compared to those of other subunits with similar molecular weights. This suggests that they may be substoichiometric components or that they interact in a dynamic manner.

Most recently, the characterization of murine cardiac 26S proteasome isolated by a conventional chromatographic method has been reported using 2D electrophoresis and LC−MS/MS (*29*). In comparison, four of the 19S subunits identified here (i.e., Gankyrin, Rpn13/ADRM1, Rpn14/PAAF1, and Rpn15/DSS1) are not present in murine heart proteasomes. Three interferon-γ inducible β subunits, a spliced form of Rpn10 and S5b (PSMD5) identified in murine heart 26S proteasomes, are absent in our and others' purifications from human cells (*22*, *39*). The variation in some proteasome subunits from different origins suggests that these subunits may be specific to different types of cells and/or purification strategies (*29*).

Deubiquitinating enzymes (DUBs), notably Rpn11, UCH37/Uch2, and Ubp6/Usp14, are known to be associated with the 26S proteasome (*49*). It is suggested that they assist in the removal of the polyubiquitin moiety from protein substrates before or during translocation into the catalytic chamber of the proteasome for degradation. In this work, UCH37, the main deubiquitinating enzyme in human cells (*39*−*42*), was captured as one of the major components in the purified human 26S proteasomes. Usp14, on the other hand, was not identified in this study, although Ubp6, the yeast ortholog of human Usp14, was found to be the main deubiquitinating enzyme in yeast proteasomes (*27*). Recently, it has been shown that Usp14 did not appear to be critical for deubiquitinating activity in human cells (*39*, *41*), and its interaction with the proteasome may not be stable during the purification procedures (*39*, *50*).

In addition to the identification of the proteasome subunit composition, N-terminal processing of the 26S proteasome subunits was investigated. Compared to previous results from yeast and murine heart proteasomes (*20*, *29*, *46*), the N-terminal modifications of 12 subunits (Rpt1, Rpt4, Rpt5, Rpn2, Rpn8, Rpn9, Gankyrin, Rpn13/ADRM1, α4, α6, β5, and β6) have been identified for the first time for mammalian proteasomes. Perhaps the most intriguing finding here is the discovery of two different N-terminally processed forms of human Rpn2, either acetylated or free at the N-terminus. This novel finding suggests the presence of a heterogeneous subunit population due to different N-terminal processing. The biological function of this unusual phenomenon needs to be further explored.

Phosphorylation of the 26S proteasome complex was also studied. Although several α subunits of the 20S were reported to be phosphorylated (*22*, *29*, *51*, *52*), most of the in vivo phosphorylation sites have not been identified. In this study, a total of 16 phosphorylation sites including 12 novel sites

Table 2: Comparison of N-Terminal Peptides of Human 26S Proteasome Subunits Identified by LC−MS/MS to Those of Yeast 26S Proteasome Subunits

	subunit name	accession number	*human N-terminal sequence in database*, **N-terminal sequence identified by MS/MS**	*yeast N-terminal sequence in database*, **N-terminal processing** (*20*, *46*)
19S	Rpt1[a]	P35998	*MPDYLGADQRK*, **PDYLGADQRK**	*MPPKEDWEK*, **unknown**
	Rpt2	P62191	*MGQSQSGGHGPGGGK*, **Myr-GQSQSGGHGPGGGK**	*MGQGVSSGQDK*, **Myr-GQGVSS...**
	Rpt3	P43686	*MEEIGILVEK*, **acetyl-MEEIGILVEK**	*MEELGIVTPVEK*, **acetyl-MEELGIV...**
	Rpt4[a]	P62333	*MADPRDKALQDYR*, **acetyl-ADPRDKALQDYR**	*MSEEQDPLLAGLGETSGDNH...*, **acetyl-SEEQDPL...**
	Rpt5[a]	P17980	*MNLLPNIESPVTR*, **acetyl-MNLLPNIESPVTR**	*MATLEELDAQTLPGDDELDQ*, **acetyl-ATLEELD...**
	Rpt6	P62195	*MALDGPEQMELEEGK*, **acetyl-ALDGPEQMELEEGK**	*MTAAVTSSNIVLETHESGIK*, **acetyl-TAAVTSS**
	Rpn1	Q13200	*MEEGGRDKAPVQPQQSPAAAPGGTDEKPSGK*, **acetyl-MEEGGRDKAPVQPQQSPAAAPGGTDEKPSGK**	*MVDESDKKQQTIDEQSQISP*, **VDESDKK**
	Rpn2[a]	Q99460	*MITSAAGIISLLDEDEPQLK*, **acetyl-MITSAAGIISLLDEDEPQLK** or **MITSAAGIISLLDEDEPQLK**	*MSLTTAAPLLALLRENQDSV*, **acetyl-SLTTAAP...**
	Rpn6	O00231	*MAAAAVVEFQR*, **acetyl-AAAAVVEFQR**	*MSLPGSKLEEARRLVNEKQY*, **acetyl-SLPGSKL**
	Rpn7	Q15008	*MPLENLEEEGLPKNPDLR*, **PLENLEEEGLPKNPDLR**	*MVDVEEKSQEVEYVDPTVNR*, **VDVEEKS**
	Rpn8[a]	P51665	*MPELAVQK*, **PELAVQK**	*MSLQHEKVTIAPLVLLSALD*, **acetyl-SLQHEKV**
	Rpn9[a]	Q9UNM6	*MKDVPGFLQQSQSSGPGQPAVWHR*, **MKDVPGFLQQSQSSGPGQPAVWHR**	*MFNNHEIDTILSTLRMEADP*, **MFNNHEI**
	Rpn10	P55036	*MVLESTMVCVDNSEYMR*, **VLESTMVCVDNSEYMR**	*MVLEATVLVIDNSEYSRNGD*, **VLEATVL**
	Gankyrin[a]	O75832	*MEGCVSNLMVCNLAYSGK*, **acetyl-MEGCVSNLMVCNLAYSGK**	*MSNYPLHQACMENEFFK*, **unknown**
	Rpn13/ ADRM1[a]	Q16186	*MTTSGALFPSLVPGSR*, **acetyl-TTSGALFPSLVPGSR**	*MSMSSTVIKF*, **unknown**
	α4[a]	O14818	*MSYDRAITVFSPDGHLFQVE*, **acetyl-SYDRAITVFSPDGHLFQVE**	*MSGYDRALSIFSPD*, **acetyl-SGYDRALSIFSPD**
	α5	P28066	*MFLTRSEYDRGVNTFSPEGRLFQVE*, **acetyl-MFLTRSEYDRGVNTFSPEGRLFQVE**	*MFLTRSEYDRGVSTFSPEGRLFQVE*, **acetyl-MFLTRSEYDRGVSTFSPEGRL-FQVE**
	α6[a]	P25786	*MFRNQYDNDVTVWSPQGRIHQIEYA*, **acetyl-MFRNQYDNDVTVWSPQGRIHQIEYA**	*MFRNNYDGDTVTFSPTGRLFQVEYA*, **acetyl-MFRNNYDGDTVTFSPTGRLFQ-VEYA**
	α7	P25788	*MSSIGTGYDLSASTFSPDGR*, **acetyl-SSIGTGYDLSASTFSPDGR**	*MTSIGTGYDLSNSVFSPDGR*, **acetyl-TSIGTGYDLSNSVFSPDGR**
20S	β1	P28072	*TTIMAVQFDGGVVLGADSR*, **TTIMAVQFDGGVVLGADSR**	*TSIMAVTFKDG*, **TSIMAVTFKDG**
	β2	Q99436	*TTIAGVVYKDGIVLGADTR*, **TTIAGVVYKDGIVLGADTR**	*TTIVGVKFNN*, **TTIVGVKFNN**
	β3	P49720	*MSIMSYNGGAVMAMK*, **acetyl-SIMSYNGGAVMAMK**	*MSDPSSINGGIVVAMTGK*, **acetyl-MSDPSSINGGIVVAMTGK**
	β4	P49721	*MEYLIGIQGPDYVLVASDR*, **acetyl-MEYLIGIQGPDYVLVASDR**	*MDIILGIRVQDSVILASSK*, **acetyl-MDIILGIRVQDSVILASSK**
	β5[a]	P28074	*TTTLAFKFRHGVIVAADSR*, **TTTLAFKFRHGVIVAADSR**	*TTTLAFRF*, **TTTLAFRF**
	β6[a]	P20618	*RFSPYVFNGGTILAIAGEDFAIVASDTR*, **RFSPYVFNGGTILAIAGEDFAIVASDTR**	*QFNPYGDN*, **QFNPYGDN**

[a] The N-termini of these subunits have not been determined before in mammals. The yeast ortholog of human Gankyrin is named Nas6 (accession number P50086).

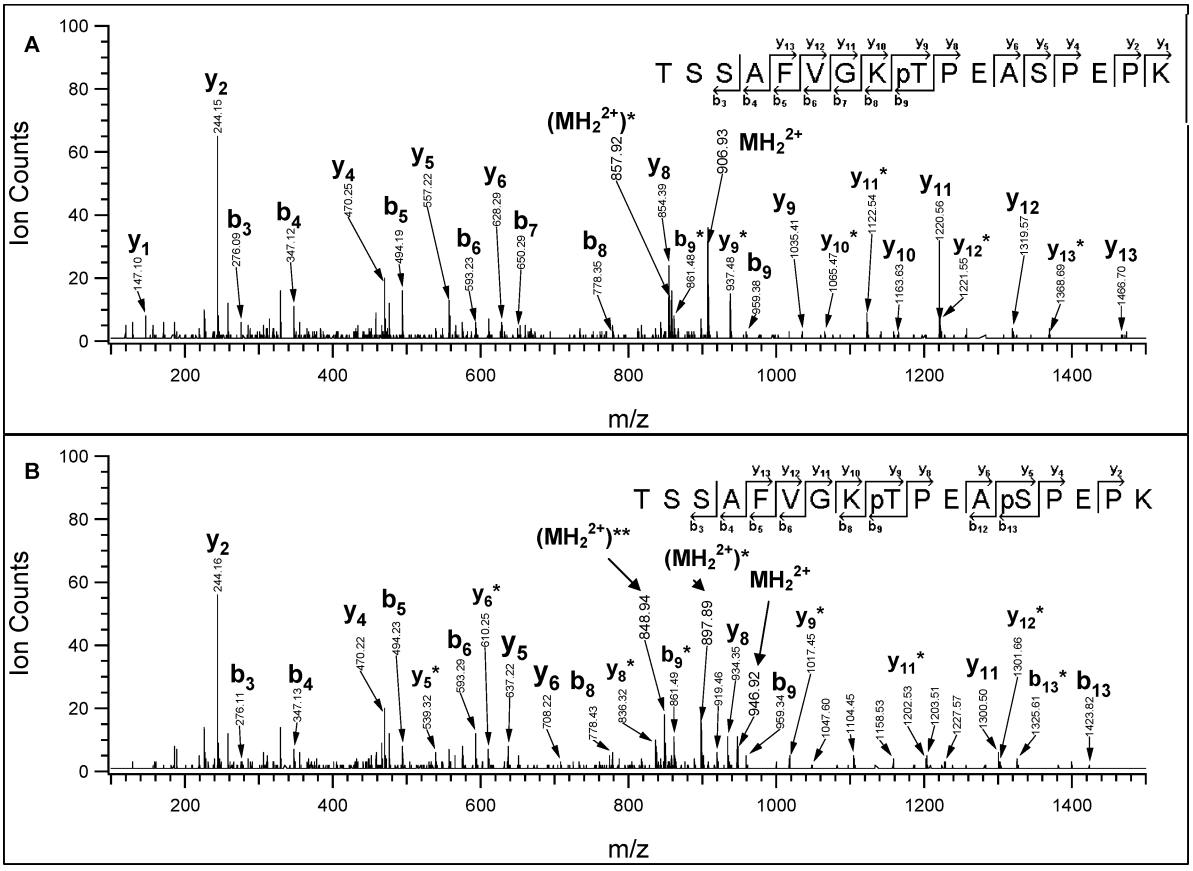

FIGURE 5: ESI-MS/MS spectra of (A) singly phosphorylated (MH_2^{2+} 906.93, [303]TSSAFVGKpTPEASPEPK[319]) and (B) doubly phosphorylated (MH_2^{2+} 946.92, [303]TSSAFVGKpTPEApSPEPK[319]) peptides, which match subunit Rpn2. The assignments of fragment ions are given in Table 3 of the Supporting Information. One asterisk denotes ions after a single neutral loss ($-H_3PO_4$, −98 Da); two asterisks denote ions after double neutral losses ($-2H_3PO_4$, −196 Da).

Table 3: Summary of Phosphorylated Peptides from the 26S Proteasome Subunits Identified by LC−MS/MS

subunit name	accession number	phosphorylated peptide identified by MS/MS[a]
Rpt5	P17980	acetyl-[1]MNLLPNIEpSPVTR[13]
Rpn1	Q13200	[9]APVQPQQpSPAAAPGGTDEKPSGK[31]
		[358]FGGpSGSQVDSAR[369]
Rpn2[b]	Q99460	[303]TSSAFVGKpTPEASPEPK[319]
		[303]TSSAFVGKpTPEApSPEPK[319]
		[270]TVGpTPIASVPGSTNTGTVPGSEK[292]
Rpn6	O00231	[11]AQpSLLSTDR[19]
		[71]YVRPFLNpSISK[81]
Rpn8	P51665	[178]DIKDTTVGpTLSQR[190]
Rpn9	Q9UNM6	[106]pSSDEAVILCK[115]
Rpn10	P55036	[263]MTIpSQQEFGR[272]
		[354]NAMGpSLASQATK[365]
		[354]NAMGpSLApSQATK[365]
Rpn11	O00487	[224]pSWMEGLTLQDYSEHCK[239]
α3	P25789	[68]LNEDMACpSVAGITSDANVLTNELR[91]
α7[b]	P25788	[241]ESLKEEDEpSDDDNm[254]
other		
hHR23B	P54727	[145]QEKPAEKPAETPVATpSPTATDSTSGDSSR[173]
		[145]QEKPAEKPAEpTPVATpSPTATDSTSGDSSR[173]

[a] pS and pT represent phosphorylated serine and threonine, respectively. [b] Only phosphorylation sites reported here that have been identified from the subunit.

have been identified. Among the eight subunits with novel phosphorylation sites identified, only a few subunits were previously suggested to be phosphorylated (*16*, *53*, *54*). For example, α3 (C9) has been shown to be phosphorylated in vivo by ^{32}P labeling (*53*, *54*), but the modified site was not known. The newly identified phosphorylation site of α3 at Ser75, which is well-conserved in mammals, indicates that this site may have important biological relevance. Previously, a large scale analysis of HeLa cell nuclear phosphoproteins suggested three phosphorylation sites (i.e., Thr273, Thr311, and Ser315) on subunit Rpn2 (*55*). We show that these phosphorylation sites actually occur in the purified proteasome complex. Almost all of the sites identified here are conserved in mammalian proteasome orthologs. Interestingly,

approximately half of the identified phosphorylated residues are serine or threonine residues followed by proline, which is indicative of phosphorylation by mitogen-activated or cyclin-dependent kinases (MAP kinases and CDKs). Given the key role MAP kinases and CDKs play in cell proliferation and the importance of the ubiquitin−proteasome pathway in cell cycle progression, it is tempting to speculate that proteasome function might be a target for cell cycle-dependent regulation. Although the functional significance of phosphorylation on the 19S subunits remains to be established, identification of novel phosphorylation sites is the first critical step toward an improved understanding of its role in regulating the function of the 26S proteasome complex.

In summary, a new purification strategy for rapidly isolating the human 26S proteasome complex under native conditions for detailed structural analysis using mass spectrometry has been developed. The results obtained in this study provide new and detailed proteomic information about the human 26S proteasome complex. This knowledge is significant to the design of future experiments for further elucidation of assembly and regulation mechanisms of the human 26S proteasome complex.

ACKNOWLEDGMENT

We thank Prof. A. L. Burlingame for allowing us to use the development version of the Protein Prospector and Aenoch Lynn for his help in operating Protein Prospector. We also thank the Huang lab members, especially Cortnie Guerrero, for their help during this work.

SUPPORTING INFORMATION AVAILABLE

Detailed peptide report for proteins identified in this work (Table 1), ion assignments of fragment ions matched to the identified sequences for Figure 2 (Table 2) and Figure 5B (Table 3), a plot of false positive rate versus expectation value (Figure 1), and ESI-MS/MS spectra of all the identified phosphorylated peptides (Figure 2A−P). This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Voges, D., Zwickl, P., and Baumeister, W. (1999) The 26S proteasome: A molecular machine designed for controlled proteolysis, *Annu. Rev. Biochem. 68*, 1015−1068.
2. Pickart, C. M., and Cohen, R. E. (2004) Proteasomes and their kin: Proteases in the machine age, *Nat. Rev. Mol. Cell Biol. 5*, 177−187.
3. Löwe, J., Stock, D., Jap, B., Zwickl, P., Baumeister, W., and Huber, R. (1995) Crystal structure of the 20S proteasome from the archaeon T. acidophilum at 3.4 Å resolution, *Science 268*, 533−539.
4. Groll, M., Ditzel, L., Lowe, J., Stock, D., Bochtler, M., Bartunik, H. D., and Huber, R. (1997) Structure of 20S proteasome from yeast at 2.4 Å resolution, *Nature 386*, 463−471.
5. Verma, R., Chen, S., Feldman, R., Schieltz, D., Yates, J., Dohmen, J., and Deshaies, R. J. (2000) Proteasomal proteomics: Identification of nucleotide-sensitive proteasome-interacting proteins by mass spectrometric analysis of affinity-purified proteasomes, *Mol. Biol. Cell 11*, 3425−3439.
6. Schmidt, M., Haas, W., Crosas, B., Santamaria, P. G., Gygi, S. P., Walz, T., and Finley, D. (2005) The HEAT repeat protein Blm10 regulates the yeast proteasome by capping the core particle, *Nat. Struct. Mol. Biol. 12*, 294−303.
7. Verma, R., Oania, R., Graumann, J., and Deshaies, R. J. (2004) Multiubiquitin chain receptors define a layer of substrate selectivity in the ubiquitin-proteasome system, *Cell 118*, 99−110.
8. Elsasser, S., and Finley, D. (2005) Delivery of ubiquitinated substrates to protein-unfolding machines, *Nat. Cell Biol. 7*, 742−749.
9. Verma, R., Aravind, L., Oania, R., McDonald, W. H., Yates, J. R. I., Koonin, E. V., and Deshaies, R. J. (2002) Role of Rpn11 Metalloprotease in Deubiquitination and Degradation by the 26S Proteasome, *Science 298*, 611−615.
10. Hoffman, L., Pratt, G., and Rechsteiner, M. (1992) Multiple forms of the 20S multicatalytic and the 26S ubiquitin/ATP-dependent proteases from rabbit reticulocyte lysate, *J. Biol. Chem. 267*, 22632−22638.
11. Ustrell, V., Hoffman, L., Pratt, G., and Rechsteiner, M. (2002) PA200, a nuclear proteasome activator involved in DNA repair, *EMBO J. 21*, 3516−3525.
12. Rechsteiner, M., Realini, C., and Ustrell, V. (2000) The proteasome activator 11 S REG (PA28) and class I antigen presentation, *Biochem. J. 345* (Part 1), 1−15.
13. Rechsteiner, M., and Hill, C. P. (2005) Mobilizing the proteolytic machine: Cell biological roles of proteasome activators and inhibitors, *Trends Cell Biol. 15*, 27−33.
14. Tanahashi, N., Murakami, Y., Minami, Y., Shimbara, N., Hendil, K. B., and Tanaka, K. (2000) Hybrid proteasomes. Induction by interferon-γ and contribution to ATP-dependent proteolysis, *J. Biol. Chem. 275*, 14336−14345.
15. Adams, J. (2004) The development of proteasome inhibitors as anticancer drugs, *Cancer Cell 5*, 417−421.
16. Mason, G. G., Murray, R. Z., Pappin, D., and Rivett, A. J. (1998) Phosphorylation of ATPase subunits of the 26S proteasome, *FEBS Lett. 430*, 269−274.
17. Rivett, A. J., Bose, S., Brooks, P., and Broadfoot, K. I. (2001) Regulation of proteasome complexes by γ-interferon and phosphorylation, *Biochimie 83*, 363−366.
18. Zhang, F., Su, K., Yang, X., Bowe, D. B., Paterson, A. J., and Kudlow, J. E. (2003) O-GlcNAc modification is an endogenous inhibitor of the proteasome, *Cell 115*, 715−725.
19. Sumegi, M., Hunyadi-Gulyas, E., Medzihradszky, K. F., and Udvardy, A. (2003) 26S proteasome subunits are O-linked N-acetylglucosamine-modified in *Drosophila melanogaster*, *Biochem. Biophys. Res. Commun. 312*, 1284−1289.
20. Kimura, Y., Takaoka, M., Tanaka, S., Sassa, H., Tanaka, K., Polevoda, B., Sherman, F., and Hirano, H. (2000) N^{α}-acetylation and proteolytic activity of the yeast 20S proteasome, *J. Biol. Chem. 275*, 4635−4639.
21. Froment, C., Uttenweiler-Joseph, S., Bousquet-Dubouch, M. P., Matondo, M., Borges, J. P., Esmenjaud, C., Lacroix, C., Monsarrat, B., and Burlet-Schiltz, O. (2005) A quantitative proteomic approach using two-dimensional gel electrophoresis and isotope-coded affinity tag labeling for studying human 20S proteasome heterogeneity, *Proteomics 5*, 2351−2363.
22. Claverol, S., Burlet-Schiltz, O., Girbal-Neuhauser, E., Gairin, J. E., and Monsarrat, B. (2002) Mapping and structural dissection of human 20S proteasome using proteomic approaches, *Mol. Cell. Proteomics 1*, 567−578.
23. Huang, L., and Burlingame, A. L. (2005) Comprehensive mass spectrometric analysis of the 20S proteasome complex, *Methods Enzymol. 405*, 187−236.
24. Schmidt, F., Dahlmann, B., Janek, K., Kloss, A., Wacker, M., Ackermann, R., Thiede, B., and Jungblut, P. R. (2006) Comprehensive quantitative proteome analysis of 20S proteasome subtypes from rat liver by isotope coded affinity tag and 2-D gel-based approaches, *Proteomics 6*, 4622−4632.
25. Zong, C., Gomes, A. V., Drews, O., Li, X., Young, G. W., Berhane, B., Qiao, X., French, S. W., Bardag-Gorce, F., and Ping, P. (2006) Regulation of murine cardiac 20S proteasomes: Role of associating partners, *Circ. Res. 99*, 372−380.
26. Huang, L., Jacob, R. J., Pegg, S. C., Baldwin, M. A., Wang, C. C., Burlingame, A. L., and Babbitt, P. C. (2001) Functional assignment of the 20S proteasome from *Trypanosoma brucei* using mass spectrometry and new bioinformatics approaches, *J. Biol. Chem. 276*, 28327−28339.
27. Leggett, D. S., Hanna, J., Borodovsky, A., Crosas, B., Schmidt, M., Baker, R. T., Walz, T., Ploegh, H., and Finley, D. (2002) Multiple associated proteins regulate proteasome structure and function, *Mol. Cell 10*, 495−507.
28. Guerrero, C., Tagwerker, C., Kaiser, P., and Huang, L. (2006) An integrated mass spectrometry-based proteomic approach: Quantitative analysis of tandem affinity-purified in vivo cross-linked protein complexes (QTAX) to decipher the 26S proteasome-interacting network, *Mol. Cell. Proteomics 5*, 366−378.

29. Gomes, A. V., Zong, C., Edmondson, R. D., Li, X., Stefani, E., Zhang, J., Jones, R. C., Thyparambil, S., Wang, G. W., Qiao, X., Bardag-Gorce, F., and Ping, P. (2006) Mapping the murine cardiac 26S proteasome complexes, *Circ. Res. 99*, 362−371.
30. Chen, P. L., Chen, Y., Shan, B., Bookstein, R., and Lee, W. H. (1992) Stability of retinoblastoma gene expression determines the tumorigenicity of reconstituted retinoblastoma cells, *Cell Growth Differ. 3*, 119−125.
31. Leggett, D. S., Glickman, M. H., and Finley, D. (2005) Purification of proteasomes, proteasome subcomplexes, and proteasome-associated proteins from budding yeast, *Methods Mol. Biol. 301*, 57−70.
32. Chalkley, R. J., Baker, P. R., Huang, L., Hansen, K. C., Allen, N. P., Rexach, M., and Burlingame, A. L. (2005) Comprehensive analysis of a multidimensional liquid chromatography mass spectrometry dataset acquired on a QqTOF mass spectrometer: 2. New developments in protein prospector allow for reliable and comprehensive automatic analysis of large datasets, *Mol. Cell. Proteomics* (in press).
33. Lynn, A. J., Chalkley, R. J., Baker, P. R., Segal, M. R., and Burlingame, A. L. (2006) Protein Prospector and Ways of Calculating Expectation Values, *Proc. 54th ASMS Conf. Mass Spectrom.*, 351.
34. Elias, J. E., Haas, W., Faherty, B. K., and Gygi, S. P. (2005) Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations, *Nat. Methods 2*, 667−675.
35. Tagwerker, C., Flick, K., Cui, M., Guerrero, C., Dou, Y., Auer, B., Baldi, P., Huang, L., and Kaiser, P. (2006) A tandem affinity tag for two-step purification under fully denaturing conditions: Application in ubiquitin profiling and protein complex identification combined with in vivo cross-linking, *Mol. Cell. Proteomics 5*, 737−748.
36. Park, Y., Hwang, Y. P., Lee, J. S., Seo, S. H., Yoon, S. K., and Yoon, J. B. (2005) Proteasomal ATPase-associated factor 1 negatively regulates proteasome activity by interacting with proteasomal ATPases, *Mol. Cell. Biol. 25*, 3842−3853.
37. Glickman, M. H., Rubin, D. M., Fried, V. A., and Finley, D. (1998) The regulatory particle of the *Saccharomyces cerevisiae* proteasome, *Mol. Cell. Biol. 18*, 3149−3162.
38. Washburn, M. P., Wolters, D., and Yates, J. R. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology, *Nat. Biotechnol. 19*, 242−247.
39. Yao, T., Song, L., Xu, W., DeMartino, G. N., Florens, L., Swanson, S. K., Washburn, M. P., Conaway, R. C., Conaway, J. W., and Cohen, R. E. (2006) Proteasome recruitment and activation of the Uch37 deubiquitinating enzyme by Adrm1, *Nat. Cell Biol. 8*, 994−1002.
40. Qiu, X. B., Ouyang, S. Y., Li, C. J., Miao, S., Wang, L., and Goldberg, A. L. (2006) hRpn13/ADRM1/GP110 is a novel proteasome subunit that binds the deubiquitinating enzyme, UCH37, *EMBO J.* (in press).
41. Hamazaki, J., Iemura, S., Natsume, T., Yashiroda, H., Tanaka, K., and Murata, S. (2006) A novel proteasome interacting protein recruits the deubiquitinating enzyme UCH37 to 26S proteasomes, *EMBO J. 25*, 4524−4536.
42. Jorgensen, J. P., Lauridsen, A. M., Kristensen, P., Dissing, K., Johnsen, A. H., Hendil, K. B., and Hartmann-Petersen, R. (2006) Adrm1, a putative cell adhesion regulating protein, is a novel proteasome-associated factor, *J. Mol. Biol. 360*, 1043−1052.
43. Sone, T., Saeki, Y., Toh-e, A., and Yokosawa, H. (2004) Sem1p is a novel subunit of the 26S proteasome from *Saccharomyces cerevisiae*, *J. Biol. Chem. 279*, 28807−28816.
44. Peng, J., Schwartz, D., Elias, J. E., Thoreen, C. C., Cheng, D., Marsischky, G., Roelofs, J., Finley, D., and Gygi, S. P. (2003) A proteomics approach to understanding protein ubiquitination, *Nat. Biotechnol. 21*, 921−926.
45. Chen, L., and Madura, K. (2006) Evidence for distinct functions for human DNA repair factors hHR23A and hHR23B, *FEBS Lett.* (in press).
46. Kimura, Y., Saeki, Y., Yokosawa, H., Polevoda, B., Sherman, F., and Hirano, H. (2003) N-Terminal modifications of the 19S regulatory particle subunits of the yeast proteasome, *Arch. Biochem. Biophys. 409*, 341−348.
47. Utsumi, T., Sato, M., Nakano, K., Takemura, D., Iwata, H., and Ishisaka, R. (2001) Amino acid residue penultimate to the amino-terminal gly residue strongly affects two cotranslational protein modifications, N-myristoylation and N-acetylation, *J. Biol. Chem. 276*, 10505−10513.
48. Bose, S., Stratford, F. L., Broadfoot, K. I., Mason, G. G., and Rivett, A. J. (2004) Phosphorylation of 20S proteasome α subunit C8 (α7) stabilizes the 26S proteasome and plays a role in the regulation of proteasome complexes by γ-interferon, *Biochem. J. 378*, 177−184.
49. Guterman, A., and Glickman, M. H. (2004) Deubiquitinating enzymes are IN/(trinsic to proteasome function), *Curr. Protein Pept. Sci. 5*, 201−211.
50. Borodovsky, A., Kessler, B. M., Casagrande, R., Overkleeft, H. S., Wilkinson, K. D., and Ploegh, H. L. (2001) A novel active site-directed probe specific for deubiquitylating enzymes reveals proteasome association of USP14, *EMBO J. 20*, 5187−5196.
51. Wehren, A., Meyer, H. E., Sobek, A., Kloetzel, P. M., and Dahlmann, B. (1996) Phosphoamino acids in proteasome subunits, *Biol. Chem. 377*, 497−503.
52. Iwafune, Y., Kawasaki, H., and Hirano, H. (2002) Electrophoretic analysis of phosphorylation of the yeast 20S proteasome, *Electrophoresis 23*, 329−338.
53. Mason, G. G., Hendil, K. B., and Rivett, A. J. (1996) Phosphorylation of proteasomes in mammalian cells. Identification of two phosphorylated subunits and the effect of phosphorylation on activity, *Eur. J. Biochem. 238*, 453−462.
54. Castano, J. G., Mahillo, E., Arizti, P., and Arribas, J. (1996) Phosphorylation of C8 and C9 subunits of the multicatalytic proteinase by casein kinase II and identification of the C8 phosphorylation sites by direct mutagenesis, *Biochemistry 35*, 3782−3789.
55. Beausoleil, S. A., Jedrychowski, M., Schwartz, D., Elias, J. E., Villen, J., Li, J., Cohn, M. A., Cantley, L. C., and Gygi, S. P. (2004) Large-scale characterization of HeLa cell nuclear phosphoproteins, *Proc. Natl. Acad. Sci. U.S.A. 101*, 12130−12135.

BI061994U