Enhanced Purification of Ubiquitinated Proteins by Engineered Tandem Hybrid Ubiquitin-binding Domains (ThUBDs)*[§]

Yuan Gao^{‡§}, Yanchang Li^{‡§}, Chengpu Zhang[‡], Mingzhi Zhao[‡], Chen Deng[‡], Qiuyan Lan[¶], Zexian Liu^{||}, Na Su[‡], Jingwei Wang[¶], Feng Xu[‡], Yongru Xu[‡], Lingyan Ping[¶], Lei Chang[‡], Huiying Gao[‡], Junzhu Wu[¶], Yu Xue^{||}, Zixin Deng[¶], Junmin Peng^{**‡‡}, and Ping Xu^{¶‡‡}

Ubiquitination is one of the most common post-translational modifications, regulating protein stability and function. However, the proteome-wide profiling of ubiquitinated proteins remains challenging due to their low abundance in cells. In this study, we systematically evaluated the affinity of ubiquitin-binding domains (UBDs) to different types of ubiquitin chains. By selecting UBDs with high affinity and evaluating various UBD combinations with different lengths and types, we constructed two artificial tandem hybrid UBDs (ThUBDs), including four UBDs made of DSK2p-derived ubiquitin-associated (UBA) and ubiquilin 2-derived UBA (ThUDQ2) and of DSK2p-derived UBA and RABGEF1-derived A20-ZnF (ThUDA20). ThUBD binds to ubiquitinated proteins, with markedly higher affinity than naturally occurring UBDs. Furthermore, it displays almost unbiased high affinity to all seven lysine-linked chains. Using ThUBD-based profiling with mass spectrometry, we identified 1092 and 7487 putative ubiquitinated proteins from yeast and mammalian cells, respectively, of which 362 and 1125 proteins had ubiquitin-modified sites. These results demonstrate that ThUBD is a refined and promising approach for enriching the

ubiquitinated proteome while circumventing the need to overexpress tagged ubiquitin variants and use antibodies to recognize ubiquitin remnants, thus providing a readily accessible tool for the protein ubiquitination research community. *Molecular & Cellular Proteomics* 15: 10.1074/mcp.O115.051839, 1381–1396, 2016.

Ubiquitination, a universal post-translational modification, refers to the covalent attachment of ubiquitin to lysine residues or the N terminus of proteins (1). It is a cascade process catalyzed by ubiquitin-activating enzyme (E1), ubiquitin-conjugating enzyme (E2), and ubiquitin ligase (E3) (2). Ubiquitination is reversible by the action of deubiquitinating enzymes, which cleave ubiquitin moieties from the substrates (3, 4). Proteins with ubiquitin-binding domains (UBDs)¹ can bind to ubiquitin, thereby expanding the functional diversity of the ubiquitination modification signaling network (5, 6). In addition, ubiquitination plays pivotal regulatory roles in a broad spectrum of cellular processes such as protein stability, gene transcription, cell cycle progression, DNA damage, and the immune response (7–10).

Ubiquitin can be conjugated to substrate proteins in different forms, including monoubiquitin, multiple monoubiquitin, and polyubiquitin. All seven lysine residues and the N terminus of ubiquitin can be involved in the formation of polyubiquitin chains (11). Different lengths and linkages of ubiquitin modifications are linked to distinct physiological functions in cells. For example, monoubiquitin regulates DNA repair and receptor endocytosis (12–14), and polyubiquitin chains, such as the well studied Lys-48-linked and Lys-63-linked polyubiq-

From the [‡]State Key Laboratory of Proteomics, National Engineering Research Center for Protein Drugs, Beijing Proteome Research Center, National Center for Protein Sciences Beijing, Beijing Institute of Radiation Medicine, Beijing 102206, China; the [¶]School of Basic Medical Science, Key Laboratory of Combinatorial Biosynthesis and Drug Discovery of Ministry of Education, School of Pharmaceutical Sciences, Wuhan University, Wuhan 430071, China; the ^{||}Department of Biomedical Engineering, College of Life Science and Technology, Huazhong University of Science and Technology, Wuhan, Hubei 430074, China; and the ^{**}Departments of Structural Biology and Developmental Neurobiology, St. Jude Proteomics Facility, St. Jude Children's Research Hospital, Memphis, Tennessee 38105

Received May 19, 2015, and in revised form, January 12, 2016

Published, MCP Papers in Press, February 3, 2016, DOI 10.1074/mcp.O115.051839

Author contributions: Y.G., J. Wu, J.P., and P.X. designed the research; Y.G., Y.L., M.Z., C.D., Q.L., J. Wang, Y. Xu, L.P., L.C., and H.G. performed the research; C.Z., M.Z., Y. Xue, Z.D., and J.P. contributed new reagents or analytic tools; Y.L., C.Z., C.D., Z.L., N.S., F.X., Y. Xue, and P.X. analyzed the data; and Y.G., Y.L., and P.X. wrote the paper.

¹ The abbreviations used are: UBD, ubiquitin-binding domain; hUBD, hybrid UBD; ThUBD, tandem hybrid UBD; UBA, ubiquitin-associated domain; LTQ, linear trap quadrupole; SPR, surface plasmon resonance; GSH, glutathione; NHS, *N*-hydroxysuccinimide; ACN, acetonitrile; SILAC, stable isotope labeling with amino acids in cell culture; AQUA, absolute quantification; Ub, ubiquitin; RP-LC, reversed-phase liquid chromatography; qUBA, the four tandem ubiquilin-1 UBA domain; Ni-NTA, nickel-nitrilotriacetic acid; FA, formic acid; HCC, human hepatocellular carcinoma.

ubiquitin chains, function in proteasomal degradation (15), endosomal trafficking to the lysosome, intracellular signaling, and DNA repair (16). The specific N-terminal linear polyubiquitin chain activates NF- κ B kinase (17). Recent studies also suggest that Lys-11 linkage plays important roles in the endoplasmic reticulum-associated protein degradation (pathway (18) and cell cycle (19), whereas Lys-33 linkage regulates cell surface receptor-mediated signal transduction (20) and post-Golgi transport (21). However, the biological functions of other atypical ubiquitinated chains still remain poorly understood (22).

Large scale profiling of ubiquitin conjugates and mapping of ubiquitination sites are instrumental to understanding the ubiquitin regulatory system. However, profiling is still technically challenging due to the low abundance and rapid degradation of the ubiquitin conjugates. To this end, several different strategies have been used to enrich the ubiquitinated proteome or ubiquitinome (23). One well established method is purifying tagged ubiquitinome from yeast (24). Later, this approach has been employed for ubiquitinome profiling in mammalian cells (25–27). However, this method is not applicable in animal tissues or pathological specimens, because of the technical challenge of expressing tagged ubiquitin in those samples. In addition, overexpressing ubiquitin in mammalian cells may interfere with normal cellular functions.

High affinity antibodies are a successful alternative approach to enriching ubiquitin conjugates (28–30). However, off-target antibody-interacting proteins or the high background of the antibody itself in liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis can be difficult to exclude. Xu *et al.* (31) developed a di-Gly-lysine-specific antibody against the signature di-glycine residues that are the remnants of ubiquitin after trypsin digestion, to directly enrich the ubiquitinated peptides in a human cell line. This method has been improved and is the primary strategy in ubiquitinome research (32–34). However, proteins modified by neural precursor cell-expressed developmentally down-regulated 8 (NEDD8) or interferon-stimulated gene 15 (ISG15) can have identical di-Gly remnants on the modified Lys residues, making it impossible for LC-MS/MS to distinguish among these modifications. In addition, the di-Gly antibody has varied affinity for different epitopes, causing biases for different ubiquitinated peptides during purification (35, 36).

UBDs are small protein module families that can recognize and bind to ubiquitin modifications and are used as alternative tools to enrich ubiquitinated proteomes. To date, more than 20 different UBD families have been identified (5, 37). The affinity of different UBDs to ubiquitin spans a wide range, from 2 to 500 μ M (6), with various preferences for different types of ubiquitin linkages (38). For example, the ubiquitin-associated (UBA) domains of hHR23A and MUD1 selectively bind to Lys-48-linked chains (39, 40), whereas the Npl4 zinc finger domain of TGF β -activated kinase 1 (TAK1)-binding protein 2 prefers Lys-63-linked chains (41), and the UBAN domain of

the NF- κ B essential modulator (NEMO) specifically binds to linear ubiquitin chains (42, 43). Although single and tandem UBDs have been used to purify ubiquitinated proteins from mammalian cells (44–46), they are usually restricted to well studied UBA domains (such as UBA from ubiquilin 1 and hHR23A) with ubiquitin chain preference, rather than reflecting the global ubiquitination landscape.

To address the bias in UBDs and efficiently recover the ubiquitinome during purification, we systematically evaluated the affinity of a number of UBDs, and we constructed an artificial tandem hybrid UBD (ThUBD) that exhibited no strong bias toward any of the seven tested polyubiquitin chains. By analyzing complex samples of yeast and mammalian cells with ThUBD, we found a dramatic improvement in binding affinity compared with previously reported UBD methods. Our study demonstrates that ThUBDs are promising tools for the efficient and unbiased analysis of the ubiquitinated proteome, especially for animal tissues or pathological specimens. This approach may aid in the discovery of novel biomarkers or therapeutic strategies against human diseases.

EXPERIMENTAL PROCEDURES

Cloning, Protein Purification, and Immobilization—The UBDs used in this study were DSK2p-derived UBA (DSK2, 327–371 amino acids), ubiquilin 1-derived UBA (UQ1, 541–586 amino acids), ubiquilin 2-derived UBA (UQ2, 575–624 amino acids), RABGEF1-derived A20-ZnF (A20, 9–73 amino acids), and HDAC6-derived ZnF-UBP (HDAC6, 985–1152 amino acids). Five different UBDs were cloned by PCR amplification and inserted into the pGEX-4T-2 vector (GE Healthcare, Chalfont St Giles, UK) using BglIII and EcoRI sites. Tandem repeats of UBDs were constructed using BglIII and BamHI restriction sites as designed by Hjerpe *et al.* (47). hUBDs (DSK2 and UQ2 (termed as UDQ2) and DSK2 and A20 (termed as UDA20)) and ThUBDs (ThUDQ2 and ThUDA20) were constructed in the same way. All of the proteins were overexpressed in *Escherichia coli* BL21 (DE3) cells, which were induced with 0.5 mM isopropyl 1-thio- β -D-galactopyranoside for 4 h at 30 °C. The harvested cells were lysed by sonication in lysis buffer (1 mM DTT, 1% Triton X-100 in PBS). GST-UBD fusion proteins were purified from cell lysates using glutathione-Sepharose (GSH) 4B beads (Qiagen, Valencia, CA) according to the manufacturer's instructions. The purified GST-UBDs were competitively eluted by reduced glutathione and then resuspended in coupling buffer (0.2 M NaHCO₃, 0.5 M NaCl, pH 8.3) before being coupled to NHS-activated Sepharose (GE Healthcare, Munich, Germany) following the manufacturer's instructions. The UBD-conjugated agarose was stored at 4 °C in PBS supplemented with 30% glycine.

Yeast Strains and Affinity Purification—The yeast strain SUB592 (48) was grown at 30 °C in yeast extract peptone dextrose (YPD) medium (1% yeast extract, 2% Bacto-peptone, 2% dextrose) and harvested in the early log phase ($A_{600} = 1.0$). His-Myc-tagged ubiquitin conjugates were purified from yeast under denaturing conditions with Ni-NTA agarose (24) or under native conditions by UBDs (18). For native conditions, cells were lysed with glass beads in 200 μ l of buffer A (50 mM Na₂HPO₄, pH 8.0, 500 mM NaCl, 0.01% SDS, 5% glycerol). Protein extracts were centrifuged at 70,000 $\times$ g for 30 min and then incubated with immobilized GST-UBD beads at 4 °C for 30 min. After incubation, the beads were washed with buffer A and then buffer B (50 mM NH₄HCO₃ and 5 mM iodoacetamide), followed by 50 mM NH₄HCO₃ to remove iodoacetamide. Bound ubiquitin conjugates were eluted by boiling in 1 $\times$ SDS-PAGE loading buffer (50 mM Tris-

HCl, pH 6.8, 2% SDS, 10% glycerol, 0.1% bromophenol blue, 1% β -mercaptoethanol).

Cell Culture and Affinity Purification—Human hepatocellular carcinoma (HCC) MHCC97-H cells (49) were cultured to confluence before harvest. Ubiquitin conjugates were purified from cells under native conditions by ThUBDs (ThUDA20) as described above (45). Finally, the purified ubiquitin conjugates were released by boiling in 1× SDS-PAGE loading buffer.

SDS-PAGE and Western Blot Analysis—Affinity-purified proteins from yeast cells were separated on 10% SDS-polyacrylamide gels, transferred onto nitrocellulose membranes (Bio-Rad), and then probed with antibodies against Myc (Abcam, Cambridge, UK) or visualized by silver staining.

Sample Preparation for MS Analysis—Yeast or human ubiquitin conjugates were separated on SDS-polyacrylamide gels or under-went off-line high pH reverse phase LC separation (24, 50). The ubiquitin conjugates separated by SDS-PAGE were sliced into different gel pieces based on molecular weight markers and digested with trypsin overnight. The tryptic peptides were extracted with extraction buffer (5% total cell lysates (FA) + 45% acetonitrile (ACN)) and then ACN and were finally dried using a vacuum dryer (Labconco Centri-Vap, Kansas City, MO).

Peptides were analyzed using an ultra-performance LC-MS/MS platform of hybrid LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher Scientific, San Jose, CA) equipped with a Waters nano-AQUITY ultra-performance liquid chromatography system (Waters). The LC separation was performed on an in-house packed capillary column (75 μ m inner diameter $\times$ 15 cm) with 3- μ m C18 reverse-phase fused silica (Michrom Bioresources, Inc., Auburn, CA). Then the sample was eluted with a 60–140-min nonlinear gradient ramped from 8 to 40% of mobile phase B (phase B, 0.1% FA in ACN; phase A, 0.1% FA + 2% ACN in water) at a flow rate of 0.3 μ l/min. Eluting peptides were analyzed using an LTQ-Orbitrap Velos mass spectrometer. The MS1 was analyzed with a mass range of 300–1600 at a resolution of 30,000 at m/z 400. The automatic gain control was set as 1×10^6 , and the maximum injection time was 150 ms. The MS2 was analyzed in data-dependent mode for the 20 most intense ions subjected to fragmentation in the linear ion trap (LTQ). For each scan, the automatic gain control was set at 1×10^4 , and the maximum injection time was set at 25 ms. The dynamic range was set at 45–60 s to suppress repeated detection of the same ion peaks.

Data Analysis—The raw data files were searched by the Sorcerer™-QUEST® (version 4.0.4 build, Sage-N Research, Inc.) and SEQUEST HT algorithms embedded in Proteome Discovery (version 1.4.1.14, Thermo Fisher) against the Swiss-Prot reviewed database (version released in 2013.10 for yeast, contains 6652 entities; version 2014.04 for human, contains 20,266 entities). MS search parameters consisted of full tryptic restriction, fixed modification of cysteine carbamidomethylation, methionine oxidation, and di-Gly-lysine as variable modifications. Peptides were allowed up to three cleavage sites. Precursor mass tolerance was set at 20 ppm and that of the MS₂ fragments was set at 0.5 Da. Peptide matches were filtered by a minimum length of seven amino acids. For ubiquitination identification, a re-searching strategy was used for protein identification; di-Gly-lysine was added in the second search. The C-terminal di-Gly-lysine site identifications were removed. The peptides and proteins were filtered until a false discovery rate lower than 1% was estimated using a target-decoy search strategy.

Targeted Quantitative MS of Polyubiquitin Linkages Using the Stable Isotope Labeling by Amino Acids in Cell Culture-Absolute Quantification (SILAC-AQUA) Approach—To accurately quantify the abundance of the seven polyubiquitin chains from each ubiquitin conjugate purified with different UBDs, ubiquitin conjugates labeled with heavy isotope ¹³C₆-lysine (Lys-6) and ¹⁵N₄-¹³C₆-arginine (Arg-10) were puri-

fied with nickel beads as a quantification standard to mix equally with each sample (18). The ubiquitin conjugates were digested as described above, and the resulting peptides were dissolved in a sample buffer (1% FA, 1% ACN) and analyzed by LC-MS/MS. Eluted peptides were detected on the Orbitrap mass spectrometer in a survey scan (300–1600 m/z , resolution 30,000) followed by selective reaction monitoring scans for seven ubiquitin chains ions in the LTQ. The peptide intensity for quantification was manually analyzed by ion chromatograms using Xcalibur version 2.0 software (Thermo Finnigan, San Jose, CA).

Determination of the Amount of Immobilized UBDs Using the SILAC-AQUA Approach—To accurately compare the binding affinity to ubiquitin of different UBDs, the same amount of each UBD was needed. We selected two typical peptides (supplemental Fig. S1) from glutathione S-transferase (GST) protein as internal standards for quantifying the amount of GST-UBDs coupled to agarose beads. The GST protein was overexpressed and completely labeled with [¹³C₆]lysine (+6.0201 Da) using the SILAC-labeled *E. coli* BL21 (DE3) cells and was then purified using GSH 4B beads. The amount of purified heavy-labeled GST was measured using the BCA protein assay kit (Thermo Scientific, Rockford, IL) and Coomassie staining (51). A total of 10 μ g of heavy-labeled GST was digested with Lys-C, and the resulting peptides were used as quantitative standards. To quantify the GST-UBD proteins, on-bead digestion with Lys-C was performed to release peptides from GST-UBDs coupled to agarose beads. The resulting peptides were desalted by passing through a C₁₈ microcolumn as reported previously (52), and then mixed with the heavy-labeled and Lys-C-digested GST peptides followed by nano-LC-MS/MS analysis. The detailed quantification method was described as reported previously (53). According to the light and heavy intensities of the same peptides from GST proteins, we calculated the amount of light GST in samples, which represented the amount of GST-UBD proteins.

Surface Plasmon Resonance Imaging Measurements and Analysis—The surface plasmon resonance imaging binding measurements were performed on the PlexArray® HT system (Plexera LLC, Washington, WA). Experiments were performed in PBS buffer at 25 °C. Purified GST-UBD proteins were resuspended in PBS buffer and then immobilized onto NanoCapture Gold™ SPR-active slides following the standard amine coupling chemistry, as described previously (54). Then different concentrations (10, 5, 2.5, 1.25, and 0.625 μ g/ μ l) of Lys-48-linked tetra-ubiquitin (Boston Biochem, Cambridge, MA) were injected at a flow rate of 2 μ l/s for 300 s. All of the injections were performed as randomized duplicates, and the entire experiment was repeated three times. The binding level was tracked by measuring the change in response units, according to the different injection concentrations. K_D values were calculated using BIAevaluation software.

Virtual Western Blot—With the extracted ion chromatograms of each protein in different bands and the experimental molecular weight of each band, we reconstructed the experimental lanes using an in-house-developed MATLAB script (55). The reconstructed images revealed the different purification methods with regard to molecular weight and abundance.

RESULTS

hUBDs Showed Improved Binding Affinity to Ubiquitinated Proteins—UBDs, which directly interact with ubiquitin, have been utilized for purifying ubiquitinated proteins (44–47). Certain UBA domains from DSK2 and UQ1 are preferred to enrich ubiquitinated proteins, because they appear to bind equally to different types of ubiquitin linkages (39). However, these results were obtained by *in vitro* pulldown assays with radiolabeled Ub₄ linked through different lysine residues, lacking

TABLE I
UBDs used in this study

Name	Source	Type	Binding affinity K_D μM
DSK2	<i>Saccharomyces cerevisiae</i> Dsk2p (DSK2)	UBA	14.8 + 5.3 ^a
UQ1	<i>Homo sapiens</i> ubiquitin 1 (UBQLN1)	UBA	20 ^b
UQ2	<i>H. sapiens</i> ubiquitin 2 (UBQLN2)	UBA	ND ^c
A20	<i>H. sapiens</i> RAB guanine nucleotide exchange factor (GEF) 1 (RABGEF1)	ZnF-A20	12–22 ^{d,e}
HDAC6	Mouse (HDAC6)	ZnF-UBP	0.06 ^f
UDQ2	This study		
UDA20	This study		

^a See Ref. 61.

^b See Ref. 62.

^c See Ref. 63.

^d See Ref. 60.

^e See Ref. 64.

^f See Ref. 65.

further systematic verification. In this study, we quantitatively compared the binding characteristics of seven UBDs (Table I and supplemental Table 1) to endogenous ubiquitinated proteins in the cell lysates by Western blotting and LC-MS/MS.

We first compared the relative binding efficiency of seven UBDs to ubiquitinated proteins (Fig. 1A). Purified GST-UBDs (Fig. 1B) were coupled to NHS beads as the bait. The amount of GST-UBDs coupled to NHS beads was analyzed by SILAC-AQUA (supplemental Fig. S1). Equal amounts of cell lysates from SUB592 were incubated with the same molar concentrations of GST-UBDs for affinity experiments. Purified ubiquitin conjugates by different UBDs were split into two portions, and 5% (v/v) ubiquitin conjugates were resolved by SDS-PAGE for Western blot analysis. The rest of the ubiquitin conjugates were mixed with equal amounts of the respective heavy isotope-labeled ubiquitin conjugates purified with Ni-NTA under denaturing conditions, followed by separation on SDS-polyacrylamide gels, digestion with trypsin, and analysis by LC-MS/MS. Compared with the A20_mut, which does not bind to ubiquitin (56), all seven UBDs pulled down ubiquitinated proteins from total cell lysates. However, different amounts of ubiquitinated proteins were captured by the UBDs (Fig. 1D), suggesting that different UBDs have diverse binding efficiencies to ubiquitinated proteins. Among these, A20, DSK2, UQ2, and two hUBDs (UDQ2 and UDA20) enriched more ubiquitinated proteins than the other UBDs. Two hUBDs, especially UDA20, enriched the most ubiquitinated proteins, suggesting that a combination of different UBDs can significantly increase the binding efficiency to ubiquitinated proteins.

Next, we investigated the binding affinity of the UBDs to seven ubiquitin linkages. The heavy isotope-labeled ubiquitinated proteins purified by Ni-NTA-agarose were used as a reference (Fig. 1A), representing the proportion of different polyubiquitin linkages in yeast cells. The relative amounts of seven polyubiquitin chains purified by different UBDs were

quantitatively measured with Ub-AQUA (18, 55). As a result, Lys-6-, Lys-11-, Lys-27-, Lys-29-, Lys-33-, Lys-48-, and Lys-63-linked polyubiquitin chains were all detected from the proteins pulled down by the UBDs (Fig. 1E), demonstrating the efficient enrichment of ubiquitinated proteins. In addition, LC-MS/MS analyses further confirmed that different UBDs exhibit diverse interaction properties with polyubiquitin chains. For instance, the A20 domain preferred Lys-63-linked chains and UQ1 purified more Lys-11-linked chains, whereas the hUBD (especially UDA20) apparently had a higher binding affinity to all seven polyubiquitin chains. These results are not completely consistent with previous studies using either SPR analyses or pulldown assays (39). These differences may be due to binding affinity that was normalized based on the proportion of different polyubiquitin linkages purified by Ni-NTA-agarose under denaturing conditions. It also provides a new way to systematically analyze the binding efficiency of UBDs to polyubiquitin chains. These data suggest that hUBDs not only show increased binding affinity to endogenous ubiquitinated proteins but also have no bias toward different polyubiquitin chains. Additional studies are needed to fully understand the binding mechanism of hUBDs to ubiquitin and different linkages of ubiquitin chains.

ThUBDs Have Superior Binding to Ubiquitinated Proteins—To achieve higher binding affinity to ubiquitin (45, 47), ThUBDs were also constructed (Fig. 2A). Five UBDs, including A20, DSK2, UQ2, and two hUBDs (UDQ2 and UDA20), were selected based on their higher binding affinities to ubiquitin (Fig. 1D). Purified ThUBDs (Fig. 2B) were coupled to NHS beads and quantified by LC-MS/MS as described previously. All five ThUBDs efficiently purified ubiquitinated proteins from total cell lysates (Fig. 2C). The different ubiquitination protein patterns of ThUBDs suggest that these UBDs may have specific preferences for ubiquitinated substrates. Furthermore, two ThUBDs (4_UDQ2 and 4_UDA20) also enriched more ubiquitinated proteins than the other three ThUBDs, as dem-

FIG. 1. Quantitative comparison of the binding affinity among seven different UBDs to ubiquitin and seven types of ubiquitin chains. A, flowchart for evaluating the binding affinity of recombinant UBDs to ubiquitin and ubiquitin chains with Western blot analysis and SILAC-AQUA. B, quantification of purified seven types of recombinant GST-UBDs on 10% SDS-polyacrylamide gels. A20_mut with the replacement of C35A and C38A was saved as the negative control for enrichment analysis. Same copy numbers of GST-UBDs were loaded on the gel and displayed with Coomassie Blue staining. C and D, comparison of the ubiquitin binding affinity of different UBDs. Equimolar UBDs were used to enrich ubiquitinated proteins from the same amount of yeast total cell lysates. The enriched ubiquitinated proteins were resolved by SDS-PAGE and probed with an anti-Myc antibody. The relative amount of ubiquitinated proteins was calculated based on intensity of the Western blot signal. E, quantitative comparison of the binding affinity of seven types of UBDs to seven types of ubiquitin chains. The amount of seven ubiquitin chains was measured by SILAC-AQUA. The same amount of heavy isotope-labeled ubiquitin chains purified by Ni-NTA under denaturing conditions was saved as the internal standard added to each reaction. The result was shown as relative amount compared with that from the enrichment of ubiquitinated proteins by Ni-NTA under denaturing conditions. Data are represented as mean and S.E.

onstrated in our pulldown experiments. This was especially the case for proteins with higher molecular masses (>50 kDa) (Fig. 2, C and D).

We also compared the binding affinity of TUBDs to seven polyubiquitin chains. By AQUA analysis, we found that the relative binding affinity of TUBDs to different polyubiquitin chains was significantly higher than that of single UBDs (Fig. 2E), in accordance with data from previous reports (45, 47). In addition, most TUBDs possess diverse binding selectivity to different ubiquitin linkages. Particularly, two ThUBDs

(4_UDQ2 and 4_UDA20) showed no obvious bias to seven polyubiquitin chains. Hence, they were selected for further characterization.

Binding Affinity of the Dimer of ThUBDs Reaches to Saturation—To optimize UBDs for the ubiquitinated proteomics study, we investigated the effects of the tandem units on ubiquitin binding affinity. We constructed the respective mono-, di-, and tetra-ThUBDs (Fig. 3A), and we found that the di-ThUBDs purified more ubiquitinated proteins than the mono-ThUBDs (Fig. 3, B and C), suggesting that repeats of

FIG. 2. Quantitative comparison of the binding affinity to ubiquitin and seven ubiquitin chains with five different types of TUBD recombinants selected from the first run of screening. **A**, schematic representation of tetra-TUBDs. UBD unit was linked through polyglycine linker and fused with GST on the N terminus. **B**, quantification of purified five types of recombinant GST-TUBDs on a 10% SDS-polyacrylamide gel. Same copy numbers of GST-TUBDs were loaded on the gel and displayed with Coomassie Blue staining. **C** and **D**, comparison of the ubiquitin binding affinity of different TUBDs. The equimolar TUBDs are used to enrich ubiquitinated proteins from the same amount of yeast total cell lysates. The enriched ubiquitinated proteins are separated by SDS-PAGE and probed with anti-Myc antibody. The relative amount of ubiquitinated proteins was calculated based on the intensity of Western blot signal. **E**, quantitative comparison of the binding affinity of seven types of ubiquitin chains to five types of TUBDs through SILAC-AQUA strategy as mentioned above. Same amount of heavy isotope-labeled ubiquitin chains purified by Ni-NTA under denaturing condition was saved as the internal standard added to each reaction. The result was shown as relative amount compared with that from the enrichment of ubiquitinated proteins by Ni-NTA in denature condition. Data are represented as mean and S.E.

the ThUBD monomer can increase ubiquitin binding affinity. However, the purified tetra-ThUBDs had lower solubility and stability, due to their high molecular weight, which might explain why there was no apparent increase in ubiquitin binding affinity to tetra-ThUBDs compared with di-ThUBDs (Fig. 3, **B** and **C**). In addition, the di-ThUBDs, especially 2_UDA20 repeats, showed no apparent bias to all seven polyubiquitin chains (Fig. 3D), further demonstrating the superiority of this hybrid UBD on the ubiquitinome study.

To further validate the pulldown results, quantitative SPR analysis was applied on mono-, di-, and tetra-UDA20. The di-ThUBD showed the highest binding affinity to Lys-48-Ub₄

(Table II), in agreement with the Ub-AQUA results (Fig. 3D). The binding affinity of 2_UDA20 to Lys-48-Ub₄ was also better than the other three tandem UBDs (Table II), including ubiquitin 1 TUBE (47), HR23A TUBE (47), and qUBA (45), which have been successfully used to purify ubiquitinated proteins in many cell lines. These results suggest that 2_UDA20 might be an optimal tool for the global ubiquitinated proteome research.

ThUBD (2_UDA20) Shows Superior Binding to Ubiquitinated Proteins than qUBA—To evaluate the effectiveness of ThUBD, we compared the relative binding efficiency of 2_UDA20 and qUBA to ubiquitinated proteins from yeast cell

FIG. 3. Evaluation and optimization of ThUBD recombinant. A, quantification of purified two types of hybrid GST-hUBDs recombinant on a 10% SDS-PAGE. Dimer, trimer, and tetramer of hUBDs were recombined and fused with GST. Same copy numbers of GST-hUBDs were loaded on the gel and displayed with Coomassie Blue staining. B and C, comparison of the ubiquitin binding affinity of different types of ThUBDs with varied units. Equimolar ThUBDs were used to enrich ubiquitinated proteins from the same amount of total cell lysates. The enriched ubiquitinated proteins were separated by SDS-PAGE and probed with an anti-Myc antibody. The relative amount of ubiquitinated proteins was calculated based on the intensity of the Western blot signal. D, quantitative comparison of the binding affinity of seven types of ubiquitin chains to two types of ThUBDs with varied number of units through SILAC-AQUA strategy as mentioned above. Same amount of heavy isotope-labeled ubiquitin-chains purified by Ni-NTA under denaturing condition was saved as the internal standard added to each reaction. The result was shown as the relative amount compared with that from the enrichment of ubiquitinated proteins by Ni-NTA in denature condition. Data are represented as mean and S.E.

TABLE II
 Equilibrium dissociation constants (K_D) of the interaction between Lys-48 tetra-ubiquitin and UBDs

Name	Binding affinity K_D	Source
	μM	
UDA20	8.09	This study
2_UDA20	4.46	This study
4_UDA20	6.41	This study
qUBA	6.8	This study
Ubiquitin 1 TUBE	8.94	47
HR23A TUBE	6.86	47

lysates (Fig. 4A). After being resolved on SDS-polyacrylamide gels, proteins captured by ThUBD and qUBA exhibited similar patterns (Fig. 4B), suggesting that both UBDs are capable of efficiently pulling down ubiquitinated proteins from yeast

cells. Interestingly, we also noticed that the amount of ubiquitinated proteins enriched by the ThUBD was more than the amount for qUBA, even for proteins in the high molecular weight range. These results strongly support the fact that the ThUBD binds more efficiently to ubiquitinated proteins than the qUBA.

These results were also corroborated by LC-MS/MS analysis (Fig. 4C) and quantitative SPR (Table II). In the high molecular mass region (>50 kDa), the superiority of ThUBD binding to ubiquitin was more evident (Fig. 4C). Protein ubiquitination, especially polyubiquitination, can shift a protein's molecular weight to a higher molecular mass region, indicating that ThUBD prefers binding to polyubiquitinated proteins. In fraction 18, which was under 20 kDa and mainly represented the monomer and dimer ubiquitin regions, we found extremely large amounts of ubiquitin purified by ThUBD rather

FIG. 4. Characterization of ThUBD and qUBA for the enrichment of polyubiquitinated proteins and ubiquitin from yeast total cell lysates. **A**, flowchart for evaluation of enrichment efficiency for ubiquitinated proteins by ThUBD and qUBA with quantitative proteomics approach. **B**, comparison of ubiquitin conjugates purified with ThUBD and qUBA by SDS-PAGE. Ubiquitinated proteins were purified from the same amount of total cell lysates, resolved on a 10% SDS-polyacrylamide gel, stained with Coomassie Blue, excised into 18 gel bands, digested by trypsin, and analyzed by LC-MS/MS. **C**, comparison of the ubiquitin purified by ThUBD and qUBA in each gel band, reflecting quantitative data in all gel bands in the purified ubiquitinated proteins. The ubiquitin abundance was represented by the intensity signal on MS of Ser-57 peptide (TLSDYNIQK) extracted from each individual gel band. Data are represented as mean and S.E. **D**, comparison of identified ubiquitinated proteins with ubiquitin-modified sites enriched by ThUBD ($n = 362$) or qUBA ($n = 260$). **E**, log₂ ratio distribution of the spectrum counts of identified ubiquitinated proteins with the ubiquitin-modified site enriched by ThUBD and qUBA. The ubiquitinated protein abundance was represented by the spectrum count of identified peptides extracted from each individual gel band.

than by qUBA. Within this fraction, we extracted the spectra of ubiquitin and found that almost 50% of the spectra purified by ThUBD contained ubiquitin-modified sites, suggesting that most of these ubiquitins were derived from polyubiquitinated rather than monoubiquitinated proteins. However, the opposite results were shown in fraction 18, which was purified by qUBA (supplemental Fig. S2A). Together, these results indicate that ThUBD captures more ubiquitinated proteins, especially polyubiquitinated proteins, than qUBA from yeast cell lysates.

Through LC-MS/MS analysis, we identified 3837 and 3594 proteins, including 362 and 260 ubiquitinated proteins with

ubiquitin-modified sites purified by ThUBD and qUBA (Fig. 4D), respectively. Based on the virtual Western blot strategy for validating ubiquitinated proteins as described previously (55), we further screened out 905 and 538 proteins with a 10-kDa increase in molecular mass (supplemental Fig. S2, B and C). Combining proteins with ubiquitin-modified sites and molecular weight increases, we identified 1092 and 690 proteins purified by ThUBD and qUBA as reliable ubiquitinated proteins (supplemental Tables 2 and 3). The number of overlapped ubiquitinated proteins was 509, representing 46.5 and 73.8% of the identified ubiquitinated proteins purified by ThUBD and qUBA, respectively (supplemental Fig. S2D). Not

only was the number of ubiquitinated proteins purified by ThUBD higher than that purified by qUBA, the relative amounts of these proteins enriched by ThUBD were also significantly enhanced (Fig. 4E). The same trend was also shown for the ubiquitinated proteins identified from samples enriched by ThUBD and qUBA (supplemental Fig. S2E), consistent with our gel imaging data.

Compared with qUBA, a known purification material that is also based on the binding affinity of UBD to ubiquitin, ThUBD purified more ubiquitinated proteins in the large scale purification from yeast cells, demonstrating that ThUBD is a promising purification reagent for global ubiquitinated proteomic research.

Comparison of Enrichment Efficiency between the Optimized ThUBD (2_UDA20) and Ni-NTA—To further evaluate the efficiency of ThUBD as a reagent for global ubiquitinated proteomic research, we also compared it with the Ni-NTA affinity purification method. Ubiquitinated proteins were purified by ThUBD under native conditions or by Ni-NTA under denaturing conditions from the JMP024 strain, which expresses His-Myc-tagged ubiquitin (24), and were then digested by trypsin after separation by SDS-PAGE, followed by LC-MS/MS analysis. Both reagents successfully purified the ubiquitinated proteins and displayed an apparent smear pattern in the high molecular mass region, compared with total cell lysates (Fig. 5A). ThUBD purified even more ubiquitinated proteins than Ni-NTA, indicating its higher efficiency. Quantification of peptides representing ubiquitin ions (Ser-57, Lys-48, and Lys-63) in gel fraction 2 (Fig. 5B) also showed that significantly more ubiquitin was captured by ThUBD than by Ni-NTA, consistent with our gel imaging results. The same trend was shown in the gel fractions in the high molecular mass region (>50 kDa) (Fig. 5C), suggesting that ThUBD prefers binding to polyubiquitinated proteins. However, in the low molecular mass region, especially in the 10-kDa region (which often represents the ubiquitin pool), more ubiquitin was captured by Ni-NTA than by ThUBD. The overall higher enrichment of the larger molecular weight ubiquitinated proteins could be attributed to the protection of ThUBDs for polyubiquitinated proteins from deubiquitinase or proteasomal degradation (47), showing the superiority of ThUBD in ubiquitinated protein purification.

Based on the same criteria used in ThUBD, we finally identified 710 ubiquitinated proteins, including 268 ubiquitinated proteins with ubiquitin-modified sites purified by Ni-NTA under denaturing conditions (supplemental Tables 2 and 3). The shared number of ubiquitinated proteins with ubiquitin-modified sites was 118, representing 32.6 and 44% of the identified ubiquitinated proteins purified by ThUBD and Ni-NTA, respectively. By reconstructing virtual Western blots for ZEO1 (YOL109W), which was identified by ThUBD and Ni-NTA as ubiquitinated proteins (Fig. 5E), we found that ZEO1 was modified by ubiquitin, indicating the reliability of identification. However, the profile of ZEO1 purified by ThUBD and Ni-NTA

was quite different. The ZEO1 protein purified by ThUBD showed a distinct molecular weight increase to the high molecular mass region, suggesting polyubiquitination. These results further confirmed the preference of ThUBD to polyubiquitinated proteins.

The total of 1092 candidate ubiquitinated proteins identified by ThUBD represents one of the largest datasets publicly available for yeast ubiquitinome study in a single experiment, compared with the 710 proteins identified by Ni-NTA (this study), 1075 proteins identified by Ni-NTA without filtration (24), and 1108 proteins identified by di-Gly antibody in yeast (57). The overlap of ubiquitinated proteins identified in these four datasets increased reliability of the identified ubiquitinated proteins (supplemental Fig. S3A). Among the overlapped proteins, about 39.7% (434/1092) of the ubiquitinated proteins were uniquely identified by ThUBD, which suggests the high binding efficiency of ThUBD in the purification of ubiquitinated proteins. Among the identified ubiquitinated proteins with ubiquitin-modified sites, there were shared proteins across all four datasets (Fig. 5D). More than 64.4% (233/362) of identified proteins by ThUBD are novel ubiquitinated proteins that are not identified by Ni-NTA. These results imply that a potentially large pool of ubiquitinated proteins remains to be explored.

The relative amount of ubiquitinated proteins enriched by ThUBD was also significantly enhanced, compared with that purified by Ni-NTA (supplemental Fig. S3B). Together, these results demonstrate that ThUBD can efficiently capture ubiquitinated proteins, especially polyubiquitinated proteins, under native conditions.

Large Scale Identification of the Endogenous Ubiquitinated Proteins in MHCC97-H Cells Using ThUBD (2_UDA20)—MHCC97-H cells are a highly metastatic HCC cell line that has high invasion ability (49). To investigate possible misregulation in the ubiquitination system of MHCC97-H cells, we performed global ubiquitinome analysis in these cells based on ThUBD. As expected, ubiquitinated proteins from MHCC97-H cells were successfully purified by ThUBD (Fig. 6A) and displayed the characteristic smear pattern on the SDS gel. Quantification of peptides representing ubiquitin ions (Ser-57, Lys-48, and Lys-63) in gel fraction 3 showed an apparent increase in the samples purified by ThUBD, compared with that in total cell lysates (Fig. 6B). Additionally, a significant portion of proteins had a clear shift toward higher molecular weight regions in purified ubiquitinated proteins, compared with that in total cell lysates (Fig. 6C), both of which demonstrated the effectiveness of using ThUBD to purify ubiquitinated proteins.

We identified 1663 credible ubiquitinated proteins from MHCC97-H cells based on the same criteria used in yeast, which was a greater number than the ubiquitinated proteins identified by other UBDs (Fig. 6D) (supplemental Tables 4 and 5) (45, 46). However, only 19.2% (319/1663) of the identified proteins had ubiquitin-modified sites, which was

FIG. 5. Comparison of the enriched ubiquitinated proteins from yeast total cell lysates by ThUBD and Ni-NTA beads under denaturing condition (Ni_denature). A, comparison of ubiquitinated proteins purified with ThUBD and Ni-NTA beads under denaturing condition by SDS-PAGE. Ubiquitin conjugates were purified from the same amount of yeast total cell lysates, resolved on a 10% SDS-polyacrylamide gel, stained with Coomassie Blue, excised into 24 gel bands, digested by trypsin, and analyzed by LC-MS/MS. B, base-peak chromatogram of three MS runs (upper) and elution profiles of peptides Ser-57, Lys-48 (LIFAGK ↓ QLEDGR), and Lys-63 (TLSDYNIQK ↓ ESTLHLVLR) (bottom) resulting from trypsin digestion of ubiquitin or ubiquitin chains for the samples of the gel band 2 on A for total cell lysates (TCL), Ni_denature, and ThUBD, respectively. Peptides derived from the gel band 2 were loaded on the column and then eluted with a 2–30% gradient of buffer B over 50 min. The peak area (AA) corresponds to the relative ion intensity of the analyzed peptide. C, comparison of ubiquitin in total cell lysates (TCL) or purified by ThUBD and Ni-NTA under denaturing condition in each gel band, reflecting quantitative data in all gel bands in the total cell lysates or ubiquitin conjugates. The ubiquitin abundance was represented by the intensity signal on MS of Ser-57 peptide extracted

lower than we expected. The current gel separation method might be responsible for the low identification of ubiquitin-modified sites, because the proteins were separated on SDS-polyacrylamide gel by molecular weight, and polyubiquitinated proteins may disperse in different molecular weight regions (Fig. 7A), resulting in a lower abundance in each gel fraction.

To further improve the identification of ubiquitinated proteins by mass spectra, we introduced an off-line high pH RP-LC separation strategy. Purified ubiquitinated proteins by ThUBD from MHCC97-H cells were digested with trypsin, and separated by high pH RP-LC (supplemental Fig. S4A). We collected 20 fractions by high pH RP-LC in a 60-min gradient (supplemental Fig. S4B), after which each fraction was analyzed by MS. One advantage of off-line RP-LC separation is that peptides are separated based on physicochemical properties, and hence the peptides from high abundance proteins (such as ubiquitin) will be aggregated in the same fraction (Fig. 7A). This prevents signals of the targeted GG peptides from being diluted and suppressed by adjacent highly abundant ions in LC-MS/MS analysis. Indeed, we found that proteins showed different distribution patterns in high pH RP-LC fractions, compared with that observed by gel separation. Taking ubiquitin as an example, it dispersed with moderate intensity in all gel fractions, while mainly being concentrated in one fraction with extremely high intensity in high pH RP-LC fractions (Fig. 7B). This also held true for the ubiquitin-modified peptide (ITITNDK↓GR) from HSPA8 (Fig. 7C), demonstrating the advantage of the high pH RP-LC separation strategy in protein identification, especially for low abundance proteins.

A total number of 7345 candidate ubiquitinated proteins were identified by the high pH RP-LC separation method, including 939 ubiquitinated proteins with ubiquitin-modified sites, which is more than that identified by gel separation (supplemental Tables 4 and 5). Up to 91.5% (1521/1663) of the identified ubiquitinated proteins by gel separation were also identified by high pH RP-LC separation (Fig. 7D), indicating the reliability of the identified candidate ubiquitinated proteins. A total of 133 identified ubiquitinated proteins with ubiquitin-modified sites was shared by the two methods, which means that more than 85.8% (806/939) of the identified proteins with ubiquitin-modified sites by the high pH RP-LC separation method were uniquely identified (Fig. 7E). These results further demonstrate the advantage of the high pH RP-LC separation method for identifying ubiquitinated proteins and ubiquitination sites. By comparing our dataset with

previously published datasets that used di-Gly antibodies (32, 33), the number of shared proteins for all three datasets was 2294 (supplemental Fig. S4C), which only represents 31.2% of those identified from this study, 51.9% of identified proteins by Kim *et al.* (32), and 55.8% of identified proteins by Wagner *et al.* (33), suggesting the high quality of identification performed in the high pH RP-LC separation method. Several previously validated ubiquitinated proteins were also identified in our study, including RHOA (58) and EPN1 (supplemental Fig. S5, A and B) (59).

A functional annotation of ubiquitinated proteins with ubiquitin-modified sites, identified from the ubiquitinated proteins enriched with ThUBD (combined gel separation and high pH RP-LC separation methods), was assigned by ingenuity pathway analysis. The identified proteins involved a large variety of biological processes including but not limited to protein translation, ubiquitination, localization, cell cycle, and apoptosis as reflected in biological processes (supplemental Fig. S5C) and signaling pathway analyses (Fig. 7F). Ingenuity pathway analysis integrates all of the identified proteins into 63 canonical pathways ($p < 0.01$), of which the top 15 pathways are shown in Fig. 7F. The canonical pathways describing the translational process as well as post-translational modification such as EIF2 signaling, mechanistic target of rapamycin signaling, and ubiquitination pathway are significantly enriched. In agreement with the high metastatic ability characteristic of MHCC97-H cells (49), a significant proportion of captured proteins was involved in adherents junctions and actin cytoskeleton signaling, suggesting that ubiquitination may play important roles in a metastasis-related process.

Ubiquitination is a conserved post-translational modification from yeast to mammalian cells. To examine the properties of amino acids surrounding ubiquitination sites, the frequencies of neighboring amino acid residues were analyzed based on the enriched ubiquitination peptides from 97H and yeast cells, respectively. We observed an obvious preference for Leu and Glu adjacent to ubiquitination sites both in mammalian and yeast cells (Fig. 7G). However, in contrast to phosphorylation motifs, we found no apparent sequence motif (supplemental Fig. S6), suggesting that the ThUBD enrichment technique used here has no bias to particular sequences surrounding the ubiquitination sites. In addition, based on the homologous search, we found 305 pairs of conserved proteins enriched from mammalian cells and yeast cells (supplemental Table 6). GO annotation (Fig. 7H) showed that these conserved proteins are involved in almost identical pathways in both mammalian and yeast cells, suggest-

from each individual gel band. Data are represented as mean and S.E. D, comparison of identified ubiquitinated proteins with ubiquitin-modified site enriched by ThUBD ($n = 362$), Ni_denature (this study, $n = 268$), and data referenced from Peng *et al.* (24) ($n = 72$). E, virtual Western blot reconstructed from proteomics data for protein ZEO1, reflecting quantitative data in all gel bands in the ubiquitin conjugates purified by ThUBD and Ni-NTA under denaturing conditions. The protein abundance was represented by the darkness and thickness of the bands, and the molecular weight information was extracted from the one-dimensional SDS gel. The arrow shows the theoretical molecular mass of ZEO1 in yeast.

FIG. 6. Large scale profiling of ubiquitinated proteins enriched by ThUBD from liver cancer cell MHCC97-H. A, comparison of MHCC97-H total cell lysates (TCL) and ubiquitinated proteins purified by ThUBD displayed on SDS-PAGE. Both samples are resolved on a 10% SDS-polyacrylamide gel, stained with Coomassie Blue, excised into 18 gel bands, digested by trypsin, and analyzed by LC-MS/MS. B, base-peak chromatogram of two MS runs (upper) and elution profiles of peptides Ser-57, Lys-48, and Lys-63 (bottom) resulting from trypsin digestion of ubiquitin or ubiquitin chains for the samples of the gel band 3 on A for total cell lysates and ubiquitinated proteins (ubc), respectively. Peptides derived from the gel band 3 were loaded on the column and then eluted in a 2–30% gradient of buffer B over 50 min. The peak area (AA) corresponds to the relative ion intensity of the analyzed peptide. C, ΔMW (the molecular weight difference between the experimental and theoretical one) histogram of the proteins identified in the total cell lysates ($n = 4518$) and ubiquitinated proteins (ubc) ($n = 2551$). The dataset was produced through SDS-PAGE separation and then analyzed by LC-MS/MS. D, comparison of ubiquitinated proteins purified by ThUBD, qUBA, and TUBE from mammalian cells.

FIG. 7. pH RP-LC separation efficiently increases the identification of ubiquitinated proteins and ubiquitination sites, compared with SDS-polyacrylamide gel separation. **A**, comparison of the characteristics of gel and pH RP-LC separation for the ubiquitin conjugates' identification through LC-MS/MS analysis. **B** and **C**, distribution of the certain peptide from ubiquitin (**B**) or HSPA8 (**C**) in different fractions separated by gel and pH RP-LC. The K# in the HSPA8 peptide stands for ubiquitin modified on the lysine. **D** and **E**, identified potential ubiquitinated proteins and ubiquitinated proteins with ubiquitin-modified sites separated by pH-RP-LC and SDS-polyacrylamide gel, respectively. All proteins are converted to UniProtKB ID for comparison. **F**, top 15 pathways integrate the ubiquitinated proteins with ubiquitin-modified sites purified by ThUBD ($n = 1125$) from MHCC97-H cells. **G**, amino acid sequence properties of ubiquitination sites enriched in MHCC97-H cells and yeast, respectively. **H**, top 10 pathways enriched by the conserved proteins ($n = 305$) purified from MHCC97-H cells and yeast.

ing that these pathways are highly conserved and ubiquitinated across different species. These results further demonstrated the conservation of ubiquitination in different species, from essential amino acid sequences to intricate pathway levels.

DISCUSSION

One of the challenges of ubiquitinome research is the enrichment of ubiquitinated proteins, which are in low abundance in cells and prone to rapid degradation by the 26S proteasome. In this study, we developed a recombination of

hybrid UBDs to effectively purify ubiquitinated proteins with no apparent biases under native conditions. Using ThUBD (2_UDA20), we identified one of the largest ubiquitinated protein datasets from yeast and mammalian cells in a single experiment to date.

In this study, we systematically compared the binding characteristics of different UBDs to ubiquitin and seven polyubiquitin chains. We found that different UBDs showed diverse binding affinity to ubiquitin and linkages, such that A20 appears to prefer Lys-63-linked chains, and UQ1 specificity binds to Lys-11-linked chains. Among the seven evaluated UBDs, DSK2 and A20 showed the highest binding affinity to ubiquitin (Fig. 1C). This is also corroborated by the calculated equilibrium dissociation constants (K_D) (Table I). To the best of our knowledge, this is the first study to identify the linkage specificity of UBDs by LC-MS/MS. This approach provides a new method to systematically evaluate the linkage specificity of UBDs under natural conditions.

In cells, many proteins carry multiple copies of UBDs or several types of UBDs. For example, protein RAD23 in yeast contains UBL (ubiquitin-like) and UBA domains at the N and C termini, respectively. It is thought that the increase of UBD copies and types in one protein might contribute to the high affinity of ubiquitin. Based on this assumption, we designed and artificially constructed hybrid UBDs, comprising two different UBDs, to increase the binding affinity to ubiquitin. The top three UBDs that showed high binding affinity to ubiquitin were chosen to construct two hybrid UBDs (*i.e.* UDA20 and UDQ2) (Fig. 1C). We found that hUBDs not only increased binding affinity to ubiquitin but also to the seven polyubiquitin chains by quantitative comparison. This might due to the different binding mechanism of UBDs to ubiquitin and linkages. For example, most UBAs contact the Ile-44-containing surface of ubiquitin (6), although A20 in Rabex-5 binds to a polar region centered on Asp-58 of ubiquitin (60). Therefore, the recombined hUBDs may utilize both binding sites to increase the binding affinity to ubiquitinated proteins. More importantly, this structure may reduce the bias to polyubiquitin chains, showing the effective binding to all seven polyubiquitin chains. With the increasing discovery of novel UBDs, the combination of these UBDs may provide better hUBDs with higher binding affinity to ubiquitin and polyubiquitin chains.

Compared with qUBA and Ni-NTA, which are two powerful approaches for the purification of ubiquitinated proteins, the ThUBD (2_UDA20) captured the largest number of candidate ubiquitinated proteins from yeast, especially polyubiquitinated proteins, demonstrating that ThUBD is a promising reagent for the global ubiquitinome research. Comparing the surrounding sequences of ubiquitination sites purified by these three methods (supplemental Fig. S6A), we found qUBA and ThUBD are more similar to each other than Ni-NTA, demonstrating that different purification methods might con-

tribute to the differences of the identified ubiquitinated proteins.

One of the advantages of using UBD in ubiquitinated proteome research is that it does not rely on ubiquitin engineering or overexpression, which are not applicable for ubiquitination profiling in mammalian cells or physiological and pathological tissues. Using ThUBD, we successfully identified 7487 endogenous ubiquitinated proteins and 1125 ubiquitinated proteins with ubiquitin modified sites from MHCC97-H liver cancer cells by two different separation methods (Fig. 7, D and E), representing the largest ubiquitinated protein dataset of mammalian cell line to date. It is worth noting that the high pH RP-LC separation method shows an apparent superiority on protein identification, especially the low abundance proteins, as well as the specific modified sites (such as ubiquitination on Lys or phosphorylation on Ser/Thr/Tyr). The ThUBD purification method also suits for quantitative ubiquitination studies, as it can be easily combined with stable isotopic labeling.

Furthermore, ThUBD could be used to study the ubiquitinated proteome of tissues or pathological specimens of human beings, which may help to identify the novel ubiquitinated proteins related to disease. This may contribute to the discovery of new biomarkers and therapy methods. In addition, ThUBD may also be a powerful tool for post-translational modification cross-talk research, such as ubiquitination and phosphorylation and ubiquitination and acetylation, to provide information on the novel regulation of proteins. This will allow us to have an even deeper understanding of post-translational modifications.

In conclusion, using ThUBD, we identified the largest ubiquitinated protein dataset of mammalian cells, demonstrating that ThUBD is a promising reagent for the global ubiquitinated proteomic research.

Acknowledgments—We thank James H. Hurley, Cecile M. Pickart, Jun Qin, and Tso-Pang Yao for generous gifts of their reagents. We also thank Baoqing Ding for critical reading and editing.

* This work was supported by Chinese National Basic Research Program Grants 2013CB911201, 2015CB910700, and 2011CB910600, National Natural Science Foundation of China Grants 31400698, 31170780, 31470809, and 31400697, National High-Tech Research and Development Program of China Grants SS2012AA020502, 2011AA02A114, 2014AA020900, and 2014AA020607, International Collaboration Program Grant 2014DFB30020, National Natural Science Foundation of Beijing Grants 5152008 and 5144027, National Mega-projects for Key Infectious Diseases Grant 2016ZX10003003, Key Projects in the National Science and Technology Pillar Program Grant 2012BAF14B00, the Unilevel 21st Century Toxicity Program Grant MA-2014-02409, and Foundation of State Key Lab of Proteomics Grant SKLPYB201404.

§ This article contains supplemental materials.

§ Both authors contributed equally to this work.

‡‡ To whom correspondence should be addressed: E-mail: junmin.peng@stjude.org, Tel./Fax: 901-595-3032; E-mail: xupingghy@gmail.com, Tel./Fax: 8610-61777113.

REFERENCES

1. Hershko, A., Ciechanover, A., and Varshavsky, A. (2000) The ubiquitin system. *Nat. Med.* **6**, 1073–1081
2. Scheffner, M., Nuber, U., and Huibregtse, J. M. (1995) Protein ubiquitination involving an E1–E2–E3 enzyme ubiquitin thioester cascade. *Nature* **373**, 81–83
3. Komander, D., Clague, M. J., and Urbé, S. (2009) Breaking the chains: structure and function of the deubiquitinases. *Nat. Rev. Mol. Cell Biol.* **10**, 550–563
4. Reyes-Turcu, F. E., and Wilkinson, K. D. (2009) Polyubiquitin binding and disassembly by deubiquitinating enzymes. *Chem. Rev.* **109**, 1495–1508
5. Dikic, I., Wakatsuki, S., and Walters, K. J. (2009) Ubiquitin-binding domains—from structures to functions. *Nat. Rev. Mol. Cell Biol.* **10**, 659–671
6. Hicke, L., Schubert, H. L., and Hill, C. P. (2005) Ubiquitin-binding domains. *Nat. Rev. Mol. Cell Biol.* **6**, 610–621
7. Kerscher, O., Felberbaum, R., and Hochstrasser, M. (2006) Modification of proteins by ubiquitin and ubiquitin-like proteins. *Annu. Rev. Cell Dev. Biol.* **22**, 159–180
8. Schwartz, A. L., and Ciechanover, A. (1999) The ubiquitin-proteasome pathway and pathogenesis of human diseases. *Annu. Rev. Med.* **50**, 57–74
9. Mukhopadhyay, D., and Riezman, H. (2007) Proteasome-independent functions of ubiquitin in endocytosis and signaling. *Science* **315**, 201–205
10. Haglund, K., and Dikic, I. (2005) Ubiquitylation and cell signaling. *EMBO J.* **24**, 3353–3359
11. Welchman, R. L., Gordon, C., and Mayer, R. J. (2005) Ubiquitin and ubiquitin-like proteins as multifunctional signals. *Nat. Rev. Mol. Cell Biol.* **6**, 599–609
12. Di Fiore, P. P., Polo, S., and Hofmann, K. (2003) When ubiquitin meets ubiquitin receptors: a signalling connection. *Nat. Rev. Mol. Cell Biol.* **4**, 491–497
13. Hicke, L. (2001) Protein regulation by monoubiquitin. *Nat. Rev. Mol. Cell Biol.* **2**, 195–201
14. Haglund, K., Sigismund, S., Polo, S., Szymkiewicz, I., Di Fiore, P. P., and Dikic, I. (2003) Multiple monoubiquitination of RTKs is sufficient for their endocytosis and degradation. *Nat. Cell Biol.* **5**, 461–466
15. Thrower, J. S., Hoffman, L., Rechsteiner, M., and Pickart, C. M. (2000) Recognition of the polyubiquitin proteolytic signal. *EMBO J.* **19**, 94–102
16. Ikeda, F., and Dikic, I. (2008) Atypical ubiquitin chains: new molecular signals. *EMBO Rep.* **9**, 536–542
17. Tokunaga, F., Sakata, S., Saeki, Y., Satomi, Y., Kirisako, T., Kamei, K., Nakagawa, T., Kato, M., Murata, S., Yamaoka, S., Yamamoto, M., Akira, S., Takao, T., Tanaka, K., and Iwai, K. (2009) Involvement of linear polyubiquitylation of NEMO in NF- κ B activation. *Nat. Cell Biol.* **11**, 123–132
18. Xu, P., Duong, D. M., Seyfried, N. T., Cheng, D., Xie, Y., Robert, J., Rush, J., Hochstrasser, M., Finley, D., and Peng, J. (2009) Quantitative proteomics reveals the function of unconventional ubiquitin chains in proteasomal degradation. *Cell* **137**, 133–145
19. Matsumoto, M. L., Wickliffe, K. E., Dong, K. C., Yu, C., Bosanac, I., Bustos, D., Phu, L., Kirkpatrick, D. S., Hymowitz, S. G., Rape, M., Kelley, R. F., and Dixit, V. M. (2010) K11-linked polyubiquitination in cell cycle control revealed by a K11 linkage-specific antibody. *Mol. Cell* **39**, 477–484
20. Huang, H., Jeon, M.-S., Liao, L., Yang, C., Elly, C., Yates, J. R., 3rd, and Liu, Y.-C. (2010) K33-linked polyubiquitination of T cell receptor- ζ regulates proteolysis-independent T cell signaling. *Immunity* **33**, 60–70
21. Yuan, W.-C., Lee, Y.-R., Lin, S.-Y., Chang, L.-Y., Tan, Y. P., Hung, C.-C., Kuo, J.-C., Liu, C.-H., Lin, M.-Y., Xu, M., Chen, Z. J., and Chen, R. H. (2014) K33-linked polyubiquitination of coronin 7 by Cul3-KLHL20 ubiquitin E3 ligase regulates protein trafficking. *Mol. Cell* **54**, 586–600
22. Kulathu, Y., and Komander, D. (2012) Atypical ubiquitylation—the unexplored world of polyubiquitin beyond Lys-48 and Lys-63 linkages. *Nat. Rev. Mol. Cell Biol.* **13**, 508–523
23. Shi, Y., Xu, P., and Qin, J. (2011) Ubiquitinated proteome: ready for global? *Mol. Cell. Proteomics* **10**, R110.006882
24. 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
25. Meierhofer, D., Wang, X., Huang, L., and Kaiser, P. (2008) Quantitative analysis of global ubiquitination in HeLa cells by mass spectrometry. *J. Proteome Res.* **7**, 4566–4576
26. Franco, M., Seyfried, N. T., Brand, A. H., Peng, J., and Mayor, U. (2011) A novel strategy to isolate ubiquitin conjugates reveals wide role for ubiquitination during neural development. *Mol. Cell. Proteomics* **10**, M110.002188
27. 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
28. Newton, K., Matsumoto, M. L., Wertz, I. E., Kirkpatrick, D. S., Lill, J. R., Tan, J., Dugger, D., Gordon, N., Sidhu, S. S., Fellouse, F. A., Komuves, L., French, D. M., Ferrando, R. E., Lam, C., Compagnon, D., et al. (2008) Ubiquitin chain editing revealed by polyubiquitin linkage-specific antibodies. *Cell* **134**, 668–678
29. Matsumoto, M., Hatakeyama, S., Oyama, K., Oda, Y., Nishimura, T., and Nakayama, K. I. (2005) Large-scale analysis of the human ubiquitin-related proteome. *Proteomics* **5**, 4145–4151
30. Vasilescu, J., Smith, J. C., Ethier, M., and Figeys, D. (2005) Proteomic analysis of ubiquitinated proteins from human MCF-7 breast cancer cells by immunoaffinity purification and mass spectrometry. *J. Proteome Res.* **4**, 2192–2200
31. Xu, G., Paige, J. S., and Jaffrey, S. R. (2010) Global analysis of lysine ubiquitination by ubiquitin remnant immunoaffinity profiling. *Nat. Biotechnol.* **28**, 868–873
32. Kim, W., Bennett, E. J., Huttlin, E. L., Guo, A., Li, J., Possemato, A., Sowa, M. E., Rad, R., Rush, J., Comb, M. J., Harper, J. W., and Gygi, S. P. (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome. *Mol. Cell* **44**, 325–340
33. Wagner, S. A., Beli, P., Weinert, B. T., Nielsen, M. L., Cox, J., Mann, M., and Choudhary, C. (2011) A proteome-wide, quantitative survey of *in vivo* ubiquitylation sites reveals widespread regulatory roles. *Mol. Cell. Proteomics* **10**, M111.013284
34. Udesi, N. D., Svinkina, T., Mertins, P., Kuhn, E., Mani, D. R., Qiao, J. W., and Carr, S. A. (2013) Refined preparation and use of anti-diglycine remnant (K- ϵ -GG) antibody enables routine quantification of 10,000s of ubiquitylation sites in single proteomics experiments. *Mol. Cell. Proteomics* **12**, 825–831
35. Sylvestersen, K. B., Young, C., and Nielsen, M. L. (2013) Advances in characterizing ubiquitylation sites by mass spectrometry. *Curr. Opin. Chem. Biol.* **17**, 49–58
36. Wagner, S. A., Beli, P., Weinert, B. T., Schölz, C., Kelstrup, C. D., Young, C., Nielsen, M. L., Olsen, J. V., Brakebusch, C., and Choudhary, C. (2012) Proteomic analyses reveal divergent ubiquitylation site patterns in murine tissues. *Mol. Cell. Proteomics* **11**, 1578–1585
37. Harper, J. W., and Schulman, B. A. (2006) Structural complexity in ubiquitin recognition. *Cell* **124**, 1133–1136
38. Randles, L., and Walters, K. J. (2011) Ubiquitin and its binding domains. *Front. Biosci.* **17**, 2140–2157
39. Raasi, S., Varadan, R., Fushman, D., and Pickart, C. M. (2005) Diverse polyubiquitin interaction properties of ubiquitin-associated domains. *Nat. Struct. Mol. Biol.* **12**, 708–714
40. Trempe, J. F., Brown, N. R., Lowe, E. D., Gordon, C., Campbell, I. D., Noble, M. E., and Endicott, J. A. (2005) Mechanism of Lys-48-linked polyubiquitin chain recognition by the Mud1 UBA domain. *EMBO J.* **24**, 3178–3189
41. Komander, D., Reyes-Turcu, F., Licchesi, J. D., Odenwaelter, P., Wilkinson, K. D., and Barford, D. (2009) Molecular discrimination of structurally equivalent Lys 63-linked and linear polyubiquitin chains. *EMBO Rep.* **10**, 466–473
42. Rahighi, S., Ikeda, F., Kawasaki, M., Akutsu, M., Suzuki, N., Kato, R., Kensche, T., Uejima, T., Bloor, S., Komander, D., Randow, F., Wakatsuki, S., and Dikic, I. (2009) Specific recognition of linear ubiquitin chains by NEMO is important for NF- κ B activation. *Cell* **136**, 1098–1109
43. Lo, Y.-C., Lin, S.-C., Rospigliosi, C. C., Conze, D. B., Wu, C.-J., Ashwell, J. D., Eliezer, D., and Wu, H. (2009) Structural basis for recognition of diubiquitins by NEMO. *Mol. Cell* **33**, 602–615
44. Layfield, R., Tooth, D., Landon, M., Dawson, S., Mayer, J., and Alban, A. (2001) Purification of poly-ubiquitinated proteins by S5a-affinity chromatography. *Proteomics* **1**, 773–777
45. Shi, Y., Chan, D. W., Jung, S. Y., Malovannaya, A., Wang, Y., and Qin, J.

- (2011) A data set of human endogenous protein ubiquitination sites. *Mol. Cell. Proteomics* **10**, M110.002089
46. Lopitz-Otsoa, F., Rodriguez-Suarez, E., Aillet, F., Casado-Vela, J., Lang, V., Matthiesen, R., Elortza, F., and Rodriguez, M. S. (2012) Integrative analysis of the ubiquitin proteome isolated using tandem ubiquitin binding entities (TUBEs). *J. Proteomics* **75**, 2998–3014
 47. Hjerpe, R., Aillet, F., Lopitz-Otsoa, F., Lang, V., England, P., and Rodriguez, M. S. (2009) Efficient protection and isolation of ubiquitylated proteins using tandem ubiquitin-binding entities. *EMBO Rep.* **10**, 1250–1258
 48. Spence, J., Gali, R. R., Dittmar, G., Sherman, F., Karin, M., and Finley, D. (2000) Cell cycle-regulated modification of the ribosome by a variant multiubiquitin chain. *Cell* **102**, 67–76
 49. Li, Y., Tang, Z.-Y., Ye, S.-L., Liu, Y.-K., Chen, J., Xue, Q., Chen, J., Gao, D.-M., and Bao, W.-H. (2001) Establishment of cell clones with different metastatic potential from the metastatic hepatocellular carcinoma cell line MHCC97. *World J. Gastroenterol.* **7**, 630–636
 50. Ding, C., Jiang, J., Wei, J., Liu, W., Zhang, W., Liu, M., Fu, T., Lu, T., Song, L., Ying, W., Chang, C., Zhang, Y., Ma, J., Wei, L., Malovannaya, A., et al. (2013) A fast workflow for identification and quantification of proteomes. *Mol. Cell. Proteomics* **12**, 2370–2380
 51. Xu, P., Duong, D. M., and Peng, J. (2009) Systematical optimization of reverse-phase chromatography for shotgun proteomics. *J. Proteome Res.* **8**, 3944–3950
 52. Zhai, L., Chang, C., Li, N., Duong, D. M., Chen, H., Deng, Z., Yang, J., Hong, X., Zhu, Y., and Xu, P. (2013) Systematic research on the pretreatment of peptides for quantitative proteomics using a C18 microcolumn. *Proteomics* **13**, 2229–2237
 53. Ping, L., Zhang, H., Zhai, L., Dammer, E. B., Duong, D. M., Li, N., Yan, Z., Wu, J., and Xu, P. (2013) Quantitative proteomics reveals significant changes in cell shape and an energy shift after IPTG induction via an optimized SILAC approach for *Escherichia coli*. *J. Proteome Res.* **12**, 5978–5988
 54. Lausted, C., Hu, Z., and Hood, L. (2008) Quantitative serum proteomics from surface plasmon resonance imaging. *Mol. Cell. Proteomics* **7**, 2464–2474
 55. Seyfried, N. T., Xu, P., Duong, D. M., Cheng, D., Hanfelt, J., and Peng, J. (2008) Systematic approach for validating the ubiquitinated proteome. *Anal. Chem.* **80**, 4161–4169
 56. Kalesnikoff, J., Rios, E. J., Chen, C.-C., Alejandro Barbieri, M. A., Tsai, M., Tam, S.-Y., and Galli, S. J. (2007) Roles of RabGEF1/Rabex-5 domains in regulating FcεRI surface expression and FcεRI-dependent responses in mast cells. *Blood* **109**, 5308–5317
 57. Tong, Z., Kim, M.-S., Pandey, A., and Espenshade, P. J. (2014) Identification of candidate substrates for the Golgi Tul1 E3 ligase using quantitative DiGly proteomics in yeast. *Mol. Cell. Proteomics* **13**, 2871–2882
 58. Wang, H.-R., Zhang, Y., Ozdamar, B., Ogunjimi, A. A., Alexandrova, E., Thomsen, G. H., and Wrana, J. L. (2003) Regulation of cell polarity and protrusion formation by targeting RhoA for degradation. *Science* **302**, 1775–1779
 59. Chen, H., Polo, S., Di Fiore, P. P., and De Camilli, P. V. (2003) Rapid Ca²⁺-dependent decrease of protein ubiquitination at synapses. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 14908–14913
 60. Lee, S., Tsai, Y. C., Mattera, R., Smith, W. J., Kostelansky, M. S., Weissman, A. M., Bonifacio, J. S., and Hurley, J. H. (2006) Structural basis for ubiquitin recognition and autoubiquitination by Rabex-5. *Nat. Struct. Mol. Biol.* **13**, 264–271
 61. Ohno, A., Jee, J. G., Fujiwara, K., Tenno, T., Goda, N., Tochio, H., Kobayashi, H., Hiroaki, H., and Shirakawa, M. (2005) Structure of the UBA domain of Dsk2p in complex with ubiquitin: molecular determinants for ubiquitin recognition. *Structure* **13**, 521–532
 62. Zhang, D., Raasi, S., and Fushman, D. J. (2008) *Mol. Biol.* **377**, 162–180
 63. Bennett, E. J., Shaler, T. A., Woodman, B., Ryu, K. Y., Zaitseva, T. S., Becker, C. H., Bates, G. P., Schulman, H., and Kopito, R. R. (2007) Global changes to the ubiquitin system in Huntington's disease. *Nature* **448**, 704–708
 64. Penengo, L., Mapelli, M., Murachelli, A. G., Confalonieri, S., Magri, L., Musacchio, A., Di Fiore, P. P., Polo, S., and Schneider, T. R. (2006) Crystal structure of the ubiquitin binding domains of rabex-5 reveals two modes of interaction with ubiquitin. *Cell* **124**, 1183–1195
 65. Boyault, C., Gilquin, B., Zhang, Y., Rybin, V., Garman, E., Meyer-Klaucke, W., Matthias, P., Müller, C. W., and Khochbin, S. (2006) HDAC6-p97/VCP controlled polyubiquitin chain turnover. *EMBO J.* **25**, 3357–3366