Supplementary Figure 1

MAPP Immuno-purification protocol robustly captures proteasome complexes

(a) Proteasomes were immunoprecipitated from HEK293 cells using a dose curve of an antibody targeting PSMA1. The precipitate was separated on SDS-PAGE and analyzed by western blot probed with a pan α-20s proteasome antibody. Shown is representative of two individual dose curves. (b) PSMA1 was immunoprecipitated according to the MAPP protocol and the immune complex was separated by SDS-PAGE and analyzed by western blot probed with antibodies against the immunoprecipitated PSMA1, as well as co-precipitated 20S subunits of the proteasome PSMA1-7 PSMB5 and PSMB7 and against the 19S subunits PSMD6 and PSMD11. Mock: empty beads; TCE: Total cell extract. (c) MAPP conditions were calibrated to evaluate non-specific binding of cellular peptides to empty beads (mock), versus α-PSMA1-conjugated beads (MAPP), or without addition of protease inhibitors (- Protease Inhibitors). The number of unique peptides significantly decreased when using empty beads and significantly increased when no protease inhibitors was used. (d) The intensity and identity of peptides identified from an immunoprecipitation using α-PSMA1 antibody and α-PSMA4, an antibody against a different subunit of the proteasome, correlated with one another (R^2 = 0.738). n = 4375 peptides. (e) A representative pull down of GAPDH was separated by SDS-PAGE and analyzed by western blot. TCE: total cell extract.
Peptides captured and identified by MAPP are produced in a proteasome dependent manner.

(a) Extracted ion chromatograms of three peptides identified by MAPP from cells under basal conditions or following treatment with the proteasome inhibitor epoxomicin (1µM, 4 hours). M+1 and M+2 represent first and second isotopes of the precursor ion respectively. (b) The fold changes in intensity of peptides identified by MAPP between untreated cells (UT) and cells treated with Velcade (50nM, 4 hours) were ranked and displayed as a bar diagram. Those with a 2-fold change or greater were marked in orange, and those that decreased by 2-fold or more were marked in blue. Of the 3505 peptides identified by MAPP, 2463 peptides or 70% of the total were reduced by at least 2 fold in intensity. (c) The abundance of PSMA1 in the MAPP protein fraction normalized to the mean abundance (Log 2 normalized ratio of LFQ intensities; Normalization was done to the median of the isotype control; Bars: mean ± range of the triplicates). (d) Protein abundance was determined by standard proteomics in untreated cells (UT) or following cycloheximide treatment (100µg/ml, 2 hours). Bars: mean ± standard deviation.
Peptides from known proteasomal substrates are detected by MAPP. 

(a) Cells were transfected with ZsGreen fused to the proteasomal degradation domain of mouse ornithine decarboxylase, and were either treated with the proteasome inhibitor epoxomicin (1µM, 4 hours) or left untreated, and analyzed by MAPP. As control, untransfected cells were analyzed. Shown are extracted fragment ion chromatograms of the parallel reaction monitoring mass spectrometry for two peptides which were ZsGreen degradation products. The peaks were not detected in cells treated with epoxomicin or in untransfected cells (UT).
Supplementary Figure 4

Bioinformatics processing of MAPP identified peptides.

(a) Bioinformatics cutoff of raw MAPP peptides allowed peptides that had at least two valid intensities out of three independent biological replicates ("replicates filtering"), and included razor peptides which belong to a unique MaxQuant "protein group". To eliminate background peptides, we considered only peptides that were detected at least 2-fold higher in the PSMA1 immunoprecipitate than in the mock beads control ("beads filtering"). 440 peptides were removed at this stage. The shown filtering process was performed on HEK293 in biological triplicates. (b) Twenty-two percent of the proteins identified by MAPP were not detected in parallel analysis of the sample using standard proteomics.
Proteins specifically identified by MAPP differ in biochemical properties from those identified by standard proteomics. (a) Identified proteins were subdivided into three categories: those identified solely by MAPP (denoted as [A]), those identified both in standard proteomics and MAPP (denoted as shared [B]) and those identified solely through standard proteomics (denoted [C]). The expected average protein abundance and turnover are listed for each group. (b) Protein turnover rate calculations adapted from Schwanhäusser et al. where $P$ indicates protein abundance and $\delta$ is the protein degradation rate. The amount of protein degraded per hour, or Protein Turnover, is the product of $\delta$ and $P$. (c) Protein abundance inferred from the deep proteomics of HEK293 cells deposited in the MaxQB database were applied to the three groups described in A ($n = 65$ [A], 779 [B], 3250 [C] proteins). (d) The protein turnover calculated from a dataset generated by Larance et al. was applied to the three groups described in a using calculations described in b ($n = 32$ [A], 753 [B], 2509 [C] proteins). (e) Endogenous proteasome-cleaved peptides demonstrated a varied distribution along the sequence of their source protein, with higher peptide frequencies of the N- and C-termini, while tryptic-digestion showed a uniform frequency (protein length was normalized to 100%). (f) Correlation of amino acid composition within peptides identified by either MAPP or standard proteomics analysis (39,700 and 330,127 peptides, respectively).

Perturbations of cellular steady state change MAPP-identified proteins and intensities.

HEK293 cells were treated with TNFα + IFNγ (20ng/ml, 24 hours, denoted T+I), or left untreated (denoted UT) and analyzed using both MAPP and standard proteomics. (a) The proteins identified by both approaches were annotated into 23 different protein classes (based on Panther annotations). Similar percentages of protein classes were classified to the proteins identified by either MAPP or standard proteomics analyses. (b) The percent of MAPP-identified proteins in each protein class (as annotated in Panther) which were also identified by standard proteomics. A red line is indicated at the 70% threshold used. (c) The percent of proteins in each class that differed by 2-fold or more in intensity between cells treated with TNFα + IFNγ to untreated cells was determined. In most of the protein classes, the percentage of the proteins that changed in intensity between the examined conditions was greater for proteins identified by MAPP in comparison to standard proteomics. (d) The ranked fold change between cells treated with TNFα + IFNγ (T+I) and untreated cells (UT) in MAPP (top, n = 1172) or standard proteomics (bottom, n = 5009). (e+f) HEK293 cells were treated with TNFα + IFNγ (20ng/ml, 24 hours, denoted T+I), or left untreated (denoted UT) and analyzed using both MAPP and standard proteomics. Proteins that are annotated in the CORUM database to be part of the spliceosome complex (f) ribosome complex are marked in green or blue, respectively. (g) Proteins annotated to be part of the ubiquitin proteasome system (UB components) in the KEGG database are marked in teal.

Supplementary Figure 7

MAPP highlights changes in the proteolytic landscape in response to pro-inflammatory cytokine stimulation.

(a) Schematic of experiment. HEK293 cells were grown for 42 hours after seeding and then stimulated with TNFα + IFNγ for the time indicated. All cells were collected 48 hours after seeding. (b-f) HEK293 cells were treated with TNFα + IFNγ (20ng/ml, T+I) for 1, 2 or 6 hours; or left untreated (UT) and analyzed using both MAPP and standard proteomics. (b) Directed clustering of MAPP-identified protein intensities after stimulation of cells with TNFα + IFNγ for 1, 2, or 6 hours normalized to 0 hours. Only proteins which changed in intensity by at least 2-fold in at least one time point are displayed (city-block distance function on median values distance matrix, Normalized Log2 transformed intensity ratios). (c) Directed clustering of the changes in intensity of proteins identified by MAPP (Log 2 transformed ratio of LFQ intensities at 1 or 2 hours of treatment normalized to 0 hours) reveals four groups of proteins that either increase or decrease over two hours of treatment with TNFα + IFNγ (T+I). (d) The abundance of selected proteins from these four clusters, known to be involved in inflammatory signals1-7, as identified by standard proteomics and MAPP (left; Log2 ratio of standard proteomics LFQ intensities normalized to 0h; right, Log 2 normalized ratio of MAPP LFQ intensities after 1, 2 or 6 hours of TNFα + IFNγ to 0 hours). (e) Examples of the intensities of two of the proteins listed in (d), ZFAND5 and CNBP, as identified by either standard proteomics or MAPP (Log2 Normalized LFQ intensity ratios between the indicated time point and 0h). (f) A network of the 127 proteins found in the four clusters in (d). Proteins are organized by functional groups, node color indicates the highest differential ratio of abundance between MAPP after 1 or 2 hours of TNFα + IFNγ stimulation (T+I) to 0 hours (UT) (Log 2 normalized LFQ intensity ratio). Node border thickness indicates abundance of the given proteins as detected by standard proteomics (Log 10 normalized LFQ intensity). Interactions are indicated by lines between nodes (STRING interaction confidence score is indicated by line width).

4. Lee, E. et al. CNBP acts as a key transcriptional regulator of sustained expression of interleukin-6. Nucleic Acids Res. 45,
Identification of patient-specific proteolytic signatures of T-cell exhaustion in the PD-1 immune-checkpoint.

(a) IL-7Rα and PD-1 levels of CD8+ T-cells purified from human blood were evaluated by flow cytometry analysis after activation with CD3/CD28 and following 7 days of co-culture with a Melanoma-derived cell line. Cells were first gated for lymphocytes followed by a single cell gate (grey box). n = 7700 cells (PD1), 12,249 cells (CD127).

(b) Pearson correlation coefficients between MAPP proteins across the examined time points (lower triangle, at exhaustion, T0, and 30 minutes or 6 hours after PD-1 blockade), and across individuals (upper triangle, P1, P2, P3). n = 1293 proteins for each pairwise comparison.

(c) Clustering the intensities of MAPP proteins (which were detected in at least two individuals) across the temporal response to PD-1 blockade (0, 30', 6h) revealed common proteolytic signatures (P1, P2, P3); Euclidean distance function, log10 intensity values)

(d) Volcano plot representation of the relative degradation intensities of proteins detected by MAPP in cytotoxic CD8+ T cells, either at time zero (T0, exhaustion) and 6 hours after reactivation by PD-1 blockade. Median values represent at least 2 individuals, and color codes for the degree of fold change between T0 and 6 hours; Two sided t test, no correction for multiple comparisons.

(e) Schematic representation of PD-1 related signaling pathways, adapted from the T-cell receptor signaling (KEGG) pathway. Proteins that were detected by MAPP as actively degraded, in either of the time points, are labeled in a blue oval including components of T cell activation, cytoskeleton organization and ubiquitination

(f) Changes detected in protein abundance and degradation signatures of selected proteins, related to the PD-1 signaling pathways, at either exhaustion, or after 30 minutes or 6 hours of PD-1 blockade. Values represent mean intensities (log 10) error bars indicate standard deviation.

Supplementary Figure 9

MAPP separates PBMC samples from SLE and Healthy Patients more robustly than Standard Proteomics.

(a) Principal component analysis of the 250 most intense proteins from PBMCs of healthy individuals or SLE patients, as identified by standard proteomics. (b) 149 proteins were detected both by MAPP and standard proteomics. Principal component analysis of the protein intensities of these 149 proteins from PBMCs of healthy individuals or SLE patients, as identified by standard proteomics. (c) The average distance between all samples within a group (within) and each intergroup sample distance (between) was calculated for principle component analysis performed on a randomly selected group of 249 proteins from the proteomics identified set. This was performed 10,000 times and the distribution of the between/within ratio is plotted. The between within ratio of the 24 protein intensities identified by MAPP is significantly higher (p < 0.001), signifying the dataset characterizes the samples better than the corresponding standard proteomics sets.
Examples of proteins which differ in MAPP of SLE Patients and Healthy Individuals.

(a + b) Scatterplot of protein intensities in the proteomics and MAPP datasets for (a) healthy individuals, n = 148 proteins or (b) SLE patients, n=143 proteins. The proteins which differ by 2-fold or greater in MAPP of SLE and healthy PBMCs are marked in blue/red. (c) The total protein abundance of histones as measured by standard proteomics, was not changed between healthy and SLE PBMCs. n = 26 (histones), 5067 (other). Line indicates mean. (d+e) The peptides identified in MAPP of PBMCs of healthy individuals (top) or SLE patients (bottom) which map to (c) HIST1H4A or (d) CHTF8 (residues 330-430), examples of histones or other nuclear proteins, respectively. Color scale indicates peptide intensity (Log10 transformed intensities).
Supplementary Figure 11

Different pathways are expressed in the proteomes of SLE Patients and Healthy Individuals.

Gene Set Enrichment Analysis (GSEA)\(^1\) was performed using the protein abundances in the proteomes of SLE Patients and Healthy Individuals as determined by standard proteomics. The HALLMARK pathways which passed the default GSEA FDR cutoff (25%) are presented. The normalized enrichment score indicates if the pathway is enriched in SLE patients (red bars, > 0) or healthy individuals (blue bars, < 0).