Chemical Science

ROYAL SOCIETY OF CHEMISTRY

EDGE ARTICLE

View Article Online
View Journal | View Issue

Cite this: *Chem. Sci.*, 2025, **16**, 16979

All publication charges for this article have been paid for by the Royal Society of Chemistry

SpectraSage unveils specific proteolytic patterns of 20S on mono-ubiquitylated Tau proteoforms involved in neurodegeneration

Gabriele Antonio Zingale, *[a] Irene Pandino, [a] Daniele Trivellato,[b] Dario Cavaterra,[c] Francesca Munari, [b] Giuseppe Grasso, [d] Peter A. Bell,[e] Francesco Oddone,[a] Alessio Bocedi,[c] Massimiliano Coletta,[a] Michael Assfalg, [b] Mariapina D'Onofrio, [b] Grazia Raffaella Tundo[af] and Diego Sbardella*[a]

The ubiquitin proteasome system is a critical regulator of proteostasis and shows altered activity and composition in neurodegenerative diseases affecting both the brain (*e.g.*, Alzheimer's disease) and the retina/optic nerve (*e.g.*, age-related macular degeneration, glaucoma). A common feature of neurodegeneration is the progressive accumulation of amyloidogenic proteins such as beta-amyloid and tau protein (MAPT gene). There is compelling evidence that the aggregation propensity of tau protein is regulated by post-synthetic modifications including phosphorylation and ubiquitylation. These alterations are gaining increasing pathological relevance not only for brain tauopathies but also for the retinal/optic nerve degenerative diseases. In this regard, site-specific mono-ubiquitylated (Ub) tau proteoforms, have been recently identified in neurodegenerative brains. In this work, the cleavage patterns of the uncapped 20S proteasome acting on mono-Ub regio-isomers of tauK18, which covers the 4RD domain, have been unveiled by using SpectraSage, a novel proteomics software conceived for the MS1 identification of complex branched peptide and here introduced for the first time. Ub position was found to affect regio-isomers susceptibility to proteolysis and unexpectedly long Ub-tauK18 branched peptides have been identified, proving distinct catalytic preferences. These findings show that the 20S digests mono-Ub proteins through specific enzymatic mechanisms and the implications of the latter on neurodegeneration are discussed.

Received 10th June 2025
Accepted 16th August 2025

DOI: 10.1039/d5sc04240b

rsc.li/chemical-science

Introduction

The proteostasis network (PN) defines the vast range of pathways and factors deputed to protein synthesis, folding, trafficking and degradation, thereby regulating intracellular proteome homeostasis. A balanced PN is critical for post-mitotic cells (*e.g.*, neurons and retinal ganglion cells) and nervous tissues, which are indeed more vulnerable to proteotoxic insults. Nonetheless, PN dysregulation and proteotoxicity (*e.g.*, protein aggregates, plaques) are hallmarks of neurodegenerative diseases, such as Alzheimer's (AD), Parkinson's diseases (PD), as well as several disorders affecting the retina and the optic nerve such as age-related macular degeneration (AMD) and glaucoma.[1–3]

Within the PN framework, degradation and turnover of proteins is primarily regulated by a major intracellular proteolytic pathway called Ubiquitin Proteasome System (UPS).[4–8]

The UPS operates through two consecutive steps: (a) ubiquitin (Ub) conjugation to a substrate protein *via* an E1–E2–E3 enzymatic cascade; (b) degradation of the ubiquitylated substrate by the 26S proteasome.[9,10]

Ubiquitylation is a major post-translational modification (PTM) that shows a diversified heterogeneity. Proteins can be decorated with mono-ubiquitylation (one Ub conjugated to a lysine [Lys, K] residue) and poly-ubiquitylation (a chain of two or more Ub molecules). Moreover, poly-Ub chains show diverse topologies (*i.e.*, M1, K6, K11, K27, K29, K33, K48, K63) determined by the amino acid residue of the first Ub involved in the isopeptide bond with the next Ub along the chain.[9,11]

The Ub code dictates protein fate, with each distinct Ub configuration conferring unique biological outcomes.[11,12] In

[a]*IRCCS Fondazione Bietti, Rome, Italy. E-mail: gabriele.zingale@fondazionebietti.it; diego.sbardella@fondazionebietti.it*

[b]*Department of Biotechnology, University of Verona, Strada Le Grazie 15, 37134 Verona, Italy*

[c]*Department of Chemical Sciences and Technologies, University of Rome Tor Vergata, 00133 Rome, Italy*

[d]*Chemical Sciences Department, University of Catania, Viale Andrea Doria 6, 95125, Catania, Italy*

[e]*Department of Oral Biological and Medical Sciences, University of British Columbia, Vancouver, Canada*

[f]*Department of Clinical Sciences and Translational Medicine, University of Rome Tor Vergata, 00133 Rome, Italy*

this regard, efficient proteasomal degradation typically requires at least four K48-linked Ub moieties.[12–14]

The 26S holoenzyme comprises a 20S catalytic core particle and a 19S regulatory particle which recognizes the poly-Ub substrate and performs its ATP-dependent unfolding and translocation into the 20S for degradation.[13–16] Indeed, the 20S is composed of four stacked rings ($\alpha_7\beta_7\beta_7\alpha_7$) surrounding a central narrow catalytic channel housing three proteolytic specificities: chymotrypsin-like (β5), trypsin-like (β2) and caspase-like (β1).[17–19]

Several studies have suggested that the 20S exists as a free uncapped particle and that it may have specific catalytic activities and biological roles.[20–23] The 20S has been shown to degrade *in vitro* damaged and oxidized proteins as well as intrinsically disordered proteins (IDP).[24,25] Moreover, the 20S was shown to degrade amyloidogenic proteins, including full-length tau *in vitro* and in cell,[26,27] suggesting it plays crucial role in protein quality control regardless of ubiquitylation.[20,24,25]

A recent study has introduced the perspective that the human 20S cleaves a model mono-Ub substrate (*i.e.*, cyclin B1) *in vitro*, generating branched fragments stretching across the iso-peptide bond and covering up to four C-terminal residues of Ub (*i.e.*, -GG, -GGR, -GGRL).[22]

Since this catalytic property may introduce novel biological insights into the mechanisms of targeted proteolysis, herein we explored further this mechanism taking advantage of synthetic and semi-synthetic mono-Ub regio-isomers of the four-repeat domain of tau protein (MAPT gene), also known as tauK18. Altered turnover, dysregulated pattern of post-synthetic modifications and accumulation/aggregation propensities of tau protein show a prominent pathological role in different neurodegenerative conditions. They include primary and secondary tauopathies (*e.g.*, Pick's disease, AD among the others), but also neurodegenerative processes of the retina and the optic nerve, such as AMD and glaucoma where retinal deposits of Ub-tau were found.[2,26–28] The Ub-tau species (referred to as proteoforms throughout the manuscript) here investigated were tracked in post-mortem brain biopsies of subjects with different tauopathies and were reported to shape the conformational diversity of proteotoxic aggregates through a disease-specific pattern.[28] Accordingly, the mono-Ub regio-isomer were described to modulate their assembly into filamentous aggregates and liquid condensates clarifying the impact of site-specific mono-Ub on structural and aggregating properties of tau.[29–32]

Herein, we performed biochemical and tailored proteomic studies which allowed us to discover that the uncapped 20S degrades Ub along with the substrate, providing mechanistic insights on this process, namely: (i) Ub is required for tauK18 catalysis and its position affects the susceptibility of tauK18

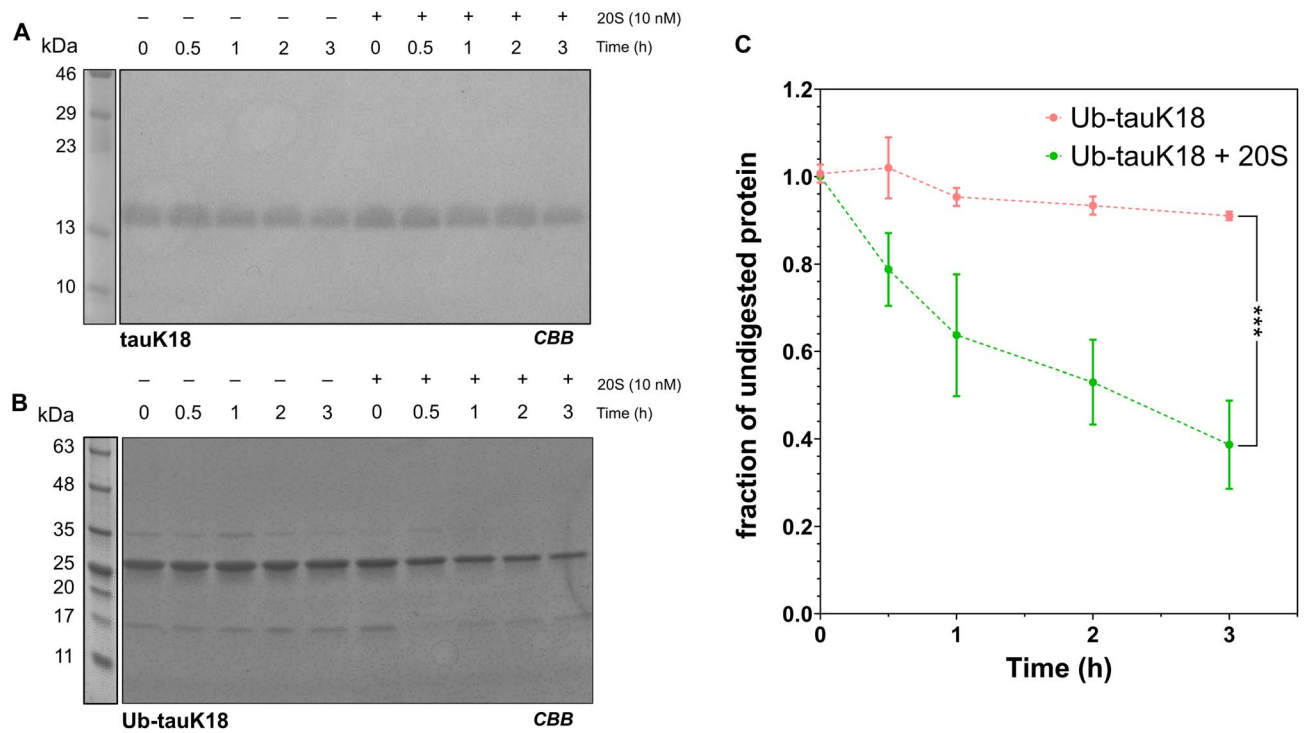

Fig. 1 (A) TauK18 (5 μM) and (B) Ub-tauK18 (5 μM) were incubated with 10 nM h20S in Tris–HCl 25 mM, pH 7.5, 37 °C. The digestion of tau species was monitored over 3 h, until linearity was observed. Proteins harvested at each time-point were heat-denatured, reduced, and separated by SDS-PAGE. Proteins were stained by CBB. The very low h20S concentration used precluded its staining. (C) Band intensity was then calculated by ImageJ software and the rate of substrate cleavage, that is expressed as relative ratio of nmol of substrate left undigested, calculated by subtracting the intensity of the band to that obtained at time 0 of analysis. A nominal value 1 was attributed to band intensity at time 0. A representative SDS-PAGE is shown. Data at each time point show the mean ± SD ($n = 3$ values) computed across independent experiments. ***$p \leq 0.001$ computed by lmstats package in R.

substrates to the h20S digestion; (ii) digestion of mono-Ub substrates generates Ub remnants on branched peptide products (hereafter referred to as Ub-branched fragments) much longer than initially thought and unidentifiable by current MS/MS approaches; (iii) the h20S has aminoacidic preferences different for tauK18 and Ub.

SpectraSage, a proteomic software that works at MS1 level and here introduced for the first time, was central to the realization of this study. It applies to enzymological studies (enzyme:substrate) *in vitro*, overcoming the limitations for the identification of complex, non-tryptic cross-linked peptides by conventional MS/MS analysis.

Moreover, we validated SpectraSage performance using a middle-down MS/MS workflow based on Tandem Mass Tag (TMT) proZero labelling of neo N-termini generated by the h20S digestion of Ub-tauK18 proteoforms.

Results

Enzymatic Ub-tauK18, but not tauK18, is a h20S substrate *in vitro*

We first set up an assay to determine the susceptibility of tauK18 and Ub-tauK18 preparations (5 μM) to digestion by the h20S proteasome (10 nM). The substrate concentration used throughout the study was a compromise for visualizing the proteins and preventing their concentration-dependent tendency to form higher molecular weight species *in vitro*.

Therefore, a time-course assay was carried out over 0.5 h, 1 h, 2 h and 3 h of incubation. As internal control, the same proteins were left in the absence of h20S (under the same experimental conditions) or fed to a h20S pre-incubated with epoxomicin, an irreversible inhibitor of chymotrypsin-like activity ($IC_{50} \approx 4$ nM).[33] Proteins at each time point were resolved by SDS-PAGE (CBB staining) (Fig. 1). The SDS-PAGE pattern indicated that h20S did not digest tauK18, since the band intensity (~15 kDa) over time was comparable between the presence and the absence of proteasome (Fig. 1A).

Conversely, the electrophoretic pattern of Ub-tauK18 provided a different outcome (Fig. 1B).

In the presence of h20S, a progressive drop over time of the intensity of the band corresponding to Ub-tauK18 (~25 kDa) was documented. This decrease was slowed down by epox (data shown in uncropped gel, see SI Fig. S4), suggesting that, unlike tauK18, Ub-tauK18 was a h20S substrate (Fig. 1C). A closer inspection of the SDS-PAGE pattern highlighted the absence of h20S digestion fragment. The lack of intermediate species with MW > 8.5 kDa (Ub MW) indeed suggested that Ub was proteolyzed along with tauK18.

Based on this evidence, we performed a pilot MS search of Ub fragments generated by the h20S. Effectively, along with several tauK18 fragments, we detected three distinct Ub proteolytic fragments covering the middle portion of Ub sequence (*i.e.* $_{2}$QIFVKTL$_{8}$, $_{11}$KTITLEVEPSD$_{21}$, $_{22}$TIENVKAKIQD$_{32}$ with XCorr = 0.98, 0.99 and 1.41, respectively) that were identified also by Sahu and co-workers.[22]

Site-specific ubiquitylation of tauK18 affects proteolytic digestion by h20S

The unexpected finding that Ub was necessary for digestion stimulated the interest in testing how Ub position impacts the susceptibility of tauK18 to h20S proteolysis. To this aim, tauK18 regio-isomers ubiquitylated at site-specific positions by disulfide coupling chemistry (*i.e.* Ub-tau254, Ub-tau311 and Ub-tau353) were enrolled in the study.[26] These Ub sites were among those of the Ub-tauK18 preparation and reproduce Ub-proteoforms identified in post-mortem brain biopsies of subjects with tauopathies.[28] To check out whether Cys residues contributed to tauK18 resistance to h20S proteolysis, a mutant called tauΔC, wherein endogenous Cys residues were replaced by Ala (C291A/C322A) was enrolled in the analysis. All these proteins (5 μM) were incubated with h20S (10 nM) over 3 h of incubation, replicating the experimental workflow discussed for tauK18 and Ub-tauK18.

As observed for tauK18, the SDS-PAGE study indicated that tauΔC did not behave as a h20S substrate (SI Fig. S4), suggesting that Cys removal had no effects on tauK18 behaviour in solution in accordance with previous physico-chemical studies.[29–31,34] Conversely, SDS-PAGE data reported in Fig. 2 highlighted a significant influence of Ub position along the tauK18 sequence on the rate of digestion of the individual proteoforms by the h20S.

Ub-tau254 turned out to be the h20S preferential substrate, exhibiting, compared to time 0 and the absence of h20S, a remarkable decrease of band intensity after 3 hours of digestion (60% of protein digested on average).

Ub-tau353 displayed a less pronounced but significant reduction over time in the presence of h20S (around 25% digested after 3 hours). By contrast, Ub-tau311 showed a low rate of degradation in the presence of h20S, suggesting that this proteoform was characterized by a greater resistance to h20S-mediated proteolysis, at least under the experimental conditions tested (Fig. 2C). All the regio-isomers proved stability in solution in the absence of h20S, whereas epox slowed down digestion (this can be appreciated for Ub-tau254, and Ub-tau353), envisaging that, like for Ub-tauK18, catalytic subunits others than the chymotrypsin-like were engaged in proteolysis (Fig. 2B). In accordance with SDS-PAGE studies on Ub-tauK18 no accumulation of proteolytic fragments was detectable in the low MW range of the gel (Fig. 2B).

Development of SpectraSage a software for enzymological studies by top-down and middle-down proteomics

The results commented above stimulated the interest in resolving the repertoire of h20S proteolytic fragments. Based on previous findings by Sahu and co-workers, to accomplish the identification of the repertoire of linear and branched fragments, which is a challenging topic in MS, especially for non-tryptic digests, we here developed SpectraSage, a proprietary software designed to identify and quantify peptides resolved at MS1 level. SpectraSage is here introduced for the first time.

A full description of the software principles is provided in the SI appendix, along with its validation using a calibrating

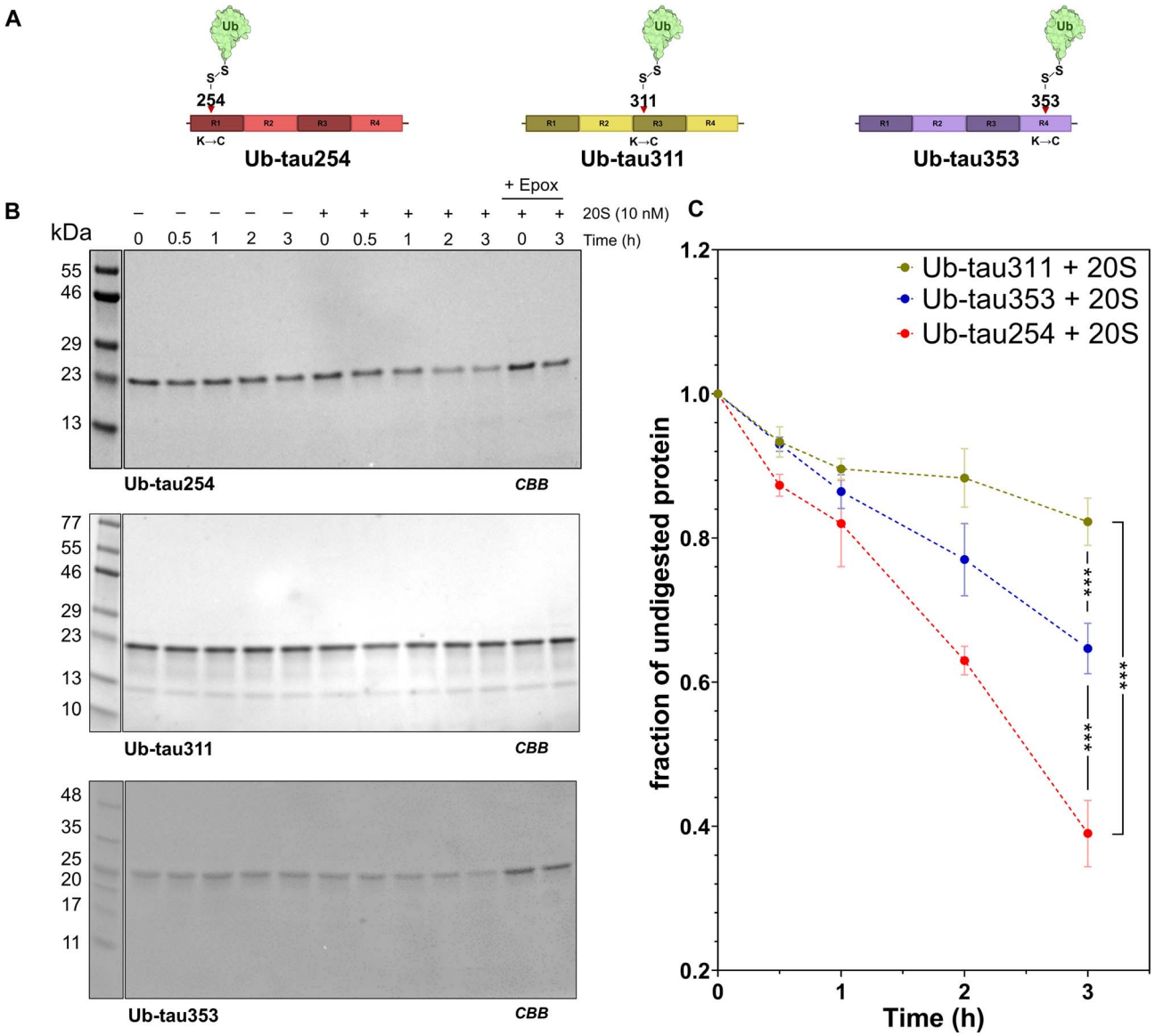

Fig. 2 (A) Ub-tau species (−254, −311, −353) (5 μM) (B) were incubated with 10 nM h20S in 25 mM Tris–HCl, pH 7.5, 37 °C. The digestion of tau species was monitored over 3 h, until linearity was observed. Proteins harvested at each time-point were heat-denatured, separated by SDS-PAGE and stained by CBB. (C) Band intensity was then calculated by ImageJ software and the rate of substrate cleavage, that is expressed as relative ratio of nmol of substrate left undigested, calculated by subtracting the intensity of the band at the specific time-point in the absence of h20S to that in the presence of h20S. A nominal value 1 was attributed to band intensity at time 0. A representative SDS-PAGE is shown. Data at each time point show the mean ± SD ($n = 3$ values) computed across independent experiments. ***$p \leq 0.001$ significantly different slopes, linear model computed by lmstats package in R.

solution of cytochrome *c* digest (supplied by Thermofisher Scientific). A summary schematic is provided in Fig. 3.

It is important to underline that the approach here proposed is tailored for digestion assays *in vitro* and universally applicable to any combination of protein/enzyme of interest. It is worth pointing out that a high degree of purity of both preparations is required to maximize the confidence of the output.

Briefly, SpectraSage operates by generating *in silico* and matching a customized protein-specific list of theoretical proteolytic fragments with the list of potential peptides effectively identified during the LC-MS run and deconvoluted by a standard algorithm (*i.e.*, Xtract). The list of theoretical proteolytic fragments generated *in silico* includes all the possible combinations of peptides, either cross-linked or not, originating from two peptides/proteins joined through a covalent bond. Thereafter, the list of identified proteolytic fragments across a spectrum is matched to the theoretical ones. Coupled with high resolution instruments, SpectraSage allows for discrimination between peptides, either branched or not, with ambiguous masses. Whilst peptides with ambiguous masses are shown in the software output, greatest confidence should be given to unambiguous peptides.

Herein the following results are based only on peptides with unambiguous masses. Moreover, regardless of ambiguity, mass

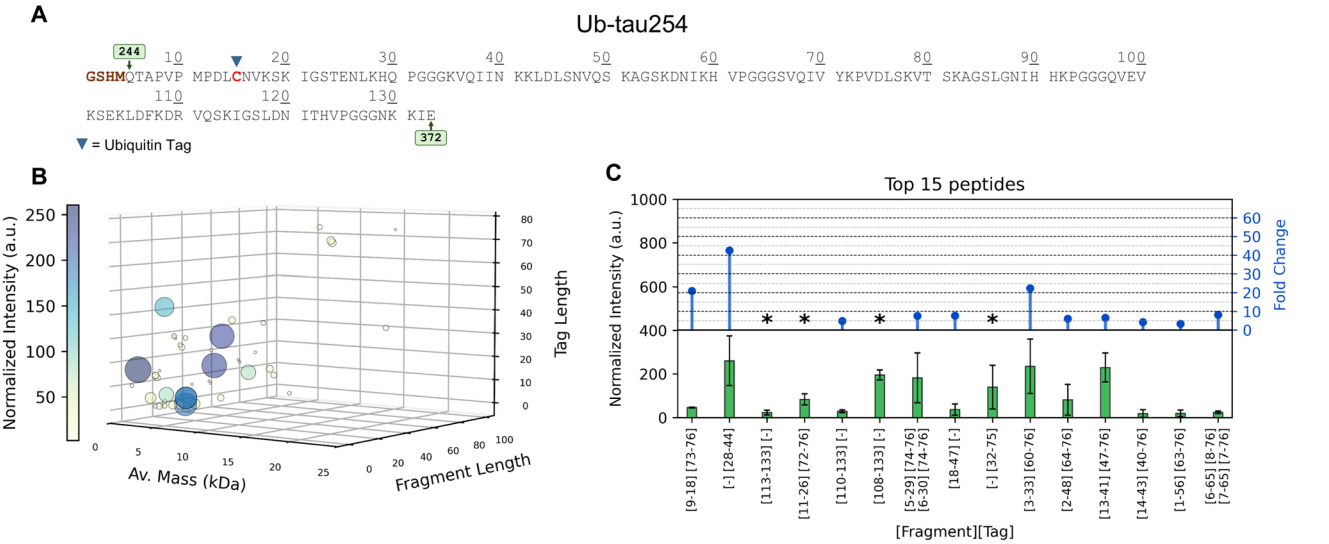

Fig. 3 SpectraSage output panels for the analysis of Ub-tau254 digested by h20S. (A) Primary sequence of Ub-tau254, a red character highlights the Cys residue where Ub is installed. Green boxes at the N- and C-termini indicate the amino acid numbering of the full-length tau protein, providing a reference for readers. (B) 3D scatter plot illustrating the distribution of molecular mass, tau fragment length, and Ub tag length for identified proteolytic fragments. Each dot represents a unique peptide. The *x*-axis corresponds to the molecular mass of the peptide, the *y*-axis represents the length of the tau portion within the branched peptide, and the *z*-axis indicates the length of the residual Ub portion attached to the tau peptide fragment. Dot size and colour are proportional to the intensity of the corresponding proteolytic fragment detected after 3 h of incubation, providing a visual representation of the relative abundance of each peptide. (C) Top 15 identified proteolytic fragments ranked by normalized intensity in green, left axis. Bars show the mean ± SD computed at 3 h in the presence of h20S. Fold change in peptide intensity, calculated as the ratio of intensity at 3 h to the intensity at time 0 (right axis, blue). *Peptides absent at time 0 (intensity = 0). In the case of tau sequence, numbers in brackets refers to numbering shown in panel "a" (1 to 133).

spectral intensities were normalized using the median intensities across the spectrum.[35]

In our case, the following results were previously filtered for peptides with no mass ambiguity at resolution 180 000.

Identification of h20S-generated branched fragments of tauK18 regio-isomers (−254, −311, −353)

The complete repertoire of branched and unbranched fragments generated by the h20S digestion of synthetic regio-isomers (−254, −311, −353) was resolved by SpectraSage. Spectra were collected at time 0 and after 3 h of incubation in the presence and absence of h20S and epox (1 μM, 15 minutes of pre-incubation).

To this aim, .raw Orbitrap files were deconvoluted and analysed setting the following filtering criteria for identification of h20S digestion peptides, namely: (a) an increase of over 2-fold in intensity from time 0 to 3 hours of incubation in the presence of h20S (notably, h20S was added before harvesting the aliquot of sample at time 0 and immediately quenched); (b) the intensity of the proteolytic fragment in the presence of epox was less than 80% of the intensity of the same proteolytic fragment in the absence of the inhibitor; (c) identification of the peptides in at least two out of the three replicates.

SpectraSage analysis, reported in Fig. 3 and SI Fig. S2 and S3, highlighted notable different aspects along the pattern of digestion of Ub-tau254, Ub-tau311, Ub-tau353. Distinct proteolytic fragment size distributions were observed among the three regio-isomers. Specifically, Ub-tau311 evidenced the presence of several proteolytic fragments meeting the criteria (a–c) introduced above ($n = 85$). However, they all had quite low intensities and failed to densely populate the three-dimensional space (SI Fig. S2) defined by peptide molecular mass, tauK18 and Ub peptide lengths. Conversely, digestion of Ub-tau353 shared a similar quantity of proteolytic fragments ($n = 106$) but exhibited a considerably broader dispersion in the 3D space (SI Fig. S3), indicating an enhanced heterogeneity in the resulting proteolytic fragments.

Particularly striking is the scenario observed in the case of Ub-tau254 (Fig. 3). Despite presenting fewer proteolytic fragments ($n = 48$), compared to the other proteoforms, the identified peptides showed higher intensities and a well-distributed presence across the 3D space (Fig. 3B). This observation was interpreted as a confirmation of SDS-PAGE data about the h20S preference for Ub-tau254. Regardless of the effective rate of digestion, the following discussion is based on the top 15 proteolytic fragments, identified by their normalized intensity and reported in Fig. 3C, and SI Fig. S2c and S3c. Most notably, in accordance with one out of the main goals of SpectraSage development, the unprecedented identification of branched fragments of variable length was here achieved. Inspection of the results retrieved cleavage sites within the (branched) Ub predominantly located near residues 74 to 72 and 60 to 47 of Ub.

Fig. 4 provides compelling evidence for the proteolytic cleavage of Ub-tau254 by the h20S. The XICs clearly show a significant increase in the number of proteolytic fragments upon enzyme treatment, suggesting extensive processing of the protein with extensive release of branched peptide fragments.

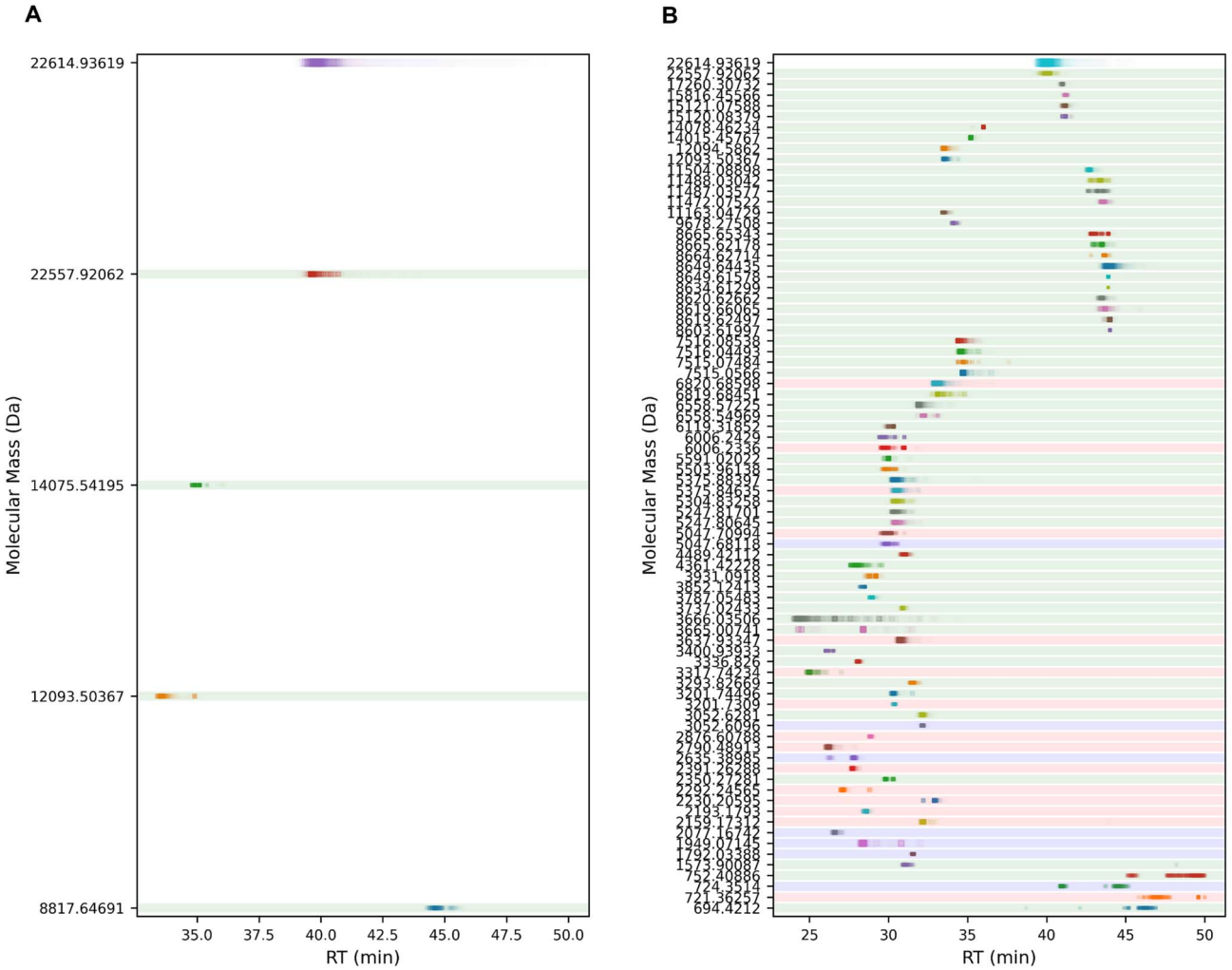

Fig. 4 2D Extracted Ion Chromatograms (XICs) of selected masses following filtering criteria described in the text. Panels (A) and (B) display the molecular mass (*y*-axis) *versus* retention time (*x*-axis) for Ub-tau254 (alone) and Ub-tau254 + h20S at 3 hours of incubation, respectively. TauK18, Ub and Ub-tauK18 branched peptides are highlighted by red, blue and green background lines, respectively. These XICs provide a rapid and intuitive visualization of the h20S degradation efficacy towards Ub-tau254.

Although the overall rate of digestion was lower, branched fragments were documented also for Ub-tau311 and Ub-tau353 (SI Fig. S2 and S3, respectively). In particular, in the case of Ub-tau353, branched fragments were found to span from the -GG motif (residues$_{76-75}$) to amino acids Asn$_{60}$ and up to Lys$_{33}$ of Ub. In the case of Ub-tau311, very few branched fragments were observed in the Top 15, with Ub cleaved in proximity of the 75th residue, and leaving a -GG remnant, or, else at Glu$_{51}$.

Distinct ubiquitylated tauK18 proteoforms exhibit different preferential cleavage sites

To rationalize the data and to gain deeper mechanistic insights into the phenomenon, upset plots and heatmaps were generated to visualize the frequency of each amino acid at specific cleavage sites separately for tauK18 and Ub (Fig. 5).

While all the investigated regio-isomers released proteolytic fragments of tauK18 and Ub alone, as well as branched peptides, their identities were almost regio-isomer specific (Fig. 5B and C).

However, cleavage site analysis, which included fragments of tauK18 and Ub alone together with branched fragments, identified conserved h20S preferences for residues surrounding the scissile bond, typically denoted as P1, P1′, P2, P2′, and so forth, based on the nomenclature established by Schechter and Berger,[36] which indicates the amino acids preceding (P1) and following (P1′) the cleavage site (Fig. 5A).

In fact, in the case of tauK18, amino acid residues including Val, Ser, Lys, and Gly showed the greatest frequency at positions P1 and P1′. Conversely, in the case of Ub a distinct preference for Leu residues, followed by Arg and Gly, was observed over the same catalytic positions.

Identification of branched fragments by MS/MS through a TMTpro zero workflow designed to label h20S-generated neo N-termini

To validate our detection of branched degradation products and the cleavage specificities for Ub and tauK18 by MS/MS approaches we developed a workflow based on TMTpro

Zero (hereafter called TMT) labelling of $-NH_2$, groups in solution, including the ε-NH_2 of Lys side chain, before introducing a trypsin digestion step or, as control, without delivering trypsin (SI Fig. S6). Notably, trypsin cleavage is blocked at TMT-labelled Lys side chains. Consequently, neo N-termini including proteolytic branched fragments generated by

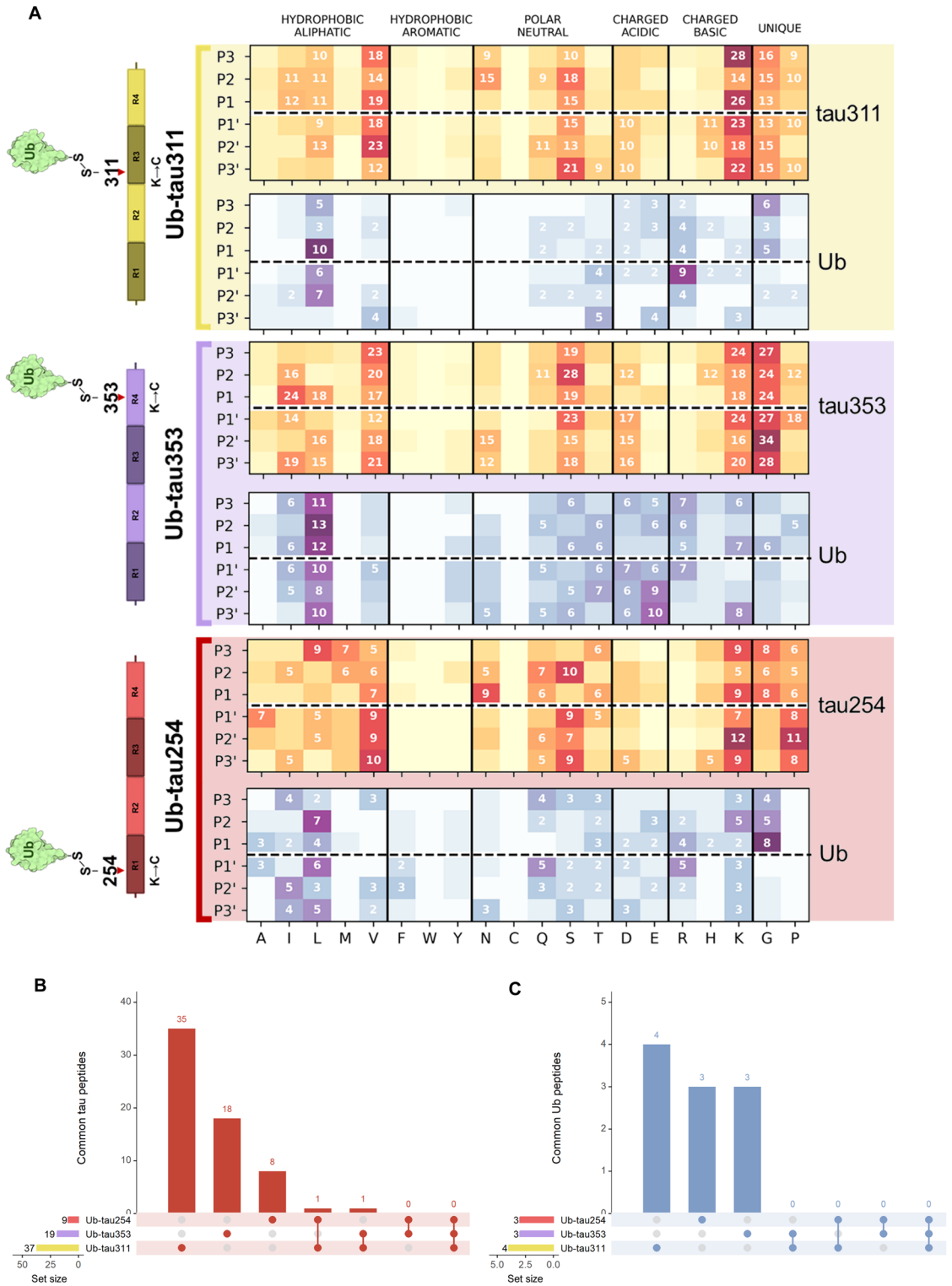

Fig. 5 (A) SpectraSage output heatmaps depicting amino acid frequencies at positions −3 to +3 relative to the cleavage site (dashed line) for Ub-tau311, Ub-tau353, Ub-tau254 for both tau and Ub sequences. (B and C) Comparative UpSet plots showing the overlap of identified proteolytic peptides across the three Ub-tau regio-isomers. (B) Distribution of unbranched tau peptides. (C) Distribution of Ub peptides. Red bars indicate the size of intersections for tau peptides; blue bars indicate the size of intersections for Ub peptides.

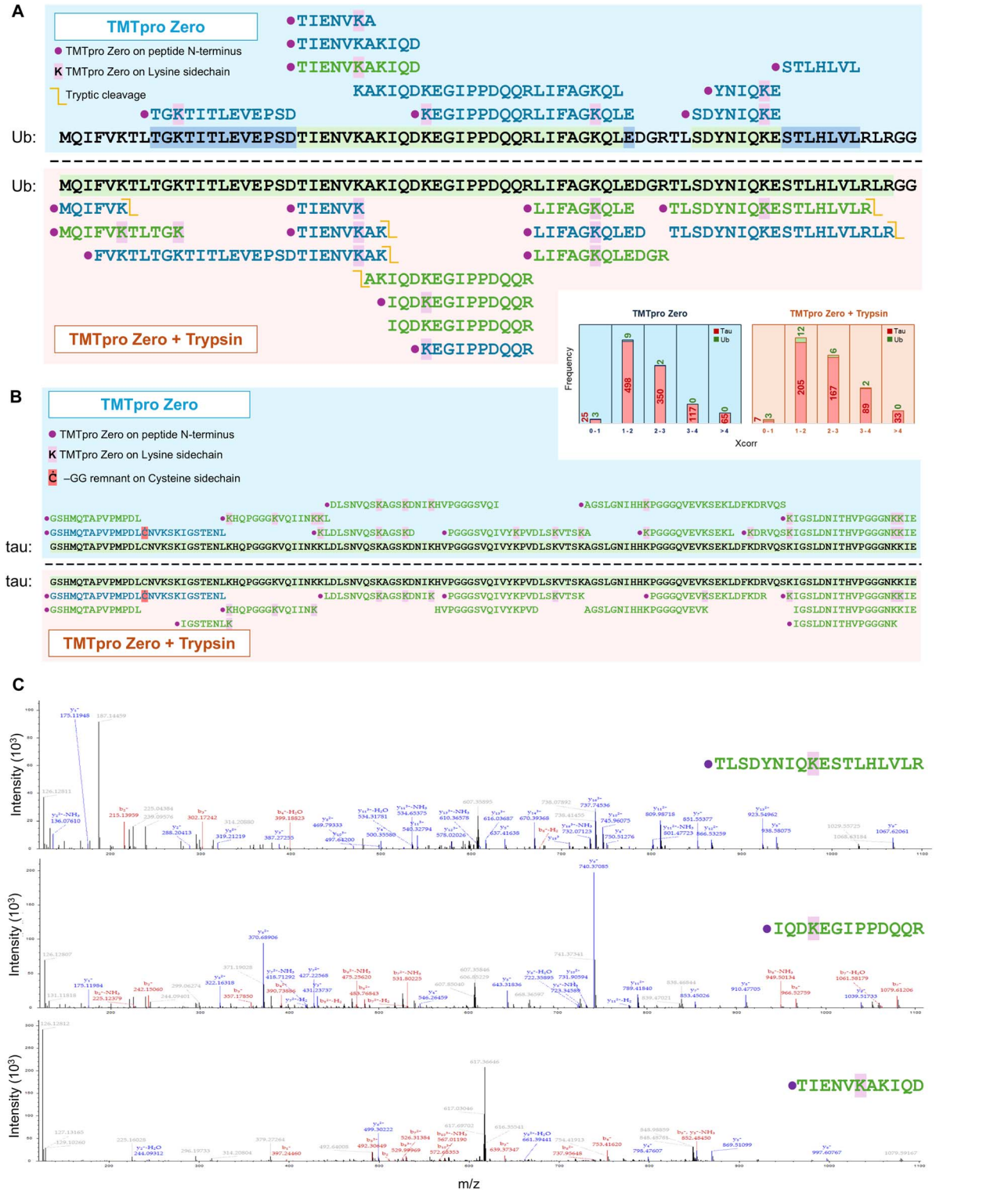

Fig. 6 Sequence coverage and distribution patterns of peptides of Ub (A) and tauK18 (B) identified in both TMT labelled and tryptic digested samples following 3 h of incubation of Ub-tau254 with h20S proteasome. Peptide identification confidence is visually represented by color. Green peptides exhibit high confidence while blue peptides indicate medium confidence. TMTproZero labels are indicated by a purple circle on peptide N-termini and a purple background on Lys residues. The -GG remnant on Cys side chains is marked by a C with a dot on top. Inset: frequency distribution of SequestHT XCorr values, binned into ranges (*x*-axis: 1–2, 2–3, 3–4, >4), showing the distribution of Ub (green) and tau (red) peptides with and without trypsin treatment. A comprehensive list of peptides is provided in SI Data 1 and 2. (C) Peptide Spectrum Matches (PSMs) of selected Ub peptides.

the h20S, are labelled and more unequivocally detected by MS/MS.

It is important to recall, based on previous data, that h20S proteolysis is expected to release oligopeptides of tauK18 alone (tauK18-only), Ub alone (Ub-only), and Ub-tauK18 with branches of variable length (see SpectraSage results). With these premises, if the long Ub-tauK18 branched peptides were truly released by the h20S, the following conditions should be met:

(a) In the absence of trypsin digestion, largest identification of TMT labelled Ub proteolytic fragments covering the N-terminal portion of Ub rather than the C-terminal one. This is based on the assumption that, in branched peptides, the C-terminal portion of Ub is still anchored to the protein (tauK18 in this case) after h20S cleavage.

(b) Analogously, there must be an increased detection of TMT labelled Ub peptides covering the C-terminal portion of the protein in the presence (*vs.* absence) of trypsin. These originally branched peptides can be released by trypsin cut at the Arg residues on the C-terminal portion of Ub.

(c) Ideally, one or more peptides should be identified, showing the features discussed in the previous point and covering one out of the first two Arg cleavage sites of trypsin at the very C-terminus of Ub (Arg residues in 74th and 72nd position), otherwise not identifiable in the absence of trypsin.

In accordance with these considerations, we set up this assay for Ub-tau-254, based on its higher susceptibility to h20S digestion discussed above. Regarding the informatic pipeline, .raw files of samples not exposed to trypsin digestion were analysed by setting "no-enzyme" with unspecific digestion in PD panel, whereas .raw files of samples digested with trypsin were analysed by setting "trypsin" with "semi-tryptic" digestion search in the same panel. All additional features were left constant between methods and, to circumstantiate the findings, data reported below were filtered for: (i) peptides PSMs validated by the Fixed Validator node in PDv2.5, enriched for high and medium-confidence identifications, reflecting that over 75% of peptides had XCorr > 1.5; (ii) missed identification in the absence of h20S and identification with lower intensity (with respect to the presence of h20S) in the presence of epox.

Mass spectrometry analysis results are presented in Fig. 6 (comprehensive data is provided in the SI Data 1 and 2). The outcome of the analysis met the assumption discussed above (see points a–c) for which a point-by-point reply is provided:

(a) In the absence of trypsin different peptides covering the full length of Ub were documented; despite high sequence coverage, the -RLRGG motif remained undetectable, supporting the hypothesis that this remnant, or a shortened form, persists anchored to the tau sequence (Fig. 6A).

(b) In the presence of trypsin, a net increase in peptides covering the C-terminal portion of Ub was observed.

(c) In the presence of trypsin, we identified a proteotypic peptide (see Fig. 6A) ($_{55}$TLSDYNIQKESTLHLVLR$_{72}$) with a double TMT label at the N-terminus and at the ε-NH_2 group of Lys, that was undetectable in the absence of trypsin.

As evident from Fig. 6B and the comprehensive peptide list in the SI (Data 1 and 2), numerous tauK18 peptides were identified (a complete list is provided in the SI Data 1 and 2). Notably, the presence of a tauK18 peptide decorated with a -GG remnant was detected exclusively in the presence of the h20S both in the presence and absence of trypsin digestion (XCorr = 1.86 and 2.13, respectively).

To confirm that branched peptides, including the peptide discussed above, are released also in the presence of a native isopeptide linkage, we set up and recapitulated the analytical workflow introduced for Ub-tau254 above, using Ub-tauK18 (enzymatic preparation).

Remarkably, in this case, the analysis (SI Fig. S6) retrieved an outcome very similar to that obtained for Ub-tau254 highlighting that:

(i) Based on TMT labelling, Ub and branched peptides were documentable also in the presence of a native isopeptide linkage.

(ii) The ($_{55}$TLSDYNIQKESTLHLVLR$_{76}$) peptide was identified in the presence of trypsin but it was undetectable in the absence of this further digestion step (a complete list is provided in the SI Data 3 and 4). Nonetheless, a manual inspection of SpectraSage outcome confirmed detection of this branched fragment in the presence of h20S which is expected to cover the tauK18 sequence covering residues Ser_2-Asn_{47} (see SI Data 5).

(iii) Notably, the only -GG remnant detected in the absence of trypsin was in proximity of Lys_{254}, indirectly confirming that, among the different mono-Ub regio-isomers of the enzymatic preparation, Ub-tau254 was the preferential substrate.

Importantly, a frequency heatmap of tau digestion peptides of Ub-tau254 (SI Fig. S7) identified by MS/MS, confirmed the catalytic preferences already highlighted by SpectraSage. Also, in the case of Ub, for which, compared to SpectraSage, a lower number of peptides was retrieved to compute an heatmap, Leu turned out to be robustly represented across P1 and P1′ positions (this outcome can be deduced by inspection of Fig. 6A).

Discussion

In this study, we provide experimental evidence that the h20S digests *in vitro* mono-ubiquitylated tauK18 proteoforms with relevant role in neurodegeneration. Given the growing enzymatic and biological role of the uncapped h20S and the relevance of tau degradation and turnover for neurodegenerative processes of the brain, retina and optic nerve, this study provides new mechanistic insights on targeted proteolysis.

Specifically, we report that, under the experimental conditions adopted:

(a) The h20S cleaves synthetic enzymatic (native isopeptide linkage) and (Ub conjugated by a disulfide bridge) proteoforms of Ub-tauK18, including three distinct regio-isomers (Ub-tau254, −311, −353) that were found to shape tau intracellular aggregates through a disease-specific pattern;[28]

(b) Unlike full-length tau, whose four-repeat domain was reported to be preferentially digested by t20S,[26] we found that Ub is necessary for tauK18 cleavage; therefore, in the absence of ubiquitylation, it is likely that the four-repeat domain needs to be embedded into a longer sequence to become a h20S substrate. Alternatively, it is worth pointing out that we monitored digestion under reasonable steady-state conditions (*i.e.*,

large excess of substrate), using very low h20S concentration (10–50 nM) and over a limited time of incubation (3 h) to explore the mechanism under conditions of limited proteolysis and to maximize the patho-physiological translation of our findings;

(c) The position of Ub along the tauK18 sequence affects the catalytic preference. Our data support the preference of h20S for tauK18 with Ub tag placed at the N-terminus (Ub-tau254), followed by the C-terminus (Ub-tau353) and the middle portion (Ub-tau311). Importantly, this preference was seemingly confirmed also by digestion of the enzymatic Ub-tauK18 and by previous data on full-length tau,[26] whose digestion fragments in most cases did fall within the R3 domain and were retrieved also by us. Indeed, in the case of Ub-tauK18, among peptides released by the h20S, a diGLY remnant was identified with a reasonable confidence only in correspondence of Lys_{254}. Regarding the debated h20S preference for the N- over C-terminus of unstructured proteins, it is likely that a certain degree of protein specificity exists, and this study was not intended to clarify this feature;

(d) Ub is cleaved alongside the substrate generating branched fragments that extend beyond the C-terminal -GGRL motif described in the previous study by Sahu and co-workers.[22] In this regard, digestion of all Ub-tauK18 regio-isomers released diverse Ub remnants on branched peptide products extending beyond the canonical four C-terminal residues and up to the 17th amino acid. The main implication of this finding is that the h20S is likely endowed with a pulling force sufficient to completely unfold the tightly packed Ub monomer;

(e) Digestion of Ub shows catalytic preferences different from that of tauK18. Specifically, in the majority of identified and quantified peptides a Leu residue in P1, P2 or P1′, P2′ position was required for Ub cleavage, whereas Gly, Ser and Lys residues were preferred in P1 or P1′ position for tauK18. The hypothesis regarding the catalytic preference is strengthened by an additional SpectraSage search carried out by removing the filter concerning the identification of peptides with lower intensity in the presence of epox. Thus, by removing this filter, a robust increase of peptides with Leu in P1 and P1′ position was documented also for tauK18. Since, at the concentration tested, epox is a weak inhibitor of trypsin-like (β_2 subunit) and caspase-like (β_1 subunit) activities, this observation is consistent with a role of these catalytic subunits, over the canonical chymotrypsin-like, in cleaving Ub. Interestingly, although the identity of tauK18 and Ub fragments (without branches) was in most cases regio-isomer-specific, the catalytic preferences were conserved.

Therefore, we suppose that distinct catalytic subunits drive the enzymatic processing of ubiquitylated substrates, revealing a mechanism where Ub and substrate enzymatic cleavage may be spatially and/or chronologically separated. As a matter of fact, h20S cleavage site preferences, particularly at P1 and P1′, remain controversial in literature. Studies have shown that catalytic subunits (β1, β2, β5) exhibit preferences for small, uncharged residues (*e.g.*, Ala, Leu, Gly, Ser), while bulky or charged residues (*e.g.*, Lys, Arg) are generally disfavoured due to interference with the catalytic mechanism.[37–39] In particular, Leu is favoured due to its compatibility with hydrophobic pockets within the proteasome active sites, enhancing substrate binding and cleavage.[10,40] While Lys and Arg are typically less preferred, exceptions exist where they contribute to substrate specificity through interactions with the active sites.[41] Notably, neighbouring residues and substrate structural features significantly influence proteasomal cleavage specificity, as demonstrated by Nussbaum *et al.*[42] and Groll *et al.*[17]

It is important to underline that SpectraSage has been fundamental to realize this study. SpectraSage is a proprietary software specifically designed to identify and quantify peptides separated by nanoUHPLC and analysed at MS1 level. This software was developed to advance *in vitro* enzymology studies (enzyme:substrate interactions) using MS and, remarkably, to overcome, at least for *in vitro* studies, the computational challenge associated with the identification of complex cross-linked peptides of non-tryptic origin, such as the branched Ub-peptides here identified. Moreover, for middle-down (or top-down) studies, SpectraSage shows key advantages including a less complex sample preparation and a rapid and automated bioinformatic analysis. In this regard, SpectraSage offers faster and automated data processing, eliminating the need for manual inspection of fragment tables and reconstruction of sequence coverage. Additionally, it automatizes the counting of amino acid occurrences at specific positions relative to putative cleavage sites.

Nonetheless, although MS1 analysis does not allow PSM validation, employing MS1-based precursor ion current quantification is a validated method, which is commonly used for middle-down experiments by virtue of its reliability and accuracy especially for these specific scientific aims.[43]

However, given the biological relevance that Ub digestion and release of long branched peptides may have in biology, we here adopted a TMT labelling workflow coupled with limited trypsin digestion (SI Fig. S6) to validate the SpectraSage ability and to identify, though indirectly, Ub proteolytic fragments and Ub-tau branched peptides. This workflow is based on labelling of free N-termini, such as those of peptides generated by h20S activity, before trypsin digestion. Moreover, TMT labelling of the ε-amino group of the Lys side chain introduces a missed cleaved site,[44] thereby restricting the search to trypsin cut at Arg residues. This protocol turns out very useful for Ub conjugates, since two Arg residues lie in the very proximity of the diGLY (-GG) motif anchored to the substrate.

Remarkably, the TMT workflow has played a key role in validating the outcomes of SpectraSage analysis and identifies a prototypical branched digestion fragment (see SI Fig. S6). Based on the comparison with SpectraSage output, the proteolytic fragment may be coherent with a branched product encompassing Ser_{241}-Asn_{286} on Ub-tau254 and Gly_{76}-Tyr_{55} of Ub. Remarkably, this proteolytic fragment was identified either in the case of h20S digestion of both (synthetic) Ub-tau254 and (enzymatic) Ub-tauK18, revealing that the catalytic process occurs regardless of the presence of a disulfide bridge or a native isopeptide linkage.

The biological role of the uncapped 20S is matter of intense debate across the scientific community. Although its

"existence" has been validated by cryo-EM studies on living neurons, the possibility that it plays proteolytic activities, distinguishable from those of the 26S, has never been addressed.[23] However, recent research indicated that free 20S possesses distinct biological functions and is subject to unique regulatory mechanisms (*e.g.*, DJ-1 and Insulin Degrading Enzyme).[21,45–51]

First, the catalytic preference of the free 20S for oxidized and unfolded proteins was proposed.[21,45] More recently, the possibility that the h20S is able to degrade mono-Ub proteins has been further advanced, an hypothesis strengthened by targeted proteomics in living cells.[22]

In this framework, the enzymatic and semi-synthetic Ub-tau species enrolled in this study, offered the possibility to investigate how the position of Ub within the domain of a protein affects its structural properties and the susceptibility to proteasome digestion and to deeply explore the catalytic mechanism.[29,34] In this context, previous studies demonstrated that enzymatic ubiquitylation of tauK18 inhibits its oligomerization and aggregation propensity.[29]

While a similar effect was observed for tauΔC, site-specific mono-ubiquitylation at different positions imparted distinct structural properties. Thioflavin-T fluorescence and transmission electron microscopy revealed that, in the presence of heparin induction, both Ub-tau254 and Ub-tau353 retained the ability to form cross-β structures, albeit with significantly slower kinetics compared to tauK18. In contrast, Ub-tau311 did not exhibit this behaviour.[29] Whether these structural and aggregating-prone properties have a patho-physiological link with the susceptibility to h20S digestion cannot be addressed at this stage. However, it is unlikely that protein conformation, at least at the concentration tested (5 μM), has prominent roles, since aggregation was not induced by heparin in our system.

In conclusion, the findings here reported can push forward research, and the debate around the biological role of h20S towards natural substrates, mono-ubiquitylated or not, which is worth being explored for its biological and pathological implications. This is particularly relevant for the comprehension of the processes underscoring the onset and progression of degenerative diseases of the brain (primary and secondary tauopathies), retina (AMD) and optic nerve (glaucoma) for which altered metabolism of tau protein has been widely documented. Nonetheless, a detailed mechanistic understanding of these processes is key to envision and foster new therapeutic interventions.

Furthermore, this study introduces SpectraSage a software specifically conceptualized to assist the scientific community in exploring the repertoire of proteolytic fragments generated from simple reaction mixture, allowing to uncover more and new additional mechanistic insights into the enzymology of proteolytic systems and their substrates.

Materials and methods

Protein synthesis and expression

For this study, we produced the:

(a) Construct of tau spanning residues Q244-E372 plus initial Met (hereafter tauK18);

(b) The variants: tauK18(C291A/C322A) (here named tauΔC), tauK18(C291A/C322A/K254C), tauK18(C291A/C322A/K311C), and tauK18(C291A/C322A/K353C);

(c) Ub-tauK18 that is tau-K18 mono-ubiquitylated at Lys residues by an enzymatic reaction (see below);

(d) Ub-S-S-tau proteoforms, namely Ub-tau254, Ub-tau311, Ub-tau353, that are proteoforms with Ub installed at specific positions (see below).

All tau proteins were expressed and purified (>95% purity) as previously described.[29] Typical chromatograms of Ub-S-S-tau proteoforms are shown in SI (Fig. S1–S3). Ub was expressed and purified following a previously described protocol both in wild type and N-terminal (histidine)6-tagged forms.[29,52] Ub-tauK18 was obtained applying an enzyme-based method employing recombinant enzymes E1, UBE2N, and CHIP E3, as previously described.[29] Ub-tauK18 is a mixture of mono-ubiquitylated tauK18 species, where Ub is linked to different Lys residues. The relative amount of the different mono-Ub-proteoforms cannot be precisely assessed, but they were proven to be all present to a significant amount through a proteomic approach based on detection of a diGLY remnant released by trypsin digestion.[29,34] Ub-S-S-Tau proteoforms were obtained using a semi-synthetic approach based on a disulfide-directed reaction. TauΔC was mutated to introduce a single cysteine in place of Lys254, Lys311 or Lys353. Ub bearing an aminoethanethiol C-terminal group (hereafter Ub-SH) for the disulfide-coupling reaction, was produced and purified as described previously.[29] Ub-SH was activated with 5,5 dithio-bis-(2-nitrobenzoic acid) and then incubated with tauΔC variants bearing lysine to cysteine mutations at selected sites. The Ub-S-S-Tau proteoforms were then purified using ion exchange chromatography.[29] The site-specificity of ubiquitin conjugation achieved by the semi-synthetic approach has been confirmed in prior studies using site-resolved NMR spectroscopy: ^{1}H–^{15}N HSQC spectra revealed localized chemical shift changes, which were corroborated using $^{15}N/^{13}C$ isotope labelling and multidimensional heteronuclear NMR experiments.[30,53] TauK18 in the construct Ub-tau254, contains additional amino acids (GSHM) at the N-terminus deriving from the cleavage of a tag used for protein purification.

Digestion assays

TauK18 (5 μM) and Ub-tauK18 (5 μM) were incubated with 10 nM affinity-purified h20S proteasome (Boston Biochem, MA, USA) in 25 mM Tris–HCl, pH 7.5, at 37 °C. Results were confirmed also using another affinity purified h20S (a generous gift by Prof. M. Glickman). In all cases, contamination by 19S was excluded by SDS-PAGE. In analogy with studies on tau-K18, semi-synthetic Ub-tauK18 (−254, −311, −353) species (5 μM), as well as tauΔC were incubated with 10 nM h20S in 25 mM Tris–HCl, pH 7.5, at 37 °C. The digestion of tau species was monitored for 3 h, over a range of linearity in terms of rate of substrate digestion, as measured by SDS-PAGE (see below).

Aliquots of reaction mixtures were harvested at each indicated time-points.

SDS-PAGE electrophoresis

Harvested aliquots were heat-denatured in 1x Laemmli buffer, separated by SDS-PAGE (in-house prepared 15% and 4–20% acrylamide pre-cast gel, BioRad, Hercules, CA) and stained by Coomassie Brilliant Blue (CBB).

Gel lanes were acquired in a gel documentation system (Azure Gel Imager C280, Azure Biosystem, Sierra CT Suites AB, Dublin, USA, and iBright Imager, Thermo Fisher Scientific, Waltham, Massachusetts, USA). Band intensity was calculated by ImageJ software.

Gels shown in figures have been slightly edited or cropped to increase band visibility or resolution. Raw uncropped gels are provided in SI Fig. S4.

Mass spectrometry studies

In accordance with the heterogeneity of proteins studied and the specific aims of the analysis, different approaches were undertaken. Middle-down mass spectrometry (MS) studies were set up for tauK18, Ub-tauK18 (enzymatic), tauΔC, and Ub-S-S species. All proteins enrolled in the study (5 μM), were incubated in the absence or presence of h20S or in the presence of h20S inhibited through pre-incubation with 1 μM epoxomicin (epox), 15 min, 37 °C (Boston Biochem, MA, USA). All studies were run and analysed in experimental replicate ($n = 3$). In all cases the reaction was stopped by adding 0.1% TFA.

Samples were then injected into an Orbitrap Exploris 240 mass spectrometer (Thermo Fisher Scientific, MA, USA) online with an Ultra High-Performance Liquid Chromatography (UHPLC) Ultimate 3000 system (Thermo Fisher Scientific, MA, USA) with a C18 prefiltration column installed.

In the case of Ub-tauK18, full length protein (~23 kDa) was filtered out by flowing the sample through 10 kDa-cutoff filter (Merck Millipore, MA, USA) at the end of degradation assay. Thereafter, eluted peptides were cleaned by Pierce™ C18 Tips (Thermo Fisher Scientific, Waltham, Massachusetts, USA), dried in a speed Vacuum system (Labconco, Kansas City, MO, USA) and resuspended in loading solvent (see below) before injection.

Regarding the UHPLC-MS parameters, chromatographic separation was performed with a 2 μm particle size column (ES904, Thermo Fisher Scientific, MA, USA) for peptides generated from the proteasome digestion of Ub-tauK18 and a 3 μm particle size column (ES900, Thermo Fisher Scientific, MA, USA) for the digestion of the semi-synthetic tau proteins.

Both columns had a 100 Å pore size, 75 μm internal diameter, and 15 cm length. Loading solvent: 2% ACN 0.05% TFA; solvent A: 100% H_2O 0.1% FA; solvent B: 80% ACN 0.1% FA. UHPLC gradient (88 min) used for Ub-TauK18: 0–62 min 6–31% B, 62–67 min 31–50% B, 67–72 min 50–99% B, 72–80 min 99% B, 80–82 min 99–6% B. UHPLC gradient (75 min) used for semi-synthetic tau proteins: 0–9 min 4.5% B, 9–62 min 4.5–50.6% B, 62–65 min 50.6–99% B, 65–69 min 99% B, 69–70 min 99–4.5% B.

The mass spectrometry parameters were set as follows: ionization mode: nano-electrospray ionization (nano-ESI); ionization voltage: 1500 V; capillary temperature: 30 °C; Scan range (m/z): 375–1650; resolution: 120 000 for Ub-tauK18 and 180 000 for semi-synthetic tau proteins; AGC target: 3×10^6; microscans: 3. For Data-Dependent Acquisition (DDA) analysis, the following filters were employed: Precursor ion intensity: 5×10^3, included charge states: 2–8; dynamic exclusion duration: 45 s; exclusion mass width: 10 ppm. MS/MS parameters: fragmentation method: HCD; isolation window (m/z): 2; normalized HCD collision energy: 30%; resolution: 15 000; AGC target: 5×10^4.

Bioinformatic analysis

Raw files of Ub-tauK18 samples coming from MS/MS analysis were submitted to Proteome Discoverer software (PD v2.5, Thermo Scientific, MA, USA) using the Sequest algorithm and searching against a FASTA containing the sequence of tauK18 and Ub (including aminoacidic modifications where necessary, such as for tau254). In accordance with a previous study on full-length tau digestion by the Thermoplasma acidophilum 20S (t20S), the PD "Fixed PSM Validator node" was used for PSM validation.[26] Here, spectra were further searched including a PD contaminant database (containing human, murine and bacterial contaminants) and using Percolator for PSM validation.

Raw files of tauK18, tauΔC, and Ub-S-S-tau species were processed by SpectraSage Software, whose development and properties are described in detail in the Appendix section of the SI. Briefly, .raw files underwent initial processing including spectra deconvolution, followed by peak extraction and deisotoping. Concurrently, a theoretical fragments table was generated based on the expected linear/branched peptide/protein structures. These processed experimental data were then used for matching against the theoretical fragments, leading to identification and subsequent visual representation of the results. While a mass error tolerance of 20 ppm was applied for initial peptide identification to accommodate potential unmodeled modifications, it is noteworthy that approximately 30% of the key Ub-tau branched peptides discussed in this study exhibit a mass error tolerance below 1 ppm, demonstrating high confidence in their identification. Although we explored the use of retention time prediction based on peptide hydrophobicity (using the Sequence Specific Retention Calculator (SSRC) method[54]) as an orthogonal validation for peptide identification, we found its applicability limited for the unusually long and complex branched peptides discussed in this study, likely due to structural effects not fully captured by these models.

Identification of branched fragments and catalytic preferences by MS/MS approaches

To validate the long-branched fragments released by the h20S digestion, Ub-tau254 (5 μM), chosen as reference by virtue of its extensive digestion, was incubated for 3 h, at 37 °C in the presence of 50 nM h20S in 25 mM HEPES (Sigma-Aldrich, St-Louis, Co, USA), pH 7.5, upon performing a buffer exchange procedure. Preliminary assays indicated that increasing h20S concentration maximized high confidence MS/MS identification of proteolytic fragments.

Aliquots were collected and all generated proteolytic fragments were separated from the full-length (undigested) protein (~23 kDa) using 10 kDa cutoff filters (Merck Millipore, MA, USA).

Eluted peptides were then labelled by Tandem Mass Tag (TMT) pro Zero label reagent (Fisher scientific, MA, USA) addition (10 : 1 TMT : protein ratio) for 1 h at room temperature (r.t.). The workflow ensured a labelling efficiency ~95%, which was preferred to rule out eventual over labelling of samples. TMT was then quenched with an excess hydroxylamine, for 20 min, r.t.

Finally, aliquots were divided into two equal ones: the first one was addressed to tryptic digestion (1 : 10 trypsin : protein ratio), for 1 h at 37 °C, the second one was left undigested (by trypsin) but treated under the same experimental conditions (1 h, at 37 °C). At the end of the incubation time, the reaction was stopped by adding 0.1% TFA (final concentration) and all peptide aliquots were cleaned up by Pierce™ C18 Tips before injection into the UHPLC-MS system.

In this case, chromatographic separation was performed with a 2 μm particle size column (ES904, Thermo Fisher Scientific, MA, USA) 100 Å pore size, 75 μm internal diameter, and 15 cm length. Loading solvent: 2% ACN 0.05% TFA; solvent A: 100% H_2O 0.1% FA; solvent B: 80% ACN 0.1% FA. UHPLC gradient (88 min): 0–62 min 6–31% B, 62–67 min 31–50% B, 67–72 min 50–99% B, 72–80 min 99% B, 80–82 min 99–6% B.

The mass spectrometry parameters were set as follows: ionization mode: nano-electrospray ionization (nano-ESI); ionization voltage: 1500 V; capillary temperature: 40 °C; scan range (m/z): 375–1200; resolution: 120 000; AGC target: 3×10^6. For Data-Dependent Acquisition (DDA) analysis, the following filters were employed: precursor ion intensity: 1×10^4, included charge states: 2–7; dynamic exclusion duration: 45 s; exclusion mass width: 10 ppm. MS/MS parameters: fragmentation method: HCD; isolation window (m/z): 2; normalized HCD collision energy: 30%; resolution: 15 000; AGC target: 5×10^4.

Author contributions

Conceptualization: GAZ, MC, MD, GRT, DS. Methodology: GAZ, IP, DT, DC, FM, MA, MD, GRT, DS. Investigation: GAZ, IP, GRT, DS. Visualization: GAZ, DS. Supervision: GAZ, GG, FO, AB, MC, GRT, DS. Resources: GG, FO, AB, MA, MD, GRT, DS. Writing—original draft: GAZ, DS. Writing—review & editing: GAZ, IP, GG, PAB, FO, AB, MC, MA, MD, GRT, DS.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

All data needed to evaluate the conclusions in the paper are present in the paper and/or the SI. See DOI: **https://doi.org/10.1039/d5sc04240b**.

Acknowledgements

The authors received financial support from LazioInnova (grant: A0375-2020-36591). This study was supported by Next Generation Promising (NGP) and PRIN-MIUR n. P2022AW2H9. The authors further acknowledge MUR Prin 2022 n. 2022R9WCZS_001. The authors acknowledge the Ministry of Health and Fondazione Roma for the support. G. A. Zingale and I. Pandino were supported by the PhD program in Chemical Sciences, University of Catania. The authors are grateful to Prof. Michael Glickman (Technion Institute, Haifa, Israel) for the generous gift of h20S and for his supportive collaboration.

References

1 I. Pandino, S. Giammaria, G. A. Zingale, G. Roberti, M. Michelessi, M. Coletta, G. Manni, L. Agnifili, A. V. Vercellin, A. Harris, F. Oddone and D. Sbardella, *Mol. Aspects Med.*, 2023, **94**, 101226.
2 M. Chiasseu, J. L. Cueva Vargas, L. Destroismaisons, C. Vande Velde, N. Leclerc and A. Di Polo, *J. Neurosci.*, 2016, **36**, 5785–5798.
3 A. Boccaccini, D. Cavaterra, C. Carnevale, L. Tanga, S. Marini, A. Bocedi, P. M. Lacal, G. Manni, G. Graziani, D. Sbardella and G. R. Tundo, *Mol. Aspects Med.*, 2023, **94**, 101225.
4 A. Ciechanover and Y. T. Kwon, *Front. Neurosci.*, 2017, **11**, 185.
5 C. L. Klaips, G. G. Jayaraj and F. U. Hartl, *J. Cell Biol.*, 2018, **217**, 51–63.
6 A. J. Sala, L. C. Bott and R. I. Morimoto, *J. Cell Biol.*, 2017, **216**, 1231–1241.
7 G. R. Tundo, D. Sbardella, A. M. Santoro, A. Coletta, F. Oddone, G. Grasso, D. Milardi, P. M. Lacal, S. Marini, R. Purrello, G. Graziani and M. Coletta, *Pharmacol. Ther.*, 2020, **213**, 107579.
8 G. R. Tundo, D. Sbardella, F. Oddone, A. A. Kudriaeva, P. M. Lacal, A. A. Belogurov, G. Graziani and S. Marini, *Cancers*, 2021, **13**, 4852.
9 Y. T. Kwon and A. Ciechanover, *Trends Biochem. Sci.*, 2017, **42**, 873–886.
10 M. H. Glickman and A. Ciechanover, *Physiol. Rev.*, 2002, **82**, 373–428.
11 D. Komander and M. Rape, *Annu. Rev. Biochem.*, 2012, **81**, 203–229.
12 T. Tenno, K. Fujiwara, H. Tochio, K. Iwai, E. H. Morita, H. Hayashi, S. Murata, H. Hiroaki, M. Sato, K. Tanaka and M. Shirakawa, *Genes Cells*, 2004, **9**, 865–875.
13 A. Martin, T. A. Baker and R. T. Sauer, *Nat. Struct. Mol. Biol.*, 2008, **15**, 1147–1151.
14 J. A. M. Bard, E. A. Goodall, E. R. Greene, E. Jonsson, K. C. Dong and A. Martin, *Annu. Rev. Biochem.*, 2018, **87**, 697–724.
15 R. Beckwith, E. Estrin, E. J. Worden and A. Martin, *Nat. Struct. Mol. Biol.*, 2013, **20**, 1164–1172.
16 G. C. Lander, E. Estrin, M. E. Matyskiela, C. Bashore, E. Nogales and A. Martin, *Nature*, 2012, **482**, 186–191.

17 M. Groll, L. Ditzel, J. Löwe, D. Stock, M. Bochtler, H. D. Bartunik and R. Huber, *Nature*, 1997, **386**, 463–471.
18 M. Groll, M. Bajorek, A. Köhler, L. Moroder, D. M. Rubin, R. Huber, M. H. Glickman and D. Finley, *Nat. Struct. Biol.*, 2000, **7**, 1062–1067.
19 E. M. Huber, W. Heinemeyer, X. Li, C. S. Arendt, M. Hochstrasser and M. Groll, *Nat. Commun.*, 2016, **7**, 1–10.
20 R. Raynes, L. C. D. Pomatto and K. J. A. Davies, *Mol. Aspects Med.*, 2016, **50**, 41–55.
21 K. J. A. Davies, *Biochimie*, 2001, **83**, 301–310.
22 I. Sahu, S. M. Mali, P. Sulkshane, C. Xu, A. Rozenberg, R. Morag, M. P. Sahoo, S. K. Singh, Z. Ding, Y. Wang, S. Day, Y. Cong, O. Kleifeld, A. Brik and M. H. Glickman, *Nat. Commun.*, 2021, **12**, 6173.
23 S. Asano, Y. Fukuda, F. Beck, A. Aufderheide, F. Förster, R. Danev and W. Baumeister, *Science*, 2015, **347**, 439–442.
24 F. Kumar Deshmukh, D. Yaffe, M. A. Olshina, G. Ben-Nissan and M. Sharon, *Biomolecules*, 2019, **9**, 190.
25 X. Zhao and J. Yang, *ACS Chem. Neurosci.*, 2010, **1**, 655–660.
26 T. Ukmar-Godec, P. Fang, A. Ibáñez de Opakua, F. Henneberg, A. Godec, K.-T. Pan, M.-S. Cima-Omori, A. Chari, E. Mandelkow, H. Urlaub and M. Zweckstetter, *Sci. Adv.*, 2020, **6**, eaba3916.
27 T. Grune, D. Botzen, M. Engels, P. Voss, B. Kaiser, T. Jung, S. Grimm, G. Ermak and K. J. A. Davies, *Arch. Biochem. Biophys.*, 2010, **500**, 181–188.
28 T. Arakhamia, C. E. Lee, Y. Carlomagno, M. Kumar, D. M. Duong, H. Wesseling, S. R. Kundinger, K. Wang, D. Williams, M. DeTure, D. W. Dickson, C. N. Cook, N. T. Seyfried, L. Petrucelli, J. A. Steen and A. W. P. Fitzpatrick, *Cell*, 2021, **184**, 6207–6210.
29 F. Munari, C. G. Barracchia, C. Franchin, F. Parolini, S. Capaldi, A. Romeo, L. Bubacco, M. Assfalg, G. Arrigoni and M. D'Onofrio, *Angew. Chem., Int. Ed.*, 2020, **59**, 6607–6611.
30 D. Trivellato, F. Floriani, C. G. Barracchia, F. Munari, M. D'Onofrio and M. Assfalg, *Bioorg. Chem.*, 2023, **132**, 106347.
31 F. Parolini, R. Tira, C. G. Barracchia, F. Munari, S. Capaldi, M. D'Onofrio and M. Assfalg, *Int. J. Biol. Macromol.*, 2022, **201**, 173–181.
32 F. Munari, L. Mollica, C. Valente, F. Parolini, E. A. Kachoie, G. Arrigoni, M. D'Onofrio, S. Capaldi and M. Assfalg, *Angew. Chem., Int. Ed.*, 2022, **61**, e202112374.
33 L. Meng, R. Mohan, B. H. B. Kwok, M. Elofsson, N. Sin and C. M. Crews, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 10403–10408.
34 F. Munari, C. G. Barracchia, F. Parolini, R. Tira, L. Bubacco, M. Assfalg and M. D'Onofrio, *Int. J. Mol. Sci.*, 2020, **21**, 4400.
35 J. E. Wulff and M. W. Mitchell, *Adv. Biosci. Biotechnol.*, 2018, **9**, 339–351.
36 I. Schechter and A. Berger, *Biochem. Biophys. Res. Commun.*, 1967, **27**, 157–162.
37 A. Craiu, M. Gaczynska, T. Akopian, C. F. Gramm, G. Fenteany, A. L. Goldberg and K. L. Rock, *J. Biol. Chem.*, 1997, **272**, 13437–13445.
38 O. Coux, K. Tanaka and A. L. Goldberg, *Annu. Rev. Biochem.*, 1996, **65**, 801–847.
39 A. F. Kisselev, T. N. Akopian, V. Castillo and A. L. Goldberg, *Mol. Cell*, 1999, **4**, 395–402.
40 J. Löwe, D. Stock, B. Jap, P. Zwickl, W. Baumeister and R. Huber, *Science*, 1995, **268**, 533–539.
41 M. Gaczynska, K. L. Rock and A. L. Goldberg, *Nature*, 1993, **365**, 264–267.
42 A. K. Nussbaum, T. P. Dick, W. Keilholz, M. Schirle, S. Stevanović, K. Dietz, W. Heinemeyer, M. Groll, D. H. Wolf, R. Huber, H.-G. Rammensee and H. Schild, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 12504–12509.
43 A. Palomba, M. Abbondio, G. Fiorito, S. Uzzau, D. Pagnozzi and A. Tanca, *J. Proteome Res.*, 2021, **20**, 3497–3507.
44 K. Yu, M. Niu, H. Wang, Y. Li, Z. Wu, B. Zhang, V. Haroutunian and J. Peng, *J. Am. Soc. Mass Spectrom.*, 2021, **32**, 936–945.
45 T. Grune, B. Catalgol, A. Licht, G. Ermak, A. M. Pickering, J. K. Ngo and K. J. A. Davies, *Free Radicals Biol. Med.*, 2011, **51**, 1355–1364.
46 A. M. Pickering, A. L. Koop, C. Y. Teoh, G. Ermak, T. Grune and K. J. A. Davies, *Biochem. J.*, 2010, **432**, 585–595.
47 D. Sbardella, G. R. Tundo, A. Coletta, J. Marcoux, E. I. Koufogeorgou, C. Ciaccio, A. M. Santoro, D. Milardi, G. Grasso, P. Cozza, M.-P. Bousquet-Dubouch, S. Marini and M. Coletta, *Cell. Mol. Life Sci.*, 2018, **75**, 3441–3456.
48 O. Moscovitz, G. Ben-Nissan, I. Fainer, D. Pollack, L. Mizrachi and M. Sharon, *Nat. Commun.*, 2015, **6**, 6609.
49 G. R. Tundo, D. Sbardella, C. Ciaccio, A. Bianculli, A. Orlandi, M. G. Desimio, G. Arcuri, M. Coletta and S. Marini, *J. Biol. Chem.*, 2013, **288**, 2281–2289.
50 D. Sbardella, G. R. Tundo, F. Sciandra, M. Bozzi, M. Gioia, C. Ciaccio, U. Tarantino, A. Brancaccio, M. Coletta and S. Marini, *PLoS One*, 2015, **10**, e0132455.
51 G. R. Tundo, D. Sbardella, C. Ciaccio, G. Grasso, M. Gioia, A. Coletta, F. Polticelli, D. Di Pierro, D. Milardi, P. Van Endert, S. Marini and M. Coletta, *Crit. Rev. Biochem. Mol. Biol.*, 2017, **52**, 554–582.
52 R. Varadan, M. Assfalg, S. Raasi, C. Pickart and D. Fushman, *Mol. Cell*, 2005, **18**, 687–698.
53 G. Viola, D. Trivellato, M. Laitaoja, J. Jänis, I. C. Felli, M. D'Onofrio, L. Mollica, G. Giachin and M. Assfalg, *Proc. Natl. Acad. Sci. U. S. A.*, 2025, **122**, e2425831122.
54 O. V. Krokhin, R. Craig, V. Spicer, W. Ens, K. G. Standing, R. C. Beavis and J. A. Wilkins, *Mol. Cell. Proteomics*, 2004, **3**, 908–919.