In-Depth Characterization of Apoptosis N-terminome Reveals a Link Between Caspase-3 Cleavage and Post-Translational N-terminal Acetylation

Rawad Hanna, Andrey Rozenberg, Daniel Ben-Yosef, Tali Lavy, and Oded Kleifeld*

Faculty of Biology, Technion-Israel Institute of Technology, Technion City, Haifa 3200003, Israel.

Correspondence: okleifeld@technion.ac.il

Key words:

Peptidyl-Lys metalloendopeptidase; N-terminomics; N-terminal acetylation; Caspase-3; degradomics

Summary

Proteins' N-termini contain information about their biochemical properties and functions, and they can undergo co- or post-translational modifications, as well as be processed by proteases. To expand the coverage of the N-terminome, we have developed LATE (LysN Amino Terminal Enrichment), a method that uses selective chemical derivatization of α-amines to isolate the N-terminal peptides. We applied LATE in combination with other N-terminomic method to study caspase-3 mediated proteolysis both *in vitro* and during apoptosis in cells. This enable us to identify many unreported caspase-3 cleavages, including some that cannot be identified by other methods. Furthermore, we find direct evidence that some of the caspase-3 generated neo-N-termini can be further modified by Nt-acetylation. This provides a global picture of the caspase-3 degradome and uncovers previously unrecognised functional crosstalk between post-translational Nt-acetylation and caspase proteolysis pathways.

Introduction

Proteases are typically described as degradative enzymes that disassemble unwanted or damaged proteins to generate amino acids for the cell (López-Otín and Bond, 2008; Lopez-Otin and Overall, 2002). Many proteases, however, are very precise and specific enzymes, and their activity produces new polypeptides for different purposes (Lopez-Otin and Overall, 2002; Schlage et al., 2017). Such specific proteolytic cleavages are irreversible post-translational modifications. They are crucial for cellular processes and therefore tightly regulated (Marino et al., 2015; Sukharev et al., 1997; Tafani et al., 2002; Zeeuwen, 2004). Thus, characterization of protease activity and substrate specificity are essential for completing the functional annotation of any proteome.

Mass-spectrometry (MS) based proteomics is one of the most used methods for system-wide protein studies. Despite that, there are plentiful challenges to overcome when MS is used to identify and quantify proteolytic substrates (Rogers and Overall, 2013). These challenges mainly stem from the low abundance of proteolytic fragments that can be easily masked by other peptides. Moreover, background *in vivo* proteolysis events and the transient state of protein synthesis and degradation also contribute to the system complexity when studying a specific protease activity or treatment-directed proteolysis. Over the past decades, various approaches have been developed for reliable detection of proteolytic fragments by separating them from the background proteome (Luo et al., 2019). N-terminomics methods highlight the N-terminal fragment of every protein (N-terminome) usually

by utilizing a series of chemical reactions to manipulate the desired subgroup. Negative selection methods such as combined fractional diagonal chromatography (COFRADIC) (Gevaert et al., 2002), terminal amine isotopic labelling of substrates (N-TAILS) (Kleifeld et al., 2010), and hydrophobic tagging-assisted N-termini enrichment (HYTANE) (Chen et al., 2016) are based on the depletion of internal peptides to enhance the identification of proteolytic products neo-N-termini and original N-termini including naturally modified ones. Positive selection techniques are based on the enrichment of protein fragments containing unblocked N-terminal after tagging them with biotin (Mahrus et al., 2008), and therefore they cannot be used to study the majority of eukaryotic N-terminal peptides that are naturally blocked (Griswold et al., 2019; Mahrus et al., 2008).

Each method has advantages and limitations weighted by its reliability, sensitivity, availability, cost, and time consumption. For example, COFRADIC can be applied to study the N-terminome with low sequence biases using commercially available reagents, but it requires a long analysis time and therefore extensive instrumental usage which can make it difficult to analyze many samples at once (Gevaert et al., 2002; Luo et al., 2019). N-TAILS, on the other hand, is simpler and more straightforward but requires the use of a specific polymer (Kleifeld et al., 2010, 2011). A recently developed method based on hydrophobic tagging akin to HYTANE, termed high-efficiency undecanal-based N-termini enrichment (HUNTER) (Weng et al., 2019) utilizes cheaper reagents and has shown to be useful for small samples. Most of the negative selection N-terminomics approaches require primary amine labelling and thus involve blocking of lysine ε-amines. As a result, these methods cannot use proteases with lysine cleavage specificity, which significantly restricts their application in annotation of the N-terminome.

The N-terminus of the nascent polypeptide chain as translated from its open reading frame (ORF) initiation site is the first to undergo modifications. Such modifications include N-terminal methionine excision and acetylation (Nt-acetylation) which usually take place co-translationally when the protein is still short and bound to the ribosome (Giglione et al., 2015). For the majority of proteins, the N-terminal methionine is removed by methionine aminopeptidases (Jonckheere et al., 2018). Nt-acetylation is catalysed by N-terminal acetyltransferases that utilizes acetyl-CoA as the donor for the transfer of the activated acetyl moiety to the N-terminus of the protein (Ree et al., 2018). It is estimated that ~80% of human proteins are modified by N-terminal acetylation (Nt-acetylation). By altering the N-terminal charge of a protein, Nt-acetylation can affect a range of its properties including stability, folding, protein-protein interactions, and subcellular targeting (Ree et al., 2018).

To improve the coverage of the N-terminome and to better chart proteolytic processing and protein N-terminal modifications, we developed LysN Amino Terminal Enrichment (LATE), an N-terminal enrichment strategy based on LysN digestion and specific N-terminal α-amine isotopic labelling (Figure 1B). We employed LATE in parallel to the HYTANE N-terminomics method to study the central cellular apoptosis protease, caspase-3. By applying this approach to a set of cell-based apoptosis experiments and controlled *in vitro* studies, we identified many reported as well as novel caspase-3 cleavages. Using this comprehensive N-terminal peptide mapping, we

found that caspase-3 cleavage can generate new sites for post-translational N-terminal acetylation. This is the first demonstration of functional crosstalk between neo-Nt-acetylation and proteolytic pathways.

Results

Most proteomics studies aimed at studying proteolytic processing and modification of protein N-termini rely on negative selection methods. These N-terminomics methods are based on blocking/labelling of lysine residues which prevents the use of proteases with lysine cleavage specificity, such that even though they utilize trypsin digestion, the resulting peptides are only those derived from cleavage at arginine residues (thus resembling ArgC digestion). This restriction limits the identification potential considerably: for example identification of ORF N-termini is possible for only less than 50% of the human proteome (Figure S1). While repeating the N-terminomics experiments with proteases other than trypsin can improve the coverage, there are still a few thousand proteins' N-termini that can only be identified by a lysine-specific protease (Figure S1). A major use of N-terminomics methods is for the identification of proteolytic cleavage sites. For example, in the human proteome, there are 134 putative cleavages of caspase-3 the apoptosis executioner protease whose canonical cleavage motif is DEVD↓X (Walsh et al., 2008). Of these, only 73 (54%) can be identified by N-terminomics methods that are based only on cleavages at arginine residues. The number of identified caspase-3 cleavage sites can be considerably expanded to 106 (~80%) by also using peptides generated from cleavage at lysine residues following LysN digestion (Figure 1A). Aside from the sequence coverage, it was also shown in previous studies that peptides with a basic residue closer to the N-termini or the C-termini have different chromatographic characteristics, which contributes to the peptide diversity between LysN and trypsin (ArgC-like) peptides (Tsiatsiani et al., 2016). Based on these considerations, we developed LATE, a new method utilizing LysN digestion followed by specific N-terminal α-amines blocking while minimizing lysine modification, then exploiting the free lysine ε-amine to deplete internal peptides (Figure 1B). After protein extraction and digestion, α-amines are targeted using reductive dimethylation in acidic buffer (Qin et al., 2012). At this point, the samples can optionally be mixed or run individually which can be useful for quantifying naturally modified N-termini. Internal peptides containing lysine residues are depleted by adding undecanal that reacts with lysine ε-amine and aid to separate them on a C18 column based on the added hydrophobicity (Chen et al., 2016; Weng et al., 2019).

A key element in the efficiency of LATE is the specific dimethylation of peptides' N-termini (α-amines) without the labelling of lysine residues (ε-amines). The conditions for this step were optimized following recent reports (Koehler et al., 2013; Qin et al., 2012). LATE labelling efficiency was tested on the pre-enriched sample where we observed labelling of free N-terminal α-amines with >95% efficiency, while only <5% of lysine side chain amines were dimethylated (Figure 1C and Table S1). Internal peptides that have Number of Enzymatic Termini (NET) = 2, those being full LysN peptides, dominated the pre-enrichment sample with more than 80% (Figure 1D) accompanied by a low percentage of ORF N-terminal (i.e. that begin at positions 1 or 2 of the predicted

ORF) and neo-N-terminal peptides. Around 10-fold enrichment of ORF N-terminal and neo-N-terminal was achieved after applying LATE, with the incidence of internal peptides decreasing to less than 10%. The same enrichment trend was observed in our HYTANE sample (Figure S2), similarly to the previous reports (Chen et al., 2016; Weng et al., 2019). As predicted, application of LATE allowed identification of a significant number of dimethylated N-terminal peptides that were not identified by HYTANE (Figure 1E). A similar trend albeit with a reduced magnitude was observed for peptides with acetylated N-termini (Figure 1E). Similar results were obtained when LATE and HYTANE were applied to the *E. coli* N-terminome (Figure S3). In these two experiments LATE improved the overall coverage of the N-terminome, but the number of identified peptides with LATE was slightly lower than the number of peptides identified by HYTANE. We noticed that this reduction was more pronounced for acetylated N-termini compared to dimethylated ones, as this may be due to the stronger hindrance of N-terminal amine ionization by the covalently bound acetyl group (Cho et al., 2016). We concluded that LATE is better suited for studying proteolysis rather than the natural acetylation of N-terminome, yet it provides valuable information and unique identifications on both.

Next, we applied LATE to study caspase-3 proteolytic activity *in vitro*. The intracellular proteome of HCT116 cells was extracted under native conditions and incubated with recombinant caspase-3 or buffer control at 37°C. Equal aliquots were sampled at 1, 6, or 18 hours and were subjected to N-terminal enrichment using LATE and HYTANE (Figure 2A). Caspase-3-treated samples and control samples were labelled with heavy and light dimethylation, respectively. Successful enrichment of the N-terminome was obtained with both LATE and HYTANE (Figure S4). Peptides with a high relative abundance ratio (caspase-3/control) are putative products of caspase-3 cleavage. The distribution of the caspase-3/control abundance ratio in the $\log_2$ scale was approximately normally distributed around zero (i.e., no change), with the frequency of peptides with a high caspase-3/control abundance ratio increasing overtime in both methods (Figure 2B). To distinguish the background proteolysis from the caspase-3 cleaved proteins, we considered only peptides that had a caspase-3/control ratio higher than 2 (Log_2=1) as an indication for putative caspase-3 cleavages. Following this, we identified over 900 different cleavage sites after 18 hours of incubation in each one of the methods with more than 200 sites identified at all time points (Figure 2C, and Tables S2 and S3). Caspase-3-mediated proteolysis has been studied extensively and its specificity is well defined (Lüthi and Martin, 2007; Walsh et al., 2008; Zhan et al., 2002). The sequence logo of the 8 amino acids (P4-P4') flanking the reported caspase-3 cleavages in the MEROPS peptidase database (Rawlings et al., 2012) shows a very similar motif to the logos obtained from our data (Figure 2E). With LATE we were able to identify 685 additional putative caspase-3 cleavage sites that could not be identified with HYTANE (Figure 2D), hence by combining the results of both methods we achieved a ~55% increase in the total number of putative cleavages reaching a total of 2049 putative cleavages enriched in the caspase-3 treated samples (cleavage after Asp and caspase-3/control abundance ratio $\geq$2).

Of these putative cleavages, 712 (35%) matched previously reported caspase-3 cleavages in TopFind (Fortelny et al., 2015), CASBAH (Lüthi and Martin, 2007), DegraBase (Crawford et al., 2013), MEROPS (Rawlings et al.,

2012) and a recent caspase-3 N-terminomics study (Araya et al., 2021) while 1337 (65%) were not reported so far. Interestingly, the proportion of unreported/novel putative caspase-3 cleavages obtained with LATE was higher than with HYTANE (74% vs 63%, respectively) (Figure 2D, Table S2 and Table S3). We were able to identify 91 additional putative caspase-3 cleavage sites with Glu at P1 (Tables S2 and S3), consistent with caspase-3 ability to cleave after Glu residues (Seaman et al., 2016). After combining LATE and HYTANE we identified 2613 putative cleavage sites with caspase-3/control abundance ratio ≥2, 816 were previously reported and 1797 sites are reported here for the first time (Figure 2F and Tables S2 and S3). To evaluate the additional contribution of LATE to the results that can be achieved by ArgC-based N-terminomics methods like HYTANE, we compared the sequences of 7 to 35 amino acids downstream to the identified cleavage sites by each method. HYTANE requires to have an arginine at one of these 7-35 amino acid segments while LATE requires a lysine at these positions. As shown in Figure 2G, more than 50% of LATE identifications consisted of sequences containing Arg residues outside positions 7 to 35 downstream to the cleavage site and a similar trend was found in the location of Lys residues in the peptides identified by each method (Figure S5). Hence LATE is very useful for the identification of cleavage sites that cannot be found with HYTANE and vice versa.

Next, we used the same combined N-terminomics approach to study caspase-3 substrates in cells. To this end, HCT116 cells were transfected with a bicistronic caspase-3/GFP plasmid and FACS-sorted to generate caspase-3 overexpressing cells (GFP-positive) and regular cells (no GFP). Caspase-3 activity was induced by ABT-199 in both populations (Figure 3A). We used LATE and HYTANE to characterise the ORF- and neo-N-terminal peptides in those cells. In this way, changes in proteolytic processing between these cell populations can be directly attributed to caspase-3. The majority of identified peptides were ORF N-terminal peptides, and most of those were acetylated (Figure 3B) with ~57% of the total identifications in HYTANE and ~43% in LATE. Yet, the number of neo-N-terminal peptides identified was similar in both methods (Figure 3B) and the sequence logos of those peptides were similar to each other with strong dominance of Asp at the P1 position (Figure 3C). The logos are very similar to those we obtained for caspase-3 *in vitro* and the caspase-3 sequence logo based on MEROPS (Figure 2E). This indicates the dominance of caspase-3 processing among the neo-N-terminal peptides. Excluding the N-terminal peptides of ORFs and N-terminal peptides resulting from known protein processing events, such as signal peptide or prodomain removals, we detected 2234 neo-cleavage sites for 1096 different proteins. Although with LATE fewer PSMs were obtained than with HYTANE (Figure 3B), it enabled the identification of more neo-N-terminal peptides (1324) than HYTANE (1105) (Figure 3D and Tables S4 and S5).

To chart the changes in the N-terminome we combined the identifications obtained by both methods and considered only those that were identified at least twice. Altogether, 1086 neo-cleavage sites were identified and from these, 508 (~47%) were cleavages after Asp or Glu corresponding to the caspase family specificity (Figure 3E). Of these putative caspase cleavages, 215 (~42%) were at sites that were not reported so far in proteomic studies of caspases (Figure 3E and Tables S4 and S5). A closer look at the Log2(caspase-3/control) abundance ratios of the 293 cleavage sites previously reported in other proteomics studies shows that while a significant

amount of these Log2-ratios is bigger than 1, their majority is within the range of 0-1 (Figure S6). Caspase-3 was activated by the ABT-199 (Souers et al., 2013) in both the caspase-3 over-expressing and control cells, hence there is expected to be no difference between the two treatments for substrates that are cleaved effectively by the endogenous caspase-3. Therefore, in this work, we also consider cleavages without large fold change as putative caspase-3 cleavages. Of note, despite the difference in background proteolysis between the cell-based (Figure 3A) and the *in vitro* experiments, when we compared the abundance ratios of the same substrate cleavage across the two experiments, they showed similar trends (Figure S7). Moreover, for substrate winnowing of caspase-3-related cleavages, we filtered the data further and considered only cleavages that show at least a 2-fold increase in the cells that overexpressed caspase-3 (Figure 3F). We identified several putative caspase-3 cleavages that have not been reported previously, either in the various studies related to apoptosis or in caspase-3-related proteomics studies (Araya et al., 2021; Crawford et al., 2013; Fortelny et al., 2015; Lüthi and Martin, 2007) (Figure 3G and 3F and Table S6). Of the 353 cleavages that show significant change, 286 (over 80%) were with caspase specificity. Of these, 164 (57%) were at sites that were previously reported, and 124 (43%) are at sites that were not reported. From these unreported 124 sites we selected several substrates and demonstrated that they are cleaved by caspase-3 *in vitro* (Figure 3G and Table S2 and S3). All of the novel substrates and their cleavage sites are shown in Table S6. Of note, one of the unreported cleavage sites identified in all of our experiments, exclusively by LATE, is a cleavage site at position 70 of caspase-3 itself (Figure 3G).

We used gene ontology (GO) enrichment analysis to compare the list of unreported putative caspase-3 substrates that we identified to previously reported substrates from apoptosis and caspase-related proteomic studies. The new putative substrates show a pattern highly similar to the proteins previously reported to be cleaved caspase-3, thus demonstrating that the new cleavages come from substrates similar to the already known ones (Figure S8). The number of potential caspase-3 cleavages and substrates identified in the cell-based experiments was smaller than that obtained in the *in vitro* experiments. In order to understand whether there is a structural basis for the differences between the cleavages obtained in the cell-based and *in vitro* experiments, we compared the secondary structure contexts and solvent accessibility around the cleavage sites. This comparison revealed very similar secondary structure contexts for the two experiments (Figure S9A) but a significant difference in the relative solvent accessibility Figure S9B), with the cleavages identified in cells occurring in more exposed regions than those obtained *in vitro*. This might reflect the impact of cell lysis in the *in vitro* experiments that alters some protein structures and disrupts cellular compartmentalization that is kept under control in the cell-based experiments. These results demonstrate the benefits of our cell-based combined approach for the discovery of novel caspase-3 cleavages.

One of the advantages of the negative selection N-terminomics methods is that in addition to the enrichment of the neo-N-terminal peptides, these methods also allow the characterization of the ORF N-terminal peptides (Figures 1E and 1F) (Kleifeld et al., 2010). We used the same approach applied to the neo-N-termini peptides and combined the ORF N-terminal peptides identification obtained by LATE and HYTANE. In our search we

included these known modifications of protein N-termini: removal of the initiator methionine (Giglione et al., 2004, 2015) and Nt-acetylation (Ree et al., 2018). Together, the ORF N-terminal peptides of 2349 proteins were identified (Table S7). Similarly to published data (Arnesen et al., 2009), 74% of the peptides were Nt-acetylated, 26% with free N-termini (Nt-free) and 70% of all peptides underwent methionine excision regardless of their Nt-acetylation status (Figure 4A). Additionally, the frequencies of the first and second amino acids in Nt-acetylated and Nt-free ORF peptides (Figure S10) were similar to those found in other proteomic studies (Arnesen et al., 2009).

Next, we checked if the nature of the ORF N-terminal modification affects the ORF N-terminal stability and susceptibility to cleavage by caspase-3. As can be seen in Figure 4B and Figure S11, the abundance ratio of Nt-acetylated ORF peptides was distributed around 0 regardless of the presence of Asp or Glu in their sequence. A similar trend was seen for Nt-free ORF peptides that do not contain potential caspase-3 cleavage sites. However, Nt-free ORF peptides with Asp or Glu in their sequence were more susceptible to cleavage by caspase-3. and many of their ratios deviated from a normal distribution and demonstrate lower abundance ratios in caspase-3 overexpressing cells relative to the control cell. This supports the notion that Nt-acetylation might have a protective and stabilizing role as was suggested in some studies (Mueller et al., 2021). Nt-acetylation is defined mostly as co-translational modification, but there are a few examples of post-translational Nt-acetylation (Aksnes et al., 2015; Drazic et al., 2018; Helbig et al., 2010; Helsens et al., 2011). To investigate the presence of post-translation Nt-acetylation that are not on protein ORF N-terminal, we checked the N-terminal peptides that were generated following known proteolytic processing of proteins. These processing events include signal, transit, or pro peptides removal and peptide or precursor protein chain cleavages that are part of protein maturation or activation. We based our analysis on protein processing annotation in Uniprot (The UniProt Consortium, 2021). By combining LATE and HYTANE results we identified 329 reported post-translation proteolytic processing sites of 262 proteins (Table S8). Of these, 11 neo-Nt-peptides from 11 proteins were found with Nt-acetylation (Figure 4C). The neo-acetylated-Nt-peptides identified were mainly from the mitochondria and relate to the removal of the mitochondria targeting sequence (mitochondria transit peptide). This type of proteolytic processing is done after the protein is imported into the mitochondria and reaches the mitochondrial matrix (Friedl et al., 2020). Next, we looked for additional hints for post-translational neo-Nt-acetylation that occur following other proteolytic cleavages. We found 371 neo-Nt-acetylation motifs in 329 different proteins (Table S9). To check if there are neo-Nt-acetylation sites in the list that are undoubtedly dependent on caspase-3 cleavage, we used the winnowing workflow shown in Figure 5A. Consequently, we focused on the proteolytic cleavage and neo-Nt-acetylation of nascent-polypeptide-associated complex A (NACA). The MS/MS identifications of this neo-Nt-acetylated peptide

by LATE and HYTANE were with high scores, high peptide sequence coverage and provide conclusive evidence for the presence of Nt-acetylation on Ser43 (Figure S12). The NACA is a subunit of the hetero-dimeric ribosome-associated (Rospert et al., 2002) nascent-polypeptide-associated complex (NAC). The NAC complex is highly

conserved among eukaryotes, and when associated with the ribosome it interacts with nascent polypeptide chains and mediates their sorting to different cellular compartments (Jomaa et al., 2022). In addition to the neo-Nt-acetylation of the peptide starting at Ser43, we found several neo-N-terminal peptides of NACA (Figure 5B upper part). The majority of these cleavages were also reported in other apoptosis-related studies (Crawford et al., 2013; Fortelny et al., 2015; Lüthi and Martin, 2007). Interestingly, in the cell-based experiments, this peptide was identified almost exclusively in its acetylated form with a total of 62 PSMs vs 2 PSMs for the free form, while in the *in vitro* experiments the total PSM numbers of both forms were similar. The location of this cleavage is within the N-terminal unstructured region of NACA which makes it suitable for caspase cleavage. Based on a recent structure of the NAC complex bound to the ribosome (Jomaa et al., 2022) this region remains accessible even when NACA is in complex with BTF3 (nascent-polypeptide-associated complex B - NACB) and the ribosome (Figure 5C). The identification of NACA Nt-acetylated-Ser43 peptide occured *in vitro* only after the addition of active caspase-3, allowing us to study its generation process. We used HCT116 cell lysate and half of it was subject to dialysis to reduce acetyl-CoA concentration and prevent the acetylation reaction. Following the addition of active caspase-3, aliquots from both dialyzed and non-dialyzed lysates were sampled at different time points and subjected to N-terminal enrichment and MS analysis. As shown in Figure 5D, the dialysis did not affect caspase-3 cleavage and the appearance of NACA neo-Nt-Ser43. This peptide was already identified 30 minutes after caspase-3 addition and was generated at an identical rate and intensity in both samples. The NACA neo-Nt-acetylated Ser43 peptide appeared at a later time point (6 hours) but only in the non-dialyzed sample (Figure 5E). Together with the rest of our results, this reveals for the first time the presence of post-translation Nt-acetylation that is directly linked to caspase-3 proteolytic processing (Figure 5F).

Discussion

The use of multiple proteolytic enzymes in bottom-up proteomics studies is known to improve the number of protein identification and sequence coverage (Giansanti et al., 2016; Swaney et al., 2010). Similar improvements were also shown in several terminomics studies that combined trypsin with other proteases while using the same terminal peptides enrichment methodology (Bell et al., 2019; Zhang et al., 2018). Here we describe LATE, a simple methodology that is based on digestion with LysN for N-terminome characterization. LysN digestion has been used for the identification of ORF Nt-acetylated peptides (Du et al., 2020; Zhang et al., 2009) but not for free ORF N-terminal peptides nor neo-N-terminal peptides generated by proteolysis. A key element of LATE is the N-terminal-specific labelling by dimethylation, which also allows the use of isotopic labels and quantitative analysis. Unlike the few other reported N-terminal specific labelling reagents, like TMPP (Bertaccini et al., 2013) or 2-pyridinecarboxyaldehydes (Griswold et al., 2019; MacDonald et al., 2015), the required reagents are inexpensive and the labelling efficiency is very high and consistent. We demonstrated the usefulness of LATE combined with HYTANE, in the investigation of caspase-3 mediated cleavages during apoptosis. This subject has been extensively studied, particularly by terminomics methodologies (Agard et al., 2012; Demon et al., 2009; Julien et al., 2016; Mahrus et al., 2008; Seaman et al., 2016). Despite this, we identified many novel caspase-3

cleavage sites and substrates as well as some known substrates whose cleavage by caspase-3 had not been previously identified by proteomic analysis. Such known substrates are desmoplakin (DSP) and plakoglobin (JUP) which have been known as caspase-3 targets for over two decades (Brancolini et al., 1998; Weiske et al., 2001) but their cleavage sites had not yet been identified by proteomics. We identified DSP and JUP cleavage sites as part of a large-scale study for the first time. The novel substrate list includes several proteins that are understudied in the context of apoptosis. For example, the unconventional prefoldin RPB5 interactor 1 (URI1) that is important to DNA stability (Parusel et al., 2006) and affects apoptosis upon its phosphorylation (Djouder et al., 2007), and Transcription factor 25 (TCF25 or NULP1) which binds the apoptosis inhibitor XIAP (Steen and Lindholm, 2008) and the overexpression of which induces cell death. Our combined system-wide *in vitro* and cell-based experiments allowed us to validate these proteins as caspase-3 substrates. In addition, this thorough N-terminomics characterization enabled us to monitor not just proteolytic cleavage but also changes in protein ORF N-termini. In this respect, our results indicate that Nt-acetylation protects proteins' N-terminal regions from proteolytic cleavage by caspase-3 during apoptosis. In mammalian proteins, Nt-acetylation can be seen in more than 70% of ORF N-termini, yet its physiological functions are not entirely understood. Nt-acetylation affects the charge of proteins' N-termini hence it would be expected to modify their function. Initially, it was suggested that a general function for Nt-acetylation is the protection of proteins from proteolytic degradation (Hershko et al., 1984; Jörnvall, 1975). Later on, it was suggested that under certain conditions, Nt-acetylation might also be a degradation signal (degron) that marks the protein for ubiquitinylation and its destruction by the proteasome (Hwang et al., 2010). We showed that in the context of apoptosis, the presence of Nt-acetylation has a protective effect and prevents cleavage around the protein N-terminal by caspase-3.

Most Nt-acetylation events are thought to be co-translational, but several post-translational Nt-acetylation events were reported too. Among them are the Nt-acetylation of secreted proteins in Apicomplexa (Nyonda et al., 2022), cleaved proteins in yeast by an unknown protease (Helbig et al., 2010; Helsens et al., 2011), transmembrane proteins in the Golgi (Aksnes et al., 2015) and cleaved actin (Drazic et al., 2018). Nt-acetylation of nuclear-coded proteins imported to the mitochondria has been described but only co-translational ORF Nt-acetylations were identified (Bienvenut et al., 2012; Vaca Jacome et al., 2015). Here we uncover the Nt-acetylation of several mitochondrial proteins that occur after their transit peptide removal in the mitochondrial matrix. Nt-acetylation of imported nuclear-encoded proteins after the cleavage of their transit peptides was identified in chloroplasts (Bienvenut et al., 2012; Huesgen et al., 2013; Zybailov et al., 2008). Both organelles are thought to have a prokaryotic origin, yet Nt-acetylation is rare in bacteria (Bienvenut et al., 2012). This raises the question of which N-acetyltransferase catalyzes the acetylation of mitochondrial proteins. Recently a family of dual lysine and N-terminal acetyltransferases was identified and several of its members were shown to reside within plastids (Bienvenut et al., 2020). Mammalian mitochondrial N-acetyltransferase has not been not identified yet, but the cytosolic N-acetyltransferase NAA30 has been linked to trafficking to organelles and has been shown to mediate

acetylation of several mitochondrial proteins and its knockdown has led to mitochondrial fragmentation (Van Damme et al., 2016).

We also revealed several caspase-3-generated neo-N-termini that undergo post-translational Nt-acetylation. We focused on and validated the caspase-3-dependence of NACA neo-Nt-acetylation that was dominant in all of our experiments and analyses. Aside from NACA's role as part of a ribosome-associated complex, it was shown to function as a transcriptional coactivator regulating bone development and hematopoiesis. Hence, its cleavage and Nt-acetylation might have implications on transcription. As part of a ribosome-associated complex, NACA binds to nascent proteins that lack a signal peptide motif as they emerge from the ribosome, while blocking their interaction with the signal recognition particle and preventing mistranslocation to the endoplasmic reticulum (Jomaa et al., 2022). Downregulation of NACA, by hypoxia or RNAi, can lead to the initiation of an ER stress response that is followed by caspase-dependent apoptosis (Hotokezaka et al., 2009, 2015). We suggest that proteolytic cleavage of NACA's N-terminal may hamper this endoplasmic reticulum gatekeeping mechanism. Considering Nt-acetylation's stabilizing effect on protein termini, NACA Nt-acetylation might maintain a malfunctioning NAC complex and lead to ER stress and apoptosis. Future studies could examine the time-course and cross-talk of these post-translational modifications of NACA during apoptosis.

Together, our data reveal an interplay between caspase-mediated proteolysis and Nt-acetylation. A similar functional interplay was described before for caspase-mediated proteolysis and phosphorylation (Dix et al., 2012; Seaman et al., 2016). It was shown that caspase cleavage can expose new sites for phosphorylation and that phosphorylation sites can enhance caspase cleavage in their vicinity (Dix et al., 2012). We found that caspase-3 cleavage can unveil new sites for post-translational Nt-acetylation and conversely that Nt-acetylation protects proteins from caspase cleavage. It would be interesting to combine N-terminomics using LATE and HYTANE with phosphoproteomics to investigate if there is tripartite crosstalk between caspase-mediated proteolysis, phosphorylation, and Nt-acetylation and if these modifications outcomes are additive or antagonistic to each other.

Significance

N-terminomics methods are proteomics methods aimed at the characterization of protein N-terminal sequences. These include the mature proteins' N-terminal sequences as well as N-termini generated following proteolysis. We developed a method to isolate N-terminal peptides that is based on proteome digestion with the lysine-specific protease LysN followed by α-amine specific labelling. Our method, called LATE, enables the identification of N-terminal peptides that can not be identified by other N-terminomics methods. LATE can be used to study cellular processes and proteases in vitro, in cell-based experiments, or in tissue-based experiments By using LATE along with HYTANE, a common N-terminomics method, we substantially improved N-terminome coverage. Applying this approach to study caspase-3 mediated proteolytic changes during apoptosis revealed many unreported caspase-3 cleavage sites that were also validated *in vitro*. Furthermore, these studies

revealed, for the first time, an interplay between caspase-mediated proteolysis and Nt-acetylation. We show that Nt-acetylation protects proteins' N-termini from cleavage by caspase-3 and on the other hand, caspase-3 cleavage can generate new sites for post-translational Nt-acetylation.

Materials and Methods

Chemicals and reagents All chemicals and reagents were purchased from Sigma-Aldrich unless specified otherwise. Undecanal was purchased from TCI, Acetonitrile was purchased from JT Baker, formic acid was purchased from Biolabs ltd, Guanidium-HCl and HEPES were purchased from the Spectrum, and ethanol was purchased from LiChorsolv. HPLC-grade water was purchased from Thermo. All solvents used are MS-grade.

Plasmids: The construct for *E.coli* expression of WT caspase-3 (pET23b-Casp3-His) was obtained from Addgene (plasmid 11821). The construct for mammalian expression of caspase-3/GFP was a gift from Dr. Yaron Fuchs Lab (Technion).

***E. coli* Proteome Preparation**: *E. coli* DH5α cells were grown in LB until an optical density of OD_{600nm}=0.5. Cells were harvested by centrifugation at 6000 g for 10 min at 4□ and washed twice with 50 mM PBS pH 7.4. The cells pellet was resuspended in 50 mM HEPES, 8M Gu-HCl pH 7.5, and lysed by heating to 95□ for 10 min. The lysate was clarified by centrifugation at 15000 g for 10 min at 4□, and the supernatant was collected.

Caspase-3 expression and purification: The expression construct of pET23b-Casp3-His6 was extracted from DH5α bacterial culture purchased from Addgene (Plasmid #11821). BL21-CodonPlus (DE3)-RIPL competent cells were transformed with the expression construct and plated on LB-agar supplemented with ampicillin (100 µg/ml) and chloramphenicol (34 µg/ml). A single colony of the transformants was inoculated into a 5 ml LB supplemented with the above antibiotics and cultured at 37°C and 200 rpm shaking, overnight. On the next morning, the overnight culture was inoculated into 50 ml LB supplemented with the above antibiotics and grown for 3 hours at 37°C and 200 rpm shaking. The culture was transferred to 1L of LB in a baffled flask and cultured. At OD_{600nm}=0.6, the temperature was lowered to 17°C and expression was induced by the addition of IPTG (1 mM). Expression was terminated after 16 hours by harvesting the cells (centrifugation for 20 min at 4000g at 4°). The harvested cells were resuspended in 50 mL ice-cold buffer A (200 mM NaCl, 20 mM Tris-HCl, pH 7.5) and stored at -80°C. The cells were lysed by three freeze-thaw cycles (liquid nitrogen/RT) followed by disruption of the cells using a cell disrupter (OS Model, Pressure Biosciences). The lysate was clarified by centrifugation (20,000 g, 4°C, 20 minutes) and the clear lysate was purified on HisTrap HP 1 ml affinity column (Cytiva) by FPLC system (NGC™□, BioRad). The protein was eluted from the column using a linear gradient (Elution buffer: 0-1M imidazole, 200 mM NaCl, 20 mM TRIS-HCl pH 7.5). The fractions were analyzed by SDS-PAGE. Fractions containing relatively pure caspase-3 were pooled together and dialyzed against 20 mM HEPES 7.5, and 80 mM NaCl. Aliquots of the protein were stored at -80°C.

Cell culture: Human colorectal carcinoma cells (HCT116) were grown in RPMI supplemented with 10% fetal calf serum (FCS), 2mM L-Glutamine, and 1% Pen-Strep solution. Immortalized aneuploid human keratinocyte cells (HaCaT) were grown in DMEM supplemented with 10% FCS, 2mM L-glutamine, and 1% Pen-Strep solution, in 10 cm plates. When the cells reached 80% confluence, the medium was discarded, and the cells were washed three times on the plate with 10 ml PBS before detachment or lysis.

For cell-based caspase-3 experiments, HCT116 were grown in a 10 cm plate until they reached 40% confluence, then the medium was changed to starvation medium (RPMI 2mM L-glutamine, 1% Pen-Strep solution without serum) for 2 hours. The transfection mix was prepared by mixing 15 µl Polyjet (SignaGen) with 6 µg of caspase-3/GFP bicistronic vector each prepared in 250 µl starvation media, the mix was left for 10 min at room temperature before adding it to the starved cells. Eight hours following transfection the media was changed back to RPMI supplemented with 10% FCS, 2 mM L-glutamine, 1% Pen-Strep solution, and GFP fluorescence was observed to estimate the efficiency. Next, the cells were split and grown to 80% confluence before treatment with or without sorting.

Cell lysis and protein extraction: For extraction under denaturative conditions, 1ml of 95□ pre-heated 50mM HEPES, 8M Gu-HCl pH 8.0 was added to each plate, and cells were scraped down and incubated 5 min at 95□. Next, the cell lysates were sonicated using UP200St with VialTweeter (Hielscher, Germany) at max amplitude, 70% cycle time for 5 minutes. The sonicated lysates were centrifuged at 15000g, 10 min, 4□, and the cleared supernatants were collected. For extraction under native conditions, cells were detached following 5 minutes of treatment with trypsin-EDTA and collected by centrifugation (5000g, 5min, 4□). Cell pellets were washed three times with PBS, and resuspended in 0.5 mL of 150mM NaCl, 20mM HEPES pH 7.5, and lysed by 3 freeze-thaw cycles of 5 minutes in liquid nitrogen followed by 20 minutes on ice. The lysate was clarified by centrifugation at 15000g for 10 min at 4□. Protein concentration was determined by BCA assay (Smith et al., 1985).

***In vitro* caspase experiments:** Clear lysate extracted under native conditions was split into two aliquots. One was left untreated (control) while the other was mixed with recombinant human caspase-3 in a 1:20 w/w ratio. Both samples were kept at 37□ with gentle mixing. After 1, 6, and 18 hours, a fraction of 200 ug protein was taken from the control and caspase-3-treated samples into a clean tube containing an equal volume of 20mM HEPES,8M Gu-HCl pH 7.5 and incubated at 95□ for 5 minutes. For checking *in vitro* Nt-acetylation, the clear lysate was split into two aliquots. One was dialysed by 3 cycles of concentration and 10X dilution at 4□ in ice-cold 20mM HEPES pH7.8, 1mM DTT,1mM EDTA, 100mM NaCl using 3k cut-off Amicon® Ultra 2 mL Centrifugal Filters (Merck). Equal amounts of protein from both dialyzed and undialyzed samples were treated with caspase-3 as described above. Aliquots were taken from each sample after 0.5, 6 and 18 hours. Nt-acetylation samples were subject only to HYTANE.

Cell-based caspase experiments: HCT-116 cells transfected caspase-3/GFP vector were detached with trypsin-EDTA, quenched by an equal volume of media, and centrifuged at 1500g for 5min. Then the cells were

resuspended in 1ml RPMI supplemented with 10% FCS, 2mM L-Glutamine, 1% Pen-Strep solution, and transferred to a 5mL polystyrene round-bottom tube with a cell-strainer cap (Falcon). Unstained negative control cells were also prepared in the same way for population determination during sorting. Sorting was done using BD FACS-Aria™ (BD Bioscience) against GFP fluorescence. Between 1.5 and 2 million cells of GFP-positive (high caspase-3) and negative populations were collected per biological repeat. After sorting, samples were centrifuged at 1500g for 5min, re-suspended in 4 ml rich media (RPMI, 20% FCS, 1% Pen-Strep, 2mM L-Glutamine), and seeded in a 6 cm plate. Cells were grown for another day while the media refreshed after 12 hours. Lastly, samples were treated with BCL-2 inhibitor ABT-199 at a final concentration of 150µM to induce caspase-3 activation for 2 hours, prior to harvesting the cells.

LysN Amino Terminal Enrichment (LATE): One hundred ug protein was re-suspended in 100mM HEPES, 4M Gu-HCl pH 7.5 and reduced with 5mM DTT at 65 for 30 minutes followed by alkylation with 12.5mM chloroacetamide (CAA) for additional 20 minutes at room temperature in the dark. CAA leftovers were quenched with another dose of DTT, and the sample was diluted 1 to 4 with 20mM HEPES pH 8 and the pH was further adjusted to 8.0 if needed using 1N NaOH. Next, LysN (Promega) was added at 1:100 w/w and the samples were digested for 12h at 37. Digestion was terminated by adding a 6% final volume of glacial acetic acid. N-terminal alpha amines specific labelling was done by adding 40mM (final) of $C^{13}D_2$ (heavy) or $C^{12}H_2$ (light) formaldehyde and 20mM (final) of sodium cyanoborohydride at 37 for 10min. Heavy and light labelled samples were quenched by addition of glycine (final concentration 100mM) dissolved in 2% acetic acid and incubation at room temperature for 10min. Next heavy- and light-labelled samples were mixed and desalted using OASIS-HLB 100mg cartridge (Waters). Peptides were eluted with 60% ACN, and 0.1% FA and dried to completeness using CentriVap Benchtop Vacuum Concentrator (Labconco). Next, samples were resuspended in 100µl of 100mM HEPES pH 7.0 with undecanal at 50:1 unedecanal:peptide w/w ratio. Undecenal was dissolved in ethanol to a final concentration of 20mg/ml. The undecanal labelling was carried out at 50 for 2 hours. Undecanal-labelled samples were centrifuged at 20000g for 5 minutes at RT and the supernatants were transferred to a new tube and dried to completeness using CentriVap Benchtop Vacuum Concentrator (Labconco, USA). Dried samples were dissolved in 2% ACN, and 0.1% FA and desalted with OASIS-HLB 100mg cartridge (Waters). Undecanal leftovers and undecanal-modified peptides should remain bound to the column after elution with 60% ACN. Samples were dried and resuspended in 2% ACN, and 0.1% FA prior to LC-MS analysis.

Hydrophobic tagging-assisted N-termini enrichment (**HYTANE)**: One hundred ug protein of each of the compared samples were resuspended in 100mM HEPES, 4M Gu-HCl pH 7.5, and reduced with 5mM DTT at 65 for 30 minutes followed by alkylation using 12.5mM CAA for additional 20 minutes at room temperature in the dark. CAA leftovers were quenched with another dose of DTT. Each sample was labelled by adding 40mM (final) of either $C^{13}D_2$ (heavy) or $C^{12}H_2$ (light) formaldehyde and 20mM (final) of sodium cyanoborohydride and incubating at 37 overnight followed by quenching with 100mM glycine (final) for 1 hour at 37C. Next, heavy and light-labelled samples were mixed and diluted 1 to 4 using 20mM HEPES pH 8.0. The mixed samples were

digested with sequencing grade trypsin (Promega) in a 1:100 w/w ratio, at 37□ overnight. The tryptic digestion was quenched by the addition of FA to a final concentration of 1%. The combined sample was desalted using OASIS-HLP and dried completely in a speedvac. Undecanal-based enrichment was done as described for LATE.

LC-MS: Desalted samples were subjected to LC-MS analysis using a Q Exactive Orbitrap Plus mass spectrometer Q Exactive Orbitrap HF coupled to Easy-nanoLC 1000 capillary HPLC. The peptides were resolved by reversed-phase using a homemade 30 cm long 75µm-diameter capillary, packed with 3.5-µm silica using ReproSil-Pur C18-AQ resin (Dr . Maisch GmbH). Peptides were eluted using a linear 120-minute gradient of 5–28% acetonitrile with 0.1% formic acid, followed by a 15-□minute gradient of 28–95% acetonitrile with 0.1% formic acid, and a 15-minute wash of 95% acetonitrile with 0.1% formic acid (at flow rates of 0.15□µl/min). MS was performed in positive mode using an m/z range of 300–1800 in a data-dependent acquisition mode, a resolution of 70,000 for MS1, and a resolution of 17,500 for MS2; repetitively full MS scans were followed by high energy collisional dissociation (HCD) of the 10 or 20 most dominant ions selected from the first MS scan for Q-Exactive Plus and Q-Exactive HF respectively.

Data analysis: Data-dependent acquisition experiments were analyzed using Trans Proteomic Pipeline (version 6.1 or 5.2) (Deutsch et al., 2010). Searches were done using COMET (version 2020_01 rev 1 or 2018_01 rev 1) (Eng et al., 2013). Searches for acetylated, dimethylated, and methylated N-termini were done separately. Comet results were further validated by PeptideProphet (Keller et al., 2002) at a false discovery rate of less than 1%. XPRESS was used for relative peptide quantification (Han et al., 2001). Mass tolerance for quantification was 20 ppm with The minimum number of chromatogram points needed for quantitation was set to 3, and the number of isotopic peaks to the sum was set to 0. Searches were conducted against *Homo sapiens* proteome containing 73248 sequences (ID:9606, downloaded from Uniprot in June 2019) and supplemented with cRAP sequences downloaded from ftp://ftp.thegpm.org/fasta/cRAP. Bioinformatic analysis of the N-terminome identifications was carried out using in-house python scripts (available for download from Github) and based on Uniprot annotations (The UniProt Consortium, 2021). These scripts were used to combine peptide identification into unified list, aggregate peptide ratios based on the sum of peak areas of different the PSM of the same peptide, map peptide location with the protein sequence and provide secondary structure and relative solvent accessibility for the peptides and its flanking amino acids. Differential amino acid usage logos were created with dagLogo v. 1.32.0 (Ou et al., 2020). Logos were plotted as the percent difference of the amino acid frequency at each site against the background of the Swiss-Port human proteome. Boxplots were created using BoxPlotR (Spitzer et al., 2014). GO term enrichment analysis was performed with topGO v. 2.46.0 (Alexa and Rahnenfuhrer, 2022) and the results were visualized by plotting terms significantly enriched with respect to the whole human proteome (FDR≤0.01) in semantic coordinates obtained with rrvgo v. 1.6.0 (Sayols, 2022). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE partner repository. Login information will be provided upon request.

Experimental Design and Statistical Rationale: Cell-based proteomic analyses were done using triplicate cell cultures (biological replicates). Biological replicates were individual cultures grown separately and prepared separately. Once these cultures reached appropriate confluence levels, they were treated as described. *In vitro* proteomics analyses were done using a protein extract from one cell culture (one biological repeat). This extract was aliquoted and each aliquot was subject to different treatments such as +/- dialysis and/or +/-incubation with caspase-3. These different treatments were sampled at different time points. The quantitative proteomics analysis was based on isotopic labelling. The abundance ratios of each sample (MS run) were normalized against the median ratio and then quantitative data from identical peptide samples were aggregated together by summation. Significant ratio changes were defined based on a fixed threshold as described before (Kleifeld et al., 2010).

Acknowledgements

This research was supported by ISF grants 1623/17 and 2167/17. While writing this manuscript, R.H. was supported by the Neubauer Fellowship.

Author contributions

Conceptualization, O.K and R.H.; Methodology, O.K and R,H.; Investigation, R.H, D.B, A.R and T.L.; Writing – Original Draft, R.H and O.K.; Writing – Review & Editing, R.H, A.R., T.L and O.K.; Supervision – O.K.

Declaration of interests

The authors declare no competing interests.

Figure legends

Figure 1. LATE workflow for improved N-terminome coverage.

A. *In silico* analysis of the "Identifiable" and "Not identifiable" putative caspase-3 DEVD↓X in the human proteome following digestion by trypsin with ArgC-like specificity (red) and LysN (green) **B.** LysN Amino Terminal Enrichment (LATE) workflow for the comparison of two conditions. Extracted proteins of each sample are digested with LysN, and the resulting peptides undergo N-terminal-specific isotopic dimethylation. The samples are then mixed and subject to hydrophobic tagging of lysines which allows for their removal and for the enrichment of the N-terminal peptides **C.** LATE N-terminal dimethylation labelling specificity following LysN digestion. Over 95% of peptides' N-terminal amines were labelled by dimethylation or naturally blocked by acetylation while less than 5% of the lysine side-chain amines were labelled **D**. Peptide-spectrum match (PSM) number of LysN internal peptides (green) versus the PSM number of ORF (burgundy) and neo N-terminal peptides (blue) before (PreLATE) and after LATE **E.** Comparison of Human proteins with free (left) and acetylated (right) N-terminal peptides identified by HYTANE (red) and LATE (green).

Figure 2. N-terminomics of caspase-3 cleavages *in vitro*.

A. *In vitro* experimental scheme. **B.** Ranked ratios of dimethylated peptides showing the accumulation trends of peptides with high Log_2 (caspase-3/control) abundance ratio (>1 as marked by the yellow dotted line) in LATE and HYTANE **C.** Venn diagram of peptides with Log_2 (caspase-3/control)>1 at each time point in LATE and in HYTANE **D.** Venn diagram of the unique cleavage sites with Log_2 (caspase-3/control) ratio >1 in at least one time point. Reported/not reported is based on a comparison to published data **E.** Comparison of Human proteins ORF N-terminal peptides in dimethylated form by HYTANE (red) and LATE (green) **F.** Comparison of Human proteins ORF N-terminal peptides in dimethylated form by HYTANE (red) and LATE (green). **G.** The residue distance distribution of the nearest arginine residue to the identified peptides sequence in LATE (purple) and HYTANE (cyan) experiments.

Figure 3. N-terminomics of caspase-3 cleavages in cells.

A. Cell-based experimental scheme to study caspase-3 mediated cleavages in HCT116 following treatment with ABT-199 **B.** Number of N-terminal peptides identifications following HYTANE and LATE enrichments **C.** Sequence Logo of neo-N-terminal peptides cleavage motifs by HYTANE (top) and LATE (bottom) **D.** Venn diagram of the combined number of neo cleavage motifs identified by HYTANE (red) and LATE (green). **E.** The number of putative caspase cleavage motifs (with P1=D or E) out of the total neo-N-terminal cleavage motifs that were identified. The categorisation of reported/not reported is based on a comparison to published data (TopFind etc). **F.** The number of putative caspase cleavage motifs (with P1=D or E) that had a high caspase-3/control

abundance ratio (Log2≥1) out of the total neo-N-terminal cleavages identified **G.** Novel caspase-3 substrates and their cleavage sites. The cleavage site (caspase-3/control) abundance ratio is indicated by color.

Figure 4. Acetylation of both ORF and neo N-terminal peptides.

A. The ORF N-terminal peptides identified in HCT116 cells overexpressing caspase-3 and their control were categorized based on the presence of initiation methionine (marked as 1st Met) or its removal (marked as 2nd position), the presence of Nt-acetylation ("Nt-Acet"; in shades of blue) or its absence ("Free"; in shades of pink). **B.** The abundance ratio distribution of the ORF N-terminal peptides in cells overexpressing caspase-3 and their control, based on their N-terminal state (free/acetylated) and the presence of Asp or Glu residue in their sequence. Center lines show the medians; box limits indicate the 25th and 75th percentiles; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles; crosses represent sample means; data points are plotted as open circles. n = 128, 808, 133, 620 sample points **C**. Post-translational neo-Nt-acetylation sites that were identified at N-termini generated following known proteolytic processing of precursor proteins. The blue diamonds represent identification in biological repeats or time points.

Figure 5. Neo-Nt-acetylation of NACA of ORF and neo-N-terminal peptides.

A**.** Winnowing of acetylated neo-N-terminal peptides identified in HCT116 cells. **B.** The N-terminal peptides of Nascent polypeptide-associated complex subunit alpha (NACA) identified in HCT116 cells. Cleaved peptides sequences and modifications are shown on the top and the PSMs numbers in the different experiments are on the bottom **C.** The position of NACA cleavage (in yellow and light blue) by caspase-3, before Ser43, lead to the formation of a neo-Nt-acetylated form of NACA. The structure model is based on the structure of the ribosome-nascent chain containing an ER signal sequence in a complex with the NAC complex (PDB:7QWR). The ribosome and RNA are in grey. The full NACA (in dark pink) and BTF3 (NACB) (in light pink) structures are based on AlphaFold (Varadi et al., 2022) prediction and were aligned on top of the original partial structures. **D.** Time-course of *in vitro* NACA cleavage at Ser43 by caspase-3 with (green) or without (orange) lysate dialysis before caspase-3 addition. **E.** Time-course of *in vitro* NACA neo-acetylation at Ser43 after caspase-3 cleavage with (green) or without (orange) lysate dialysis before caspase-3 addition. **F.** Suggest order of events for NACA neo-Nt-acetylation following caspase-3 processing.

Bibliography

Agard, N.J., Mahrus, S., Trinidad, J.C., Lynn, A., Burlingame, A.L., and Wells, J.A. (2012). Global kinetic analysis of proteolysis via quantitative targeted proteomics. Proc. Natl. Acad. Sci. U. S. A. *109*, 1913–1918. https://doi.org/10.1073/pnas.1117158109.

Aksnes, H., Van Damme, P., Goris, M., Starheim, K.K., Marie, M., Støve, S.I., Hoel, C., Kalvik, T.V., Hole, K., Glomnes, N., et al. (2015). An Organellar Nα-Acetyltransferase, Naa60, Acetylates Cytosolic N Termini of Transmembrane Proteins and Maintains Golgi Integrity. Cell Rep. *10*, 1362–1374. https://doi.org/10.1016/j.celrep.2015.01.053.

Alexa, A., and Rahnenfuhrer, J. (2022). topGO: Enrichment Analysis for Gene Ontology (Bioconductor version: Release (3.15)).

Araya, L.E., Soni, I.V., Hardy, J.A., and Julien, O. (2021). Deorphanizing Caspase-3 and Caspase-9 Substrates in and out of Apoptosis with Deep Substrate Profiling. ACS Chem. Biol. *16*, 2280–2296. https://doi.org/10.1021/ACSCHEMBIO.1C00456.

Arnesen, T., Van Damme, P., Polevoda, B., Helsens, K., Evjenth, R., Colaert, N., Varhaug, J.E., Vandekerckhove, J., Lillehaug, J.R., Sherman, F., et al. (2009). Proteomics analyses reveal the evolutionary conservation and divergence of N-terminal acetyltransferases from yeast and humans. Proc. Natl. Acad. Sci. *106*, 8157–8162. https://doi.org/10.1073/pnas.0901931106.

Bell, P.A., Solis, N., Kizhakkedathu, J.N., Matthew, I., and Overall, C.M. (2019). Proteomic and N-Terminomic TAILS Analyses of Human Alveolar Bone Proteins: Improved Protein Extraction Methodology and LysargiNase Digestion Strategies Increase Proteome Coverage and Missing Protein Identification. J. Proteome Res. *18*, 4167–4179. https://doi.org/10.1021/acs.jproteome.9b00445.

Bertaccini, D., Vaca, S., Carapito, C., Arsène-Ploetze, F., Van Dorsselaer, A., and Schaeffer-Reiss, C. (2013). An Improved Stable Isotope N-Terminal Labeling Approach with Light/Heavy TMPP To Automate Proteogenomics Data Validation: dN-TOP. J. Proteome Res. *12*, 3063–3070. https://doi.org/10.1021/pr4002993.

Bienvenut, W.V., Sumpton, D., Martinez, A., Lilla, S., Espagne, C., Meinnel, T., and Giglione, C. (2012). Comparative Large Scale Characterization of Plant versus Mammal Proteins Reveals Similar and Idiosyncratic N-α-Acetylation Features*. Mol. Cell. Proteomics *11*, M111.015131. https://doi.org/10.1074/mcp.M111.015131.

Bienvenut, W.V., Brünje, A., Boyer, J., Mühlenbeck, J.S., Bernal, G., Lassowskat, I., Dian, C., Linster, E., Dinh, T.V., Koskela, M.M., et al. (2020). Dual lysine and N□terminal acetyltransferases reveal the complexity underpinning protein acetylation. Mol. Syst. Biol. *16*, e9464. https://doi.org/10.15252/msb.20209464.

Brancolini, C., Sgorbissa, A., and Schneider, C. (1998). Proteolytic processing of the adherens junctions components beta-catenin and gamma-catenin/plakoglobin during apoptosis. Cell Death Differ. *5*, 1042–1050. https://doi.org/10.1038/sj.cdd.4400443.

Chen, L., Shan, Y., Weng, Y., Sui, Z., Zhang, X., Liang, Z., Zhang, L., and Zhang, Y. (2016). Hydrophobic Tagging-Assisted N-Termini Enrichment for In-Depth N-Terminome Analysis. Anal. Chem. *88*, 8390–8395. https://doi.org/10.1021/acs.analchem.6b02453.

Cho, K.-C., Kang, J.W., Choi, Y., Kim, T.W., and Kim, K.P. (2016). Effects of peptide acetylation and dimethylation on electrospray ionization efficiency. J. Mass Spectrom. *51*, 105–110. https://doi.org/10.1002/jms.3723.

Crawford, E.D., Seaman, J.E., Agard, N., Hsu, G.W., Julien, O., Mahrus, S., Nguyen, H., Shimbo, K., Yoshihara, H.A.I., Zhuang, M., et al. (2013). The DegraBase: a database of proteolysis in healthy and apoptotic human cells. Mol. Cell. Proteomics MCP *12*, 813–824. https://doi.org/10.1074/MCP.O112.024372.

Demon, D., Van Damme, P., Berghe, T.V., Deceuninck, A., Van Durme, J., Verspurten, J., Helsens, K., Impens, F., Wejda, M., Schymkowitz, J., et al. (2009). Proteome-wide Substrate Analysis Indicates Substrate Exclusion as a Mechanism to Generate Caspase-7 Versus Caspase-3 Specificity*. Mol. Cell. Proteomics *8*, 2700–2714. https://doi.org/10.1074/mcp.M900310-MCP200.

Deutsch, E.W., Mendoza, L., Shteynberg, D., Farrah, T., Lam, H., Tasman, N., Sun, Z., Nilsson, E., Pratt, B., Prazen, B., et al. (2010). A guided tour of the Trans-Proteomic Pipeline. PROTEOMICS *10*, 1150–1159. https://doi.org/10.1002/PMIC.200900375.

Dix, M.M., Simon, G.M., Wang, C., Okerberg, E., Patricelli, M.P., and Cravatt, B.F. (2012). Functional Interplay between Caspase Cleavage and Phosphorylation Sculpts the Apoptotic Proteome. Cell *150*, 426–440. https://doi.org/10.1016/j.cell.2012.05.040.

Djouder, N., Metzler, S.C., Schmidt, A., Wirbelauer, C., Gstaiger, M., Aebersold, R., Hess, D., and Krek, W. (2007). S6K1-mediated disassembly of mitochondrial URI/PP1gamma complexes activates a negative feedback program that counters S6K1 survival signaling. Mol. Cell *28*, 28–40. https://doi.org/10.1016/j.molcel.2007.08.010.

Drazic, A., Aksnes, H., Marie, M., Boczkowska, M., Varland, S., Timmerman, E., Foyn, H., Glomnes, N., Rebowski, G., Impens, F., et al. (2018). NAA80 is actin's N-terminal acetyltransferase and regulates cytoskeleton assembly and cell motility. Proc. Natl. Acad. Sci. *115*, 4399–4404. https://doi.org/10.1073/pnas.1718336115.

Du, X., Lu, J., Yang, L., Bao, H., Zhang, L., Yan, G., Fang, C., and Lu, H. (2020). Rapid and Easy Enrichment Strategy for Naturally Acetylated N Termini Based on LysN Digestion and Amine-Reactive Resin Capture. Anal. Chem. *92*, 8315–8322. https://doi.org/10.1021/ACS.ANALCHEM.0C00695.

Eng, J.K., Jahan, T.A., and Hoopmann, M.R. (2013). Comet: An open-source MS/MS sequence database search tool. Proteomics *13*, 22–24. https://doi.org/10.1002/pmic.201200439.

Fortelny, N., Yang, S., Pavlidis, P., Lange, P.F., and Overall, C.M. (2015). Proteome TopFIND 3.0 with TopFINDer and PathFINDer: database and analysis tools for the association of protein termini to pre- and post-translational events. Nucleic Acids Res. *43*, D290–D297. https://doi.org/10.1093/NAR/GKU1012.

Friedl, J., Knopp, M.R., Groh, C., Paz, E., Gould, S.B., Herrmann, J.M., and Boos, F. (2020). More than just a ticket canceller: the mitochondrial processing peptidase tailors complex precursor proteins at internal cleavage sites. Mol. Biol. Cell *31*, 2657–2668. https://doi.org/10.1091/mbc.E20-08-0524.

Gevaert, K., Van Damme, J., Goethals, M., Thomas, G.R., Hoorelbeke, B., Demol, H., Martens, L., Puype, M., Staes, A., and Vandekerckhove, J. (2002). Chromatographic Isolation of Methionine-containing Peptides for Gel-free Proteome Analysis: Identification Of More Than 800 Escherichia Coli Proteins. Mol. Cell. Proteomics *1*, 896–903. https://doi.org/10.1074/MCP.M200061-MCP200.

Giansanti, P., Tsiatsiani, L., Low, T.Y., and Heck, A.J.R. (2016). Six alternative proteases for mass spectrometry–based proteomics beyond trypsin. Nat. Protoc. *11*, 993–1006. https://doi.org/10.1038/nprot.2016.057.

Giglione, C., Boularot, A., and Meinnel, T. (2004). Protein N-terminal methionine excision. Cell. Mol. Life Sci. CMLS *61*, 1455–1474. https://doi.org/10.1007/s00018-004-3466-8.

Giglione, C., Fieulaine, S., and Meinnel, T. (2015). N-terminal protein modifications: Bringing back into play the ribosome. Biochimie *114*, 134–146. https://doi.org/10.1016/j.biochi.2014.11.008.

Griswold, A.R., Cifani, P., Rao, S.D., Axelrod, A.J., Miele, M.M., Hendrickson, R.C., Kentsis, A., and Bachovchin, D.A. (2019). A Chemical Strategy for Protease Substrate Profiling. Cell Chem. Biol. *26*, 901-907.e6. https://doi.org/10.1016/J.CHEMBIOL.2019.03.007.

Han, D.K., Eng, J., Zhou, H., and Aebersold, R. (2001). Quantitative profiling of differentiation-induced

microsomal proteins using isotope-coded affinity tags and mass spectrometry. Nat. Biotechnol. 2001 1910 *19*, 946–951. https://doi.org/10.1038/nbt1001-946.

Helbig, A.O., Rosati, S., Pijnappel, P.W., van Breukelen, B., Timmers, M.H., Mohammed, S., Slijper, M., and Heck, A.J. (2010). Perturbation of the yeast N-acetyltransferase NatB induces elevation of protein phosphorylation levels. BMC Genomics *11*, 685. https://doi.org/10.1186/1471-2164-11-685.

Helsens, K., Van Damme, P., Degroeve, S., Martens, L., Arnesen, T., Vandekerckhove, J., and Gevaert, K. (2011). Bioinformatics Analysis of a Saccharomyces cerevisiae N-Terminal Proteome Provides Evidence of Alternative Translation Initiation and Post-Translational N-Terminal Acetylation. J. Proteome Res. *10*, 3578–3589. https://doi.org/10.1021/pr2002325.

Hershko, A., Heller, H., Eytan, E., Kaklij, G., and Rose, I.A. (1984). Role of the alpha-amino group of protein in ubiquitin-mediated protein breakdown. Proc. Natl. Acad. Sci. U. S. A. *81*, 7021–7025. https://doi.org/10.1073/pnas.81.22.702

Hotokezaka, Y., van Leyen, K., Lo, E.H., Beatrix, B., Katayama, I., Jin, G., and Nakamura, T. (2009). αNAC depletion as an initiator of ER stress-induced apoptosis in hypoxia. Cell Death Differ. *16*, 1505–1514. https://doi.org/10.1038/cdd.2009.90.

Hotokezaka, Y., Katayama, I., van Leyen, K., and Nakamura, T. (2015). GSK-3β-dependent downregulation of γ-taxilin and αNAC merge to regulate ER stress responses. Cell Death Dis. *6*, e1719–e1719. https://doi.org/10.1038/cddis.2015.90.

Huesgen, P.F., Alami, M., Lange, P.F., Foster, L.J., Schröder, W.P., Overall, C.M., and Green, B.R. (2013). Proteomic Amino-Termini Profiling Reveals Targeting Information for Protein Import into Complex Plastids. PLOS ONE *8*, e74483. https://doi.org/10.1371/journal.pone.0074483.

Hwang, C.S., Shemorry, A., and Varshavsky, A. (2010). N-terminal acetylation of cellular proteins creates specific degradation signals. Science *327*, 973–977. https://doi.org/10.1126/science.1183147.

Jomaa, A., Gamerdinger, M., Hsieh, H.-H., Wallisch, A., Chandrasekaran, V., Ulusoy, Z., Scaiola, A., Hegde, R.S., Shan, S., Ban, N., et al. (2022). Mechanism of signal sequence handover from NAC to SRP on ribosomes during ER-protein targeting. Science *375*, 839–844. https://doi.org/10.1126/science.abl6459.

Jonckheere, V., Fijałkowska, D., and Damme, P.V. (2018). Omics Assisted N-terminal Proteoform and Protein Expression Profiling On Methionine Aminopeptidase 1 (MetAP1) Deletion *. Mol. Cell. Proteomics *17*, 694–708. https://doi.org/10.1074/mcp.RA117.000360.

Jörnvall, H. (1975). Acetylation of protein N-terminal amino groups structural observations on α-amino acetylated proteins. J. Theor. Biol. *55*, 1–12. https://doi.org/10.1016/S0022-5193(75)80105-6.

Julien, O., Zhuang, M., Wiita, A.P., O'Donoghue, A.J., Knudsen, G.M., Craik, C.S., and Wells, J.A. (2016). Quantitative MS-based enzymology of caspases reveals distinct protein substrate specificities, hierarchies, and cellular roles. Proc. Natl. Acad. Sci. U. S. A. *113*, E2001-10. https://doi.org/10.1073/pnas.1524900113.

Keller, A., Nesvizhskii, A.I., Kolker, E., and Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. *74*, 5383–5392. https://doi.org/10.1021/ac025747h.

Kleifeld, O., Doucet, A., auf dem Keller, U., Prudova, A., Schilling, O., Kainthan, R.K., Starr, A.E., Foster, L.J., Kizhakkedathu, J.N., and Overall, C.M. (2010). Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products. Nat. Biotechnol. *28*, 281–288. https://doi.org/10.1038/nbt.1611

Kleifeld, O., Doucet, A., Prudova, A., auf dem Keller, U., Gioia, M., Kizhakkedathu, J.N., and Overall, C.M. (2011). Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic

labeling of substrates. Nat. Protoc. *6*, 1578–1611. https://doi.org/10.1038/nprot.2011.382.

Koehler, C.J., Arntzen, M.O., De Souza, G.A., and Thiede, B. (2013). An approach for triplex-isobaric peptide termini labeling (triplex-IPTL). Anal. Chem. *85*, 2478–2485. https://doi.org/10.1021/ac3035508.

López-Otín, C., and Bond, J.S. (2008). Proteases: Multifunctional enzymes in life and disease. J. Biol. Chem. *283*, 30433–30437. https://doi.org/10.1074/jbc.R800035200.

Lopez-Otin, C., and Overall, C.M. (2002). Protease degradomics: a new challenge for proteomics. Nat. Rev. Mol. Cell Biol. *3*, 509–519. https://doi.org/10.1038/nrm858.

Luo, S.Y., Araya, L.E., and Julien, O. (2019). Protease Substrate Identification Using N-terminomics. ACS Chem. Biol. *14*, 2361–2371. https://doi.org/10.1021/ACSCHEMBIO.9B00398.

Lüthi, A.U., and Martin, S.J. (2007). The CASBAH: a searchable database of caspase substrates. Cell Death Differ. *14*, 641–650. https://doi.org/10.1038/SJ.CDD.4402103.

MacDonald, J.I., Munch, H.K., Moore, T., and Francis, M.B. (2015). One-step site-specific modification of native proteins with 2-pyridinecarboxyaldehydes. Nat. Chem. Biol. *11*, 326–331. https://doi.org/10.1038/nchembio.1792.

Mahrus, S., Trinidad, J.C., Barkan, D.T., Sali, A., Burlingame, A.L., and Wells, J.A. (2008). Global sequencing of proteolytic cleavage sites in apoptosis by specific labeling of protein N termini. Cell *134*, 866–876. https://doi.org/10.1016/j.cell.2008.08.012.

Marino, G., Eckhard, U., and Overall, C.M. (2015). Protein Termini and Their Modifications Revealed by Positional Proteomics. ACS Chem. Biol. *10*, 1754–1764. https://doi.org/10.1021/ACSCHEMBIO.5B00189.

Mueller, F., Friese, A., Pathe, C., da Silva, R.C., Rodriguez, K.B., Musacchio, A., and Bange, T. (2021). Overlap of NatA and IAP substrates implicates N-terminal acetylation in protein stabilization. Sci. Adv. *7*, eabc8590. https://doi.org/10.1126/sciadv.abc8590.

Nyonda, M.A., Boyer, J.-B., Belmudes, L., Krishnan, A., Pino, P., Couté, Y., Brochet, M., Meinnel, T., Soldati-Favre, D., and Giglione, C. (2022). N-acetylation of secreted proteins in Apicomplexa is widespread and is independent of the ER acetyl-CoA transporter AT1. J. Cell Sci. *135*, jcs259811. https://doi.org/10.1242/jcs.259811.

Ou, J., Liu, H., Nirala, N.K., Stukalov, A., Acharya, U., Green, M.R., and Zhu, L.J. (2020). dagLogo: An R/Bioconductor package for identifying and visualizing differential amino acid group usage in proteomics data. PLOS ONE *15*, e0242030. https://doi.org/10.1371/journal.pone.0242030.

Parusel, C.T., Kritikou, E.A., Hengartner, M.O., Krek, W., and Gotta, M. (2006). URI-1 is required for DNA stability in C. elegans. Development *133*, 621–629. https://doi.org/10.1242/dev.02235.

Qin, H., Wang, F., Zhang, Y., Hu, Z., Song, C., Wu, R., Ye, M., and Zou, H. (2012). Isobaric cross-sequence labeling of peptides by using site-selective N-terminus dimethylation. Chem. Commun. *48*, 6265. https://doi.org/10.1039/c2cc31705b.

Rawlings, N.D., Barrett, A.J., and Bateman, A. (2012). MEROPS: the database of proteolytic enzymes, their substrates and inhibitors. Nucleic Acids Res. *40*, D343-50. https://doi.org/10.1093/nar/gkr987.

Ree, R., Varland, S., and Arnesen, T. (2018). Spotlight on protein N-terminal acetylation. Exp. Mol. Med. *50*, 1–13. https://doi.org/10.1038/s12276-018-0116-z.

Rogers, L.D., and Overall, C.M.C. (2013). Proteolytic Post-translational Modification of Proteins: Proteomic Tools and Methodology. Mol. Cell. Proteomics *12*, 3532–3542. https://doi.org/10.1074/MCP.M113.031310.

Rospert, S., Dubaquié, Y., and Gautschi, M. (2002). Nascent-polypeptide-associated complex. Cell. Mol. Life Sci. CMLS *59*, 1632–1639. https://doi.org/10.1007/PL00012490.

Sayols, S. (2022). rrvgo: Reduce + Visualize GO (Bioconductor version: Release (3.15)).

Schlage, P., Egli, F.E., and auf dem Keller, U. (2017). Time-resolved analysis of matrix metalloproteinase substrates in complex samples. In Methods in Molecular Biology, pp. 185–198. https://doi.org/10.1007/978-1-4939-6863-3_9.

Seaman, J.E., Julien, O., Lee, P.S., Rettenmaier, T.J., Thomsen, N.D., and Wells, J.A. (2016). Cacidases: caspases can cleave after aspartate, glutamate and phosphoserine residues. Cell Death Differ. *23*, 1717–1726. https://doi.org/10.1038/cdd.2016.62.

Smith, P.K., Krohn, R.I., Hermanson, G.T., Mallia, A.K., Gartner, F.H., Provenzano, M.D., Fujimoto, E.K., Goeke, N.M., Olson, B.J., and Klenk, D.C. (1985). Measurement of protein using bicinchoninic acid. Anal. Biochem. *150*, 76–85. https://doi.org/10.1016/0003-2697(85)90442-7.

Souers, A.J., Leverson, J.D., Boghaert, E.R., Ackler, S.L., Catron, N.D., Chen, J., Dayton, B.D., Ding, H., Enschede, S.H., Fairbrother, W.J., et al. (2013). ABT-199, a potent and selective BCL-2 inhibitor, achieves antitumor activity while sparing platelets. Nat. Med. 2013 192 *19*, 202–208. https://doi.org/10.1038/nm.3048.

Spitzer, M., Wildenhain, J., Rappsilber, J., and Tyers, M. (2014). BoxPlotR: a web tool for generation of box plots. Nat. Methods *11*, 121–122. https://doi.org/10.1038/nmeth.2811.

Steen, H., and Lindholm, D. (2008). Nuclear localized protein-1 (Nulp1) increases cell death of human osteosarcoma cells and binds the X-linked inhibitor of apoptosis protein. Biochem. Biophys. Res. Commun. *366*, 432–437. https://doi.org/10.1016/j.bbrc.2007.11.146.

Sukharev, S.A., Pleshakova, O.V., and Sadovnikov, V.B. (1997). Role of proteases in activation of apoptosis. Cell Death Differ. 1997 46 *4*, 457–462. https://doi.org/10.1038/sj.cdd.4400263.

Swaney, D.L., Wenger, C.D., and Coon, J.J. (2010). Value of Using Multiple Proteases for Large-Scale Mass Spectrometry-Based Proteomics. J. Proteome Res. *9*, 1323–1329. https://doi.org/10.1021/PR900863U.

Tafani, M., Karpinich, N.O., Hurster, K.A., Pastorino, J.G., Schneider, T., Russo, M.A., and Farber, J.L. (2002). Cytochrome c Release upon Fas Receptor Activation Depends on Translocation of Full-length Bid and the Induction of the Mitochondrial Permeability Transition. J. Biol. Chem. *277*, 10073–10082. https://doi.org/10.1074/JBC.M111350200.

The UniProt Consortium (2021). UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Res. *49*, D480–D489. https://doi.org/10.1093/nar/gkaa1100.

Tsiatsiani, L., Giansanti, P., Scheltema, R.A., Toorn, H. van den, Overall, C.M., Altelaar, A.F.M., and Heck, A.J.R. (2016). Opposite Electron-Transfer Dissociation and Higher-Energy Collisional Dissociation Fragmentation Characteristics of Proteolytic K/R(X)n and (X)nK/R Peptides Provide Benefits for Peptide Sequencing in Proteomics and Phosphoproteomics. J. Proteome Res. *16*, 852–861. https://doi.org/10.1021/ACS.JPROTEOME.6B00825.

Vaca Jacome, A.S., Rabilloud, T., Schaeffer-Reiss, C., Rompais, M., Ayoub, D., Lane, L., Bairoch, A., Van Dorsselaer, A., and Carapito, C. (2015). N-terminome analysis of the human mitochondrial proteome. PROTEOMICS *15*, 2519–2524. https://doi.org/10.1002/pmic.201400617.

Van Damme, P., Kalvik, T.V., Starheim, K.K., Jonckheere, V., Myklebust, L.M., Menschaert, G., Varhaug, J.E., Gevaert, K., and Arnesen, T. (2016). A Role for Human N-alpha Acetyltransferase 30 (Naa30) in Maintaining Mitochondrial Integrity *. Mol. Cell. Proteomics *15*, 3361–3372. https://doi.org/10.1074/mcp.M116.061010.

Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G.,

Laydon, A., et al. (2022). AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. Nucleic Acids Res. *50*, D439–D444. https://doi.org/10.1093/nar/gkab1061.

Walsh, J.G., Cullen, S.P., Sheridan, C., Lüthi, A.U., Gerner, C., and Martin, S.J. (2008). Executioner caspase-3 and caspase-7 are functionally distinct proteases. Proc. Natl. Acad. Sci. *105*, 12815–12819. https://doi.org/10.1073/PNAS.0707715105.

Weiske, J., Schöneberg, T., Schröder, W., Hatzfeld, M., Tauber, R., and Huber, O. (2001). The Fate of Desmosomal Proteins in Apoptotic Cells*. J. Biol. Chem. *276*, 41175–41181. https://doi.org/10.1074/jbc.M105769200.

Weng, S.S.H., Demir, F., Ergin, E.K., Dirnberger, S., Uzozie, A., Tuscher, D., Nierves, L., Tsui, J., Huesgen, P.F., and Lange, P.F. (2019). Sensitive Determination of Proteolytic Proteoforms in Limited Microscale Proteome Samples. Mol. Cell. Proteomics *18*, 2335–2347. https://doi.org/10.1074/MCP.TIR119.001560.

Zeeuwen, P.L.J.M. (2004). Epidermal differentiation: The role of proteases and their inhibitors. Eur. J. Cell Biol. *83*, 761–773. https://doi.org/10.1078/0171-9335-00388.

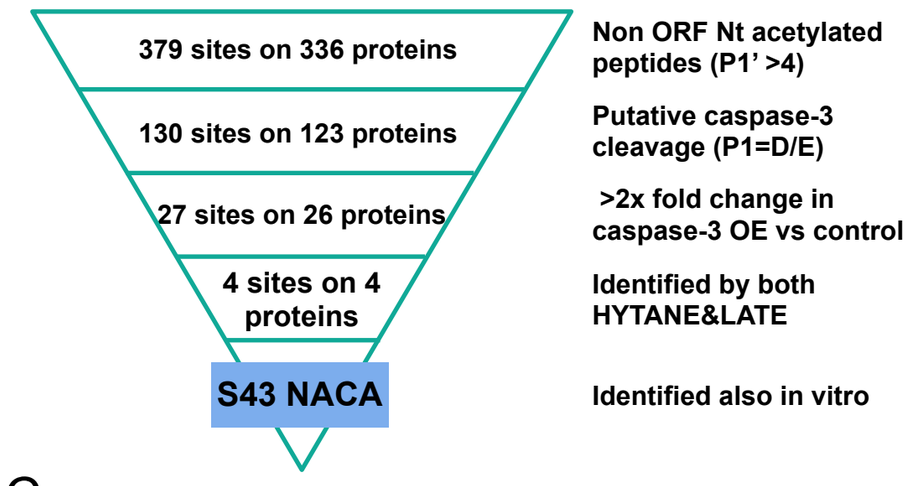
Zhan, Q., Jin, S., Ng, B., Plisket, J., Shangary, S., Rathi, A., Brown, K.D., and Baskaran, R. (2002). Caspase-3 mediated cleavage of BRCA1 during UV-induced apoptosis. Oncogene 2002 2134 *21*, 5335–5345. https://doi.org/10.1038/sj.onc.1205665.

Zhang, X., Ye, J., and Højrup, P. (2009). A proteomics approach to study in vivo protein Nα-modifications. J. Proteomics *73*, 240–251. https://doi.org/10.1016/j.jprot.2009.09.007.

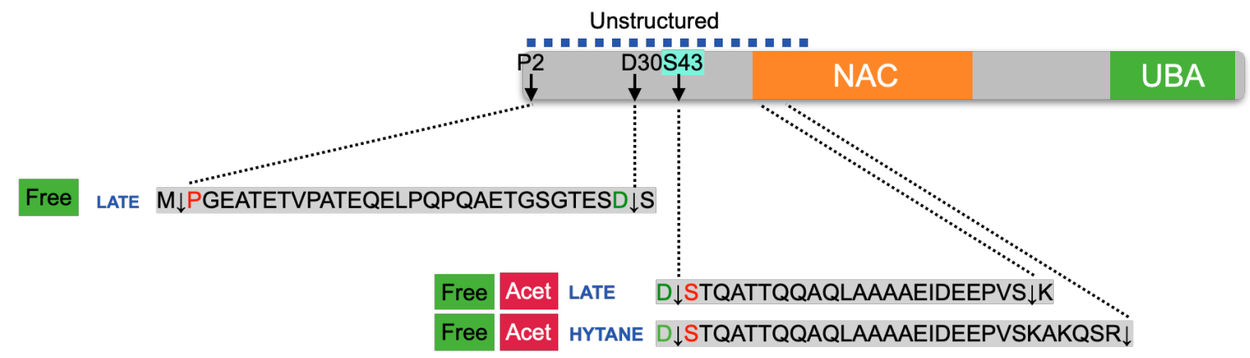
Zhang, Y., Li, Q., Huang, J., Wu, Z., Huang, J., Huang, L., Li, Y., Ye, J., and Zhang, X. (2018). An Approach to Incorporate Multi-Enzyme Digestion into C-TAILS for C-Terminomics Studies. Proteomics *18*. https://doi.org/10.1002/pmic.201700034.

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., and Wijk, K.J. van (2008). Sorting Signals, N-Terminal Modifications and Abundance of the Chloroplast Proteome. PLOS ONE *3*, e1994. https://doi.org/10.1371/journal.pone.0001994.

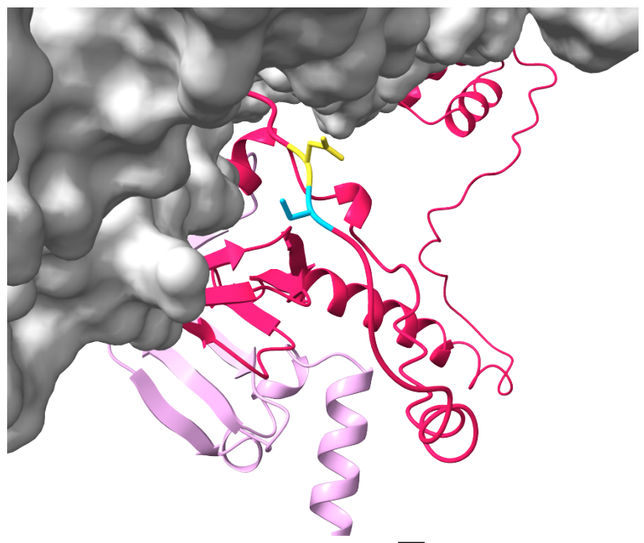
A



B

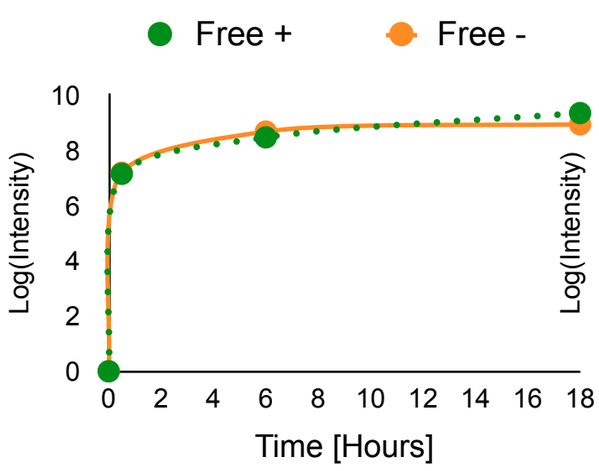


C

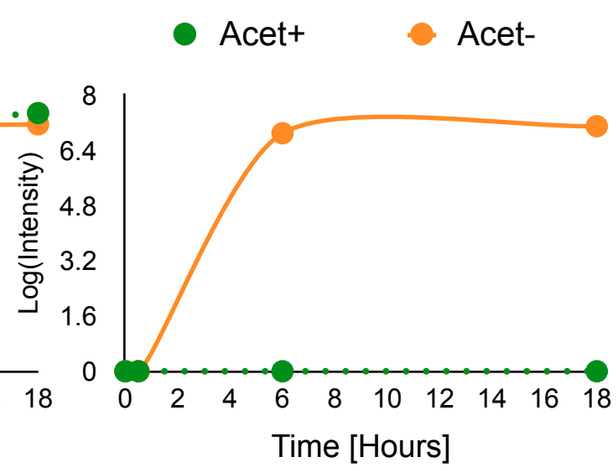


Method	Exp	Peptide	Log2(ratio(caspase3/control)) Repeat1	Repeat2	Repeat3	PSMs
HYTANE	In vitro	Acet-S43	8.1	6.5	7.8	19
LATE	In vitro	Acet-S43	ID only	ID only	ID only	11
HYTANE	In vitro	DiMt-S43	10.9	9.7	11.3	17
LATE	In vitro	DiMt-S43	9.3	5.6	6.9	39
HYTANE	In Vivo	Acet-S43	2.7	1.8	2.4	40
LATE	In Vivo	Acet-S43	ID only	ID only	ID only	22
HYTANE	In Vivo	DiMt-S43	1.5	1.1		2
LATE	In Vivo	DiMt-S43				
HYTANE	In vitro	Metyl-P2-D30			13.3	1
LATE	In vitro	Metyl-P2-D30			13.3	1
HYTANE	In Vivo	Metyl-P2-D30				
LATE	In Vivo	Metyl-P2-D30	13.3	6.5	13.3	5

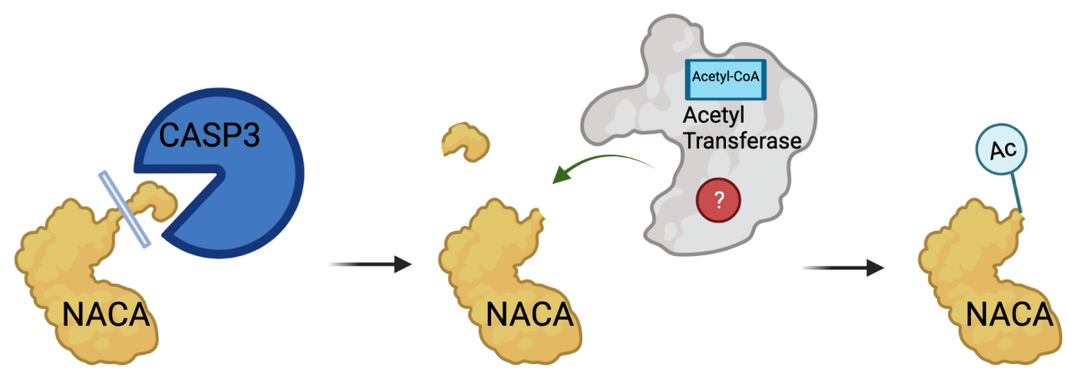
D



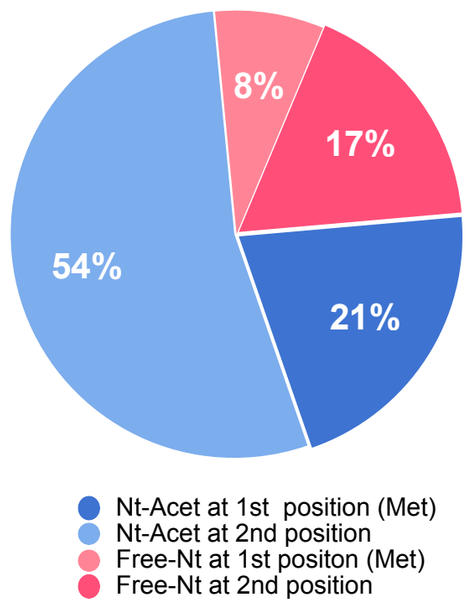
E



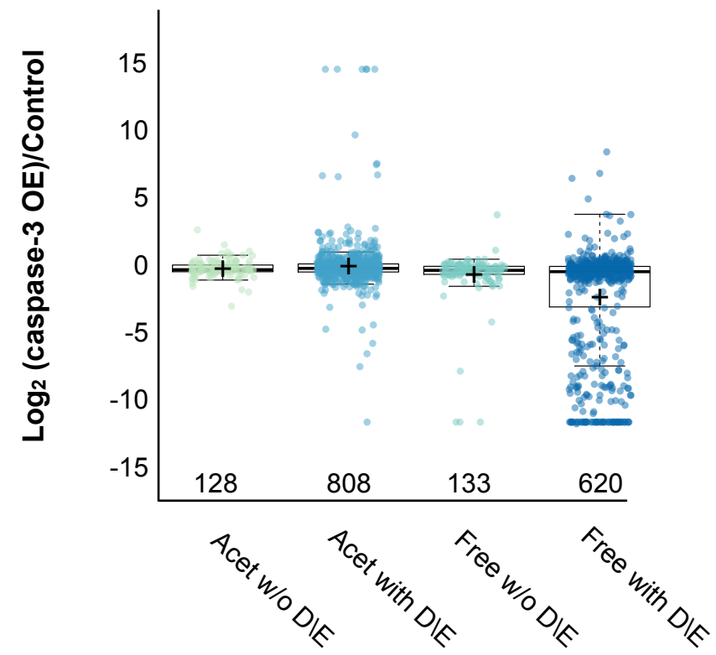
F



A

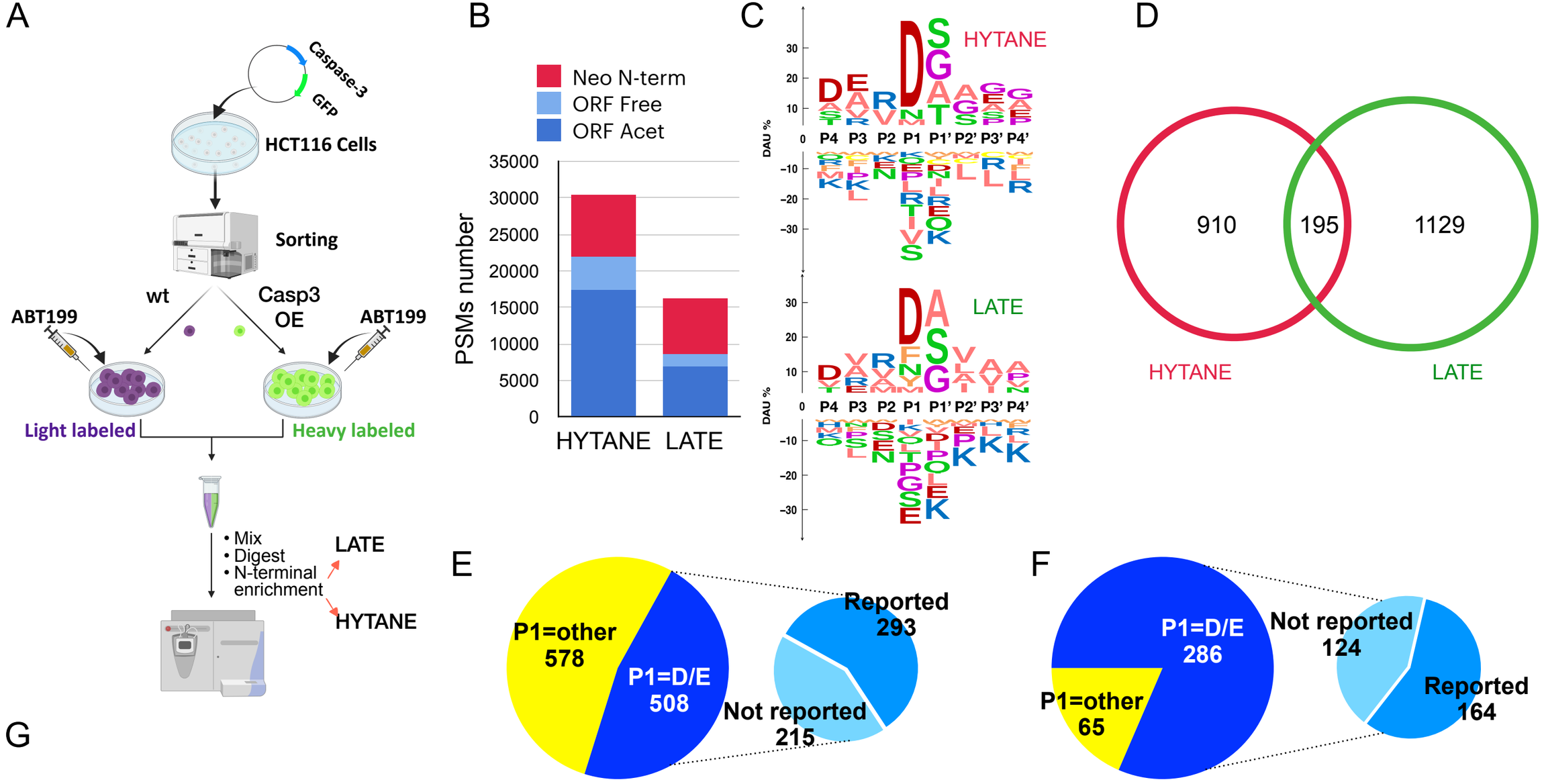


B



C

Gene	UniProt ID	Motif P4-P4'	Cell-based HYTANE	Cell-based LATE	In Vitro HYTANE 1	In Vitro HYTANE 6	In Vitro HYTANE 18	In Vitro LATE 1	In Vitro LATE 6	In Vitro LATE 18	Processing type
SERPINH1	P50454	AALA↓Ac[19]AEVK	◆	◆	◆								Signal peptide
ATP5F1D	P30049	ARAY↓Ac[23]AEAA	◆					◆	◆	◆				Transit peptide
PPA2	Q9H2U2	RRAM↓Ac[33]ALYH	◆	◆		◆	◆		◆	◆		◆	◆	◆	Transit peptide
SITR5	Q9NXA8	CLKM↓Ac[34]ARPS	◆			◆	◆					◆	◆	◆	Transit peptide
ATP5F1A	P25705	NTHL↓Ac[44]QKTG	◆	◆											Transit peptide
FH	P07954	AARM↓Ac[45]ASQN	◆	◆	◆	◆	◆		◆	◆		◆			Transit peptide
GSR	P00390	SRAM↓Ac[45]ACRQ	◆	◆	◆				◆	◆	◆				Transit peptide
PRDX3	P30048	SCHA↓Ac[63]PAVT				◆	◆								Transit peptide
C1QBP	Q07021	CGCG↓Ac[73]SLHT	◆												Transit peptide
GRN	P28799	HTVG↓Ac[73]DVKC	◆	◆	◆				◆	◆	◆				Peptide processing

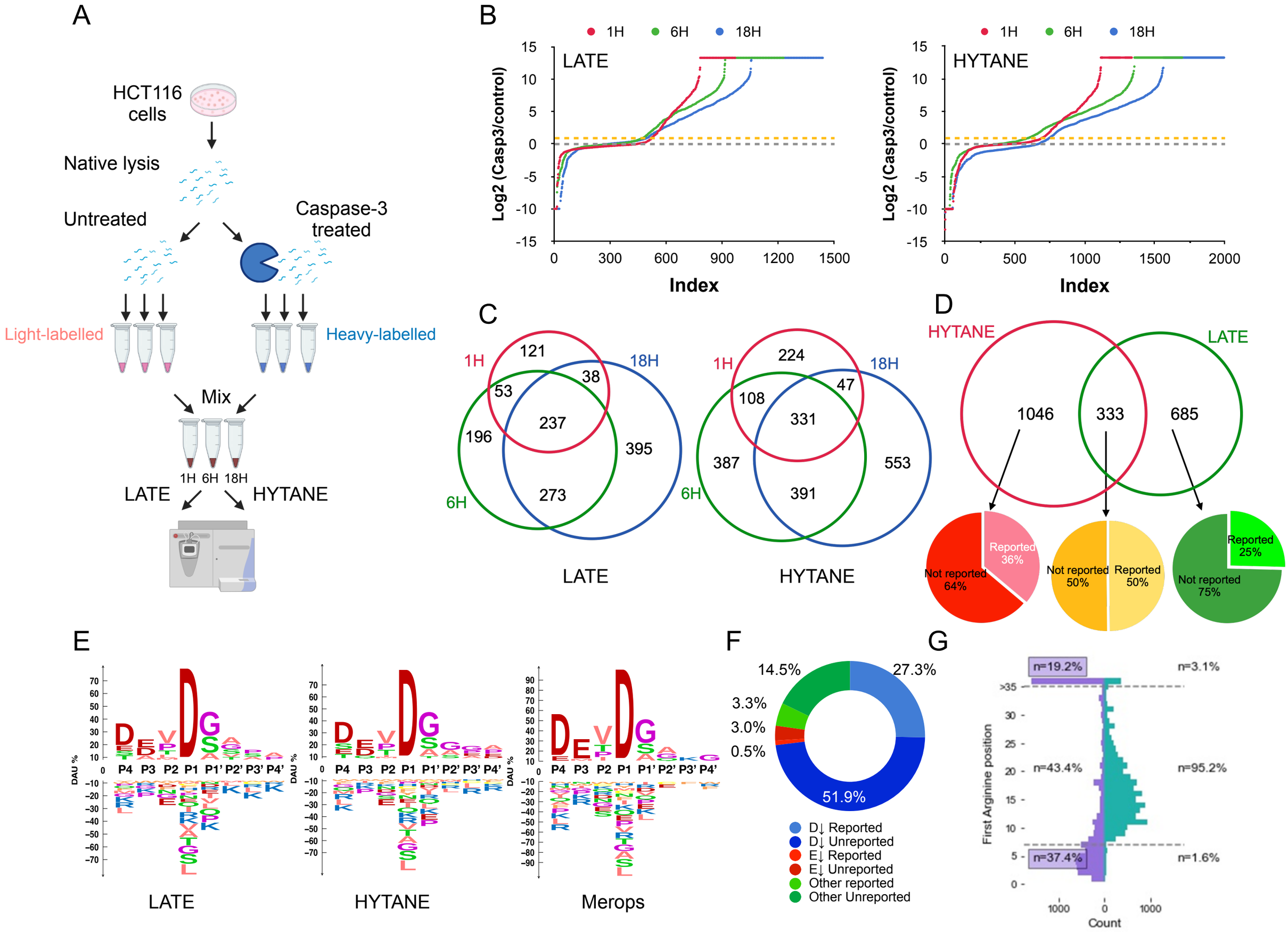


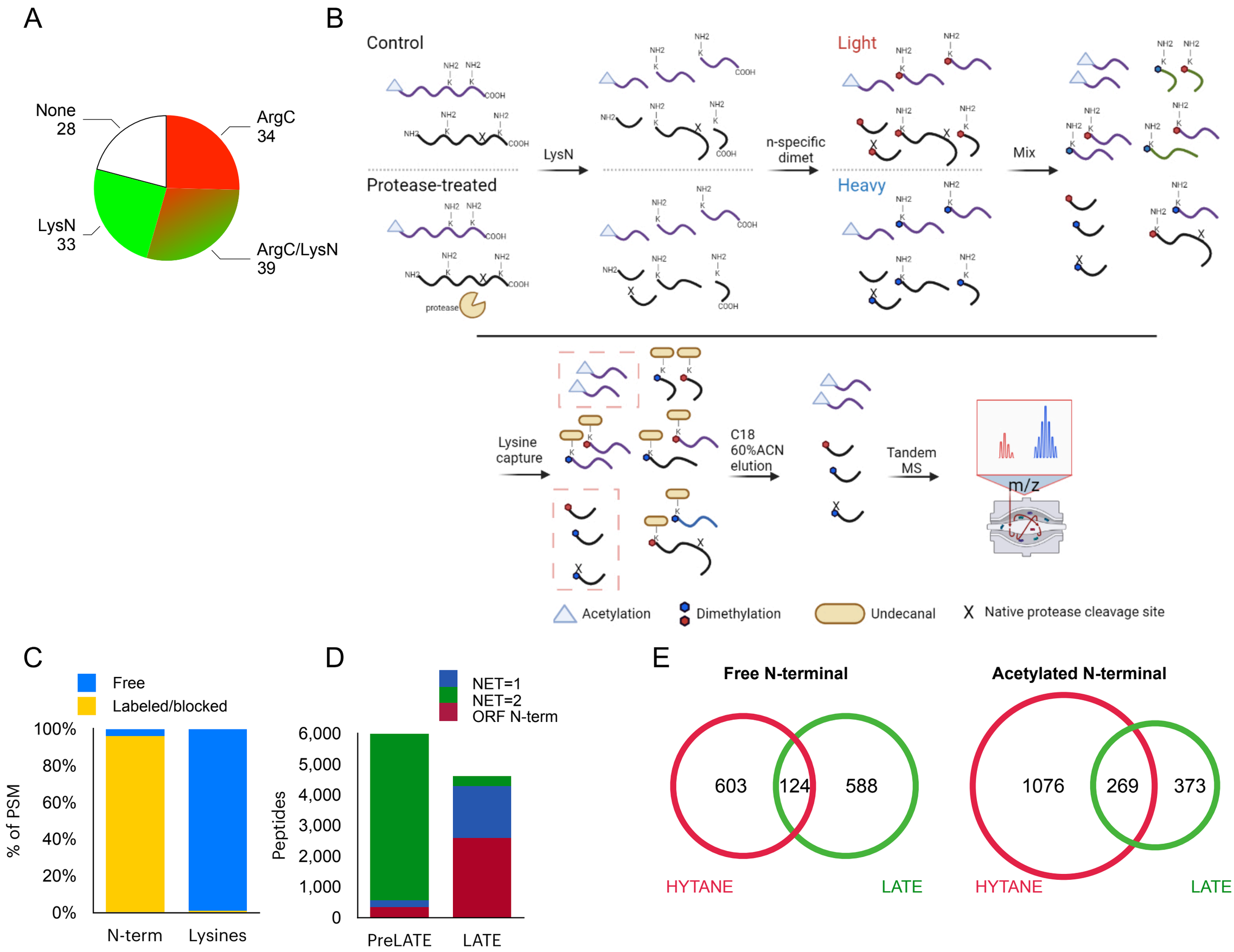
G

UniProt ID	Gene	Motif P4-P4'	Cell-based HYTANE	Cell-based LATE	In Vitro HYTANE 1	In Vitro HYTANE 6	In Vitro HYTANE 18	In Vitro LATE 1	In Vitro LATE 6	In Vitro LATE 18
O75179	ANKRD17	STAD↓192GKAF	>1	-1< >1	>1	>1				
P15924	DSP	ELVD↓2225SGIL	>1	>1		>1	>1			
Q9GZP8	IMUP	MEFD↓5LGAA	>1	>1			>1			
Q9GZV4	EIF5A2	DEID↓7FTTG		-1< >1				>1	>1	>1
P84103	SRSF3	MHRD↓5SCPL	>1	>1		>1	>1			
O94763	URI1	TCSD↓451TSES		>1					>1	
Q9BQ70	TCF25	SKTD↓108GDDT	-1< >1	>1	>1	>1	>1		>1	>1
Q9P0J1	PDP1	AHVD↓281GVDL	-1< >1	>1		>1	>1			
Q9BR76	CORO1B	VLSD↓415SRPA	>1	-1< >1	>1	>1	>1			
Q86WR7	PROSER2	DVID↓110LVQP	>1	>1		>1	>1		>1	>1
Q9UQB8	BAIAP2	TGPD↓307GEDY	>1	>1		>1	>1			>1
Q8IVT2	MISP	QVVD↓275DPGS	>1	>1		>1				
Q13542	EIF4EBP2	AISD↓27AAQL	>1	>1	>1	>1	>1			
Q6P597	KLC3	DHPD↓332VAKQ	>1	-1< >1	>1	>1				
Q9NW64	RBM22	SNSD↓130GTRP	>1	-1< >1				>1	>1	>1
O15126	SCAMP1	SDFD↓65NPF	>1	>1			>1			

UniProt ID	Gene	Motif P4'-P4	Cell-based HYTANE	Cell-based LATE	In Vitro HYTANE 1	In Vitro HYTANE 6	In Vitro HYTANE 18	In Vitro LATE 1	In Vitro LATE 6	In Vitro LATE 18
Q9GZV4	EIF5A2	DEID↓7FTTG	Identified without ratio	<-1				>1	>1	>1
C9J7I0	UMAD1	TEAD↓23GFVL	<-1	<-1	>1	>1				
P51114	FXR1	TESD↓526TASV	-1< >1	>1	-1< >1	-1< >1	-1< >1		>1	>1
Q9GZR2	REXO4	DDVD↓201PADI	>1	-1< >1	>1	>1	>1	>1	>1	>1
Q03001	DST	TDVD↓3837GQVG			-1< >1	-1< >1	-1< >1		>1	>1
Q15836	VAMP3	DRAD↓52ALQA			>1	>1			>1	>1
Q13033	STRN3	DLAD↓447LTVT			>1	>1			>1	
P14923	JUP	TQVE↓108GQAT			>1	>1	-1< >1		>1	>1
Q9BV44	THUMPD3	TESD↓30LGSE			-1< >1	>1	-1< >1		>1	>1
Q9H5V8	CDCP1	PEVD↓761TYRP				>1	-1< >1		>1	Identified without ratio
Q92621	NUP205	DSQD↓1923SLFA				>1	-1< >1			>1
Q9Y546	LRRC42	DHID↓106SLIG				-1< >1	>1		>1	>1
Q96TA1	NIBAN2	ASPD↓649GVTE				-1< >1	-1< >1		>1	>1
Q96KQ7	EHMT2	DTID↓605SSGP				-1< >1	-1< >1		>1	>1
P42574	CASP3	TDVD↓70AANL		>1	>1	>1	>1		>1	>1

Log2 (caspse-3 OE/control)
- >1
- -1< >1
- <-1
- Identified without ratio


A
HCT116 cells
Native lysis
Untreated
Caspase-3 treated
Light-labelled
Heavy-labelled
Mix
1H 6H 18H
LATE
HYTANE
B
1H 6H 18H
LATE
HYTANE
Log2 (Casp3/control)
Index
C
1H
6H
18H
121
38
53
237
196
395
273
224
47
108
331
387
553
391
LATE
HYTANE
D
HYTANE
LATE
1046
333
685
Reported 36%
Not reported 64%
Not reported 50%
Reported 50%
Reported 25%
Not reported 75%
E
P4 P3 P2 P1 P1' P2' P3' P4'
DAU %
LATE
HYTANE
Merops
F
14.5%
27.3%
3.3%
3.0%
0.5%
51.9%
D↓ Reported
D↓ Unreported
E↓ Reported
E↓ Unreported
Other reported
Other Unreported
G
n=19.2%
n=3.1%
n=43.4%
n=95.2%
n=37.4%
n=1.6%
First Arginine position
Count


A
None
28
ArgC
34
LysN
33
ArgC/LysN
39
B
Control
Protease-treated
protease
LysN
n-specific
dimet
Light
Heavy
Mix
Lysine
capture
C18
60%ACN
elution
Tandem
MS
m/z
Acetylation
Dimethylation
Undecanal
X Native protease cleavage site
C
Free
Labeled/blocked
% of PSM
100%
80%
60%
40%
20%
0%
N-term
Lysines
D
NET=1
NET=2
ORF N-term
Peptides
6,000
5,000
4,000
3,000
2,000
1,000
0
PreLATE
LATE
E
Free N-terminal
603
124
588
HYTANE
LATE
Acetylated N-terminal
1076
269
373
HYTANE
LATE