Identification of a mitochondrial targeting sequence in cathepsin D and its localization in mitochondria

Naoki Ikari and Hirofumi Arakawa

Division of Cancer Biology, National Cancer Center Research Institute, 5-1-1 Tsukiji, Chuo-ku, Tokyo, Japan

Running title

Mitochondrial localization of cathepsin D

Keywords

cathepsin D, mitochondrial targeting sequence, mitochondrial localization, live-cell imaging, dual targeting protein

*Correspondence should be addressed to:

Hirofumi Arakawa, M.D., Ph.D.

Division of Cancer Biology

National Cancer Center Research Institute

5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, JAPAN

Tel: +81-3-3547-5273, Fax: +81-3-3546-1369

E-mail: harakawa@ncc.go.jp

1

Abstract

Cathepsin D (CTSD) is a major lysosomal protease harboring an N-terminal signal peptide (amino acids 1–20) to enable vesicular transport from endoplasmic reticulum to lysosomes. Here, we report the possibility of a mitochondrial targeting sequence and mitochondrial localization of CTSD in cells. Live-cell imaging analysis with enhanced green fluorescent protein (EGFP)-CTSD indicated that CTSD localizes to mitochondria. CTSD amino acids 21–35 are responsible for its mitochondrial localization, which exhibit typical features of mitochondrial targeting sequences, and are evolutionarily conserved. A proteinase K protection assay and sucrose gradient analysis showed that a small population of endogenous CTSD molecules exists in mitochondria. These results suggest that CTSD is a dual-targeted protein that may localize in both lysosomes and mitochondria.

Introduction

CTSD is an aspartic lysosomal protease, dysfunction of which causes the most severe congenital form of neuronal ceroid lipofuscinosis (NCL), characterized by neurodegeneration and accumulation of ceroid lipofuscin containing the mitochondrial F1F0 ATPase [1, 2]. To date, the process of lysosomal localization of CTSD has been characterized as follows [3, 4]. Human CTSD, comprising 412 amino acids (pre-pro-CTSD), is translocated into endoplasmic reticulum (ER) via the ER signal peptide (amino acids 1-20), which is removed during translocation to yield pro-CTSD (392 amino acids, 52 kDa) [3]. After undergoing N-linked glycosylation (amino acids 134 and 263) in ER and subsequent modifications in Golgi [5, 6], pro-CTSD is transported to the endolysosomal compartment via mannose-6-phosphate (M6P)-dependent or independent pathways [4, 7]. Upon entering the acidic endolysosomal compartment, the N-terminal propeptide (amino acids 21-64) is proteolytically processed to a 48-kDa single-chain intermediate form. Finally, the intermediate form is processed into the mature enzymatic form, comprising heavy (34 kDa) and light (14 kDa) chains linked by non-covalent interactions, being cleaved by cathepsins B and L [8] and progranulin [9].

After lysosomal localization, CTSD is reportedly secreted to the extracellular space in breast cancer cells [10]. Other studies have found that CTSD is also released from lysosomes to the cytosol under stresses such as acetic acid [11] and staurosporine [12].

However, several studies have detected CTSD in mitochondria. i) The fruit fly ortholog of CTSD was detected in mitochondrial extracts using IEF/SDS-PAGE [13]. ii) The yeast ortholog of CTSD was detected in a 44.5-kDa form during proteomic analyses of mitochondria [14]. iii) The mouse ortholog of CTSD was detected during O-GlcNAcylome analysis of isolated mitochondria [15]. iv) Human CTSD was detected in the mitochondrial fraction of colon cancer cells with Mieap/SPATA18 expression [16], and v) human CTSD was detected in the mitochondrial fraction of gastric cancer cells in a 48-kDa form during hypoxia experiments [17]. These reports of CTSD in mitochondria cannot be explained by protein transport from the ER to lysosomes. Nevertheless, no mechanism for this phenomenon has been suggested.

In this study, to further explore subcellular localization of CTSD, we performed live-cell imaging analysis of EGFP-CTSD using various cell lines. We confirmed that EGFP-CTSD is localized not only in lysosomes, but also in mitochondria. Furthermore, sequence analysis, live-cell imaging analysis using deletion mutants of EGFP-CTSD in live cells, a proteinase K protection assay and sucrose gradient analysis for endogenous CTSD, and conventional immunofluorescence (IF) analysis suggest that there is a mitochondrial targeting sequence (MTS) in CTSD, and that a subpopulation of physiological CTSD may localize in mitochondria via its MTS. On the basis of these results, we conclude that CTSD localizes in mitochondria via its MTS.

Results

EGFP-CTSD localizes in mitochondria

To examine subcellular localization of CTSD in cells, we first constructed a plasmid to express C-terminal EGFP-tagged CTSD (EGFP-CTSD) and performed live-cell imaging of EGFP-CTSD using three cancer cell lines (HeLa, SK-BR-3, and HepG2 cells), and a non-cancer cell line (293 cells). Lysosomal localization of EGFP-CTSD was observed in all four cell lines (Hela, 45.9±7.3%; SK-BR-3, 39.7±1.4%; HepG2, 60.1±4.9%; 293, 58.7±2.6%) (Fig. 1A –D and 1I). Surprisingly, EGFP-CTSD signals were also detected in mitochondria of all four cell lines (HeLa, 42.6±6.8%; SK-BR-3, 78.7±9.5%; HepG2, 42.7±6.1%; 293, 38.3±2.2%) (Fig. 1E – H and 1J). Interestingly, we frequently observed that mitochondrial and lysosomal localization of EGFP-CTSD were mutually exclusive in individual cells. Thus, these results suggest that regardless of cellular profiles, CTSD does localize in mitochondria.

CTSD has an evolutionary conserved MTS

To identify a potential MTS in the sequence of CTSD, we performed sequence analyses. First, we used DeepLoc-1.0 and DeepLoc-2.0, because DeepLoc-1.0 can depict a hierarchical tree likelihood, mimicking a subcellular sorting pathway [21], and DeepLoc-2.0 is improved for multi-localization prediction [22]. Consistent with evidence that the N-terminal 1–20 amino acids of CTSD is an ER signal peptide for vesicular transport from the ER to lysosomes [3], DeepLoc-1.0 sorted CTSD toward lysosomes (Fig. 2A) [21], and DeepLoc-2.0 predicted that CTSD has an ER signal peptide and localizes in lysosomes and vacuoles (Fig. 2B) [22].

On the other hand, when we used MITOPROT II, it predicted amino acids 1–35 of CTSD as a putative MTS, and probable transport of CTSD to mitochondria as 0.6115 [24] (Fig. 2C). It has been suggested that dual-localized proteins often exhibit lower MITOPROT II scores, compared to typical mitochondrial proteins. The median value of dual-localized proteins is 0.603 [25]. MITOPROT II scores for several CTSD orthologs also revealed this feature (Fig. 2D).

In addition, when we used MitoFates [23], it predicted at least 21–30 amino acids of CTSD as an amphipathic alpha-helix (Fig. 2C). Thus far, the 3D structure of amino acids 21–35 has not been determined, according to the Protein Data Bank (PDB) (https://www.rcsb.org/, 6th May 2022) [26]. When we performed multiple sequence alignment for CTSD orthologs in representative eukaryotes, we found that amino acids 21–35 are evolutionarily conserved, showing enrichment of positively charged residues and a paucity of acidic residues (Fig. 2E). We also found that there is a putative mitochondrial protein peptidase (MPP) cleavage site at amino acids 35/36 of CTSD (Fig. 2C) [23, 24]. These are all characteristic features of MTS [24, 27, 28].

The predicted MTS of CTSD is required for mitochondrial localization of EGFP-CTSD

To verify the putative MTS predicted by sequence analyses for CTSD, we examined cells expressing EGFP-CTSD-full (Full-length) and four deletion-mutant forms, EGFP-CTSD Δ2–20 (without the ER signal peptide), Δ2–35 (without the ER signal peptide and the putative MTS predicted by MITOPROT

4

II [24]), Δ2–64 (without the ER signal peptide, the putative MTS, and propeptide), and Δ21–35 (without the putative MTS) (Fig. 3A).

Reproducibly as in Fig. 1E, cells expressing full-length EGFP-CTSD showed dual localization of CTSD in mitochondria and lysosomes (43.9±3.3% and 51.1±1.3% of expressing cells, respectively) (Fig. 3B, C, and M). Cells expressing Δ2–20, which lacks only the ER signal peptide, showed more intensive mitochondrial localization (96.8±1.5% of expressing cells) and loss of lysosomal localization (Fig. 3D, E, and M) of CTSD. Mitochondrial localization of CTSD was lost in cells expressing Δ2–35, Δ2–64, and Δ21–35, all of which lack amino acids 21–35 of CTSD (Fig. 3F–K, and 3M). Among them, Δ2–35 and Δ2–64 lacking the ER signal peptide were distributed all through the cytoplasm (Fig. 3F–I, and 3M), and Δ21–35 cells, retaining the ER signal peptide, were distributed to ER, but were not delivered to lysosomes (Fig. 3J–L, and 3M). These data indicate that the MTS of CTSD is limited to amino acids 21–35 of CTSD, and potentially competes with the ER signal peptide.

We performed western blot (WB) analysis for cells expressing EGFP-CTSD and deletion mutants. As shown in Fig. 3N, there were no bands equivalent to or smaller than Δ2–35 in cells expressing EGFP-CTSD full-length, suggesting that at least EGFP-tagged CTSD is not cleaved at the putative MPP cleavage site (amino acids 35/36) predicted by sequence analyses (Fig. 2C). One caveat here is that there were also no bands similar to the band of Δ2–64 (corresponding to the intermediate form) or smaller than Δ2–64 (corresponding to the mature form) in cells expressing EGFP-CTSD full-length. This implies that EGFP-tagged CTSD may be impaired in processing for transport from ER to lysosomes (Fig. 3N). In summary, CTSD amino acids 21–35 comprise a bona fide MTS, and at least EGFP-tagged CTSD is not cleaved at the putative MPP cleavage site (amino acids 35/36) predicted by sequence analyses.

A small population of endogenous CTSD exists in mitochondria

To examine whether endogenous CTSD exists in mitochondria, we performed a proteinase K protection assay with HeLa cells. In advance, we confirmed WB bands of endogenous CTSD variants using CTSD-knockdown (KD) cells and two anti-CTSD antibodies (CTD-19, mouse monoclonal and sc-6486 goat polyclonal). Multiple bands proved specific for endogenous CTSD expression, which was detected in 43~48-kDa (precursor form) and 30-kDa (active form) regions (Fig. 4A).

A crude mitochondrial fraction isolated from Hela cells was treated with proteinase K in the presence and absence of Triton X-100. TOMM20 (an outer mitochondrial membrane protein) and ATP5B (the matrix face of the inner mitochondrial membrane protein) were used as mitochondrial protein controls. A subpopulation comprising bands of 43~48-kDa were protected by proteinase K digestion, suggesting that a small proportion of endogenous CTSD may exist in mitochondria (Fig. 4B). On the other hand, the 30-kDa band was retained upon proteinase K digestion, even in the presence of Triton X-100. This suggests that this band is resistant to proteinase K (Fig. 4B), so it could not be evaluated using a proteinase K protection assay.

We also examined whether exogenous EGFP-CTSD or exogenous CTSD is protected from

5

proteinase K digestion. Both EGFP-CTSD and exogenous CTSD were much more clearly concentrated in the mitochondrial fraction and protected by proteinase K digestion, compared to endogenous CTSD (Fig. 4C–F).

To examine whether mitochondrial forms of endogenous CTSD contain the putative MTS (amino acids 21-35), we compared multiple bands of endogenous CTSD, which were protected from proteinase K digestion, with each band of full-length and Δ2–20, Δ2–35, and Δ2–64 mutants of exogenous CTSD in WB. Mitochondrial forms of endogenous CTSD are likely consistent with full-length and Δ2–20, but larger than Δ2–35 and Δ2–64 (Fig. 4G). This result is consistent with our finding in the imaging analysis with their mutants (Fig. 3) that the MTS of CTSD is amino acids 21–35 of CTSD. In addition, consistent with the results of EGFP-CTSD, mitochondrial forms of endogenous CTSD may not be cleaved at the putative MPP cleavage site (amino acids 35/36).

To validate the mitochondrial forms of endogenous CTSD in normal cells, we performed a proteinase K protection assay with 293 cells. Similar results were obtained with 293 cells, suggesting that mitochondrial localization of endogenous CTSD may be common to all cell types (Fig. 4H).

To further confirm the possible existence of endogenous CTSD in mitochondria, we performed sucrose gradient analysis with HeLa cell extract. A small population of endogenous CTSD was detected in the mitochondrial fraction (lanes 8-10), indicated by bands of TOMM20 and ATP5B (Fig. 4I). This result also supports the notion that a small part of endogenous CTSD exists in mitochondria.

Immunostaining experiments with fixed cells may fail to detect mitochondrial CTSD protein
Thus far, to our knowledge, endogenous CTSD has never been detected in immunofluorescence (IF) experiments using fixed cells. Therefore, to examine whether mitochondrial CTSD is stable in fixed cells, which are usually used in IF experiments, we performed fixed-cell imaging with cells expressing EGFP-CTSD, comparing it with imaging of live-cells (Fig. 5A).

Compared to live imaging data (Fig. 5A), localization of EGFP-CTSD seemed to be prevented by a gentle fixation procedure (4% PFA 15 min) without a permeabilization procedure (Fig. 5B). This implies that, even when we skipped the cell permeabilization procedure, subsequent general procedures required for IF may cause diffusion of mitochondrial EGFP-CTSD in the cell. When we performed a permeabilization procedure, which is essential to detect proteins in mitochondria surrounded by a double-membrane (0.1% Triton X-100 5 min), mitochondrial localization of EGFP-CTSD was more corrupted (Fig. 5C) or completely lost with immunostaining procedures, though staining of TOMM20 and cytochrome *c* was still maintained (Fig. 5D and 5E). These results suggest that immunostaining experiments with fixed cells may fail to detect mitochondrial localization of endogenous CTSD.

Discussion

Here, we show that, upon live-cell imaging, EGFP-CTSD (known as lysosomal protease) localizes in mitochondria. Amino acids 21–35 of CTSD were responsible for mitochondrial localization of CTSD, which presented typical features of an MTS and were evolutionarily conserved. Furthermore, a subpopulation of endogenous CTSD was protected from proteinase K, and detected in mitochondrial fractions by sucrose gradient analysis, indicating that a small proportion of endogenous CTSD could be localized in mitochondria. Molecular weights of mitochondrial forms of CTSD were higher than those of CTSD mutants without the MTS in WB, which suggests that mitochondrial forms are not cleavable at a putative MPP site in mitochondria. We also found that mitochondrial CTSD was unstable in fixed cells. That could explain why it has not been detected in past IF experiments.

In the present study, we identified a previously unknown mitochondrial localization of CTSD, one of the most investigated lysosomal proteases [3, 4, 29]. Although mitochondrial targeting mechanisms for subcellular proteins are conservative [30], the mitochondrial localization of CTSD has been overlooked by more than 6,400 studies in PubMed until this discovery. Since mitochondrial localization of CTSD has been confirmed in various cell types, and the MTS is conserved across all species examined, this phenomenon may be universal. In addition, since CTSD is associated with various common diseases including neurodegenerative diseases [31], cancer [32] [33], and atherosclerosis [34], our finding may stimulate a lot of interest and discussion regarding associated pathophysiology and possible therapeutic strategies.

Among 798 proteins that fulfill criteria for mitochondrial localization, including genome-wide experimental screens [35], 316 proteins are potentially dual-targeted, with a second location prone to be concealed by uneven distribution [35, 36]. However, in our study, mitochondrial CTSD was clearly visualized by live-cell imaging analysis of over-expressed EGFP-CTSD. Thus, live-cell imaging analysis of over-expressed proteins with optimized fluorescent tags could be a powerful tool for detecting potential dual-targeted proteins to overcome "eclipsed distributions", probably comprising a considerable number of proteins [35, 36].

Why has CTSD never been detected in past imaging analyses? We suggest several possible reasons. (1) Live-cell imaging analysis of over-expressed EGFP-CTSD may have not been performed in previous studies. (2) Since the amount of mitochondrial endogenous CTSD is very small, these forms may have been overlooked in past biochemical analyses, such as WB. The mitochondrial volume is much larger than the lysosomal volume in cells, suggesting that further dilution of a small amount of endogenous CTSD may have made detection difficult. (3) Immunostaining experiments with fixed cells fail to detect mitochondrial CTSD. In this study, we demonstrated the fragility of mitochondrial CTSD proteins in IF experiments by using EGFP-CTSD and fixed cells. This feature of mitochondrial CTSD in IF experiments may have made it additionally difficult to detect mitochondrial CTSD previously.

What is the rationale for mitochondrial localization of CTSD? CTSD requires an acidic pH to be proteolytically active [29]. In this regard, several reports have demonstrated proteolytic activity of CTSD for some substrates even in neutral pH environments [12, 37, 38]. In addition, interestingly,

7

the CTSD precursor was recently reported to function as phosphatase in neutral pH environments [39]. These findings may provide a rationale for mitochondrial localization of CTSD, and concurrently, may suggest novel functions of CTSD in mitochondria. Further analysis is needed.

NCL is characterized by accumulation of ceroid lipofuscin containing the mitochondrial F1F0 ATPase [1, 2]. Involvement of autophagy has been demonstrated experimentally [40, 41], and the pathophysiology of NCL is explained, in part, by incomplete digestion of mitochondrial products by autophagy [42, 43]. Therefore, NCLs are caused by lysosomal disorders [1]. Among them, patients with CTSD dysfunction (CLN10), manifest early onset in NCL [1]. If mitochondrial forms of CTSD have a regulatory role in mitochondrial functions, the most severe forms of the disease may be attributed to the dysfunction of both mitochondrial and lysosomal CTSDs. In addition, some double-membrane structures in CLN10 containing mitochondrial F1F0 ATPase may reflect not only autophagic vacuoles (AVs), but also abnormal mitochondria themselves. Consistent with this idea, Sulzer et al. listed mitochondria as the only candidate, other than AVs [42], and some autolysosome-like bodies in CLN10 models [2] [40] sharing a common appearance with onion-like mitochondria, which are an abnormal form of mitochondria [44-47].

So far, CTSD has always been viewed as a lysosomal protein. The existence of mitochondrial CTSD demonstrated here may provide many new insights and may suggest new research questions regarding classical conceptions about CTSD.

8

Experimental Procedures

Cell lines

The following cell lines were purchased from the American Type Culture Collection: HeLa (tissue, cervical cancer, female), SK-BR-3 (tissue, breast cancer, female), HepG2 (tissue, liver cancer, male), and 293 (tissue, embryonic kidney). Cells were cultured in DMEM (Sigma). All media were supplemented with 10% fetal bovine serum. Cells were maintained at 37°C in a humidified chamber with 5% CO_2.

Establishment of KD cell lines

We established a CTSD-KD cell line using HeLa cells, as previously described [16]. CTSD expression was inhibited in this cell line by retroviral expression of short-hairpin RNA (shRNA) against the CTSD sequence [18]. We also established HeLa-cont cells using an empty retroviral vector. All oligonucleotides are listed in Supplementary Table S1.

Plasmid construction

For construction of plasmids containing C-terminal enhanced green fluorescent protein-CTSD (pEGFP-CTSD), the nucleotide sequence of CTSD was PCR-amplified using the primers, CTSD-F1 and CTSD-R1. PCR products were digested with Nhe I and Hind III and ligated into pC-EGFP [19] cut with the same enzymes.

For construction of plasmids containing CTSD (pCTSD), pEGFP-CTSD was subjected to inverse PCR using the primers, CTSD-F2 and CTSD-R2, and the product was self-ligated using a KOD-Plus-Mutagenesis Kit (TOYOBO).

For construction of plasmids containing deletion mutants of C-terminal EGFP-CTSD and CTSD (Δ2-20, Δ2-35, Δ2-64, and Δ21-35), pEGFP-CTSD and pCTSD were subjected to inverse PCR using the primers, Δ-20F and Δ2-R, Δ-35F and Δ2-R, Δ-64F and Δ2-R, and Δ-35F and Δ21-R, respectively, and the product was self-ligated using a KOD-Plus-Mutagenesis Kit (TOYOBO). All primers are listed in Supplementary Table S1.

Recombinant adenovirus construction

Replication-deficient recombinant viral Ad-CTSD was generated from pCTSD and purified as described previously [20]. Briefly, DNA fragments obtained by restriction of each plasmid vector were blunted using T4 DNA polymerase, ligated into the SmiI site of the cosmid, pAxCAwtit (Takara), which contains the CAG promoter and the entire genome of type 5 adenovirus, except the E1 and E3 regions. Recombinant adenoviruses were generated using *in vitro* homologous recombination in the 293 cell line with the cDNA-inserted pAxCAwtit and the adenovirus DNA terminal–protein complex. Viruses were propagated in the 293 cell line and purified with two rounds of CsCl density centrifugation. Viral titers were determined with a limiting dilution bioassay using 293 cells.

9

Transfection

Cells seeded in 35-mm glass bottom dishes, 8-chamber slides (0.7 cm^2 culture area), or 100-mm dishes were transfected with plasmids (2 μg /dish, 0.4 μg /well, or 10 μg /dish respectively) using FuGENE6 transfection reagent (Promega), according to the manufacturer's instructions.

Adenoviral infection

For transfection, cells were seeded (2×10^6 cells/dish) in 100-mm dishes. Viral solutions were added to these cell monolayers, and cells were incubated at 37°C for 120 min with brief agitation every 20 min. This was followed by addition of culture medium, and cells were returned to the 37°C incubator.

Amino acid sequence analyses of CTSD

The amino acid sequence of CTSD was analyzed using bioinformatics tools, DeepLoc-1.0 [21], DeepLoc-2.0 [22], MitoFates [23], and MITOPROT II [24].

Analyses of confocal microscopy images

Confocal microscopy images were taken with a FLUOVIEW FV3000 confocal laser scanning microscope (Olympus). Line-scan profiles were acquired using MetaMorph ver. 7.8 (Molecular Devices).

Proteinase K Protection Assay

HeLa and 293 cells (4×10^6 cells/100-mm dish) were cultured under normal conditions for 24 h. HeLa cells (2×10^6 cells/100-mm dish) 24 h after transfection with pEGFP-CTSD, or HeLa cells (2×10^6 cells/100-mm dish) 24 h after infection with Ad-CTSD were harvested and suspended in 10 mM HEPES-KOH pH 7.5, 220 mM mannitol, 70mM sucrose. Suspended cells were homogenized using a syringe with a 27-gauge needle. Cell homogenates were centrifuged at $300 \times g$ at 4 °C for 5 min to remove pellets containing unbroken cells and nuclei. Supernatants were centrifuged at $7,000 \times g$ at 4 °C for 10 min to obtain pellets of crude mitochondria. Pellets of crude mitochondria were resuspended in 10 mM HEPES-KOH pH 7.5, 220 mM mannitol, 70 mM sucrose, and incubated with 0, 100, 200, 400 μg/mL of proteinase K, or 400 μg/mL of proteinase K with 1% Triton X-100 at 37 °C for 30 min. Subsequently, 1 mM phenylmethylsulfonyl fluoride (PMSF) was added to stop the reaction, followed by a 20 min incubation on ice. Samples were subjected to Western blotting.

Sucrose Gradient

HeLa cells (1.2×10^7 cells/100-mm dish) cultured under normal conditions for 24 h were harvested and suspended in 10 mM HEPES-KOH pH 7.5, 0.3 M sucrose containing Complete Protease Inhibitor Cocktail (Roche). Suspended cells were homogenized using a syringe with a 27-gauge needle. The cell homogenate was centrifuged at $300 \times g$ at 4 °C for 5 min to remove unbroken cells and nuclei. Supernatant (1 vol.) was layered on top of a sucrose gradient consisting of 1.8 M, 1.5M, 1.2 M, 0.9 M, 0.6 M sucrose in 10 mM HEPES-KOH pH 7.5, 5 vol. containing Complete Protease Inhibitor

Cocktail (Roche). After centrifugation at 100,000 × g at 4 °C for 16 h, 12 fractions were collected from the top.

Western blotting analysis

Proteins separated with SDS-PAGE were transferred onto polyvinylidene difluoride (PVDF) membranes (Millipore). Membranes were blocked with 5% skim milk in Tris-buffered saline with 0.05% Tween 20 for 1 h. Membranes were incubated with primary antibodies at 4 °C overnight, and subsequently with secondary antibodies for 1 h. Antibodies used were as follows: anti-CTSD mouse monoclonal (1:1000, CTD-19, NOVUS), anti-CTSD goat polyclonal (1:250, sc-6486, Santa Cruz Biotechnology), anti-TOMM20 (1:250, sc-17764, Santa Cruz Biotechnology), anti-ATP5F1B (1:400, PAB28448, Abnova), anti-SEC22B (1:200, sc-101267, Santa Cruz Biotechnology), anti-Golgin97 (1:200, sc-59820, Santa Cruz Biotechnology), anti-beta-actin (1:5000, A5316, Sigma-Aldrich), and anti-GFP (sc-9996, Santa Cruz Biotechnology), anti-rabbit IgG-HRP (1:5000, sc-2004, Santa Cruz Biotechnology), anti-mouse IgG-HRP (1:5000, sc-2005, Santa Cruz Biotechnology), and anti-goat IgG-HRP (1:5000, sc-2020, Santa Cruz Biotechnology)

Analysis of CTSD localization in fixed cells

HeLa cells transfected with pEGFP-CTSD were subjected to none, one, pairs, or all of the following procedures for immunofluorescence (IF), 24 h after transfection. i) Fixation with 4% paraformaldehyde for 15 min. ii) Permeabilization with 0.1% Triton X-100 for 5 min. iii) Subsequent procedures for antigen-antibody reaction: washing 3x with phosphate-buffered saline (PBS) at room temperature, blocking with 3% bovine serum albumin (BSA) in PBS for 1 h, incubating with primary antibody, anti-TOMM20 (1:100, sc-17764, Santa Cruz Biotechnology) or anti-cytochrome *c* (1:100, 6H2.B4, Imgenex) in 3% BSA in PBS for 2 h at room temperature, washing 3x with PBS, incubating with secondary antibody, Alexa Fluor 546 (1:200, A-11003, Thermo Fisher Scientific) in 3% BSA in PBS for 1 h at room temperature, and washing 3x with PBS. Cells were mounted with VECTASHIELD H-1000 (Vector Laboratories) and observed using a FLUOVIEW FV3000 confocal laser scanning microscope (Olympus).

Statistical analysis

Statistical analyses were performed in JMP 14.2.0 (SAS). Levels of significance were assessed using Student's two-tailed t-tests. $p < 0.05$ was considered statistically significant.

11

Data availability

All data are contained within this manuscript.

Supporting information

This article contains supporting information.

Acknowledgments

We thank T. Shibata, Y. Nakamura, K. Honjo, and Y. Sagami for their assistance.

Author contribution

N.I. and H.A. conceptualization; N.I. and H.A. data curation; N.I. and H.A. funding acquisition; N.I. formal analysis; N.I. and H.A. investigation; N.I. methodology; H.A. project administration; H.A. supervision; N.I. and H.A. writing – original draft.

Funding and additional information

Conflict of interest

The authors declare no competing interests.

References

[1] J.W. Mink, E.F. Augustine, H.R. Adams, F.J. Marshall, J.M. Kwon, Classification and natural history of the neuronal ceroid lipofuscinoses, J. Child Neurol., 28 (2013) 1101-1105.

[2] M. Koike, H. Nakanishi, P. Saftig, J. Ezaki, K. Isahara, Y. Ohsawa, W. Schulz-Schaeffer, T. Watanabe, S. Waguri, S. Kametaka, M. Shibata, K. Yamamoto, E. Kominami, C. Peters, K. von Figura, Y. Uchiyama, Cathepsin D deficiency induces lysosomal storage with ceroid lipofuscin in mouse CNS neurons, J. Neurosci., 20 (2000) 6898-6906.

[3] N. Zaidi, A. Maurer, S. Nieke, H. Kalbacher, Cathepsin D: a cellular roadmap, Biochem. Biophys. Res. Commun., 376 (2008) 5-9.

[4] O. Mijanovic, A.I. Petushkova, A. Brankovic, B. Turk, A.B. Solovieva, A.I. Nikitkina, S. Bolevich, P.S. Timashev, A. Parodi, A.A. Zamyatnin, Jr., Cathepsin D-Managing the Delicate Balance, Pharmaceutics, 13 (2021).

[5] M.L. Reitman, S. Kornfeld, UDP-N-acetylglucosamine:glycoprotein N-acetylglucosamine-1-phosphotransferase. Proposed enzyme for the phosphorylation of the high mannose oligosaccharide units of lysosomal enzymes, J. Biol. Chem., 256 (1981) 4275-4281.

[6] A. Varki, S. Kornfeld, Identification of a rat liver alpha-N-acetylglucosaminyl phosphodiesterase capable of removing "blocking" alpha-N-acetylglucosamine residues from phosphorylated high mannose oligosaccharides of lysosomal enzymes, J. Biol. Chem., 255 (1980) 8398-8401.

[7] S. Markmann, M. Thelen, K. Cornils, M. Schweizer, N. Brocke-Ahmadinejad, T. Willnow, J. Heeren, V. Gieselmann, T. Braulke, K. Kollmann, Lrp1/LDL Receptor Play Critical Roles in Mannose 6-Phosphate-Independent Lysosomal Enzyme Targeting, Traffic, 16 (2015) 743-759.

[8] V. Laurent-Matha, D. Derocq, C. Prébois, N. Katunuma, E. Liaudet-Coopman, Processing of human cathepsin D is independent of its catalytic function and auto-activation: involvement of cathepsins L and B, J. Biochem., 139 (2006) 363-371.

[9] V.J. Butler, W.A. Cortopassi, A.R. Argouarch, S.L. Ivry, C.S. Craik, M.P. Jacobson, A.W. Kao, Progranulin Stimulates the In Vitro Maturation of Pro-Cathepsin D at Acidic pH, J. Mol. Biol., 431 (2019) 1038-1047.

[10] P. Briozzo, M. Morisset, F. Capony, C. Rougeot, H. Rochefort, In vitro degradation of extracellular matrix with Mr 52,000 cathepsin D secreted by breast cancer cells, Cancer Res., 48 (1988) 3688-3692.

[11] C.S. Oliveira, H. Pereira, S. Alves, L. Castro, F. Baltazar, S.R. Chaves, A. Preto, M. Côrte-Real, Cathepsin D protects colorectal cancer cells from acetate-induced apoptosis through autophagy-independent degradation of damaged mitochondria, Cell Death Dis., 6 (2015) e1788.

[12] N. Bidère, H.K. Lorenzo, S. Carmona, M. Laforge, F. Harper, C. Dumont, A. Senik, Cathepsin D triggers Bax activation, resulting in selective apoptosis-inducing factor (AIF) relocation in T lymphocytes entering the early commitment phase to apoptosis, J. Biol. Chem., 278 (2003) 31401-31411.

[13] G. Morrow, H.J. Kim, O. Pellerito, M. Bourrelle-Langlois, M. Le Pécheur, K. Groebe, R.M. Tanguay, Changes in Drosophila mitochondrial proteins following chaperone-mediated lifespan extension confirm a role of Hsp22 in mitochondrial UPR and reveal a mitochondrial localization for cathepsin D, Mech. Ageing Dev., 155 (2016) 36-47.

[14] A. Sickmann, J. Reinders, Y. Wagner, C. Joppich, R. Zahedi, H.E. Meyer, B. Schönfisch, I. Perschil, A. Chacinska, B. Guiard, P. Rehling, N. Pfanner, C. Meisinger, The proteome of Saccharomyces cerevisiae mitochondria, Proc. Natl. Acad. Sci. U. S. A., 100 (2003) 13207-13212.

[15] J. Dontaine, A. Bouali, F. Daussin, L. Bultot, D. Vertommen, M. Martin, R. Rathagirishnan, A. Cuillerier, S. Horman, C. Beauloye, L. Gatto, B. Lauzier, L. Bertrand, Y. Burelle, The intra-mitochondrial O-GlcNAcylation system rapidly modulates OXPHOS function and ROS release in the heart, Communications biology, 5 (2022) 349.

[16] Y. Miyamoto, N. Kitamura, Y. Nakamura, M. Futamura, T. Miyamoto, M. Yoshida, M. Ono, S. Ichinose, H. Arakawa, Possible existence of lysosome-like organella within mitochondria and its role in mitochondrial quality control, PLoS One, 6 (2011) e16054.

[17] K. Okuyama, Y. Kitajima, N. Egawa, H. Kitagawa, K. Ito, S. Aishima, K. Yanagihara, T. Tanaka, H. Noshiro, Mieap-induced accumulation of lysosomes within mitochondria (MALM) regulates gastric cancer cell invasion under hypoxia by suppressing reactive oxygen species accumulation, Sci. Rep., 9 (2019) 2822.

[18] S. Su, X. Zhu, L. Lin, X. Chen, Y. Wang, J. Zi, Y. Dong, Y. Xie, Y. Zhu, J. Zhang, J. Zhu, D. Xu, N. Xu, X. Lou, S. Liu, Lowering Endogenous Cathepsin D Abundance Results in Reactive Oxygen Species Accumulation and Cell Senescence, Mol. Cell. Proteomics, 16 (2017) 1217-1232.

[19] N. Ikari, K. Honjo, Y. Sagami, Y. Nakamura, H. Arakawa, Mieap forms membraneless organelles to compartmentalize and facilitate cardiolipin metabolism. bioRxiv. 2022 doi: 10.1101/2020.10.26.354365.

[20] K. Oda, H. Arakawa, T. Tanaka, K. Matsuda, C. Tanikawa, T. Mori, H. Nishimori, K. Tamai, T. Tokino, Y. Nakamura, Y. Taya, p53AIP1, a potential mediator of p53-dependent apoptosis, and its regulation by Ser-46-phosphorylated p53, Cell, 102 (2000) 849-862.

[21] J.J. Almagro Armenteros, C.K. Sønderby, S.K. Sønderby, H. Nielsen, O. Winther, DeepLoc: prediction of protein subcellular localization using deep learning, Bioinformatics, 33 (2017) 3387-3395.

[22] V. Thumuluri, J.J. Almagro Armenteros, A.R. Johansen, H. Nielsen, O. Winther, DeepLoc 2.0: multi-label subcellular localization prediction using protein language models, Nucleic Acids Res., (2022).

[23] Y. Fukasawa, J. Tsuji, S.C. Fu, K. Tomii, P. Horton, K. Imai, MitoFates: improved prediction of mitochondrial targeting sequences and their cleavage sites, Mol. Cell. Proteomics, 14 (2015) 1113-1126.

[24] M.G. Claros, P. Vincens, Computational method to predict mitochondrially imported proteins and their targeting sequences, Eur. J. Biochem., 241 (1996) 779-786.

[25] M. Dinur-Mills, M. Tal, O. Pines, Dual targeted mitochondrial proteins are characterized by lower MTS parameters and total net charge, PLoS One, 3 (2008) e2161.

[26] S.K. Burley, C. Bhikadiya, C. Bi, S. Bittrich, L. Chen, G.V. Crichlow, C.H. Christie, K. Dalenberg, L. Di Costanzo, J.M. Duarte, S. Dutta, Z. Feng, S. Ganesan, D.S. Goodsell, S. Ghosh, R.K. Green, V. Guranović, D. Guzenko, B.P. Hudson, C.L. Lawson, Y. Liang, R. Lowe, H. Namkoong, E. Peisach, I. Persikova, C. Randle, A. Rose, Y. Rose, A. Sali, J. Segura, M. Sekharan, C. Shao, Y.P. Tao, M. Voigt, J.D. Westbrook, J.Y. Young, C. Zardecki, M. Zhuravleva, RCSB Protein Data Bank: powerful new tools for exploring 3D structures of biological macromolecules for basic and applied research and education in fundamental biology, biomedicine, biotechnology, bioengineering and energy sciences, Nucleic Acids Res., 49 (2021) D437-d451.

14

[27] Y. Abe, T. Shodai, T. Muto, K. Mihara, H. Torii, S. Nishikawa, T. Endo, D. Kohda, Structural basis of presequence recognition by the mitochondrial protein import receptor Tom20, Cell, 100 (2000) 551-560.

[28] A. Chacinska, C.M. Koehler, D. Milenkovic, T. Lithgow, N. Pfanner, Importing mitochondrial proteins: machineries and mechanisms, Cell, 138 (2009) 628-644.

[29] P. Benes, V. Vetvicka, M. Fusek, Cathepsin D--many functions of one aspartic protease, Crit. Rev. Oncol. Hematol., 68 (2008) 12-28.

[30] N. Pfanner, Protein sorting: recognizing mitochondrial presequences, Curr. Biol., 10 (2000) R412-415.

[31] J. Bunk, S. Prieto Huarcaya, A. Drobny, J.P. Dobert, L. Walther, S. Rose-John, P. Arnold, F. Zunke, Cathepsin D Variants Associated With Neurodegenerative Diseases Show Dysregulated Functionality and Modified α-Synuclein Degradation Properties, Frontiers in cell and developmental biology, 9 (2021) 581805.

[32] S. Ketterer, J. Mitschke, A. Ketscher, M. Schlimpert, W. Reichardt, N. Baeuerle, M.E. Hess, P. Metzger, M. Boerries, C. Peters, B. Kammerer, T. Brummer, F. Steinberg, T. Reinheckel, Cathepsin D deficiency in mammary epithelium transiently stalls breast cancer by interference with mTORC1 signaling, Nature communications, 11 (2020) 5133.

[33] S.U. Seo, S.M. Woo, S.S. Im, Y. Jang, E. Han, S.H. Kim, H. Lee, H.S. Lee, J.O. Nam, E. Gabrielson, K.J. Min, T.K. Kwon, Cathepsin D as a potential therapeutic target to enhance anticancer drug-induced apoptosis via RNF183-mediated destabilization of Bcl-xL in cancer cells, Cell Death Dis., 13 (2022) 115.

[34] J.K. Hakala, R. Oksjoki, P. Laine, H. Du, G.A. Grabowski, P.T. Kovanen, M.O. Pentikäinen, Lysosomal enzymes are released from cultured human macrophages, hydrolyze LDL in vitro, and are present extracellularly in human atherosclerotic lesions, Arterioscler. Thromb. Vasc. Biol., 23 (2003) 1430-1436.

[35] I. Kisslov, A. Naamati, N. Shakarchy, O. Pines, Dual-targeted proteins tend to be more evolutionarily conserved, Mol. Biol. Evol., 31 (2014) 2770-2779.

[36] N. Regev-Rudzki, O. Pines, Eclipsed distribution: a phenomenon of dual targeting of protein and its significance, Bioessays, 29 (2007) 772-782.

[37] A. Kenessey, P. Nacharaju, L.W. Ko, S.H. Yen, Degradation of tau by lysosomal enzyme cathepsin D: implication for Alzheimer neurofibrillary degeneration, J. Neurochem., 69 (1997) 2026-2038.

[38] M. Heinrich, J. Neumeyer, M. Jakob, C. Hallas, V. Tchikov, S. Winoto-Morbach, M. Wickel, W. Schneider-Brachert, A. Trauzold, A. Hethke, S. Schütze, Cathepsin D links TNF-induced acid sphingomyelinase to Bid-mediated caspase-9 and -3 activation, Cell Death Differ., 11 (2004) 550-563.

[39] Y.J. Liu, T. Zhang, S. Chen, D. Cheng, C. Wu, X. Wang, D. Duan, L. Zhu, H. Lou, Z. Gong, X.D. Wang, M.S. Ho, S. Duan, The noncanonical role of the protease cathepsin D as a cofilin phosphatase, Cell Res., 31 (2021) 801-813.

[40] M. Koike, M. Shibata, S. Waguri, K. Yoshimura, I. Tanida, E. Kominami, T. Gotow, C. Peters, K. von Figura, N. Mizushima, P. Saftig, Y. Uchiyama, Participation of autophagy in storage of lysosomes in neurons from mouse models of neuronal ceroid-lipofuscinoses (Batten disease), Am. J. Pathol., 167 (2005) 1713-1728.

[41] A.R.A. Marques, A. Di Spiezio, N. Thießen, L. Schmidt, J. Grötzinger, R. Lüllmann-Rauch, M. Damme, S.E. Storck, C.U. Pietrzik, J. Fogh, J. Bär, M. Mikhaylova, M. Glatzel, M. Bassal, U. Bartsch, P. Saftig, Enzyme replacement therapy with recombinant pro-CTSD (cathepsin D) corrects defective proteolysis

15

and autophagy in neuronal ceroid lipofuscinosis, Autophagy, 16 (2020) 811-825.

[42] D. Sulzer, E. Mosharov, Z. Talloczy, F.A. Zucca, J.D. Simon, L. Zecca, Neuronal pigmented autophagic vacuoles: lipofuscin, neuromelanin, and ceroid as macroautophagic responses during aging and disease, J. Neurochem., 106 (2008) 24-36.

[43] D.N. Palmer, L.A. Barry, J. Tyynelä, J.D. Cooper, NCL disease mechanisms, Biochim. Biophys. Acta, 1832 (2013) 1882-1893.

[44] V. Soubannier, J. Vaillier, P. Paumard, B. Coulary, J. Schaeffer, J. Velours, In the absence of the first membrane-spanning segment of subunit 4(b), the yeast ATP synthase is functional but does not dimerize or oligomerize, J. Biol. Chem., 277 (2002) 10739-10745.

[45] P. Paumard, J. Vaillier, B. Coulary, J. Schaeffer, V. Soubannier, D.M. Mueller, D. Brèthes, J.P. di Rago, J. Velours, The ATP synthase is involved in generating mitochondrial cristae morphology, EMBO J., 21 (2002) 221-230.

[46] L. Zhang, P. Yang, X. Bian, Q. Zhang, S. Ullah, Y. Waqas, X. Chen, Y. Liu, W. Chen, Y. Le, B. Chen, S. Wang, Q. Chen, Modification of sperm morphology during long-term sperm storage in the reproductive tract of the Chinese soft-shelled turtle, Pelodiscus sinensis, Sci. Rep., 5 (2015) 16096.

[47] R. Anand, A.K. Kondadi, J. Meisterknecht, M. Golombek, O. Nortmann, J. Riedel, L. Peifer-Weiß, N. Brocke-Ahmadinejad, D. Schlütermann, B. Stork, T.O. Eichmann, I. Wittig, A.S. Reichert, MIC26 and MIC27 cooperate to regulate cardiolipin levels and the landscape of OXPHOS complexes, Life science alliance, 3 (2020).

Figure legends

Fig. 1. Mitochondrial localization of EGFP-CTSD.

(A–J) Live-cell imaging showing the subcellular localization of EGFP-CTSD expressed in HeLa **(A and E)**, SK-BR-3 **(B and F)**, HepG2 **(C and G)**, and 293 cells **(D and H)** compared with a lysosome probe, LysoTracker Red **(A–D)** and a mitochondria probe, MitoTracker Red **(E–H)**.

(I and J) Co-localization rates of EGFP-CTSD with LysoTracker Red **(I)** and MitoTracker Red **(J)** in HeLa, SK-BR-3, HepG2, and 293 cells, as in **(A–J)**. n = 103 – 192 cells from three independent experiments. Data shown are means ±SD. Raw data are provided in Supplementary Table S2.

Fig. 2. CTSD has a predicted MTS containing an evolutionarily conserved region (amino acids 21-35).

(A) Prediction of subcellular localization of CTSD by DeepLoc-1.0, depicted in a hierarchical tree diagram.

(B) Prediction of subcellular localization of CTSD by DeepLoc-2.0. The predicted localization is highlighted in green (upper panel). Regions in the sequence responsible for the predicted localization are shown as high values in the Logo-like plot (lower panel).

(C) Prediction of mitochondrial localization and the MTS of CTSD by MITOPROT II and MitoFates.

(D) Prediction of mitochondrial localization of CTSD orthologs in representative eukaryotes by MITOPROT II.

(E) Multiple sequence alignment for CTSD orthologs in representative eukaryotes. Black and gray boxes indicate 100% and 70% identical residues among eukaryotes, respectively. Pink letters indicate positively charged residues. Blue letters indicate negatively charged residues.

Fig. 3. CTSD amino acids 21-35 comprise a bona-fide MTS, responsible for mitochondrial localization

(A) A schematic of EGFP-CTSD and deletion mutants (Δ2–20, Δ2–35, Δ2–64, and Δ21–35). Numbers indicate amino acids.

(B–K) Live-cell imaging showing subcellular localization of EGFP-CTSD full-length **(B and C)**, Δ2–20 **(D and E)**, Δ2–35 **(F and G)**, Δ2–64 **(H and I)**, and Δ21–35 **(J and K)** expressed in HeLa cells, compared with a mitochondrial probe, MitoTracker Red **(B, D, F, H, J)** and a lysosomal probe, LysoTracker Red **(C, E, G, I, K)**.

(L) Live-cell imaging showing subcellular localization of EGFP-CTSDΔ21–35 (left) and Δ2–35 (right) compared with an endoplasmic reticulum (ER) probe, ER Tracker Blue-White DPX in HeLa cells.

(M) Co-localization rates of EGFP-CTSD full-length, Δ2–20, Δ2–35, Δ2–64, and Δ21–35 expressed in HeLa cells with MitoTracker Red (left) and LysoTracker Red (right). n = 104 – 165 cells from three independent experiments. Data shown are means ±SD. Asterisks indicate significance levels as follows: ****, p < 0.0001.

(N) Western blot to evaluate cleavage of EGFP-CTSD expressed in HeLa cells, in comparison with

17

molecular weights of deletion mutants (Δ2–20, Δ2–35, Δ2–64, and Δ21–35) expressed in HeLa cells. Raw data are provided in Supplementary Table S3. Full-gel images are provided in Supplementary Fig. S1.

Fig. 4. A small amount of endogenous CTSD is localized in mitochondria

(A) Verification of specific bands of endogenous CTSD in Western blots. The HeLa parent, control-KD, and two CTSD-KD cells expressing different short-hairpin RNAs against the CTSD sequence (shRNA#1 and shRNA#2) are compared. Two anti-CTSD antibodies (CTD-19, mouse monoclonal; sc-6486 goat polyclonal) were used for accuracy. Arrowheads indicate specific bands of endogenous CTSD.

(B) Proteinase K protection assay for HeLa cells. The crude mitochondrial fraction (7,000 g pellet) isolated from HeLa cells was treated with increasing concentrations of proteinase K in the presence and absence of Triton X-100, followed by Western blotting. The remaining 7,000 g supernatant and whole cell lysates were applied to adjacent lanes for comparison. Arrowheads indicate specific bands of endogenous CTSD verified in **(A)**.

(C and D) For HeLa cells expressing EGFP-CTSD, verification of bands of EGFP-CTSD and endogenous CTSD **(C)** and proteinase K protection assay **(D)** were performed as in **(A and B)**. Arrows and arrowheads indicate bands of EGFP-CTSD and endogenous CTSD, respectively.

(E and F) For HeLa cells expressing exogenous CTSD, verification of bands of CTSD **(E)** and proteinase K protection assay **(F)** were performed as in **(A and B)**. Arrowheads indicate bands of CTSD.

(G) Comparison of ~50kDa bands that contained proteinase K-inaccessible CTSD in **(B)** with bands of CTSD-deletion mutants (Δ2–20, Δ2–35, Δ2–64, and Δ21–35) in Western blots.

(H) For 293 cells, the proteinase K protection assay was performed as in **(B)**.

(I) Sucrose gradient analysis for HeLa cells. HeLa cell extract was fractionated by sucrose density gradient centrifugation. Fractions were collected and subjected to Western blot analysis, using antibodies against indicated proteins. Full-gel images are provided in Supplementary Fig. S2–S7.

Fig. 5. Mitochondrial EGFP-CTSD cannot be detected by IF with fixed cells

(A–E) Live-cell imaging **(A)** and fixed-cell imaging **(B–E)** of HeLa cells expressing EGFP-CTSD with the following procedures. Fixation with 4% PFA for 15 min **(B)**, with permeabilization in 0.1% Triton X-100 for 5 min **(C)**, and immunoreaction **(D and E)** using anti-TOMM20 antibody **(D)** and anti-cytochrome *c* antibody **(E)**.

18

Figure 1

A HeLa
EGFP-CTSD
LysoTracker
50 μm   10 μm   Intensity

B SK-BR-3
EGFP-CTSD
LysoTracker
50 μm   10 μm   Intensity

C HepG2
EGFP-CTSD
LysoTracker
50 μm   10 μm   Intensity

D 293
EGFP-CTSD
LysoTracker
50 μm   10 μm   Intensity

E HeLa
EGFP-CTSD
MitoTracker
50 μm   10 μm   Intensity

F SK-BR-3
EGFP-CTSD
MitoTracker
50 μm   10 μm   Intensity

G HepG2
EGFP-CTSD
MitoTracker
50 μm   10 μm   Intensity

H 293
EGFP-CTSD
MitoTracker
50 μm   10 μm   Intensity

I Positive cells for co-localization (%)
HeLa   SK-BR-3   HepG2   293

J Positive cells for co-localization (%)
HeLa   SK-BR-3   HepG2   293

Figure 2

A

Secretory
Non-Secretory Pathway

0.997   0.003

Intracellular/
Extracellular   N-terminal
sequences

0.803   0.195   0.002   0.0

Intermediate   PTS signals   Mitochondrion/
compartment   Extracellular   Chloroplast

0.777   0.026   0.0   0.002   0.0   0.0

Cell membrane/   ER/Golgi   NLS   Chloroplast
Lysosome   Peroxisome   Mitochondrion

0.008   0.769   0.026   0.0   0.0   0.002

Cell   Lysosome   ER   Golgi   Nucleus   Cytoplasm
membrane   apparatus

B

NP_001900.1

Predicted localizations: Lysosome/Vacuole

Predicted signals: Signal peptide

Localizations	Signals	Cytoplasm	Nucleus	Extracellular	Cell membrane	Mitochondrion	Plastid	Endoplasmic reticulum	Lysosome /Vacuole	Golgi apparatus	Peroxisome
Lysosome /Vacuole	Signal peptide	0.1946	0.0620	0.6046	0.1358	0.1269	0.0031	0.4553	0.6586	0.4546	0.0042

Sequence
Predicted Signals: Signal peptide

C

CTSD

| S | Pro | Light chain | Heavy chain |

1   21   65   169   412

MQPSSLLPLALCLLAAPASA LVRIPLHKFTSIRRTMSEVGGSVEDLIAKGP

1   21   30   35

Amphypathic alpha-helix
in N-terminal 30 residues
(MitoFates)

Probability of export to mitochondria: 0.6115
(MITOPROT II)

MPP cleavage site
(MitoFates)(MITOPROT II)

D

Probability of export to mitochondria (MITOPROT II)

(bar chart with species: H. sapiens, P. troglodytes, M. musculus, G. gallus, X. tropicalis, D. rerio, D. melanogaster)

E

Putative mitochondrial targeting sequence (MITOPROT II)

ER signal peptide (UniProt)

Amphypathic alpha-helix in N-terminal 30 residues (MitoFates)

Light chain (UniProt)

```
H. sapiens       1 ------MQPSSLLPLALCLLAAPASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGPVSKYSQAVPAVTEGPIP.VLANY 74
P. troglodytes   1 ------MQPSSLLLLALCLLAAPASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGPVSKYSQAVPSVTAGPIP.VLANY 74
M. musculus      1 ------MKTPGVLLLILGLLASSFAILRIPLHKFTSIRRTMTEVGGSVEDLIAKGPITKYSMQQSPKTTEPVS.LLANY 74
G. gallus        1 ------MAPGLLVLLLLALVGPCAALIRIPLTKFTSIRRMLTEVGGSEIPDMNAITQFLKFKLGF-ADLAEPTP.ILANY 73
X. tropicalis    1 ---MASAPVWALLALCCVMQPGSSLVRIPLKKFTSIRRAMSETDQDALKLSGNEAATKYS-AF-LNSKNPTP.ILANY 73
D. rerio         1 ------MRIAFLLLVVVAFFCTSDALVRIPLKKFRTLRRTLSDSGRSLEELVSSSNSLKYNLGF-PASNDPTP.VLSNY 71
D. melanogaster  1 MQKVALLLVAFLAAAVAHPNSQEKPGLLVPLHKFQSARRHFADVGTELQQL-----RIRY----------GGGDVP.PLSNY 68
```

Figure 3

Figure 4

Figure 5

A EGFP-CTSD
50 µm
10 µm

B EGFP-CTSD
50 µm
10 µm

C EGFP-CTSD
50 µm
10 µm

D EGFP-CTSD
TOMM20
50 µm
10 µm

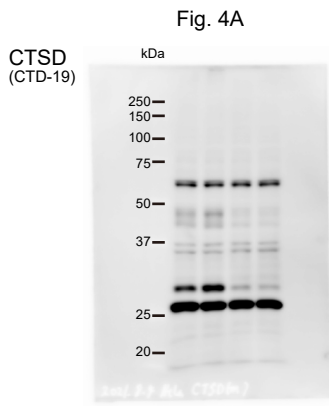E EGFP-CTSD
Cytochrome c
50 µm
10 µm

Fig. S1: Source data of western blot in Fig. 3N
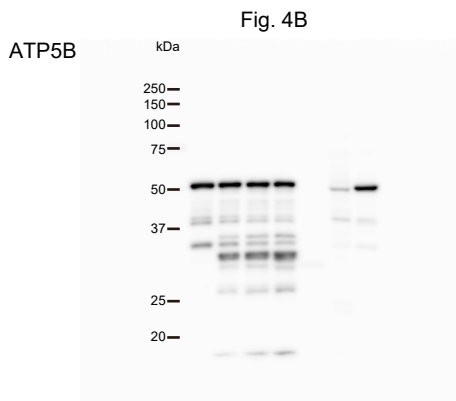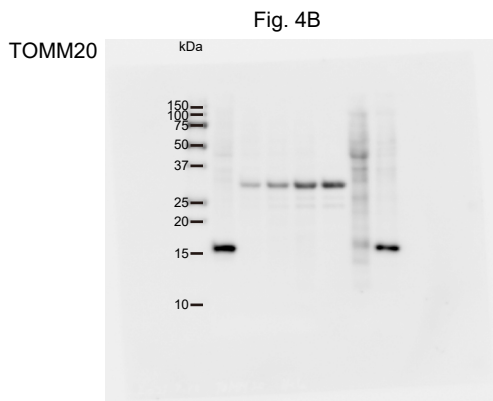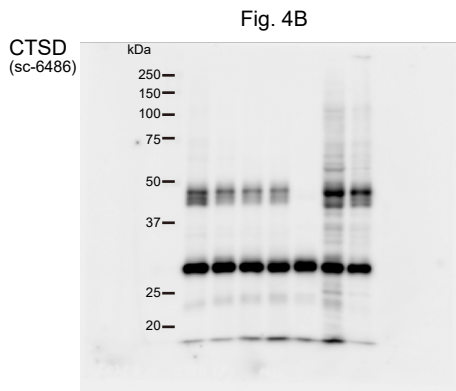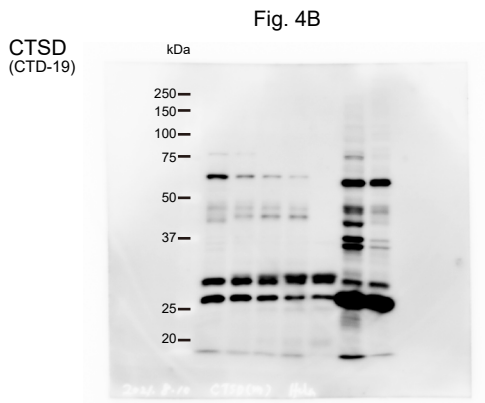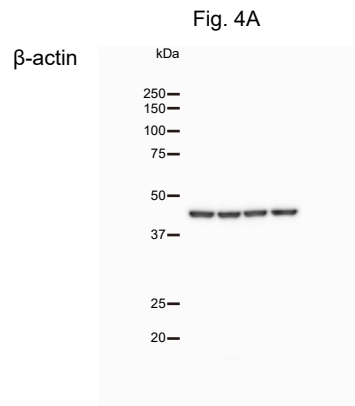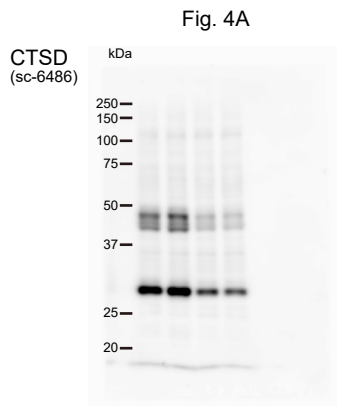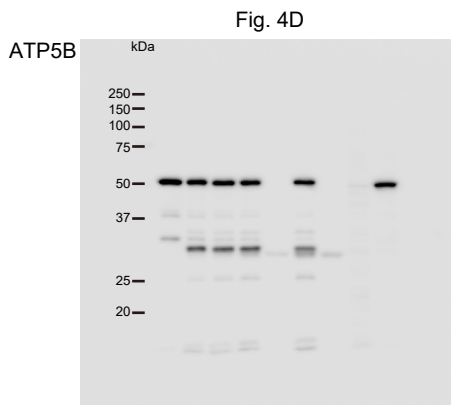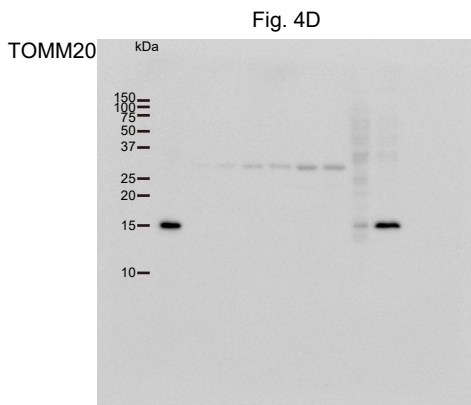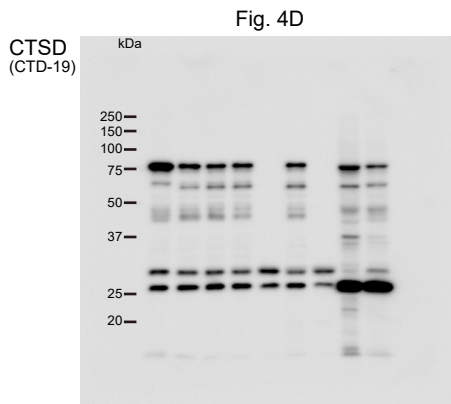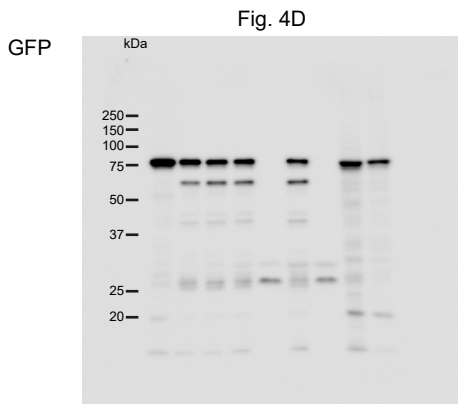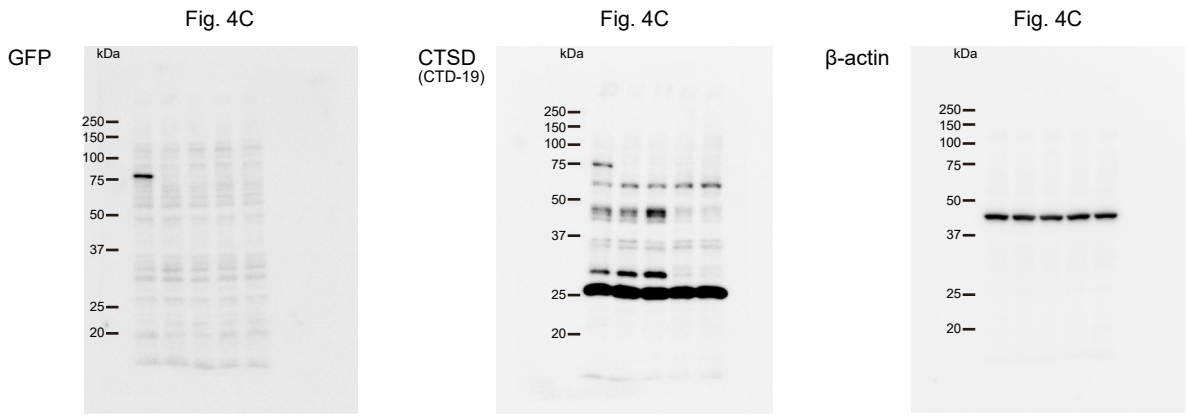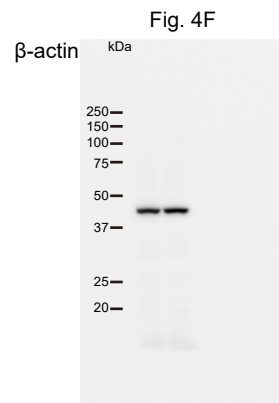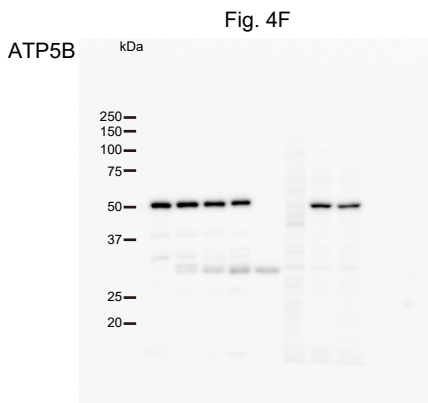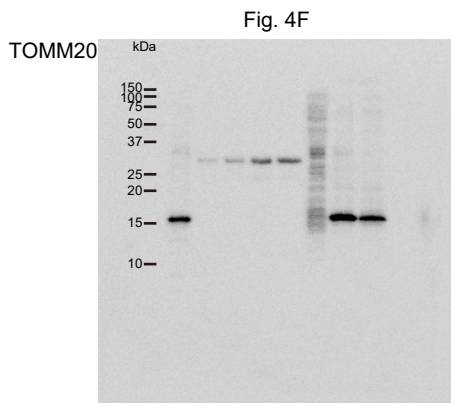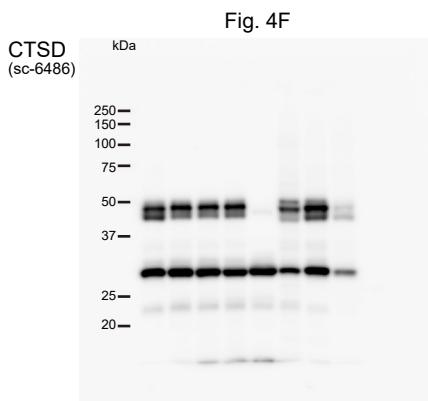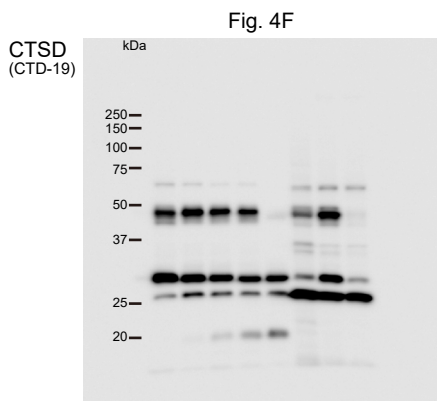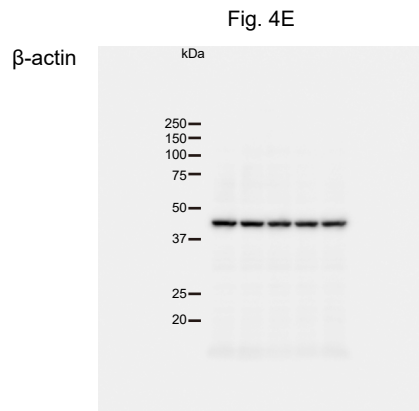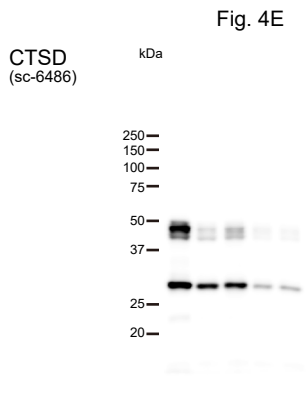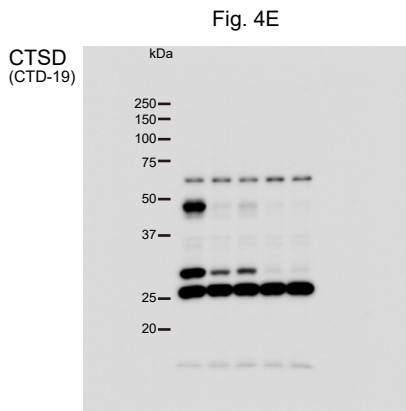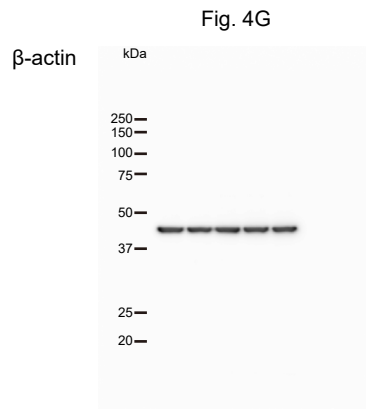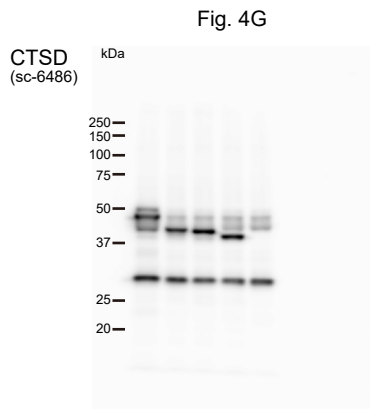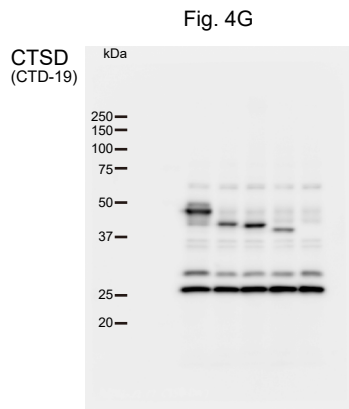
Fig. S2: Source data of western blot in Fig. 4A and 4B

Fig. 4C

GFP
kDa

250—
150—
100—
75—
50—
37—

25—
20—

Fig. 4C

CTSD
(CTD-19)
kDa

250—
150—
100—
75—
50—
37—

25—

20—

Fig. 4C

β-actin
kDa

250—
150—
100—
75—
50—
37—

25—

20—

Fig. 4D

GFP
kDa

250—
150—
100—
75—
50—
37—

25—
20—

Fig. 4D

CTSD
(CTD-19)
kDa

250—
150—
100—
75—
50—
37—

25—

20—

Fig. 4D

TOMM20
kDa

150—
100—
75—
50—
37—

25—
20—

15—

10—

Fig. 4D

ATP5B
kDa

250—
150—
100—
75—

50—

37—

25—

20—

Fig. S3: Source data of western blot in Fig. 4C and 4D

Fig. 4E

CTSD
(CTD-19)

CTSD
(sc-6486)

β-actin

Fig. 4F

CTSD
(CTD-19)

CTSD
(sc-6486)

Fig. 4F

TOMM20

ATP5B

β-actin

Fig. S4: Source data of western blot in Fig. 4E and 4F
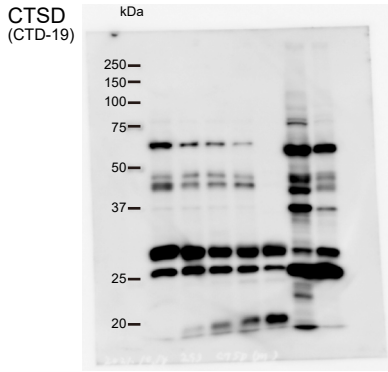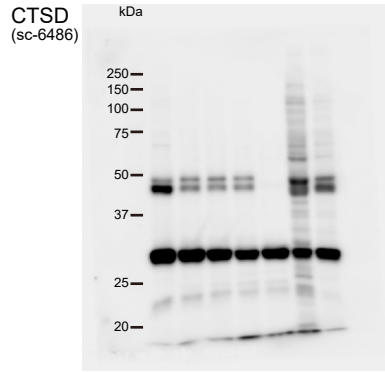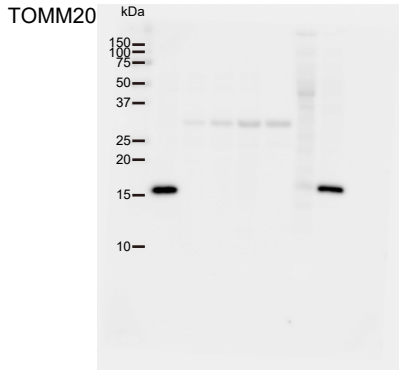
Fig. S5: Source data of western blot in Fig. 4G
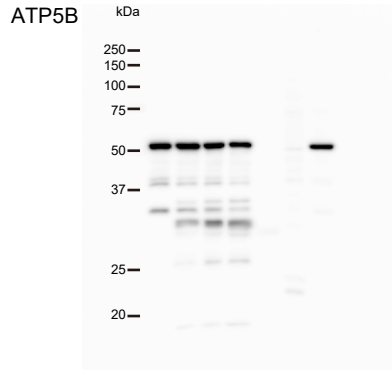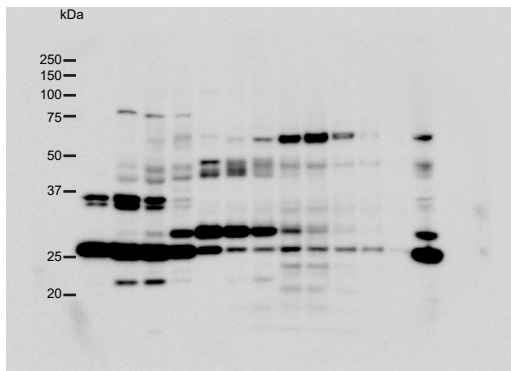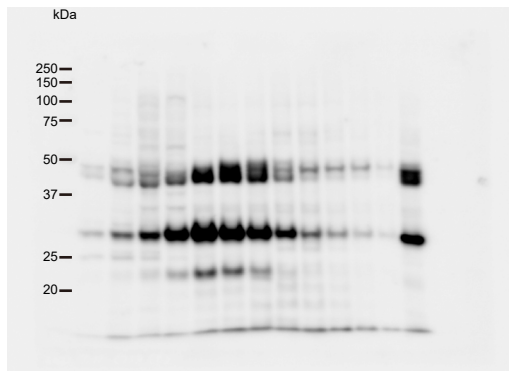
Fig. S6: Source data of western blot in Fig. 4H
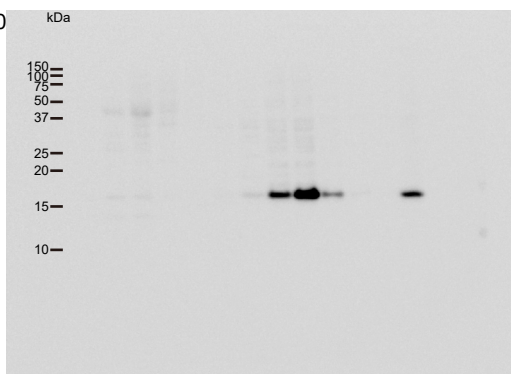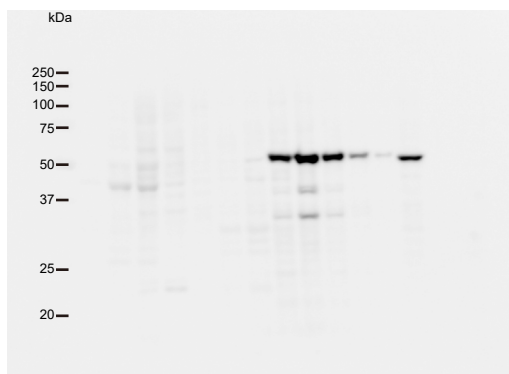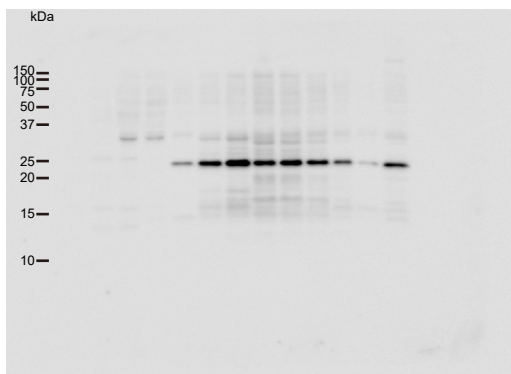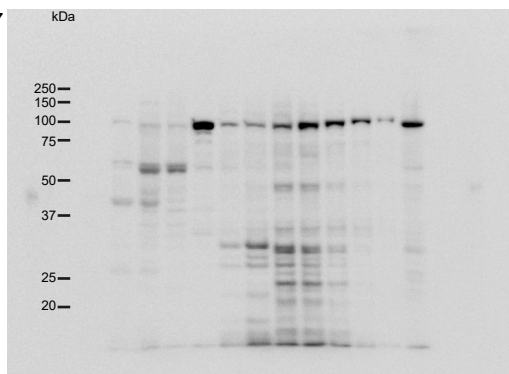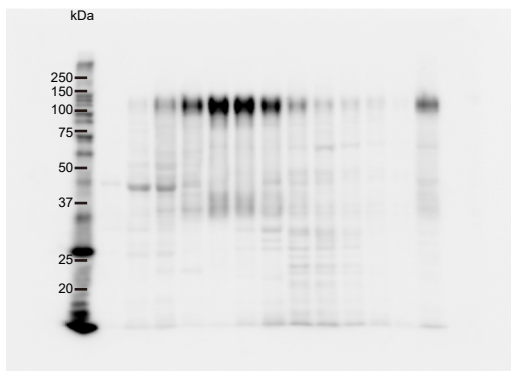
Fig. S7: Source data of western blot in Fig. 4I

Table S1. Oligonucleotides used in this paper

Primers

Name	Sequence
CTSD-F1	5′-AAATAAATGCTAGCGCCACCATGCAGCCCTCCAGCCTTCTGCCGC-3′
CTSD-R1	5′-AAATAAATAAGCTTGAGGCGGGCAGCCTCGGCGAAGCCC-3′
CTSD-F2	5′-TAAGCGGCCGCTCGAGTCTAGAGGG-3′
CTSD-R2	5′-GAGGCGGGCAGCCTCGGCGAAG-3′
Δ-20F	5′-CTCGTCAGGATCCCGCTGCACAAGTTC-3′
Δ2-R	5′-CATGGTGGCGCTAGCCAGCTTGGG-3′
Δ-35F	5′-ATGTCGGAGGTTGGGGGCTCTGTGG-3′
Δ-64F	5′-GGGCCCATTCCCGAGGTGCTCAAG-3′
Δ21-R	5′-CGCGGAGGCGGGTGCAGCCAG-3′

shRNA

Name	Sequence
shRNA CTSD_#1-F	5'-GATCCCCGAATGGTACCTCGTTTGACTTCAAGAGAGTCAAACGAGG TACCATTCTTTTTGGAAA-3'
shRNA CTSD_#1-R	5'-AGCTTTTCCAAAAAGAATGGTACCTCGTTTGACTCTCTTGAAGTCAA ACGAGGTACCATTCGGG-3'
shRNA CTSD_#2-F	5'-GATCCCCGCTACTACACTGTGTTTGATTCAAGAGATCAAACACAGT GTAGTAGCTTTTTGGAAA-3'
shRNA CTSD_#2-R	5'-AGCTTTTCCAAAAAGCTACTACACTGTGTTTGATCTCTTGAATCAA ACACAGTGTAGTAGCGGG-3'

Table S2. Subcellular localization of EGFP-CTSD expressed in various cell lines.

Biological replicate	Cells	Dyes used for co-localization analysis	Number of co-localization positive cells	Number of co-localization negative cells	Positive rate of co-localization (%)
1	HeLa	LysoTracker Red	47	75	38.52
2			61	54	53.04
3			74	86	46.25
1		MitoTracker Red CMXRos	61	66	48.03
2			49	91	35.00
3			68	84	44.74
1	SK-BR-3	LysoTracker Red	51	82	38.35
2			61	87	41.22
3			76	116	39.58
1		MitoTracker Red CMXRos	79	35	69.30
2			128	17	88.28
3			138	38	78.41
1	HepG2	LysoTracker Red	108	72	60.00
2			101	82	55.19
3			67	36	65.05
1		MitoTracker Red CMXRos	70	77	47.62
2			52	93	35.86
3			59	73	44.70
1	293	LysoTracker Red	64	48	57.14
2			79	49	61.72
3			64	48	57.14
1		MitoTracker Red CMXRos	48	73	39.67
2			54	83	39.42
3			39	70	35.78

Table S3. Subcellular localization of EGFP-CTSD and the deletion mutants expressed in HeLa cells.

Biological replicate	Expression vector	Dyes used for co-localization analysis	Number of co-localization positive cells	Number of co-localization negative cells	Positive rate of co-localization (%)
1			56	78	41.79
2	pEGFP-CTSD		48	66	42.11
3			61	67	47.66
1			119	6	95.20
2	pEGFP-CTSDΔ2-20		152	3	98.06
3			139	4	97.20
1		MitoTracker Red CMXRos	0	154	0.00
2	pEGFP-CTSDΔ2-35		0	104	0.00
3			0	150	0.00
1			0	139	0.00
2	pEGFP-CTSDΔ2-64		0	142	0.00
3			0	128	0.00
1			0	154	0.00
2	pEGFP-CTSDΔ21-35		0	137	0.00
3			0	126	0.00
1			68	65	51.13
2	pEGFP-CTSD		65	72	47.45
3			63	52	54.78
1			0	139	0.00
2	pEGFP-CTSDΔ2-20		0	136	0.00
3			0	153	0.00
1		LysoTracker Red	0	143	0.00
2	pEGFP-CTSDΔ2-35		0	128	0.00
3			0	165	0.00
1			0	139	0.00
2	pEGFP-CTSDΔ2-64		0	112	0.00
3			0	107	0.00
1			0	165	0.00
2	pEGFP-CTSDΔ21-35		0	113	0.00
3			0	128	0.00