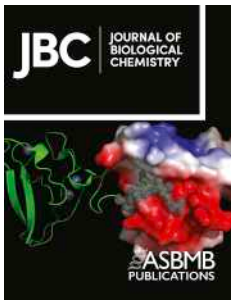


Arylsulfatase L is a Golgi chondroitin sulfatase regulating skeletal development

Marianna Maddaluno, Chiara De Leonibus, Eugenio Del Prete, Francesco Giuseppe Salierno, Daniela Intartaglia, Diego Carrella, Ivan Conte, Nicola Volpi, Carmine Settembre

PII: S0021-9258(26)01983-6

DOI: https://doi.org/10.1016/j.jbc.2026.113111

Reference: JBC 113111

To appear in: *Journal of Biological Chemistry*

Received Date: 23 October 2025

Revised Date: 23 April 2026

Accepted Date: 25 April 2026

Please cite this article as: Maddaluno M, De Leonibus C, Del Prete E, Salierno FG, Intartaglia D, Carrella D, Conte I, Volpi N, Settembre C, Arylsulfatase L is a Golgi chondroitin sulfatase regulating skeletal development, *Journal of Biological Chemistry* (2026), doi: https://doi.org/10.1016/j.jbc.2026.113111.

Arylsulfatase L is a Golgi chondroitin sulfatase regulating skeletal development

Marianna Maddaluno[1,2*], Chiara De Leonibus[1,3*], Eugenio Del Prete[1], Francesco Giuseppe Salierno[1], Daniela Intartaglia[4], Diego Carrella[1], Ivan Conte[5], Nicola Volpi[6#] and Carmine Settembre[1,2#]

Affiliation:

1: Telethon Institute of Genetics and Medicine (TIGEM), Pozzuoli, Italy

2: Department of Clinical Medicine and Surgery, Federico II University, Naples, Italy

3: Department of Health Sciences, University of Basilicata, Potenza, Italy

4: CNR, Institute of Applied Sciences and Intelligent Systems "Eduardo Caianiello" (ISASI), Pozzuoli, Italy

5: Department of Biology, Federico II University, Naples, Italy

6: Department of Life Sciences, UniMORE, University of Modena and Reggio Emilia, Italy

*: These authors contributed equally

#: Corresponding authors

#Nicola Volpi, Via Università 4, 41121 Modena, Tel. 059 2056511, volpi@unimore.it

#Carmine Settembre, Via Campi Flegrei 34, Pozzuoli (NA), Tel. +39 08119230601, settembre@tigem.it

Abstract

Sulfatases are a family of enzymes that hydrolyze sulfate esters from various substrates. Defects, in sulfatase activity, are associated with various human diseases due to the accumulation of sulfated substrates. Deficiency in ARSL, a Golgi sulfatase, is associated with X-linked recessive chondrodysplasia punctata (CDPX), a disorder characterized by defects in cartilage and bone development. However, until now, ARSL function has remained unknown. In this study, we demonstrate that ARSL promotes 4-O-desulfation of Chondroitin Sulfate (CS) during proteoglycan biosynthesis. Chondrocytes lacking ARSL exhibit hypersulfated CS and altered responses to TGF-β stimulation. Loss of function of ARSL orthologous in medaka fish (Ol-Arsd) results in hyper-4-O-sulfated CS, skeletal malformations, and craniofacial defects that partly resemble the human CDPX

phenotype. Our findings uncover a previously unrecognized step in glycosaminoglycan (GAG) maturation—Golgi-based desulfation—and reveal a new layer of regulatory control in skeletal development.

Keywords: Sulfatase, Proteoglycan; Chondroitin sulfate; chondrodysplasia.

Introduction

Proteoglycans are essential components of the extracellular matrix (ECM). They consist of a core protein covalently linked to highly sulfated glycosaminoglycan (GAG) chains, primarily heparan sulfate (HS), chondroitin sulfate (CS), and dermatan sulfate (DS), with CS being the predominant GAG in cartilage and bone (1). HS and CS/DS differ in disaccharide composition and sulfation patterns: HS chains are composed of N-acetylglucosamine and glucuronic or iduronic acid residues modified at defined N- and O-positions, whereas CS/DS chains consist of N-acetylgalactosamine and uronic acid residues predominantly sulfated at the 4-O and 6-O positions (Fig. S1A).

The pattern of GAG sulfation critically influences the physical properties of the matrix and modulates developmental processes by fine-tuning the distribution and activity of signaling molecules within the ECM (1). GAG sulfation is carried out in the Golgi apparatus by sulfotransferases, such as C4ST-1 and C6ST. Defects in these enzymes are associated with congenital osteochondrodysplasia (OCBMD, OMIM: 618167 Osteochondrodysplasia, brachydactyly, and overlapping malformed digits) and spondyloepiphyseal dysplasia (SED, OMIM 608637) Omani type 1 (2–4).Conversely, GAG desulfation is mediated by sulfatases. Seventeen sulfatase genes have been identified in the human genome, classified according to their cellular localization. Lysosomal sulfatases, active at acidic pH, desulfate GAGs during their stepwise degradation within lysosomes. Mutations in these enzymes lead to lysosomal accumulation of sulfated GAGs and are associated with mucopolysaccharidoses (MPS) types I–VI, inherited diseases often characterized by severe skeletal manifestations (5). GAG desulfation can also occur at the plasma membrane, where the sulfatases SULF1 and SULF2 desulfate heparan sulfate chains at the 6-O-position, hence influencing fibroblast growth factor (FGF) signaling during development (6, 7)

The functions of some sulfatases have not been well characterized. Arylsulfatase L (ARSL, formerly known as ARSE) is a Golgi-resident sulfatase that is mutated in patients with X-linked chondrodysplasia punctata (CDPX) (8, 9). CDPX is characterized by epiphyseal calcifications (“stippling”), facial dysmorphisms (severe nasal hypoplasia, depressed nasal bridge, short nasal septum), and short stature (10, 11). Interestingly, warfarin, an inhibitor of vitamin K epoxide reductase

(VKOR), induces warfarin embryopathy, a condition with clinical features overlapping CDPX (8, 12). Warfarin has been shown to inhibit ARSL activity, at least in vitro, suggesting secondary ARSL deficiency is involved in warfarin embryopathy (8). Thus, understanding the function of ARSL will not only represent a first step toward understanding of CDPX pathophysiology but more broadly will expand our knowledge of skeletal development process.

To date no murine Arsl gene has been identified. Recent bioinformatic analyses suggest a putative Arsl-coding gene in rats with the highest amino acid similarity to its human counterpart (13). In this study, we confirm that rat Arsl is a putative ortholog of the human ARSL sulfatase. Through gain- and loss-of-function experiments in rat chondrosarcoma cells, we identify CS, sulfated at 4-O-position, as a candidate substrate of ARSL. In vivo, genetic ablation of the ARSL ortholog (Ol-Arsd) in medaka fish results in hyper-sulfated CS at the 4-O-position, defective ossification and signs of naso-maxillary hypoplasia, providing *in vivo* evidence of Ol-Arsd involvement in skeletal development. Collectively, this study unveils a novel mechanism regulating GAG sulfation, which plays a critical role in skeletal development and is implicated in human disease.

Results

Human and rat ARSL are luminal Trans-Golgi sulfatases

The human ARSL is a Golgi arylsulfatase (9). Thus far, no murine counterpart of ARSL has been identified. However, bioinformatic analysis (13) identified a putative rat Arsl sequence on chromosome 2, which shares the highest amino acid similarity with the human counterpart (57%). Notably, the human ARSL (hARSL) exhibits maximal protein identity with the rat Arsl in the active site (100%) and catalytic domain (93%), along with significant identity in the binding site (86%), suggesting that it may be a putative ortholog of human ARSL (Figure S1B-D).

Prompted by these observations, we decided to characterize the putative rat Arsl in vitro. We used rat chondrosarcoma cells (RCS) that represent a relevant cell line, as patients with CDPX mainly show cartilage defects. First, we observed the expression of endogenous Arsl and confirmed its Golgi localization in RCS, whereas no colocalization with lysosomal or ER markers was observed (marked by Lamp1 and PDI, respectively) (Figure 1A-B). Subcellular fractionation on a 10–30% (w/v) Iodixanol continuous gradient, which separates the Golgi from other membrane fractions (14)confirmed complete overlap of rat Arsl with Golgi fractions (marked by Gm130), and no or partial overlap with lysosomal or ER fractions (marked by Lamp1 and Vapa, respectively) (Figure 1C). We also performed immunofluorescence analysis with Nocodazole, which interferes with microtubule polymerization and induces the dispersal of the Golgi compartment into mini-stacks distributed throughout the cytoplasm (15). Higher colocalization was observed between Arsl and Tgn38 (a marker of Trans-Golgi) compared to Gm130 (a marker of Cis-Golgi), suggesting the Trans-Golgi localization of the enzyme (Figure 1D-E).

We next generated Arsl knock-out RCS cells (RCS Arsl KO) using a CRISPR/cas9 approach. (Figure S2A). We verified the complete absence of Arsl protein by western blot analysis and immunofluorescence (Figure S2B-D). Using an established protocol to measure human ARSL activity (8), we demonstrated that Arsl enzymatic activity was completely lost in RCS Arsl KO (Figure S2E).

Sulfatase activation is mediated by the conversion of a cysteine residue into α-formylglycine (C86FGly), catalyzed by formylglycine-generating enzyme (FGE) encoded by the SUMF1 gene (16) . We generated an inactive form of the enzyme by mutating the cysteine residue (C86A) in hARSL. This mutant, overexpressed in the RCS Arsl KO background cells (RCS KO Myc-hARSL C86A), showed no enzymatic activity, while reconstitution with the WT form rescued activity (RCS KO Myc-hARSL) (Figure 1F-G and Figure S2F).

Transfection of a FLAG tagged version of hARSL resulted in increased enzymatic activity and complete colocalization with the endogenous rat Arsl (Figure 1H-I), as demonstrated by both subcellular fractionation and immunofluorescence (Figure S2G-H). The drug warfarin, which has been shown to inhibit ARSL activity (8), inhibited the activity of both rat and human ARSL in RCS WT and hARSL-overexpressing cells (Figure 1J-K).

Both human and rat ARSL model reported in AlphaFold Protein Structure Database depict that ARSL's catalytic region is positioned within the lumen of the Trans-Golgi network, suggesting it may function enzymatically in this compartment (Figure 2A). To confirm these data, we performed a protease protection assay in RCS FLAG-hARSL. This biochemical assay is used to investigate the accessibility of membrane proteins by treating samples with protease enzymes, specifically proteinase K (PK). The application of PK in conjunction with Triton X, capable of disrupting membranes, leads to the degradation of luminal proteins (Figure 2B) (17). The FLAG-hARSL protein was resistant to PK treatment but was degraded when PK was incubated in the presence of Triton X, consistent with the model that ARSL is an intraluminal Golgi enzyme (Figure 2C).

Collectively, these findings suggest that the rat Arsl is a putative ortholog of the human ARSL sulfatase, it is localized in the lumen of the Trans-Golgi network, and it is inhibited by the coumarin derivative warfarin, sharing significant structural and functional similarities with its human counterpart.

We employed the TurboID proximity labeling technique to identify proteins that interact with or are near Arsl. TurboID is a promiscuous biotin ligase that biotinylates nearby proteins, enabling their enrichment using streptavidin-coated beads (18). The biotinylated proteins are then analyzed and identified by mass spectrometry (MS) (Figure S3A and Table S1- S2).

We infected RCS cells with a pMSCV FLAG-TurboID-Arsl retroviral vector to stably overexpress the rat Arsl protein tagged with FLAG and fused to the TurboID sequence. We verified localization of FLAG-TurboID-Arsl to the Golgi apparatus (Figure S3B–C). Cells were left untreated or incubated with Biotin (12 hours, 50 µM), lysed and by using streptavidin-based immunoprecipitation followed by mass spectrometry (MS) analysis, we identified 129 proteins that were significantly enriched in biotinylated compared to non-biotinylated samples (false discovery rate (FDR) < 0.05) (Figure 2D-E

and Figure S3D). Gene ontology (GO) enrichment analysis of the proteomic dataset revealed "extracellular matrix organization" as one of the most significantly upregulated biological processes (Figure S3E). Notably, among the ARSL-identifying proximal proteins, we identified Aggrecan core protein (*Acan*) the precursor of Aggrecan, a major cartilage proteoglycan that is extensively modified with highly sulfated glycosaminoglycans (GAGs), particularly chondroitin sulfate (19)(Figure S3F).

Chondroitin sulfate (4-O-S-CS) is a candidate ARSL substrate

Next, we analyzed GAG sulfation levels in RCS upon gain- and loss-of-function of ARSL by using high-performance liquid chromatography (HPLC) analysis. This approach is considered a sensitive method for GAG quantification compared to other techniques (20). This system can quantify and differentiate GAG classes, including Heparan sulfate (HS), Chondroitin sulfate (CS), and Dermatan sulfate (DS), by identifying their distinct sulfation patterns in disaccharide chains (20). First, we investigated by HPLC analysis the sulfation state of GAGs subfamilies. The sulfation levels of the HS GAGs were not influenced by ARSL manipulation. Conversely, we observed a significant reduction in 4-O-sulfated GalNAc (CS) in RCS overexpressing the hARSL (Figure 3A). Consistently, a significant increase of 4-O-sulfation levels of CS/DS was present in RCS Arsl KO cells (Figure 3B). Reintroducing the wild-type but not the catalytically inactive hARSL (hARSL C86A) normalized 4-O-sulfation levels of CS/DS in RCS Arsl KO cells (Figure 3C-D).

To assess the ability of Arsl to directly desulfate chondroitin sulfate, we incubated exogenous CS molecules with Golgi-enriched fractions isolated from three RCS cell conditions: wild-type (WT), Arsl knockout (KO), and Arsl KO cells reconstituted with human ARSL (hARSL) (Figure 3E and Figure S4A–B). Subsequent HPLC analysis of CS sulfation patterns revealed that 4-O-sulfation levels of CS were significantly elevated following incubation with Golgi fractions from Arsl KO cells, compared to those incubated with fractions from WT or hARSL-reconstituted cells (Figure 3F). These findings indicate that ARSL activity within the Golgi is necessary for the removal of 4-O-sulfate groups from CS, and that reintroducing hARSL restores this desulfation capacity in KO cells.

Collectively, these data indicate that 4-O-sulfated chondroitin sulfate (4-O-CS) is a likely substrate of ARSL and that ARSL regulates the 4-O-sulfation state of CS within the Golgi apparatus, identifying a previously unappreciated step of glycosaminoglycan (GAG) desulfation during proteoglycan biosynthesis. ARSB can be regarded as the lysosomal counterpart of the Golgi-resident ARSL, acting on structurally related substrates in distinct intracellular compartments. Consistently, sequence alignment of ARSL and ARSB reveals a high degree of similarity within the catalytic site. The ARSL catalytic region (Ser84–Arg97), corresponding to the sequence SLCTPSRAAFLTGR, is highly conserved among arylsulfatases, with six conserved residues across the family and up to ten conserved residues when directly comparing ARSL and ARSB (Figure 3G), supporting a shared catalytic architecture (13, 21).

Chondroctyes lacking ARSL show altered TGF-β signalling

To define the functional consequences of ARSL deficiency in RCS cells, we performed RNA-seq profiling comparing WT and Arsl KO cells (Table S3, GSE301703). This analysis identified 229 differentially expressed genes, with 67 upregulated and 162 downregulated in Arsl KO cells. While no significant pathways emerged among the upregulated genes, the most strongly induced transcript was Cspg4 (Figure 4A), encoding a chondroitin sulfate proteoglycan, likely reflecting a compensatory response to altered sulfation. Notably, pathway analysis revealed significant downregulation of TGF-β signaling in Arsl KO cells (Figure 4B). To validate these findings, we assessed the response to TGF-β1 stimulation (Figure 4C). In WT cells, TGF-β1 robustly induced the expression of chondrogenic markers, including COL2A1 and SOX9, whereas this response was markedly blunted in ARSL-deficient cells. Together, these results indicate that ARSL loss impairs cellular responsiveness to TGF-β signaling and highlight a critical role for ARSL in regulating the TGF-β–dependent chondrocyte differentiation program (Figure 4D).

Ol-Arsd controls CS sulfation and skeletal development in medaka

To explore in vivo the role of ARSL during skeletal development we performed a phylogenetic analysis on sequences and structural comparison on three-dimensional (3D) structures (in silico) of ARSL gene across multiple species. In humans, the ARSL gene is part of a gene cluster on the X chromosome, which also includes ARSD, ARSH, and ARSF (13). This analysis demonstrated that the medaka fish (Oryzias latipes, Ol) contains a unique Arsd (Ol-Arsd) gene within an arylsulfatase cluster. This gene is closely related to three human sequences (ARSL, ARSD, ARSH), sharing greater than 50% identity (Table S4). The Ol-Arsd sequence from medaka clusters with ARSL, ARSD, and ARSH within the Human Arylsulfatase family (Figure S5A).

We generated an Ol-Arsd knockout (KO) medaka model using a CRISPR/Cas9 approach. Following the creation of a mosaic model, we identified medaka with a specific 200-nucleotide deletion, resulting in a loss of Ol-Arsd transcript (Figure S5B-C). Notably, Ol-Arsd KO fish exhibited an increase in 4-O-sulfated CS at stage 40, corresponding to early larval phase at 4-9 days post-fertilization, compared to controls. This elevation became more pronounced one-month post-fertilization (Figure 5A), highlighting the physiological role of Ol-Arsd in regulating CS-proteoglycan sulfation during in vivo development. Ol-Arsd KO fish showed shorter body length and smaller head size at various developmental stages (from stage 40 to 10 days post-hatching) (Figure 5B-C and Figure S5D). Alcian blue and Alizarin red staining revealed reduced size and mineralization of cranial cartilage, particularly in the rostral region of Ol-Arsd KO larvae at stage 40, compared to controls (Figure 5D-G). Although μCT analysis did not reveal punctate calcifications in the cartilage, a phenotypic feature commonly observed in CDPX patients, it confirmed smaller, dysmorphic heads in Ol-Arsd KO fish, partially

resembling clinical features seen in CDPX (Figure 5H). To further explore the ossification process, we used a transgenic medaka line in which chondrocytes and osteoblasts are marked by GFP and mCherry fluorescence (Col10a1:GFP/Osx:mCherry), respectively. Direct fluorescence analysis revealed defective perichondral ossification of the palatoquadrate and ceratohyal elements in the head, accompanied by reduced osteoblast differentiation. We also observed impaired intramembranous ossification of the operculum, with decreased Osx:mCherry-positive osteoblastic cells (Figure 5I). Collectively, these observations suggest a critical role for Ol-Arsd during both perichondral and intramembranous ossification in medaka fish.

Discussion

In this study, we identified ARSL, a sulfatase enzyme localized at the Trans-Golgi, involved in the desulfation of specific glycosaminoglycans. To our knowledge, this is the first example of proteoglycan desulfation occurring at the Golgi. Indeed, the Golgi has been recognized to date as the primary site for sulfation of GAGs, while their desulfation occurs primarily in lysosomes, where arylsulfatases play a key role in the stepwise degradation of GAGs (22). In this respect, the sulfatase ARSB represents the lysosomal counterpart of the Golgi-localized ARSL. Our results revealed that ARSL is specifically localized at the Trans-Golgi, with its active site situated in the luminal space, enabling ARSL to interact with its substrate. We identified ARSL as specifically targeting the 4-O-sulfation of CS. Our study indicates a key role of ARSL sulfatase in the fine regulation of GAG sulfation during skeletal development and it might, at least in part, explain the pathological manifestations observed in CDPX patients. Indeed, during human development, GAG sulfation patterns vary spatially and temporally. Studies have shown high levels of 4-O-sulfate CS at the growth plate in newborns, while adult cartilage contains only 6-O-sulfate CS (23),(24). The 4-O-sulfated CS plays a role in bone mineralization at the growth plate, as it has a higher capacity to bind calcium ions, potentially regulating hydroxyapatite crystal growth (25). Notably, loss of chondroitin 4-O-sulfotransferase (CHST11) in humans results in a form of osteochondrodysplasia (OCBMD, OMIM: 618167) characterized by short stature and skeletal abnormalities predominantly affecting the limbs, hands, and feet. Similarly, in mice, *Chst11* gene defects lead to severe chondrodysplasia with disorganized growth plates. Sulfated GAGs also influence growth factor diffusion and binding to their cognate receptors. Chondrocytes lacking chondroitin 4-O-sulfotransferase display a pronounced upregulation of TGF-β signaling, accompanied by a concomitant downregulation of BMP signaling, underscoring a critical role for CS sulfation in modulating key signaling pathways (2, 4). Consistently, we observed an impaired response to TGF-β stimulation in RCS lacking ARSL, suggesting that abnormal levels of 4-O-sulfated CS may contribute to the cartilage and bone defects seen in CDPX patients through altered responsiveness to morphogen stimulation. However, we cannot exclude the possibility that other proteoglycans are also regulated by ARSL, potentially expanding its involvement to additional cellular functions.

To investigate the physiological relevance of the ARSL enzyme in glycosaminoglycan metabolism and the CDPX disease pathological manifestation, we generated the medaka model with genetic Ol-Arsd deficiency. The Ol-Arsd KO fish displays reduced growth and subtle features that could be suggestive of naso-maxillary hypoplasia, showing limited overlap with aspects of the human CDPX phenotype (10). Remarkably, we did not detect epiphyseal stippling by µCT analysis, suggesting that this feature might be present only in human. Indeed, an important limitation of this study is that we cannot rule out the potential involvement of additional ARSL substrates, particularly in human. While we have identified specific GAGs targeted by ARSL, it is possible that other substrates, not yet characterized, are also impacted by the enzyme's activity.

While the medaka model provides a valuable model for the molecular understanding of how the genetic absence of the ARSL enzyme leads to CDPX, it is important to keep in mind that the model used in this study may display other phenotypic abnormalities beyond those directly linked to ARSL deficiency. This is likely due to the fact that Ol-Arsd is part of a cluster of related arylsulfatase homologs that may underlie broader biological effect. Further research will be needed to explore these possibilities and better delineate the specific contributions of ARSL and other sulfatases in the context of skeletal development.

Material and Methods

Ethical considerations

The experimental protocol on fish was approved by the Italian Ministry of Health (Directorate-General for Animal Health and Veterinary Medicinal Products (authorization No. 598/2017-PR issued on 24/7/2017). All animal procedures were conducted in accordance with the applicable legislation, and in particular with the Legislative Decree n.26 issued on 4/3/2014 by the Italian Ministry of Health and implementing the Directive 2010/63/EU on the protection of animals used for scientific purposes.

Cell culture, Transfections and Stable Cell Lines

RCS cell lines were cultured in DMEM (Euroclone) supplemented with 10% foetal bovine serum (FBS from Euroclone) and 1% penicillin/streptomycin at 37°C in 5% CO_2. The RCS cell line was a Swarm chondrosarcoma chondrocyte line (26).

For transfections, cells were transfected with Lipofectamine LTX and Plus reagent (Invitrogen) and Lipofectamine 3000 following a reverse transfection protocol according to the manufacturer's instructions.

Stable cell lines were generated using retroviral system as previously described (PMID: 12717450)(27). Retroviral particles were generated as follows: 1.3 µg of pMSCV_neo plasmid (#634401, Addgene) with 0.95 µg of Gag Pol packaging plasmid (#14887, Addgene) and 0,75 µg of VSVG envelope plasmid (#14888, Addgene) were co-transfected in HEK293T cells in a six-well plate. 48 hours post transfection the media was collected and RCS cells were incubated with the retroviral suspension containing polybrene (8 µg/ml; TR-1003-G Merck) for 24 hours. Confluent-infected cell lines were then selected with the Neomycin antibiotic at 1mg/ml for 10 days (G418, #A1720 Sigma).

Cloning Plasmids and Mutagenesis

Generation of the FLAG-hARSL plasmid:

The pcDL-SRa296 h*ARSL* plasmid was previously described (8). The h*ARSL* fragment was amplified by PCR using the following primers:

hARSL Fw: 5’ TAATAAGCTTATGTTACATCTGCACCATTC 3’;
hARSL Rev: 3’ ATTATCTAGATTGTGGGCATCTTCCCTA 5’.

The hARSL fragment was cloned into the p3XFLAG-CMV-14 plasmid (#E7908, Sigma-Aldrich) using HindIII and XbaI (New England Biolabs, NEB) restriction enzymes. The vector was treated with 5 µl of CIP (alkaline phosphatase, NEB), followed by ligation using the Quick Ligation Kit (NEB). Sequence integrity was confirmed by Sanger sequencing.

Generation of the pMSCV_Myc-hARSL retroviral plasmid:

The hARSL insert amplified from pcDL-SRa296 was first cloned into Myc-BioID2-MCS (#74223, Addgene) using the following primers:

Myc_hARSL_Fw: 5' ACCCAAGCTGGCTAGACCATGTTACATCTGCACCATTCTTG 3';
Myc_hARSL_Rev: 3' TTCCATGGTGGCTAGCCCTGCTTGTGGGTCATCTTCCCTAAGG 5'.

The Myc-BioID2-MCS plasmid was linearized with NheI (NEB) enzyme and ligation was performed using the In-Fusion Snap Assembly Master Mix (#638944, Takara).

The Myc-hARSL cassette was then amplified by PCR using:

hARSL_pMSCV_FW: 5' CTAGGCGCCGGAATTATGTTACATCTGCACCATTCTTG 3';
hARSL_pMSCV_Rev: 3' AATTAGATCTCTCGACTAGCTTCTTCTCAGGCTGAACTCGC 5'.

The fragment was cloned into the pMSCV_neo retroviral vector (#634401, Addgene), digested with EcoRI and XhoI (NEB) restriction enzymes, using the In-Fusion Snap Assembly Master Mix (#638944, Takara).

Generation of the rat FLAG-Arsl *plasmid:*

The rat Arsl cDNA was amplified from RCS cDNA using with following primers:

Rat Arsl Fw: 5'ATGGCTCCGCCCCATCGCACCTGT 3';
Rat Arsl Rev: 3' TGGAAGGCGGGAGCCGGAAGTGA 5'.

The rat Arsl sequence was then cloned into the p3XFLAG-CMV-14 plasmid (#E7908, Sigma-Aldrich) using the following primers:

FLAG – rat Arsl Fw: 5' CCCAAGCTTGGGATGGCTCCGCCC 3';
FLAG – rat Arsl Rev:3' CGGGATCCGCTTCCGGCTCCCGCC 5'.

The vector was digested using HindIII and BamHI (NEB) restriction enzymes and treated with 5 µl of CIP (alkaline phosphatase, NEB), followed by ligation using the Quick Ligation Kit.
Sequence integrity was confirmed by Sanger sequencing.

Generation of the FLAG- TurboID -Arsl retroviral plasmid:

For the generation of a retroviral vector expressing rat Arsl protein tagged with FLAG and fused to the TurboID sequence (pMSCV_FLAG-TurboID-Arsl), we used 2 steps:

1. First, we cloned the rat Arsl coding sequence, derived from the rat FLAG-Arsl plasmid, into the empty pLVX FLAG–TurboID lentiviral vector (PMID: 36007018) (28) using the following primers:

pLVX_RatArsl_Fw: 5' AGCTTCGAATTGGATCATGGCTCCGCCCCATCGCACCT 3';
pLVX_RatArsl_Rev: 3' ATATGTGCAGAACTTCCGGCTCCCGCCTTCCA 5′.

The lentiviral vector was linearized with BamHI (NEB) enzyme, and ligation was performed using the In-Fusion Snap Assembly Master Mix (#638944, Takara).

2. Next, the resulting FLAG- TurboID -Arsl cassette was subcloned into the retroviral vector pMSCV_neo (#634401, Addgene). The Arsl sequence was amplified by PCR using the following primers:

TurboID_Fw: 5′ CTCTAGGCGCCGGAATTATGGCTCCGCCCCATCGC 3′;
TurboID_Rev: 3′ AATTAGATCTCTCGAGGTTATTTATCATCGTCATCCTT 5′.

The retroviral vector was digested with EcoRI and XhoI (NEB) restriction enzymes, followed by ligation using the In-Fusion Snap Assembly Master Mix (#638944, Takara).

Generation of the Myc-hARSL C86A retroviral plasmid:

The inactive form of hARSL (C86A) was generated by site-directed mutagenesis using the Agilent QuikChange XL Site-Directed mutagenesis kit (# 200516) in the *pMSCV_Myc-hARSL retroviral plasmid* using following primers:

C86A_FW: 5' TCTGCCGCATCTTTGGCCACCCCAAGCAGAGC 3';
C86A_REV: 3' GCTCTGCTTGGGGTGGCCAAAGATGCGGCAGA 5'.

Generation of Arsl CRISPR clones

The RCS Arsl knockout cells were generated using the CRISPR–Cas9 system with the pCMV-Cas9-GFP (CAS9GFPPRO, Merk) plasmid an all-in-one vector encoding both Cas9 and a guide RNA (gRNA) targeting the *Arsl* genomic sequence.The sgRNAs were designed by CRISPR Core (Merk) to recognize exon 4, which encodes the active site of the enzyme (Rat Arsl_gDNA: TCGGTGTGCACGCCCAGTAGGG). The all-in-one vector (5 µg) was transfected by reverse transfection in a p60 plate with 1ml of OPTIMEM, 5 µl of PLUS, and 25 µl of LTX (Lipofectamine™ LTX Reagent with PLUS™ Reagent, #A12621 Thermo Fisher Scientific). After 48h the cells were collected and pelleted with PBS and the efficiency of Cas9 was controlled with the T7 Endonuclease I test. Following transfection, GFP-positive cells were isolated by fluorescence-activated cell sorting (FACS) and subjected to single-cell sorting to obtain clonal Arsl knockout cell lines. The targeted genomic region from individual clones was amplified by PCR using the following primers:

Rat Arsl Fw: 5' GGGCCCAGAAAACACCAGTA 3';
Rat Arsl Rev: 3' CCATCGCCATAGTCAGGGTA 5'.

The PCR products were then Sanger sequenced to confirm the presence of frameshift-inducing mutations predicted to abolish Arsl protein expression at the CRISPR target site (76 clones were Sanger sequenced). To predict the functional consequences of these mutations, the resulting sequences were analyzed using bioinformatic tools (ExPASy). One clone carrying predicted loss-of-function mutations in Arsl was selected for subsequent experiments (Fig. S2A).

Chemicals

Nocodazole was used at 33 µM for 3h at 37°C. Biotin (#B4639, Sigma-Aldrich) was used overnight at a concentration of 50 µM. The coumarin derivative warfarin (#45706, Sigma-Aldrich) was added to the incubation mix at a concentration of 2 mM for 3h. Human TGF-β1 was used at a final concentration of 5 µg for 48 hours to treat RCS WT and ARSL KO cells cultured in DMEM only (Human TGF-beta 1 Recombinant Protein, PeproTech; #100-21-10UG).

Immunofluorescence

Cells were seeded on coverslips at least 24 h before treatment and fixed for 10 min in 4% PFA in PBS and permeabilized for 30 min in blocking buffer (0.05% (w/v) saponin, 0.5% (w/v) BSA, 50 mM NH_4Cl and 0.02% NaN_3 in PBS, pH 7.4). Cells were incubated in a humid chamber for 1 h at room temperature with primary antibodies: Arsl (Novus Biologicals NBP1-62667, 1:100); Lamp1 (Santa Cruz Biotechnology sc-19992, 1:500); PDI (Enzo Life Sciences ADI-SPA-891-D, 1:200); Flag M2 (Sigma F1804, 1:200), Gm130 (BD Transduction laboratories, 610823 1:200) and Tgn38 (homemade, 1:100). After washing three times in PBS, they were incubated for 1 h at room temperature with the secondary

antibody (Alexa Fluor–labeled goat anti-rabbit A11011/A11008 or goat anti-mouse A11001, A11004; Life Technologies, Thermo Fisher Scientific), washed three times in PBS, incubated for 20 min with 1 μg/ml Hoechst 33342, and finally mounted in Mowiol (Sigma-Aldrich) or Vectashield (Vector Laboratories) supplemented with 4′,6-diamidino-2-phenylindole (DAPI). For Nocodazole treatment: the cells were plated on a coverslip in a 24-well plate and after 24h were put on ice for 5' and Nocodazole (33 μM) was added for 3h at 37°C. Cells were then fixed in 4% PFA for 15' and stained with primary and secondary antibodies.

Confocal and super resolution microscopy

Scanning laser confocal experiments were acquired using a Zeiss LSM 800 or Leica TCS SP5 confocal microscope equipped with a 63× 1.4 numerical aperture oil objective. Airyscan microscopy was performed using a Zeiss LSM 880 confocal microscope, equipped with a Plan-Apochromat 63×/1.4 numerical aperture oil objective and pixel size of 8.7 nm. Images were subjected to post-acquisition Airyscan processing. Image acquisition and processing were performed with Zen Blue software.
The colocalization analysis was carried out using the JACOP plugin in ImageJ which calculates the Mander's coefficient to quantitatively assess the degree of overlap between two fluorescent signals. Images were background-subtracted and thresholded to minimize noise, applying identical threshold values across all images. Manders' coefficients were calculated as the fraction of fluorescence intensity in one channel overlapping with the signal in the second channel. M1 and M2 represent the overlap of channel 1 with channel 2 and vice versa, respectively. Coefficient values range from 0 (no overlap) to 1 (complete overlap) and were calculated within identical regions of interest. The M1 coefficient was used to quantify signal overlap (Fig. 1B and 1E) (29, 30).

Western blotting

Cells were washed twice with PBS and then scraped in RIPA lysis buffer (20 mM Tris [pH 8.0], 150 mM NaCl, 0.1% SDS, 1% NP-40, 0.5% sodium deoxycholate) supplemented with PhosSTOP and EDTA-free protease inhibitor tablets 1× final concentration (Roche, Indianapolis, IN, USA). Cell lysates were incubated on ice for 20 min. Then, the soluble fraction was isolated by centrifugation at 18,000 g for 20 min at 4°C. Total protein concentration in cellular extracts was measured using the colorimetric BCA protein assay kit (Pierce Chemical Co, Boston, MA, USA). Protein extracts, separated by SDS–PAGE and transferred onto PVDF, were probed with primary antibodies overnight against: Arsl (Novus Biologicals NBP1-62667, 1:1,000 which recognizes human (100% identities) and rat (61% identities) protein); Gm130 (BD Transduction laboratories, 610823 1:1,000); Lamp1 (Novus Biologicals, NB120-19294 1:1,000); Vapa (Proteintech, 15275-1-AP 1:1,000); β-actin (Novus Biologicals, NB600-501 1:5,000); Filamin A (Cell Signaling, Technology 4762 1:1,000); Flag M2 (Sigma, F1804 1:500); Myc (Cell Signaling Technology, 2278S 1:1,000); Proteins of interest were

detected with HRP-conjugated goat anti-mouse or anti-rabbit IgG antibody (1:2,000, Vector Laboratories) and visualized with the ECL Star Enhanced Chemiluminescent Substrate (Euroclone) according to the manufacturer's protocol. The Western blotting images were acquired using the ChemiDoc-lt imaging system (UVP).

Subcellular Fractionation

The cells were washed in HB solution (0.25 M sucrose, 10 mM triethanolamine, 10 mM acetic acid, 1 mM EDTA, pH 7.8) and then in HB solution plus 20 nM of N-Ethylmaleimide (NEM, #E3876 Sigma). The cells were collected with HB-NEM-CLAP (Protease Inhibitor Cocktails) and lysed by passing them through a 25G needle for 25 passages, and then centrifuged at 1,500g for 5 min at 4°C twice to obtain the post-nuclear supernatant (PNS). The PNS were loaded below a linear 10–20% OptiPrep gradient, prepared by mixing 10% and 20% OptiPrep solutions (OptiPrep, #D1556 Sigma) using a Gradient Maker (Hoefer™ SG 50). The OptiPrep medium was diluted to obtain a 50% OptiPrep solution with Diluent B (0.25 M sucrose, 1 mM EDTA, 120 mM Tricine, pH 7.8), which was further diluted to generate the 10–20% OptiPrep solution. The PNS was added to the bottom of the centrifuge tube, and intracellular organelles were separated by overnight ultracentrifugation at 29,000 rpm in an SW50 rotor. Fractions were collected from the top of the gradient and treated with 50% trichloroacetic acid (TCA) plus 2 mg/ml sodium deoxycholate, followed by centrifugation at 13,500 rpm for 10 min at 4°C. The fractions were washed with 100% cold acetone and left to evaporate. All samples were then denatured and reduced in βME-containing sample buffer for 10 min at 95°C and separated by SDS–PAGE (14).

Golgi Fractionation

Four 15-cm plates of densely confluent cells for each condition were trypsinized and the collected pellets were washed once with HB buffer (0.25 M sucrose, 10 mM Tris-HCl, pH 7.4). The samples were then centrifuged at 1,500 rpm for 5 min at 4°C.

The resulting pellets were resuspended in HB buffer to a final volume corresponding to five times the pellet volume (20% w/v). The suspensions were then homogenized using a cell homogenizer (Isobiotec) fitted with an 8 µm sphere by passing the samples through the homogenizer 10 times.

The homogenates were adjusted to 1.4 M sucrose by adding 6 ml of ice-cold 2.3 M sucrose containing 10 mM Tris-HCl (pH 7.4). Next, 1 mM EDTA was added, and the mixture was vortexed vigorously for uniform mixing before being loaded into an SW41 tube for gradient centrifugation. A sucrose gradient was prepared with the following layers: 2.0 M (700 µl), 1.6 M (1.4 ml), homogenate adjusted to 1.4 M sucrose (2.8 ml), 1.2 M (4.2 ml), and 0.8 M (2.8 ml).

The gradients were centrifuged at 26,500 rpm for 2.5 hours at 4°C. Golgi fractions were identified at the 0.8 M/1.2 M sucrose interface and extracted using a syringe puncture, as previously described (31). The collected Golgi fractions were then analyzed by Western blotting and enzymatic assays.

Protease Protection Assay

Cells from four 10-cm dishes were harvested and washed twice with ice-cold PBS. The cell pellets were collected after centrifugation at 700 g for 5 min and resuspended in ice-cold HB Buffer (250 mM sucrose, 20 mM HEPES-KOH, pH 7.4, 1 mM EDTA, and complete EDTA-free protease inhibitor). Cells were then disrupted through a 25G needle gauge syringe. The homogenized cells were centrifuged twice at 3,000 g for 10 min to remove cell debris and undisrupted cells. To enrich for membrane fractions, the supernatant was centrifuged at 100,000 g for 60 min. The pellet fraction was subjected to proteinase K protection assay by incubation with and without 50 ug/mL proteinase K (#P6556; Sigma-Aldrich) for 60 or 90 min on ice and 1% Triton X-100 (Sigma-Aldrich) (17). Proteins in each fraction were isolated by trichloroacetic acid (TCA) precipitation as previously described. The final pellet was denatured and reduced in βME-containing sample buffer for 10 min at 95°C and separated by SDS–PAGE. Antibodies against Lamp1 targets the cytosolic portion of the lysosomal protein (Novus Biologicals, NB120-19294 1:1,000), while PDI antibody serves as a control for luminal proteins (Enzo Life Sciences ADI-SPA-891-D, 1:1,000). There were used as controls to detect cytosol-facing and intraluminal epitopes, respectively.

Biotinylation and streptavidin pulldown

RCS cells stably expressing rat Arsl fused to a FLAG tag –TurboID sequence were generated by infection with the pMSCV FLAG-TurboID-Arsl retroviral vector and subsequently selected with Neomycin (G418, #A1720 Sigma) at 1 mg/mL for 10 days. For the biotinylation experiment, cells from four 10-cm dishes were incubated with 50 µM biotin at 37 °C for 12 hours, while untreated cells were included as controls to account for nonspecific binding to streptavidin beads.

Following biotinylation, cells were harvested and washed twice with ice-cold PBS.The pellets were resuspended in ice-cold lysis buffer (1% Triton X-100, 130 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, 25 mM HEPES pH 7.4, supplemented with protease inhibitor cocktail) and incubated on a rotator at 4ºC for 30 min to ensure complete lysis. After centrifugation (13,300 rpm for 10 min at 4ºC) the supernatant was collected to check the protein concentration.

The pulldown of biotinylated proteins was performed using Dynabeads MyOne Streptavidin C1 (#65001, Life Technologies). The lysate was incubated with beads at 4ºC on a rotator for 12h. Then, the beads were washed with different buffers (Wash buffer 1: 2% SDS; Wash buffer 2: 0.1% deoxycholate, 1% Triton X-100, 500 mM NaCl, 1 mM EDTA, and 50 mM HEPES, pH 7.5; Wash

buffer 3: 250 mM LiCl, 0.5% NP-40, 0.5% deoxycholate, 1 mM EDTA, and 10 mM Tris, pH 8.1; and Wash buffer 4: 50 mM Tris, pH 7.4, 50 mM NaCl) and were prepared for MS/MS analysis.

ARSL enzymatic assay

The cells were collected in MQ water and lysed by three fast passages in dry ice/water (RT). The supernatants were collected and subjected to protein quantification using a BCA protein assay kit (Pierce Chemical). 50 µg of cell lysate were added to the incubation mix (100 µl) containing 50 mM Tris HCI buffer (pH 7.5) and 0.4 mM of 4-Methylumbelliferone sulfate (4MU-S) (#M7133, Sigma-Aldrich). Incubation was performed for 3h at 37 °C in a 24-well plate. The reaction was stopped with 1.8 ml of glycine-carbonate buffer (pH 10.7). Fluorescence was determined at 365 nm (excitation) and 460 nm (emission). The obtained fluorescence values were then normalized by a calibration curve using different concentrations of 4MU (4-Methylumbelliferone)(8).

HPLC Analysis

For HPLC analysis, the collected media from different cell lines and the Golgi-extracted fractions were treated with 3 mg/ml Collagenase B (Roche) and centrifuged at 1,000 rpm for 5 min at 4°C. The supernatants were then analyzed by HPLC-Fp-MS, as previously described in the specific protocols (32, 33).

qRT-PCR analysis

For quantitative RT-PCR, RCS WT and ARSL KO cells, either treated or untreated with TGF-β1 for 48 hours, were harvested for RNA extraction using the RNeasy Mini Kit (#74106, Qiagen) following the manufacturer's protocol. 1 µg of total RNA was used for reverse transcription using QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's instructions. qRT–PCR was performed in triplicate using LightCycler 480 SYBER Green I Master (Roche) and analyzed by LightCycler 480 (Roche). The Ct values were normalized to s16 gene expression, and the expression of each gene was represented as 2(−ddCt) relative to control.

The primers used were as follows:

Rat_S16 _Fw: 5' ACTTGTAGCTATGCCGTCCA 3';
Rat_S16 _Rev: 3' GCAGAACAGGCTCCAGTAAC 5'.

Rat_Actin _Fw: 5' GGCAGTGGCCATCTCTTGCT 3';
Rat_Actin _Rev: 3' CTCATGCCATCCTGCGTCTG 5'.

Rat_Col2a1_Fw: 5' CCGATCCCCTGCAGTACATG 3';
Rat_Col2a1_Rev: 3' TGCTCTCGATCTGGTTGTTCA 5'.

Rat_Sox9_Fw: 5' CAGTCCCAGAGAACGCACAT 3';
Rat_Sox9_Rev: 3' TCTGGTGGTCGGTCTAGTCA 5'.

Rat_Runx2_Fw: 5'GTCTCGGAGGGAACGATGAG 3';
Rat_Runx2_Rev: 3' GGGACCGACACAGCCATATA 5'.

MS-proteomics

We pulled down ARSL biotinylated proteins using Streptavidin -beads (Dynabeads MyOne Streptavidin C1, Life Technologies #65001). We eluted the proteins with 40ul of 2X Laemmli Buffer (4X stock: 250mM Tris-HCl pH 6.8; 8% (w/s) SDS, 40% (v/v) Glycerol, 200mM b-Mercaptoethanol) and precipitated them overnight at -20°C with 4 volumes of Acetone. Proteins pellets were collected via centrifugation, washed with ice-cold 70% Ethanol, reduced and alkylated heating at 90°C for 10 min in 20 ul solution of 6 M guanidine-HCl, 5 mM tris (2- carboxyethyl)phosphine (TCEP), and 20 mM chloroacetamide. Peptides were obtained by a double digesting. For the first digestion, protein solutions were diluted 1:2 in 50 mM ABC solution adding LysC 1:100 (Wako) for 3h at 37°C. For the second digestion, we further diluted the samples 1:3 in 50 mM ABC buffer with endopeptidase sequencing-grade trypsin (#V528A, Promega) overnight at 37°C. Collected peptide mixtures were concentrated and desalted using the Stop and Go Extraction (STAGE) technique (PMID: 17703201) (34).

Instruments for LC-MS/MS analysis consisted of a NanoLC 1200 coupled via a nanoelectrospray ionization source to the quadrupole-based Q Exactive HF benchtop mass spectrometer. Peptide separation was carried out according to their hydrophobicity on a homemade chromatographic column, 75-µm inside diameter, 8-µm tip, bed-packed with Reprosil-PUR (C18-AQ), 1.9-µm particle size, 120-Å pore size, using a binary buffer system consisting of solution A (0.1% formic acid) and solution B (80% acetonitrile and 0.1% formic acid). Runs of 120 min after loading were used for proteome samples, with a constant flow rate of 300 nl/min. After sample loading, the run started at 5% buffer B for 5 min, followed by a series of linear gradients, from 5 to 30% B for 90 min, then a 10-min step to reach 50% and a 5-min step to reach 95%. This last step was maintained for 10 min. Q Exactive HF settings were as follows: MS spectra were acquired using 3E6 as an AGC target, a maximal injection time of 20 ms, and a 120,000 resolution at 200 mass/charge ratio (m/z). The mass spectrometer was operated in a data-dependent Top20 mode with subsequent acquisition of higher-energy collisional dissociation fragmentation MS/MS spectra of the top 20 most intense peaks. Resolution, for MS/MS spectra, was set to 15,000 at 200 m/z, AGC target to 1E5, maximum injection time to 20 ms, and the isolation window to 1.6 Th. The intensity threshold was set at 2.0 E4 and dynamic exclusion at 30 s.

MS–based proteomics data analysis

All acquired RAW files were processed using MaxQuant (1.6.2.10) implemented with the Andromeda search engine. For protein assignment, spectra were correlated with the Uniprot Rattus Norvegicus (v.2019) including a list of common contaminants. Searches were performed with tryptic specifications and default settings for mass tolerances for MS and MS/MS spectra. Fixed modifications: Carbamidomethyl(C). Variable modifications: Oxidation, Acetyl(N-term) and Biotin(K). Digestion: Trypsin, Lys-C. Min. Peptide length=7 amino acids; Max. Peptide length=47 amino acids. False discovery rate for proteins and peptide-spectrum=1%.

For further analysis, the Perseus software (1.6.2.3) was used and first filtered for contaminants and reverse entries as well as proteins that were only identified by a modified peptide [First filter]. The LFQ Ratios were logarithmized, grouped and filtered for min.valid number (min. 3 in at least one group) [Second filter].

Missing values (only controls) have been replaced by random numbers that are drawn from a normal distribution. Two –sample t-test was performed using FDR = 0.05. Protein with Log2 Difference $\geq \pm 1$ and -Log10 p-value > 1.3 were considered significantly enriched.

Library preparation for transcriptomic analysis

Total RNA was quantified using the Qubit 4.0 fluorimetric Assay (Thermo Fisher Scientific). Libraries were prepared from 125 ng of total RNA using the NEGEDIA Digital mRNA-seq research grade sequencing service v2.0 (Next Generation Diagnostic srl) which included library preparation, quality assessment and sequencing on a NovaSeq 6000 sequencing system using a single-end, 100 cycle strategy (Illumina Inc.).

Bioinformatics workflow for transcriptomic analysis

The raw data were analyzed by Next Generation Diagnostic srl proprietary NEGEDIA Digital mRNA-seq pipeline (v2.0) which involves a cleaning step by quality filtering and trimming, alignment to the reference genome and counting by gene. The raw expression data were normalized, analysed by NEGEDIA degs analysis pipeline (v1.2.0) and visualized in a proprietary report (v2.0). Differentially expressed genes (DEGs) were selected based on a statistical significance threshold of False Discovery Rate (FDR) < 0.05 to correct for multiple testing.

Bioinformatic analysis for ARSL enzyme

AlphaFold Modelling

Starting from the list of arylsulfatase proteins, we interrogated the UniProtKB database (36) to download the structural files from the ‘Structure’ panel, selecting AlphaFold (37, 38) prediction results.

With particular interest in the sulfatase domain, where the catalytic site is located for each protein, we retrieved the range of the sub-sequence and the position of the amino acids from the InterPro database for the classification of the protein families (39).

We used the PyMOL Molecular Graphics System (Version 3.1.3, Schrödinger, LLC) to compute the Root Mean Square Deviation (RMSD) and quantify structural similarity between human ARSL and rat ARSL. The RMSD, refined over five cycles following structural superimposition, was 0.466 Å for the full-length proteins and 0.421 Å for the sulfatase domains.

To check if the catalytic site was conserved, we performed multiple sequence alignment of the arylsulfatase sequences using the ClustalW algorithm (40) with default parameters, and investigated conservation among the sequences, reconstructing a phylogenetic tree (41, 42) from the GenomeNet network of database and services (https://www.genome.jp). Finally, we used the software Chimera X (43) to show the superimposition of ARSB and ARSL structures.

We retrieved the information on the transmembrane regions of ARSL from the TmAlphaFold database (Transmembrane Protein Structure Database), available at the link https://tmalphafold.ttk.hu (44), by entering the Uniprot code in the SEARCH field.

ARSL multiple sequence alignment

We performed the protein sequence alignment with the online tool Clustal Omega (Multiple Sequence Alignment), available at the link https://www.ebi.ac.uk/jdispatcher/msa/clustalo (45). We uploaded a FASTA file with the two sequences to align:

a) ARSL from Homo sapiens (P51690) and Arsl from Rattus norvegicus (Q32KK0);

b) ARSL from Homo sapiens (P51690) and ARSB from Homo sapiens (P15848);

c) ARSL from Homo sapiens (P51690) and Arylsulfatase Multiple Sequence Alignment,

and selected ‘ClustalW with characters counts’ as OUTPUT FORMAT, keeping the default values for all the other parameters. The similarity score was extracted from the percentage identity matrix generated by Clustal Omega.

Phylogenetic analysis

The sequences of Protein Arylsulfatase Family were obtained from Ensembl using the BioMart tool by EMBL-EBI (https://www.ensembl.org/biomart/martview) (46).

Multiple sequence alignment and phylogenetic tree were performed using Clustal Omega software, version clustalo 1.2.4 (http://www.clustal.org/omega/) (47).

Medaka stocks

The Cab strain of wild-type medaka fish (Oryzias latipes) was maintained following standard conditions (i.e., 12-h/12-h dark/light conditions at 27°C). Embryos were staged according to Iwamatsu (48).
In all experiments, measurements were performed on distinct animals, and N represents independent biological replicates, each corresponding either to an individual fish or, where indicated, to a pool of medaka samples of the specified genotype and developmental stage for each assay.

Generation of Ol-Arsd KO medaka model and qRT-PCR analysis

The available medaka Ol-Arsd genomic sequences were retrieved from public databases (http://genome.ucsc.edu/) from human ARSL (NM_000047.3) transcript.
The Ol-Arsd KO medaka model was generated using CRISPR/Cas9 genome-editing technology. A single-guide RNA (Ol-Arsd_gRNA: CATCGACAGACTTGCTTCAG) was microinjected along with Cas9 into fertilized medaka embryos at the one-cell stage, producing mosaic (chimeric) founder fish. From these founders, stable mutant lines carrying frameshift mutations resulting in complete loss of function were identified. A line harboring a specific 200-nucleotide deletion within the Ol-Arsd gene was selected, and loss of Ol-Arsd expression was confirmed by quantitative RT-PCR.
Confirmed knockout lines were intercrossed to generate homozygous Ol-Arsd KO offspring, from which the embryos analyzed in the experiments were derived.
For quantitative RT-PCR, medaka fish were harvested for RNA extraction using a RNeasy Mini Kit (#74106 , Qiagen) according to the manufacturer's protocol. 1 µg of total RNA was used for reverse transcription using QuantiTect Reverse Transcription Kit (Qiagen) according to the manufacturer's instructions. qRT–PCR was performed in triplicate using LightCycler 480 SYBER Green I Master (Roche) and analyzed by LightCycler 480 (Roche). The Ct values were normalized to Vinculin gene expression, and the expression of each gene was represented as 2(−ddCt) relative to control.

The primers used were as follows:
Medaka (Oryzias latipes, Ol)
Ol-Arsd _Fw_2: 5’ TCATGATGGTGGATGACCTC 3’;
Ol-Arsd _Rev_2: 3’ CAGTAAAGGCTGGACTTTTG 5’.

Cartilage and bone staining

Staining for cartilage (Alcian Blue) and bone (Alizarin Red) in fixed medaka embryos was performed according to standard medaka skeleton phenotyping protocols (https://shigen.nig.ac.jp/medaka/medakabook/index.php). Pictures were taken using a DM6000 microscope (Leica Microsystems, Wetzlar, Germany). Measurement of both cartilage and bone length was performed using ImageJ.

Immunofluorescence in medaka fish

Col10a1:nlGFP/osx:mCherry expression in embryos and larvae was analyzed by fluorescence microscopy. Embryos and larvae were anaesthetized with 0.05% tricaine methanesulfonate (MS-222) and mounted in a glass bottom petri-dish where Z-stacks were imaged using an LMS700 microscope.

Medaka MicroCT Imaging

For microCT imaging, fish were euthanized using an overdose of tricaine methanesulfonate (MS-222), and the whole fish were fixed in 4% paraformaldehyde (PFA) for 24h at 4°C. The fixed fish were then washed in 1X phosphate-buffered saline (PBS). The specimens were then stored in 70% ethanol at 4°C until further analysis. MicroCT scans were performed using the SCANCO Medical-µCT40 (Scanco Medical AG, Bassersdorf, Switzerland). The fish were placed in an 11 mm diameter sample holder and immobilized with foam to prevent distortion. The whole 2-month-old medaka fish were imaged over five scans, covering specific skeletal regions (skull, rib-bearing abdominal vertebrae, caudal vertebrae, caudal-fin vertebrae, and caudal fin). For detailed measurement analyses, fish were imaged at an isotropic voxel size of 3 µm, using the following parameters: X-ray tube potential (E) of 70 kVp; X-ray intensity (I) of 114 µA, 8 W; integration time of 300 ms with an average data (frame averaging) of 2, per slice. To reconstruct solid 3D images, selected regions were scanned in high resolution. A lower threshold of 140 was used for the evaluation of all scans. Scans were processed using the SCANCO Software to reconstruct solid 3D images. The reconstructed images were applied for visualizing bone morphology and microarchitecture and analyzed for skeletal alterations. For scan acquisition, the µCT40 was calibrated to phantoms of known hydroxyapatite density by the manufacturer, and quality controls against an external phantom were performed weekly and monthly.

Statistics

Statistics were performed in GraphPad PRISM software. A two-tailed, paired and unpaired Student's t-test was performed when comparing the same cell population with two different treatments or cells with different genotypes, respectively. One-way ANOVA with Dunnett's post hoc test was performed when comparing more than two groups relative to a single factor (treatment) and Šídák's multiple comparisons test for multiple comparisons. Two-way ANOVA was used for comparisons involving two independent factors, followed by Šidák's, Dunnett's, or Tukey's post hoc tests as appropriate. A P-value of 0.05 or less was considered statistically significant.

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (32) partner repository with the dataset identifier PXD061172. The transcriptomic data has

been deposited to the Gene Expression Omnibus (GEO) database, specifically under the accession number GSE301703. Source data and quantifications given in the main text have associated raw data. All other data supporting the findings of this study are available from the corresponding authors on reasonable request.

Funding

This study was supported by the Italian Telethon funding agency (GSA21F005) and Italian Ministry of Health (PNRR-MAD-2022-12376672).

Supporting information

This article contains supporting information.

Acknowledgements

We thank Cathal Wilson and Nicola Brunetti Pierri for insightful comments and critical manuscript review. We thank the TIGEM medaka fish facility for generating the Ol-Arsd knockout medaka model. We also acknowledge the mass spectrometry and bioinformatics core facilities at the TIGEM Institute for their support and NEGEDIA for RNA-sequencing analysis. Medaka Col10a1:GFP/Osx:mCherry transgenic line in this study was obtained by the National Bio-Resource Project (NBRP) (https://shigen.nig.ac.jp/medaka/).

Author contributions

MM and CDL performed most of the experiments. NV performed the HPLC analysis of GAG sulfation. EDP carried out bioinformatic analyses of protein structure and sequence. IC, FGS, and DI generated the Ol-Arsd knockout medaka model and performed related experiments. DC conducted the phylogenetic analysis of ARSL. CS, MM and CDL designed the study. CS, MM and CDL prepared figures and wrote the manuscript.

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Soares da Costa, D., Reis, R. L., and Pashkuleva, I. (2017) Sulfation of Glycosaminoglycans and Its Implications in Human Health and Disorders. *Annu. Rev. Biomed. Eng.* **19**, 1–26
2. Klüppel, M., Wight, T. N., Chan, C., Hinek, A., and Wrana, J. L. (2005) Maintenance of chondroitin sulfation balance by chondroitin-4-sulfotransferase 1 is required for chondrocyte development and growth factor signaling during cartilage morphogenesis. *Development*. **132**, 3989–4003
3. Thiele, H., Sakano, M., Kitagawa, H., Sugahara, K., Rajab, A., Höhne, W., Ritter, H., Leschik, G., Nürnberg, P., and Mundlos, S. (2004) Loss of chondroitin 6-O-sulfotransferase-1 function results in severe human chondrodysplasia with progressive spinal involvement. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 10155–60
4. Shabbir, R. M. K., Nalbant, G., Ahmad, N., Malik, S., and Tolun, A. (2018) Homozygous CHST11 mutation in chondrodysplasia, brachydactyly, overriding digits, clino-symphalangism and synpolydactyly. *J. Med. Genet.* **55**, 489–496
5. Mashima, R., and Nakanishi, M. (2022) Mammalian Sulfatases: Biochemistry, Disease Manifestation, and Therapy. *Int. J. Mol. Sci.* 10.3390/ijms23158153
6. Lamanna, W. C., Frese, M.-A., Balleininger, M., and Dierks, T. (2008) Sulf loss influences N-, 2-O-, and 6-O-sulfation of multiple heparan sulfate proteoglycans and modulates fibroblast growth factor signaling. *J. Biol. Chem.* **283**, 27724–27735
7. Settembre, C., Arteaga-Solis, E., McKee, M. D., de Pablo, R., Al Awqati, Q., Ballabio, A., and Karsenty, G. (2008) Proteoglycan desulfation determines the efficiency of chondrocyte autophagy and the extent of FGF signaling during endochondral ossification. *Genes Dev.* **22**, 2645–50
8. Franco, B., Meroni, G., Parenti, G., Levilliers, J., Bernard, L., Gebbia, M., Cox, L., Maroteaux, P., Sheffield, L., Rappold, G. A., Andria, G., Petit, C., and Ballabio, A. (1995) A cluster of sulfatase genes on Xp22.3: mutations in chondrodysplasia punctata (CDPX) and implications for warfarin embryopathy. *Cell*. **81**, 15–25
9. Daniele, A., Parenti, G., d'Addio, M., Andria, G., Ballabio, A., and Meroni, G. (1998) Biochemical characterization of arylsulfatase E and functional analysis of mutations found in patients with X-linked chondrodysplasia punctata. *Am. J. Hum. Genet.* **62**, 562–72

10. Brunetti-Pierri, N., Andreucci, M. V., Tuzzi, R., Vega, G. R., Gray, G., McKeown, C., Ballabio, A., Andria, G., Meroni, G., and Parenti, G. (2003) X-linked recessive chondrodysplasia punctata: spectrum of arylsulfatase E gene mutations and expanded clinical variability. *Am. J. Med. Genet. A.* **117A**, 164–8
11. Parenti, G., Meroni, G., and Ballabio, A. (1997) The sulfatase gene family. *Curr. Opin. Genet. Dev.* **7**, 386–91
12. Pauli, R. M. (1988) Mechanism of bone and cartilage maldevelopment in the warfarin embryopathy. *Pathol. Immunopathol. Res.* **7**, 107–12
13. Sardiello, M., Annunziata, I., Roma, G., and Ballabio, A. (2005) Sulfatases and sulfatase modifying factors: an exclusive and promiscuous relationship. *Hum. Mol. Genet.* **14**, 3203–17
14. Fumagalli, F., Noack, J., Bergmann, T. J., Cebollero, E., Pisoni, G. B., Fasana, E., Fregno, I., Galli, C., Loi, M., Soldà, T., D'Antuono, R., Raimondi, A., Jung, M., Melnyk, A., Schorr, S., Schreiber, A., Simonelli, L., Varani, L., Wilson-Zbinden, C., Zerbe, O., Hofmann, K., Peter, M., Quadroni, M., Zimmermann, R., and Molinari, M. (2016) Translocon component Sec62 acts in endoplasmic reticulum turnover during stress recovery. *Nat. Cell Biol.* **18**, 1173–1184
15. Tie, H. C., Mahajan, D., and Lu, L. (2022) Visualizing intra-Golgi localization and transport by side-averaging Golgi ministacks. *J. Cell Biol.* 10.1083/jcb.202109114
16. Diez-Roux, G., and Ballabio, A. (2005) Sulfatases and human disease. *Annu. Rev. Genomics Hum. Genet.* **6**, 355–79
17. Chino, H., Hatta, T., Natsume, T., and Mizushima, N. (2019) Intrinsically Disordered Protein TEX264 Mediates ER-phagy. *Mol. Cell*. **74**, 909-921.e6
18. Branon, T. C., Bosch, J. A., Sanchez, A. D., Udeshi, N. D., Svinkina, T., Carr, S. A., Feldman, J. L., Perrimon, N., and Ting, A. Y. (2018) Efficient proximity labeling in living cells and organisms with TurboID. *Nat. Biotechnol.* **36**, 880–887
19. Kiani, C., Chen, L., Wu, Y. J., Yee, A. J., and Yang, B. B. (2002) Structure and function of aggrecan. *Cell Res.* **12**, 19–32
20. Frazier, S. B., Roodhouse, K. A., Hourcade, D. E., and Zhang, L. (2008) The Quantification of Glycosaminoglycans: A Comparison of HPLC, Carbazole, and Alcian Blue Methods. *Open Glycosci.* **1**, 31–39
21. Bond, C. S., Clements, P. R., Ashby, S. J., Collyer, C. A., Harrop, S. J., Hopwood, J. J., and Guss, J. M. (1997) Structure of a human lysosomal sulfatase. *Structure*. **5**, 277–89

22. Lübke, T., and Damme, M. (2020) Lysosomal sulfatases: a growing family. *Biochem. J.* **477**, 3963–3983
23. Mourão, P. A. (1988) Distribution of chondroitin 4-sulfate and chondroitin 6-sulfate in human articular and growth cartilage. *Arthritis Rheum.* **31**, 1028–33
24. Bayliss, M. T., Osborne, D., Woodhouse, S., and Davidson, C. (1999) Sulfation of chondroitin sulfate in human articular cartilage. The effect of age, topographical position, and zone of cartilage on tissue composition. *J. Biol. Chem.* **274**, 15892–900
25. Jiang, H., Liu, X.-Y., Zhang, G., and Li, Y. (2005) Kinetics and template nucleation of self-assembled hydroxyapatite nanocrystallites by chondroitin sulfate. *J. Biol. Chem.* **280**, 42061–6
26. King, K. B., and Kimura, J. H. (2003) The establishment and characterization of an immortal cell line with a stable chondrocytic phenotype. *J. Cell. Biochem.* **89**, 992–1004
27. Reya, T., Duncan, A. W., Ailles, L., Domen, J., Scherer, D. C., Willert, K., Hintz, L., Nusse, R., and Weissman, I. L. (2003) A role for Wnt signalling in self-renewal of haematopoietic stem cells. *Nature*. **423**, 409–14
28. Shin, H. R., Citron, Y. R., Wang, L., Tribouillard, L., Goul, C. S., Stipp, R., Sugasawa, Y., Jain, A., Samson, N., Lim, C.-Y., Davis, O. B., Castaneda-Carpio, D., Qian, M., Nomura, D. K., Perera, R. M., Park, E., Covey, D. F., Laplante, M., Evers, A. S., and Zoncu, R. (2022) Lysosomal GPCR-like protein LYCHOS signals cholesterol sufficiency to mTORC1. *Science*. **377**, 1290–1298
29. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods*. **9**, 671–5
30. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D. J., Hartenstein, V., Eliceiri, K., Tomancak, P., and Cardona, A. (2012) Fiji: an open-source platform for biological-image analysis. *Nat. Methods*. **9**, 676–82
31. Balch, W. E., Dunphy, W. G., Braell, W. A., and Rothman, J. E. (1984) Reconstitution of the transport of protein between successive compartments of the Golgi measured by the coupled incorporation of N-acetylglucosamine. *Cell*. **39**, 405–16
32. Volpi, N., and Linhardt, R. J. (2010) High-performance liquid chromatography-mass spectrometry for mapping and sequencing glycosaminoglycan-derived oligosaccharides. *Nat. Protoc.* **5**, 993–1004

33. Volpi, N., Galeotti, F., Yang, B., and Linhardt, R. J. (2014) Analysis of glycosaminoglycan-derived, precolumn, 2-aminoacridone-labeled disaccharides with LC-fluorescence and LC-MS detection. *Nat. Protoc.* **9**, 541–58

34. Rappsilber, J., Mann, M., and Ishihama, Y. (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* **2**, 1896–906

35. Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., Inuganti, A., Griss, J., Mayer, G., Eisenacher, M., Pérez, E., Uszkoreit, J., Pfeuffer, J., Sachsenberg, T., Yilmaz, S., Tiwary, S., Cox, J., Audain, E., Walzer, M., Jarnuczak, A. F., Ternent, T., Brazma, A., and Vizcaíno, J. A. (2019) The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450

36. UniProt Consortium (2023) UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res.* **51**, D523–D531

37. Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., Yuan, D., Stroe, O., Wood, G., Laydon, A., Žídek, A., Green, T., Tunyasuvunakool, K., Petersen, S., Jumper, J., Clancy, E., Green, R., Vora, A., Lutfi, M., Figurnov, M., Cowie, A., Hobbs, N., Kohli, P., Kleywegt, G., Birney, E., Hassabis, D., and Velankar, S. (2022) AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **50**, D439–D444

38. Varadi, M., Bertoni, D., Magana, P., Paramval, U., Pidruchna, I., Radhakrishnan, M., Tsenkov, M., Nair, S., Mirdita, M., Yeo, J., Kovalevskiy, O., Tunyasuvunakool, K., Laydon, A., Žídek, A., Tomlinson, H., Hariharan, D., Abrahamson, J., Green, T., Jumper, J., Birney, E., Steinegger, M., Hassabis, D., and Velankar, S. (2024) AlphaFold Protein Structure Database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res.* **52**, D368–D375

39. Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., Bileschi, M. L., Bork, P., Bridge, A., Colwell, L., Gough, J., Haft, D. H., Letunić, I., Marchler-Bauer, A., Mi, H., Natale, D. A., Orengo, C. A., Pandurangan, A. P., Rivoire, C., Sigrist, C. J. A., Sillitoe, I., Thanki, N., Thomas, P. D., Tosatto, S. C. E., Wu, C. H., and Bateman, A. (2023) InterPro in 2022. *Nucleic Acids Res.* **51**, D418–D427

40. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D.,

Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics*. **23**, 2947–8

41. Price, M. N., Dehal, P. S., and Arkin, A. P. (2010) FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One*. **5**, e9490
42. Huerta-Cepas, J., Serra, F., and Bork, P. (2016) ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol. Biol. Evol.* **33**, 1635–8
43. Meng, E. C., Goddard, T. D., Pettersen, E. F., Couch, G. S., Pearson, Z. J., Morris, J. H., and Ferrin, T. E. (2023) UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci.* **32**, e4792
44. Dobson, L., Szekeres, L. I., Gerdán, C., Langó, T., Zeke, A., and Tusnády, G. E. (2023) TmAlphaFold database: membrane localization and evaluation of AlphaFold2 predicted alpha-helical transmembrane protein structures. *Nucleic Acids Res.* **51**, D517–D522
45. Madeira, F., Madhusoodanan, N., Lee, J., Eusebi, A., Niewielska, A., Tivey, A. R. N., Lopez, R., and Butcher, S. (2024) The EMBL-EBI Job Dispatcher sequence analysis tools framework in 2024. *Nucleic Acids Res.* **52**, W521–W525
46. Kinsella, R. J., Kähäri, A., Haider, S., Zamora, J., Proctor, G., Spudich, G., Almeida-King, J., Staines, D., Derwent, P., Kerhornou, A., Kersey, P., and Flicek, P. (2011) Ensembl BioMarts: a hub for data retrieval across taxonomic space. *Database (Oxford).* **2011**, bar030
47. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., and Higgins, D. G. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539
48. Iwamatsu, T. (2004) Stages of normal development in the medaka Oryzias latipes. *Mech. Dev.* **121**, 605–18

Figure legends

Figure 1. Characterization of human and rat Arsl

A. Immunofluorescence analysis of endogenous rat Arsl (green) in RCS WT cells co-stained for different intracellular markers (red): PDI (ER marker); Lamp1 (lysosome marker) and Gm130 (Golgi marker). Nuclear staining with DAPI (blue). Rightmost panels, magnification of the boxed areas. Scale bar, 20 μm.

B. The bar graph shows co-localization of endogenous Arsl with different intracellular markers (expressed as Manders' Coefficient). Mean ± standard error of mean (SEM) of N = 3 independent experiments. n = 30 cells were analyzed. One-way analysis of variance (ANOVA) with Sidak's multiple comparisons test: ****p < 0.0001.

C. Western Blot analysis of subcellular fractions isolated from RCS WT cells using antibodies for Arsl, Gm130 (Golgi marker), Lamp1 (lysosomes marker), or Vapa (ER marker). Input: post-nuclear supernatant.

D. Immunofluorescence analysis of RCS WT with Nocodazole treatment (33 μM for 3h). Nocodazole interferes with microtubule polymerization, generating Golgi mini-stacks distributed throughout the cytoplasm. A triple staining was performed using Trans-Golgi (Tgn38, blue), Cis-Golgi (Gm130, red), and Arsl (green) antibodies. Scale bar, 5 μm.

E. The bar graph shows co-localization of Arsl with Cis- (Gm130) and Trans-Golgi (Tgn38) markers (expressed as Manders' Coefficient). Mean ± standard error of mean (SEM) of N = 3 independent experiments. n = 18 cells were counted. Student's unpaired t test: **p < 0.005.

F. Western Blot analysis of WT (CTRL), Arsl KO, and Arsl KO overexpressing either Myc-tagged wild-type (Myc-hARSL) or the inactive-mutant form of hARSL with the C86A mutation (Myc-hARSL C86A), detected using an anti-Myc Ab. Actin was used as a loading control.

G. ARSL enzymatic assay in CTRL, Arsl KO, and Arsl KO overexpressing wild-type (Myc-hARSL) and the inactive-mutant form of hARSL with C86A mutation (Myc-hARSL C86A). Mean ± standard error of mean (SEM) of N = 3 biological replicates. One-way analysis of variance (ANOVA) with Sidak's multiple comparisons test: *p < 0.05; **p < 0.005.

H. ARSL enzymatic assay in RCS overexpressing FLAG-hARSL. Mean ± standard error of mean (SEM) of N = 3 biological replicates. Student's unpaired t test: **p < 0.005.

I. Immunofluorescence staining of endogenous rat Arsl (green) and human ARSL fused to a FLAG tag (red) in RCS cells overexpressing FLAG-hARSL. Nuclear staining with DAPI (blue). Scale bar, 10 μm.

J. ARSL enzymatic assay in RCS WT with and without 2 mM warfarin treatment. Mean ± standard error of mean (SEM) of N = 3 independent experiments Student's unpaired t test: ****$p < 0.0001$.

K. ARSL enzymatic assay in RCS overexpressing FLAG-hARSL with and without 2 mM warfarin treatment. Mean ± standard error of mean (SEM) of N = 3 independent experiments Student's unpaired t test: **$p < 0.005$.

Figure 2. ARSL is a Golgi luminal enzyme that potentially modulates the composition of extracellular matrix proteoglycans

A. Putative tridimensional conformation of the ARSL enzyme retrieved from AlphaFold Protein Structure Database (AF-P51690-F1-v4). The catalytic domain (green) is located within the lumen of the Trans-Golgi network, along with the active site (red). The zoomed-in view highlights the amino acid residues of the catalytic domain and the active site. The transmembrane domains are shown in yellow.

B. Schematic overview of the protease protection assay used to assess protein localization and topology within membrane-bound compartments.

C. Protease protection assay analysis of FLAG-hARSL topology in RCS cells. Western blot detection of FLAG-hARSL and control proteins following Proteinase K treatment in the absence or presence of detergent. Differential protease accessibility distinguishes cytosolic and luminal proteins, with LAMP1 and PDI used as cytosolic and luminal topology markers, respectively. FLAG immunoblotting was used to assess the subcellular localization of ARSL.

D. Volcano plot highlights all significantly upregulated biotinylated proteins identified in RCS WT cells overexpressing the rat FLAG-TurboID-Arsl compared to non-biotinylated control cells. The red dots specifically indicate Golgi-resident proteins.

E. Number and percentage of TurboID/ streptavidin- enriched proteins identified by mass spectrometry (MS) and classified according to cellular compartmentalization based on Gene Ontology (GO) term enrichment analysis.

Figure 3. Chondroitin sulfate (4-O-S-CS) is a candidate ARSL substrate

A. Bar graphs show HPLC-based quantification of sulfation levels of heparan sulfate (HS; left) and chondroitin/dermatan sulfate (CS/DS; right) disaccharide units in RCS cells overexpressing hARSL. Values were expressed as fold change relative to untransfected RCS cells (CTRL). Mean ± standard error of mean (SEM) of N = 3 independent experiments.

Two-way analysis of variance (ANOVA) with Tukey's multiple comparisons test. ***p < 0.0005; ns: not significant.

B. HPLC analysis measuring sulfation levels of CS/DS disaccharide units at different positions in Arsl KO and control RCS cells. Values were expressed as fold change relative to RCS WT cells (CTRL). Mean ± standard error of mean (SEM) of N = 3 independent experiments. Two-way analysis of variance (ANOVA) with Sidak's multiple comparisons test: **p < 0.005; ***p < 0.0005; ns: not significant.

C. Workflow of HPLC analysis of CS/DS GAG sulfation levels. First, the collected medium from WT (CTRL), Arsl KO, KO cells expressing wild-type human ARSL (Myc-hARSL), and KO cells expressing the inactive mutant ARSL (Myc-hARSL C86A), was used to isolate the extracellular matrix (1). Next, the isolated extracellular matrix was digested with chondroitinase ABC or heparinases and heparitinases, which specifically breaks down long GAG chains into their disaccharide units. The sulfation levels of the disaccharide units were analyzed by HPLC (2), which can distinguish various types of GAG sulfation.

D. CS/DS sulfation state at different positions from media collected from Arsl KO, KO cells reconstituted with wild type ARSL (hARSL), or catalytically inactive ARSL (hARSL C86A). Values were expressed as fold change relative to RCS WT cells (CTRL). Mean ± standard error of mean (SEM) of N = 4 biological replicates. Two-way analysis of variance (ANOVA) with Dunnett's multiple comparisons test: *p < 0.05; **p < 0.005; ***p < 0.0005; ****p < 0.0001.

E. The workflow for HPLC analysis of CS GAG sulfation levels in isolated Golgi incubated with exogenous CS**.** First, Golgi fractions were isolated from WT (CTRL), *Arsl* KO, and *Arsl* KO cells overexpressing hARSL (1). The extracted Golgi fractions were then incubated with exogenous CS (2). Finally, the sulfation levels of the disaccharide units in CS were analyzed by HPLC (3).

F. Exogenous CS molecules were incubated with Golgi fractions isolated from Arsl KO, and Arsl KO overexpressing hARSL cells. Bar graph shows CS sulfation levels at indicated position expressed as fold change relative to RCS WT cells (CTRL). Mean ± standard error of mean (SEM) of N = 9 biological replicates. Two-way analysis of variance (ANOVA) with Tukey's multiple comparisons test. **p < 0.005; ***p < 0.0005; ****p < 0.0001.

G. Schematic representation of a structural comparison of ARSL (blue) and ARSB (yellow). The catalytic site of the proteins is depicted in dark blue (ARSL) and in dark red (ARSB), and the initial and final amino acids of the related sub-sequence are zoomed.

Figure 4. Chondroctyes lacking ARSL show altered TGF-β signalling

A. Volcano plot showing differentially expressed genes identified between RCS WT and Arsl KO cells. Black dashed line represents a threshold on False Discovery Rate (Benjamini–Hochberg correction) at 0.05. No threshold was applied on logarithmic Fold Change. Legend: Up, Upregulated genes (red); Down, Downregulated genes (green); NS, not statistically significant genes (grey); FC, Fold Change; FDR, False Discovery Rate.

B. Pathway enrichment analysis of downregulated genes in ARSL KO versus WT cells identified the TGF-β signaling pathway among the most significantly affected pathways. Analysis was performed using BioPlanet (Enrichr).

C. qPCR analysis of gene expression in WT and ARSL KO cells following treatment with TGF-β1. Cells were stimulated with TGF-β1 for 48h and the expression levels of chondrogenic markers (*COL2A1*, *SOX9*) and hypertrophic marker (*RUNX2*) were assessed. Gene expression levels are presented as fold change relative to WT cells. Mean ± standard error of mean (SEM) of N = 4 independent experiments. Two-way analysis of variance (ANOVA) with Sidak's multiple comparisons test. *p < 0.05; ****p < 0.00005.

D. Schematic model of ARSL function in the trans-Golgi network. ARSL, a Golgi-resident luminal enzyme, acts on 4-O-sulfated chondroitin sulfate (4-O-CS) chains during proteoglycan biosynthesis. Under physiological conditions (WT), it modulates 4-O-sulfation, supporting proper extracellular matrix organization and normal cartilage and bone development via the TGF-β pathway. In pathological conditions (ARSL KO), loss of ARSL increases 4-O-sulfation, leading to an altered extracellular matrix (hyper–4-O-sulfated CS) and impaired development through downregulation of TGF-β signaling.

Figure 5. Ol-Arsd controls CS sulfation and skeletal development in medaka

A. Bar graph shows HPLC analysis of CS/DS sulfation in Ol-Arsd KO and control (CTRL) pooled medaka samples collected at stage 40 (which corresponds to early larval phase at 4-9 days post-fertilization), and 1 month after birth. Mean ± standard error of mean (SEM) of N = 3 represents independent biological replicates, each consisting of a distinct pool of fish of the same genotype. One-way analysis of variance (ANOVA) with Sidak's multiple comparisons test: **p < 0.005; ****p < 0.0001; ns: not significant.

B. Representative images of the total length of Ol-Arsd KO and WT (CTRL) medaka at stage 40 post-birth.

C. The bar graph quantifies the total length measurements in both WT (CTRL) and Ol-Arsd KO medaka models, expressed in fold-change relative to WT. Mean ± standard error of mean (SEM) of N = 3 biological replicates, each corresponding to an individual and independent fish per genotype. Student's unpaired t test: **p < 0.005.

D-G. Representative images of Alcian Blue (cartilage, E) and Alizarin red (bone, F) skeletal staining in WT (CTRL) and Ol-Arsd KO medaka model at st.40. Bar graph (G) shows the measured lengths of ceratobranchials 1–5 (CB1 to CB5) and Meckel's cartilage (m; w: width; l: length) in WT and Ol-Arsd KO medaka. Bar graph (H) illustrates the quantification of cartilage ossification relative to the Opercle (op) and Palatoquadrate (pq) in wild-type (WT) and Ol-Arsd KO medaka model. Data are expressed in fold-change relative to WT. Mean ± standard error of mean (SEM) of N = 5 biological replicates, each representing an individual and independent fish per genotype. Two-way analysis of variance (ANOVA) with Sidak's multiple comparisons test: ***$p < 0.0005$; ****$p < 0.0001$.

H. MicroCT visualization of the rostral region in WT and Ol-Arsd KO medaka at 2 months. Bar graph (below) shows the measure of Meckel's cartilage (w: width; l: length) in WT (CTRL) and Ol-Arsd KO medaka model. Mean ± standard error of mean (SEM) of N = 3 biological replicates, each corresponding to an individual fish with the same genotype. One-way analysis of variance (ANOVA) with Sidak's multiple comparisons test: *$p < 0.05$; **$p < 0.005$.

I. Representative immunofluorescence images of the rostral region in WT and *Ol-Arsd* KO medaka expressing fluorescent reporters for chondrogenic and osteogenic markers (Col10a1:nuGFP and Osx:mCherry). Palatoquadrate (pq), ceratohyal (ch), opercle (op).

Table legends

Table S1. Proteomic analysis of biotinylated proteins in RCS WT cells overexpressing the rat FLAG-TurboID-Arsl (imputated data).

Table S2. Proteomic analysis of biotinylated proteins in RCS WT cells overexpressing the rat FLAG-TurboID-Arsl (non-imputated data).

Table S3. RNA-seq analysis of Arylsulfatase L (ARSL) deletion in rat chondrosarcoma (RCS) cell lines.

Table S4. Phylogenetic analysis of the arylsulfatase protein family in Medaka, Humans, and Mice.

Figure 1

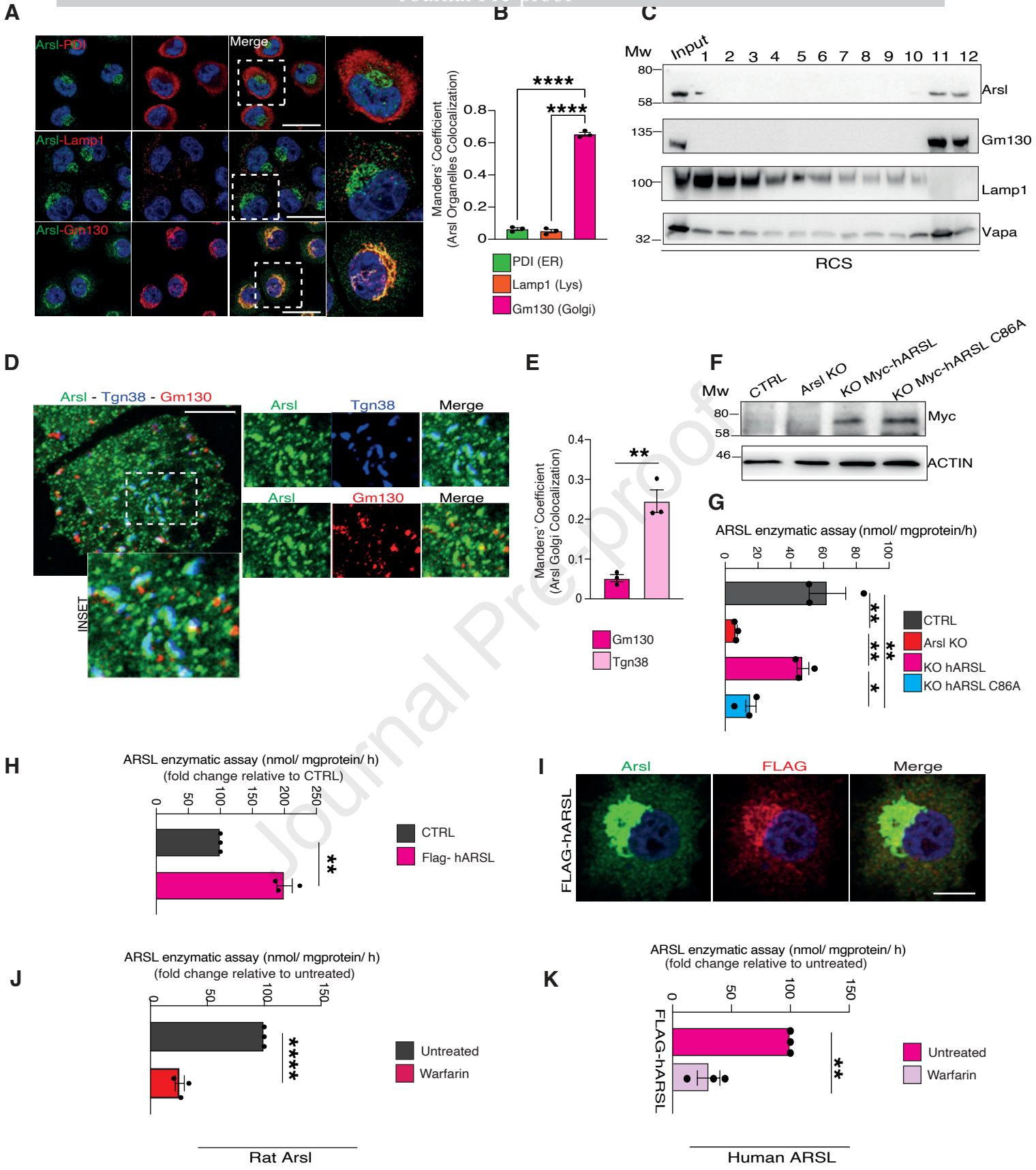


Figure 2

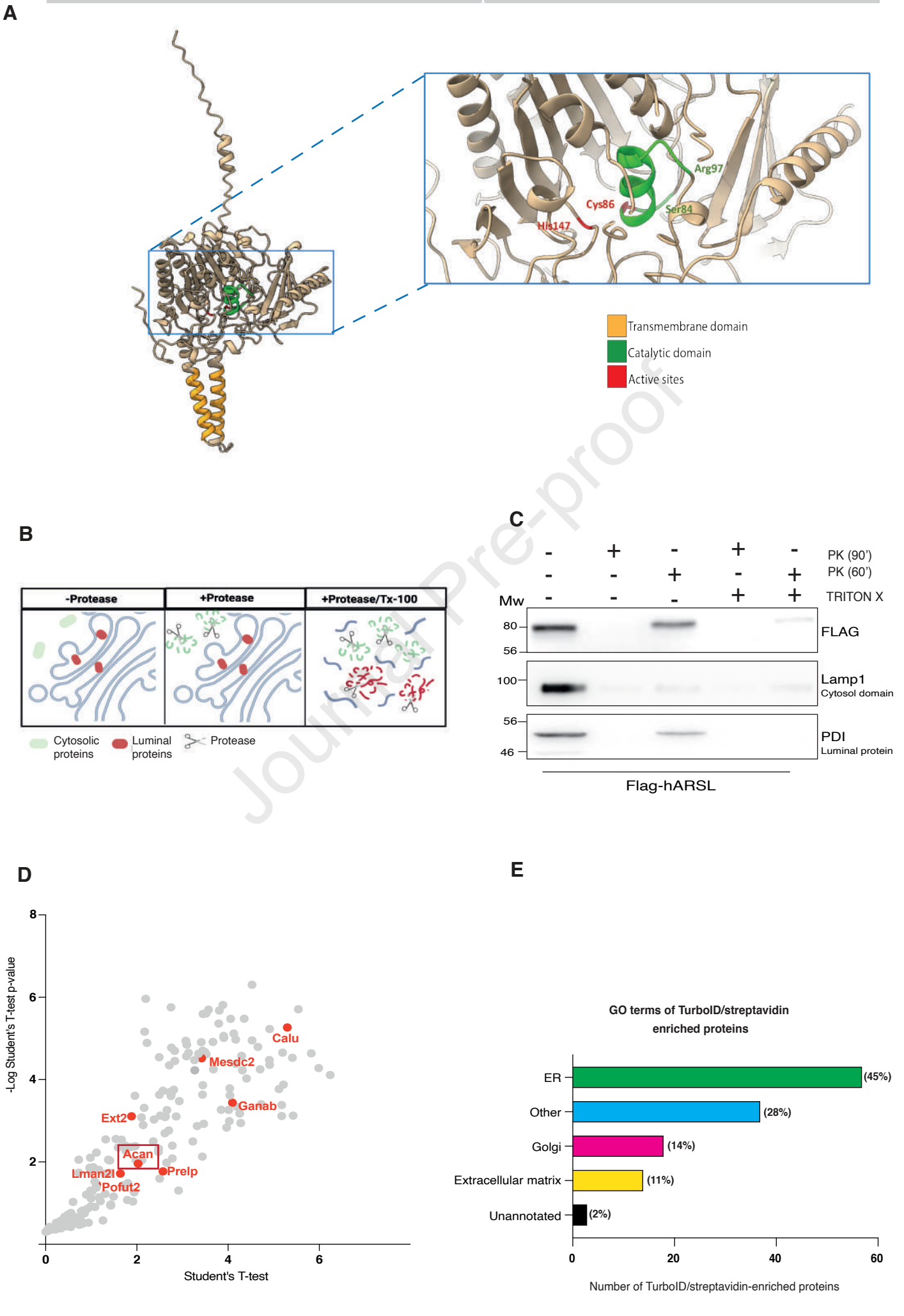


Figure 3

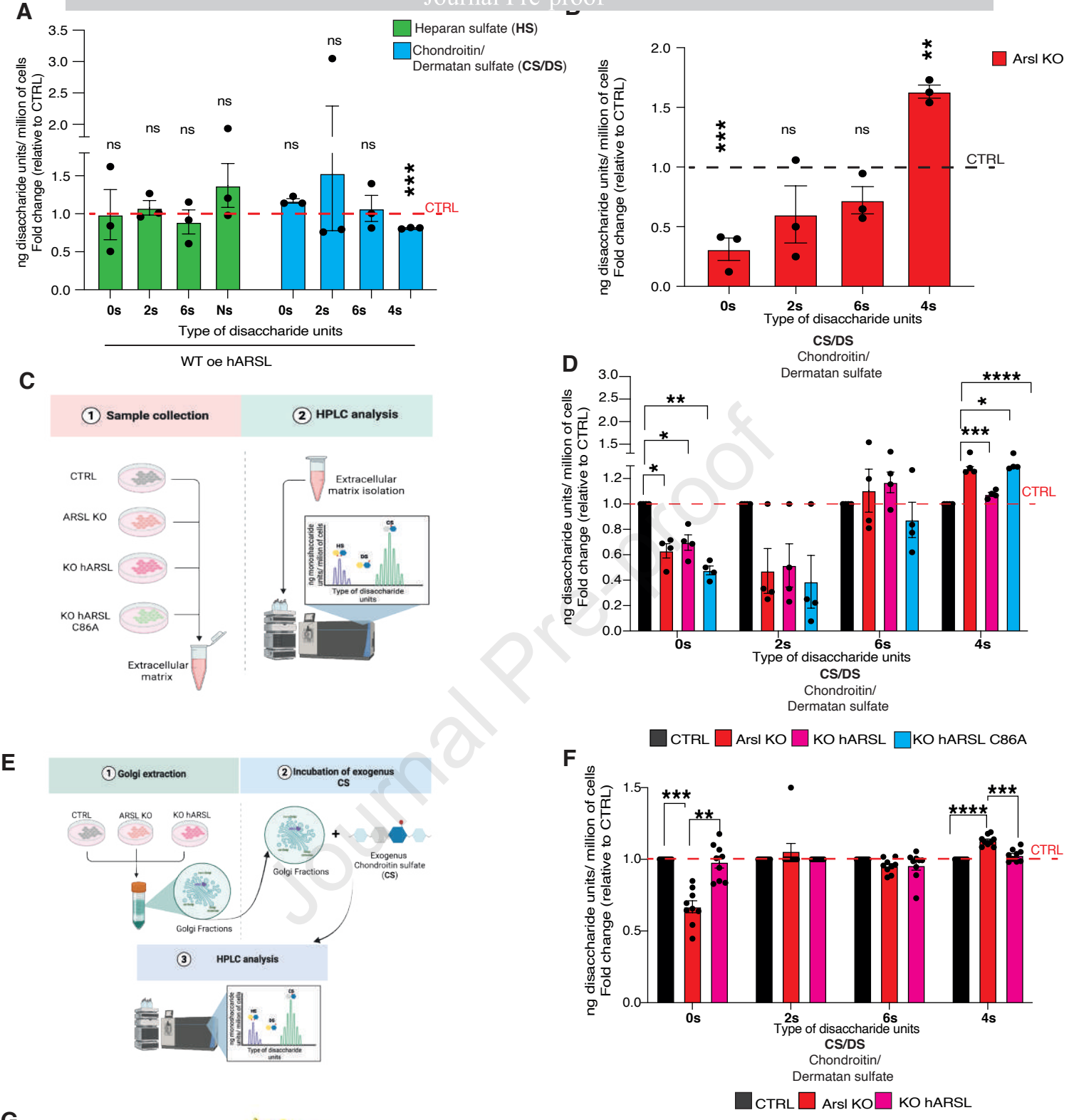


Figure 4

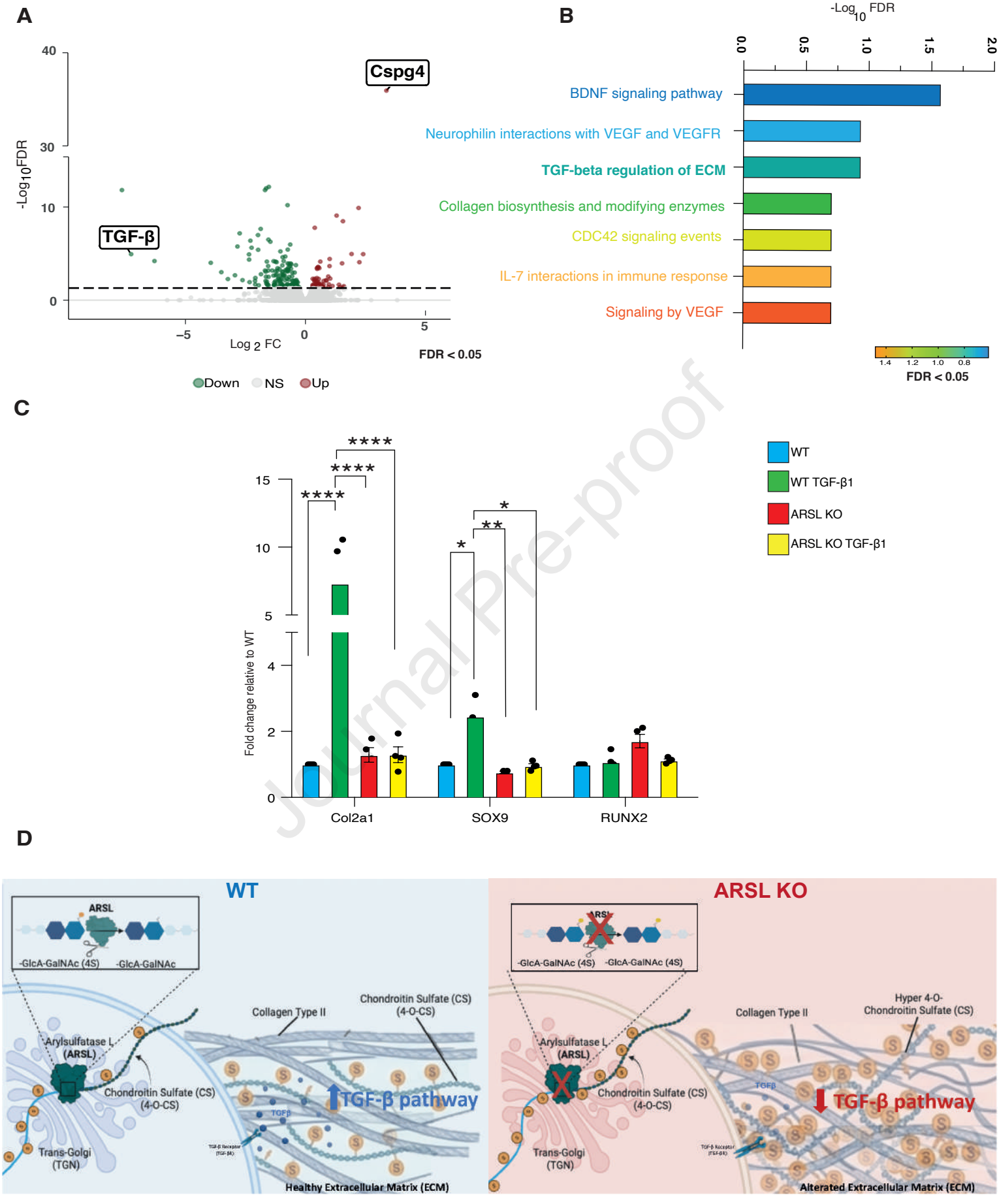
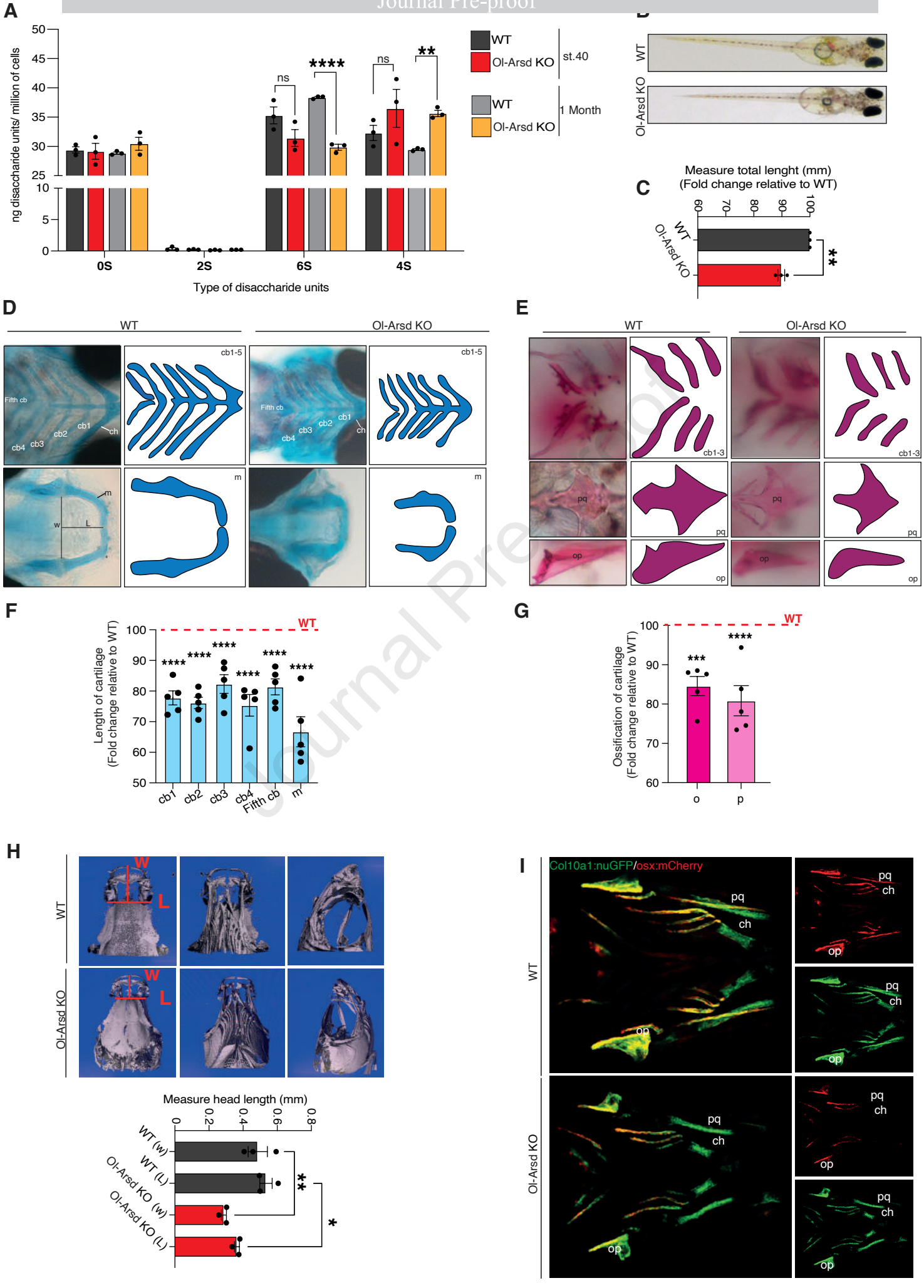


Figure 5



Declaration of Interest Statement

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

☐ The author is an Editorial Board Member/Editor-in-Chief/Associate Editor/Guest Editor for this journal and was not involved in the editorial review or the decision to publish this article.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: