THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 288, NO. 42, pp. 30019–30028, October 18, 2013

Arylsulfatase K, a Novel Lysosomal Sulfatase*

Received for publication, July 4, 2013, and revised form, August 20, 2013 Published, JBC Papers in Press, August 28, 2013, DOI 10.1074/jbc.M113.499541

Elena Marie Wiegmann[‡1], **Eva Westendorf**[‡1], **Ina Kalus**[‡], **Thomas H. Pringle**[§], **Torben Lübke**[‡], **and Thomas Dierks**[‡2]

From the [‡]*Department of Chemistry, Biochemistry I, Bielefeld University, 33615 Bielefeld, Germany and the* [§]*Sperling Foundation, Eugene, Oregon 97405*

Background: Human sulfatases play key roles in physiology and cause numerous pathological conditions upon deficiency/misregulation.
Results: ARSK is ubiquitously expressed, localizes to lysosomes, and shows arylsulfatase activity at acidic pH.
Conclusion: ARSK is a novel lysosomal sulfatase acting on a ubiquitous substrate.
Significance: ARSK functions in lysosomal degradation, possibly of glycosaminoglycans, and, in all probability, is associated with a non-classified lysosomal storage disorder.

The human sulfatase family has 17 members, 13 of which have been characterized biochemically. These enzymes specifically hydrolyze sulfate esters in glycosaminoglycans, sulfolipids, or steroid sulfates, thereby playing key roles in cellular degradation, cell signaling, and hormone regulation. The loss of sulfatase activity has been linked to severe pathophysiological conditions such as lysosomal storage disorders, developmental abnormalities, or cancer. A novel member of this family, arylsulfatase K (ARSK), was identified bioinformatically through its conserved sulfatase signature sequence directing posttranslational generation of the catalytic formylglycine residue in sulfatases. However, overall sequence identity of ARSK with other human sulfatases is low (18–22%). Here we demonstrate that ARSK indeed shows desulfation activity toward arylsulfate pseudosubstrates. When expressed in human cells, ARSK was detected as a 68-kDa glycoprotein carrying at least four *N*-glycans of both the complex and high-mannose type. Purified ARSK turned over *p*-nitrocatechol and *p*-nitrophenyl sulfate. This activity was dependent on cysteine 80, which was verified to undergo conversion to formylglycine. Kinetic parameters were similar to those of several lysosomal sulfatases involved in degradation of sulfated glycosaminoglycans. An acidic pH optimum (~4.6) and colocalization with LAMP1 verified lysosomal functioning of ARSK. Further, it carries mannose 6-phosphate, indicating lysosomal sorting via mannose 6-phosphate receptors. *ARSK* mRNA expression was found in all tissues tested, suggesting a ubiquitous physiological substrate and a so far non-classified lysosomal storage disorder in the case of ARSK deficiency, as shown before for all other lysosomal sulfatases.

Sulfatases represent an evolutionary conserved enzyme family that comprises 17 members in humans (1, 2). These enzymes catalyze the hydrolysis of sulfate esters of a variety of substrates such as glycosaminoglycans (heparin, heparan sulfate, chondroitin/dermatan sulfate, and keratan sulfate), sulfolipids (*e.g.* cerebroside-3-sulfate), and sulfated hormones (*e.g.* dehydroepiandrosteron-3-sulfate), thereby contributing either to the degradation of macromolecules and cellular components or hormone activation (3, 4). Two sulfatases act on the cell surface as editors of the sulfation status of heparan sulfate proteoglycans (5–7) and, thereby, regulate fundamental signaling pathways involving numerous heparan sulfate-dependent growth factors and morphogens (for a review, see Ref. 8).

In humans, sulfatases display functional and structural homologies but show strict specificity toward their natural substrate. Each enzyme catalyzes a precise desulfation step, hence explaining the non-redundancy of sulfatases *in vivo*. *In vitro*, however, many human sulfatases share activity against small sulfated aromatic pseudosubstrates like *p*-nitrocatechol sulfate (pNCS)[3] or *p*-nitrophenyl sulfate (pNPS) and 4-methylumbelliferyl sulfate, which was the basis for the arylsulfatase nomenclature. For enzymatic activity, all sulfatases require C_α-formylglycine (FGly) in their catalytic site (3, 9, 10). This unique amino acid functionality is introduced by the oxidation of a conserved cysteine residue that is part of a C-T/S/C/A-P-S-R motif within the so-called sulfatase signature (11, 12). FGly modification occurs during the translocation of newly synthesized sulfatase polypeptides into the endoplasmic reticulum (ER) and is catalyzed by the ER-resident FGly-generating enzyme (FGE) (13, 14). A compromised FGE function leads to the severe metabolic disorder multiple sulfatase deficiency, in which the activity of all sulfatases is severely reduced (14–16).

All human sulfatases are processed via the secretory pathway and are extensively glycosylated in the ER and Golgi during transport to their final subcellular compartment. They can be grouped into the non-lysosomal and the lysosomal sulfatases according to their subcellular localization and pH preference. The non-lysosomal group includes the ER-localized arylsulfatases C, D, and F as well as the Golgi-localized arylsulfatase E and the cell surface-localized sulfatases Sulf1 and Sulf2, which are all active at neutral pH. The second group consists of seven

* This work was supported by the Deutsche Forschungsgemeinschaft and Shire Human Genetic Therapies Inc. (Lexington, MA).
[1] Both authors contributed equally to this work.
[2] To whom correspondence should be addressed: Dept. of Chemistry, Biochemistry I, Bielefeld University, Universitätsstr. 25, 33615 Bielefeld, Germany. Tel.: 49-521-1062092; Fax: 49-521-1066014; E-mail: thomas.dierks@uni-bielefeld.de.

[3] The abbreviations used are: pNCS, *p*-nitrocatechol sulfate; pNPS, *p*-nitrophenyl sulfate; FGly, formylglycine; ER, endoplasmic reticulum; FGE, formylglycine-generating enzyme; M6P, mannose 6-phosphate; MPR, mannose 6-phosphate receptor; ARSK, arylsulfatase K.

human sulfatases (iduronate 2-sulfatase, glucosamine 6-sulfatase, galactosamine 6-sulfatase, sulfamidase, and arylsulfatases A, B, and G) that have been demonstrated to be localized in the lysosome and exhibit an acidic pH optimum (4, 17).

The importance of the human sulfatases is underlined by the existence of, so far, eight inherited diseases that are due to single sulfatase deficiencies. Loss of arylsulfatase C function leads to the skin disease X-linked ichthyosis (18). Mutations in arylsulfatase E lead to the bone disease chondrodysplasia punctata type 1 (19). Six of the seven known lysosomal sulfatases are correlated to different forms of lysosomal storage disorders. While deficiency of arylsulfatase A (cerebroside-3-sulfatase) leads to metachromatic leukodystrophy, five sulfatases, namely arylsulfatase B, galactosamine-6-sulfatase, glucosamine-6-sulfatase, sulfamidase, and iduronate-2-sulfatase, which all are involved in the degradation of glycosaminoglycans, lead to different types of mucopolysaccharidosis in case of deficiency (4). In affected patients with these lysosomal storage disorders, the degradation of a specific sulfated compound is blocked, leading to its accumulation in the lysosomes and in the extracellular fluids. Lysosomal storage finally results in an overall dysfunction of the lysosome, cellular damage, and apoptosis (20). Recently, we characterized the novel lysosomal sulfatase arylsulfatase G and showed that its inactivation in mice results in loss of heparan sulfate 3-*O*-sulfatase activity, thus leading to a new lysosomal storage disorder, mucopolysaccharidosis IIIE (17, 21). Therefore, the consistent association of all known lysosomal sulfatases with corresponding storage diseases gives reason for in-depth analyses of sulfatases of unknown function that had been identified in a genome-wide search for sulfatases in humans. In fact, for several sulfated substrates, the corresponding sulfatases and possible associated storage disorders have not yet been identified.

One of these novel sulfatases is encoded by the *ARSK* gene that is located on chromosome 5q15 in the human genome. The gene encodes a 536-amino acid protein with a predicted 22-amino acid signal peptide directing ER translocation. ARSK (earlier names are SulfX, Sulf3, TSulf, and bone-related sulfatase) displays an overall sequence identity of 18–22% (32–38% sequence similarity) to other human sulfatases (2, 22, 23) and was classified as a human sulfatase because of the presence of the sulfatase signature sequence motif CCPSR at positions 80–84 and the conservation of other catalytic residues. Conversion of the cysteine residue at position 80 into FGly was indirectly verified by demonstrating efficient *in vitro* FGly formation in the ARSK-derived peptide Sulf3-(70–91) FLNAYTNSPI*CCPSR*AAMWSGLS by purified FGE (24). ARSK lacks a transmembrane domain and a putative GPI anchor site and is predicted to be a soluble protein with multiple *N*-glycosylation sites. In this work, we demonstrate that human ARSK is a lysosomal enzyme that shows an acidic pH optimum for catalytic activity against arylsulfatase substrates and carries mannose 6-phosphate as a lysosomal sorting signal.

EXPERIMENTAL PROCEDURES

Antibodies—A rabbit polyclonal antiserum (rabbit anti-ARSK) was generated against recombinant human ARSK-RGS-His_6, expressed in *Escherichia coli* Tuner (DE3) cells using the pET-Blue system (Novagen). The antigen was purified from inclusion bodies under denaturing conditions on nickel-nitrilotriacetic acid-agarose (Qiagen) as described by the manufacturer (QIAexpressionist Handbook). Mannose 6-phosphate (M6P)-containing proteins were detected using the scFv M6P-1 single-chain antibody fragment, as described previously (25), and a rabbit anti-c-Myc antibody (catalog no. C3956, Sigma). Other antibodies used were anti-RGS-His_6-tag (Qiagen), anti-LAMP-1 (catalog name 1D4B, Developmental Studies Hybridoma Bank), and horseradish peroxidase-conjugated secondary antibodies (Invitrogen).

Expression Analysis of ARSK in Human Tissues—To identify *ARSK* mRNA transcripts, a panel of normalized cDNAs from eight different human tissues (MTC panel human I, Clontech) was amplified by PCR using *ARSK*-specific primers (forward primer 5′-TTA ATT CAT CTG GAT CCG AGG AAA G-3′ and reverse primer 5′-AAT CGT GTG GAA GCT GG-3′) to generate a 931-bp fragment. PCR was carried out for 36 cycles with an annealing temperature of 55 °C. The resulting fragment was verified by sequencing. Normalization was confirmed by amplifying a 1000-bp fragment for glyceraldehyde-3-phosphate dehydrogenase cDNA (*GAPDH*).

Cloning and Expression of ARSK—The human *ARSK* cDNA was reverse-transcribed from total mRNA of human fibroblasts. *ARSK* was amplified as a C-terminal RGS-His_6-tagged derivative by add-on PCR using a XhoI forward primer (5′-CCG CTC GAG CCA CCA TGC TAC TGC TGT GGG TG-3′) and a NotI-RGS-His_6 reverse primer (5′-ATA GTT TAG CGG CCG CTA GTG ATG GTG ATG GTG ATG CGA TCC TCT AAC TGC TCT TGG ATT CAT ATG G-3′). The *ARSK*-His_6 cDNA construct was initially cloned into the multiple cloning site of pLPCX (Clontech) and, to achieve better expression, finally moved as a blunted fragment into the pSB4.7pA vector (provided by Shire Human Genetic Therapies, Lexington MA). We inserted the C80A mutation into the *ARSK*-His_6 construct using the QuikChange site-directed mutagenesis protocol (Stratagene) with the following complementary primers: 5′-CAC AAA CTC TCC AAT TGC CTG CCC ATC ACG CG-3′ and 5′-CGC GTG ATG GGC AGG CAA TTG GAG AGT TTG TG-3′. Finally, all constructs were full length-sequenced.

HT1080 and HEK293 cells were transfected with Lipofectamine LTX (Invitrogen) as recommended by the manufacturer. To obtain stably ARSK-expressing cell lines, HT1080 and HEK293 colonies were selected with G418 for 12 days and subsequently grown in the presence of 800 μg/ml G418.

Cell Culture—If not stated otherwise, HT1080, HEK293 cells, and mouse embryonic fibroblasts were grown at 37 °C under 5% CO_2 in complete DMEM (Invitrogen) containing 10% FCS (Lonza).

Purification of Recombinant ARSK-His_6 and ARSK-C80A-His_6—HEK293 cells stably expressing ARSK-His_6 or ARSK-C80A-His_6, respectively, were shifted to DMEM with 1% FCS. Conditioned medium was collected three times every 48 h and precipitated with ammonium sulfate (50% w/v). After reconstitution in HisTrap binding buffer (20 mM imidazole, 20 mM Tris, 500 mM NaCl (pH 7.4)) and dialysis overnight at 4 °C, the dialyzed protein was cleared by centrifugation at 17 000 × *g* for 30

min, filtered through a 0.22-μm filter, and loaded onto a 1-ml HisTrap column at a flow rate of 1 ml/min using the ÄKTA Explorer purification system (GE Healthcare). After washing with washing buffer (20 mM imidazole, 20 mM Tris, 500 mM NaCl (pH 7.4)) elution of the column was performed applying a linear gradient from to 20–500 mM imidazole (in 20 mM Tris, 50 mM NaCl (pH 7.4)) over 15 column volumes (1 column volume/fraction). Fractions were analyzed by Western blotting using anti-RGS-His_6 antibodies and Roti-Blue colloidal Coomassie staining (Roth). ARSK-containing fractions (fractions 6–14) were pooled, diluted 1:2 in HiTrap SP binding buffer (20 mM MES, 20 mM NaCl (pH 6.5)), and directly loaded onto a HiTrap SP column (GE Healthcare) at a flow rate of 1 ml/min using the ÄKTA Explorer purification system. The column was washed with washing buffer (20 mM MES, 20 mM NaCl (pH 6.5)). Elution was performed with a linear gradient from 20 mM to 1 M NaCl (in 50 mM MES (pH 6.5)) over 15 column volumes. Fractions were analyzed by Western blotting and Coomassie staining as above.

Enzyme Assays—Activities of ARSK toward different pseudosubstrates like pNCS or pNPS were assayed as described before (17). Absorbances were measured at 515 nm (ϵ_{515} = 12,400 M^{-1} cm^{-1}) in the case of pNCS or at 405 nm (ϵ_{405} = 18,500 M^{-1} cm^{-1}) for pNPS. All measurements were performed using the infinite M200 microplate reader (Tecan).

SDS-PAGE and Western Blot Analysis—Standard techniques were used for SDS-PAGE and Western blot analyses with PVDF membranes (Millipore). Proteins were detected by enhanced chemiluminescence detection reagent (Pierce) and quantified using the AIDA 4.06 software package (Raytest).

Endoglucosaminidase H and Peptide N-glycosidase F Treatment—40 μl of cell lysates from ARSK-expressing cells or 40 μl of HisTrap-enriched secreted ARSK were deglycosylated by treatment with peptide *N*-glycosidase F (PNGaseF, Roche) or endoglucosaminidase H (EndoH, Roche) as described before (17) and analyzed by Western blotting.

Mannose 6-phosphate Receptor (MPR) Binding Assay—Purified ARSK and purified recombinant Scpep1 (26), respectively, were incubated overnight at 4 °C with goat-MRP46 and goat-MRP300 immobilized on a 2-ml Affi-Gel 10 matrix (Bio-Rad). Washing with glucose-6-phosphate and elution with mannose 6-phosphate were performed as described before (27). The resulting fractions were analyzed by Western blotting detecting the RGS-His_6 tag present on both proteins.

ARSK Uptake and Immunofluorescence—For uptake experiments, immortalized mouse embryonic fibroblasts were grown to 70% confluency for 24 h on poly-L-lysine-coated coverslips in 24-well plates. 1 μg of ARSK-His_6 in a total volume of 200 μl of 10 mM HEPES, 0.9% NaCl (pH 7.4) were mixed with 400 μl of medium and added to the cells for 2 h. After incubation, the cells were washed with PBS, fixed with 4% paraformaldehyde in 10 mM Na_2HPO_4 (pH 7.3) containing 3% sucrose for 20 min at room temperature and washed three times with permeabilization buffer (500 mM NaCl, 10 mM Na_2HPO_4 (pH 7.3) with 0.1% Tween 20 and 0.1% Triton X-100) prior to blocking with 2% FCS for 30 min. ARSK was detected by incubation with the polyclonal rabbit anti-ARSK antibody and LAMP-1 with the monoclonal rat anti-LAMP-1 antibody (1D4B) for 1.5 h at room temperature. After washing with immunofluorescence washing buffer (500 mM NaCl, 10 mM Na_2HPO_4, 0.1% Tween 20 (pH 7.3)), primary antibodies were detected with a goat-anti-rabbit Alexa Fluor-488 and a goat anti-rat Alexa Fluor-536 antibody (Invitrogen). Images were obtained on a Leica DM5000B microscope equipped with an HCX PL APO ×100 oil immersion objective.

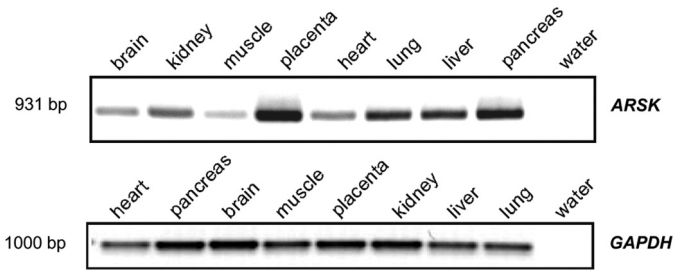


FIGURE 1. **Reverse transcription PCR analysis of *ARSK* mRNA expression in human tissues.** Normalized cDNAs from different human tissues were used to amplify a fragment of 931 bp by PCR using primers specific for human *ARSK*. Normalization was verified using primers specific for glycerol aldehyde 3-phosphate dehydrogenase (*GADPH*). A sample without cDNA was used as a negative control (water). See "Experimental Procedures" for further details.

Pulse-chase Experiments—HEK293 cells expressing ARSK and untransfected cells, respectively, were grown on 6-cm dishes to a confluency of 80%. The medium was removed, and the cells were washed two times with PBS. Starvation medium lacking methionine and cysteine with 5% dialyzed FCS was added for 1 h. Thereafter, the medium was replaced by starvation medium containing ^{35}S-labeled methionine and cysteine (PerkinElmer Life Sciences) for 1 h to achieve metabolic labeling of newly synthesized proteins (pulse). After removal of the labeling medium, the cells were incubated in normal DMEM for different time periods (chase). At the indicated chase times, the medium was removed, and cells were harvested in 500 μl of lysis buffer (0.1% Triton X-100, 1 mM EDTA, 1 mM PMSF, 5 mM iodoacetamide in 1× TBS) and stored at −20 °C. Immunoprecipitation was performed as described earlier for cathepsin D (28) with the following modifications. 10 μl of rabbit anti-ARSK was added instead of anti-cathepsin D antibody, and the pansorbin immunocomplex was extensively washed four times with 1.5 M NaCl, 0.1% Triton X-100 in 0.1× PBS. Proteins were separated by SDS-PAGE on a 15% gel. The gel was dried and analyzed by phosphorimaging.

RESULTS

Endogenous Expression of Arylsulfatase K in Human Tissues—To verify endogenous expression of human ARSK, we first analyzed its mRNA levels. We looked for tissue-specific expression by RT-PCR of normalized cDNA samples from different human tissues and found that ARSK is ubiquitously expressed (Fig. 1). High expression levels are found in placenta and pancreas, and low expression levels are found in muscle. Other tissues (lung, brain, heart, liver, and kidney) show intermediate expression levels. Because a specific signal could be found in all tissues analyzed, we conclude that ARSK is ubiquitously expressed in most, if not all, human tissues.

Expression of Recombinant Arylsulfatase K—The human ARSK-encoding cDNA was obtained by reverse transcription PCR (see "Experimental Procedures"). Its coding sequence

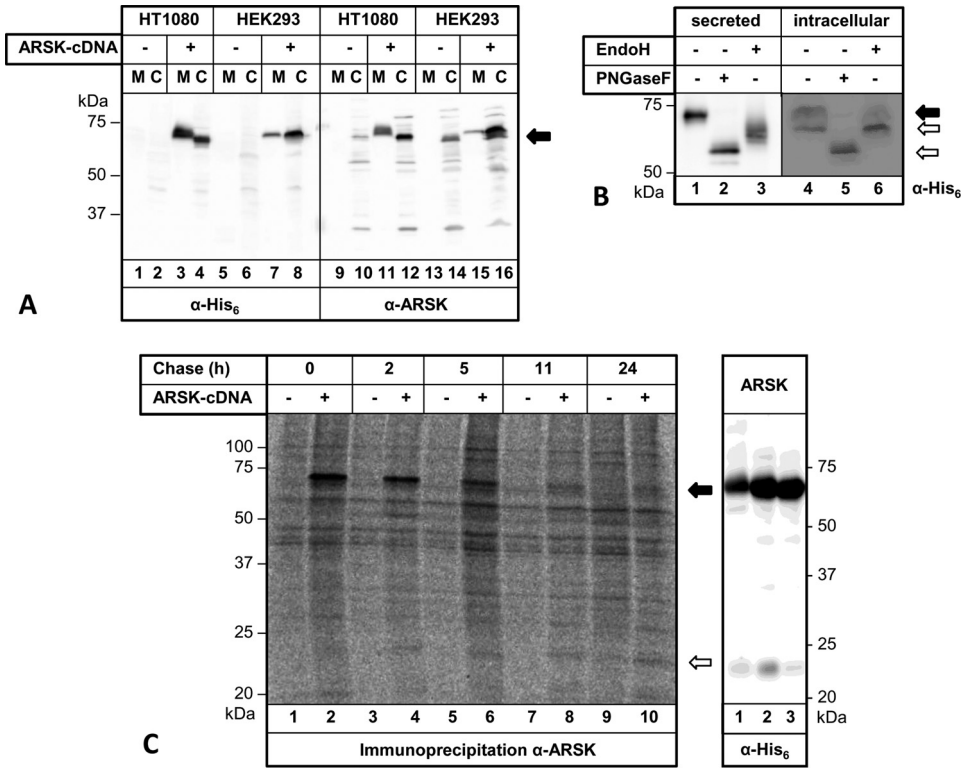


FIGURE 2. **Recombinant expression, *N*-glycosylation, and stability/processing of ARSK in human cells.** *A*, ARSK was stably expressed in HT1080 and HEK293 cells. Cell lysates (*C*) and medium (*M*) samples were analyzed for ARSK expression by Western blotting using an anti-RGS-His_6 antibody or an anti-ARSK antiserum, as indicated. Untransfected cells served as a control. The *arrow* indicates the 68-kDa form of ARSK, as detected in the cell lysates. *B*, HEK293 cells stably expressing ARSK were lysed, and the cellular protein was treated with endoglycosidases PNGaseF or EndoH, as indicated. In parallel, ARSK secreted by HEK293 cells and enriched via HisTrap chromatography was subjected to treatment with endoglycosidases. All samples were analyzed by Western blotting using the anti-RGS-His_6 antibody. The *black arrow* indicates the fully glycosylated 68-kDa form, whereas the *white arrows* indicate the partially (64-kDa) or fully deglycosylated forms (60-kDa). *C*, HEK293 cells either overexpressing ARSK or not overexpressing ARSK were metabolically labeled for 1 h with [^{35}S]methionine/cysteine and then chased for the indicated times. ARSK was immunoisolated from cell extracts using the anti-ARSK-antibody, separated by SDS-PAGE, and analyzed by autoradiography. ARSK was detected as a 68-kDa protein (*black arrow*). In addition, a 23-kDa fragment (*white arrow*) appeared during the chase, suggesting processing of the precursor (*left panel*). A corresponding C-terminal fragment was detected, albeit only weakly, by the anti-RGS-His_6 antibody when analyzing ARSK enriched from conditioned medium of producer cells by Western blotting (*right panel,* showing three elution fractions from the HisTrap column, *cf.* Fig. 3*A*).

(1608 bp) fully matched GenBank™ accession number AY358596. ARSK was stably expressed in HEK293 cells and HT1080 cells as a C-terminally RGS-His_6-tagged variant. These cells were also stably transfected with the FGE-encoding cDNA because sulfatase activity depends on posttranslational formylglycine modification. Western blot analyses of untransfected control and ARSK-expressing HEK293 and HT1080 cells using a His tag-specific antibody (Fig. 2*A*, *left panel*) as well as an ARSK-specific antibody (*right panel*) detected a protein with an apparent molecular mass of 68 kDa in transfected cells. The secreted form of ARSK present in conditioned medium from HT1080 cells exhibited a molecular mass of ~70 kDa, *i.e.* slightly higher than the cellular form (Fig. 2*A*, *lanes 3* and *11*).

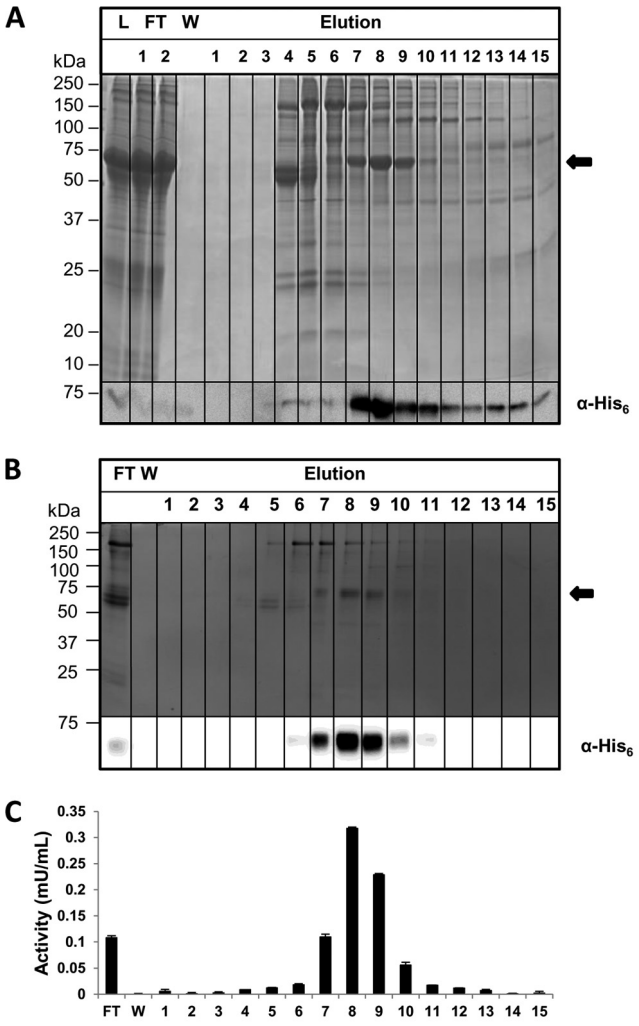
Glycosylation Pattern and Processing—Bioinformatic analysis predicts seven putative *N*-glycosylation sites with the consensus sequence N*X*S/T. To analyze the extent of glycosylation, recombinant ARSK was partially purified from HT1080 or HEK293 cells as well as from conditioned medium by chromatography on nickel-Sepharose and subjected to treatment with the endoglycosidases PNGaseF and EndoH. PNGaseF treatment resulted in a band shift from 68 kDa to 60 kDa, which corresponds to the calculated mass of the unglycosylated protein. EndoH treatment led to heterogenous products of the secreted protein from both HT1080 and HEK293 cells (Fig. 2*B*). These results indicate that ARSK from both cell lines is secreted as a multiple *N*-glycosylated protein with four to five *N*-glycans, of which some are of the high-mannose or hybrid type and some of the complex type. Intracellular ARSK is sensitive to EndoH and PNGaseF digest, leading to similar products observed for secreted ARSK with a most prominent 64-kDa product after EndoH treatment. In HEK293 cells, intracellular ARSK is detected as a double band (Fig. 2*B*, *lane 4*) of 64 kDa and 68 kDa even without EndoH treatment. The 64-kDa species is not secreted. Because full deglycosylation by PNGaseF results in a nearly homogenous product, the 64-kDa species may represent an underglycosylated form of ARSK.

Several sulfatases, in particular those residing in lysosomes, are synthesized as single-chain precursors and are proteolytically processed in the course of lysosomal transport. To analyze for processing of ARSK and to further examine its general stability, ARSK-expressing HEK293 cells were metabolically labeled with [^{35}S]methionine/[^{35}S]cysteine for 1 h and harvested after various chase periods for up to 24 h. ARSK was immunoprecipitated, separated by SDS-PAGE, and analyzed by phosphorimaging. As expected, ARSK was synthesized as a 68-kDa protein that was clearly visible in the first 5 h (Fig. 2*C*,

left panel). After 24 h, the signal dropped by 80%. This observation may reflect processing of ARSK because a specific band of 23 kDa could be immunoprecipitated with increasing chase periods (Fig. 2*C*), which corresponds to a signal detected by the anti-His_6 antibody in enriched ARSK preparations (*right panel*). Additional bands were immunoprecipitated by the antibody, which, however, could also be detected in the untransfected controls. At least one further ARSK-derived polypeptide lacking the His-tag would be expected in case of a processing event. We cannot exclude the possibility that other processed forms of ARSK failed to be immunoprecipitated and, hence, escaped detection.

Purification and Arylsulfatase Activity of ARSK—To characterize ARSK in detail, we purified the recombinant protein from the conditioned medium of stably expressing HEK293 cells, which were cultivated in medium containing 1% fetal calf serum. Medium proteins were precipitated by ammonium sulfate, dialyzed, and sequentially subjected to chromatography on nickel-Sepharose and on the strong cation exchange sulfopropyl matrix. Elution fractions from the nickel-Sepharose (Fig. 3*A*) and sulfopropyl (*B*) column were analyzed by SDS-PAGE and either Coomassie staining (*A* and *B*, *upper panels*) or Western blotting (*lower panels*). In addition, we determined arylsulfatase activity in each elution fraction (shown in Fig. 3*C* for the ion exchange chromatography) to monitor coelution of sulfatase activity with the ARSK protein band and removal of other arylsulfatases. Nickel-Sepharose chromatography resulted in partially purified ARSK with an apparent molecular mass of 68 kDa, as judged by Coomassie staining (Fig. 3*A*, *upper panel*) and Western blot analysis using the His tag antibody (*lower panel*). In the second purification step by cation exchange chromatography, ARSK eluted in fractions 7–9, as demonstrated by Coomassie staining (Fig. 3*B*, *upper panel*) and Western blot analysis (*lower panel*). Mass spectrometry peptide mass fingerprint analysis of the 68-kDa band from the Coomassie gel identified human ARSK with a Mascot score of 1907 and a sequence coverage of 54%, including N- and C-terminal regions of the mature protein after signal peptide cleavage (Fig. 3*D*). Arylsulfatase activity assays using the arylsulfate pseudosubstrate pNCS revealed arylsulfatase activity in ARSK-enriched fractions 7–10 after nickel-Sepharose chromatography (not shown) as well as in fractions 7–9 after cation exchange chromatography (Fig. 3*C*).

Purification and Characterization of the Inactive ARSK-C/A Mutant Protein—All eukaryotic sulfatases are characterized by a critical formylglycine (FGly) residue in their active site, which is generated by FGE from a conserved cysteine located in the so-called sulfatase signature sequence. In ARSK, the key motif of this signature is represented by the sequence 80-CCPSR-84, in which the first cysteine is expected to be converted to FGly. We mutated cysteine 80 to alanine to generate an enzymatically inactive form called ARSK-C/A. ARSK-C/A was also stably expressed in HEK293 cells and purified as described for the active form. As expected, ARSK-C/A showed markedly reduced activity against pNCS. The arylsulfatase activity measured in the ARSK-C/A-enriched fractions reached up to 20% of wild-type ARSK activity when measured at neutral pH. However, at its pH optimum, the specific activity of wild-type ARSK



D

```
1   MLLLWVSVVA ALALAVLAPG AGEQRRRAAK APNVVLVVSD SFDGRLTFHP
51  GSQVVKLPFI NFMKTRGTSF LNAYTNSPIC CPSRAAMWSG LFTHLTESWN
101 NFKGLDPNYT TWMDVMERHG YRTQKFGKLD YTSGHHSISN RVEAWTRDVA
151 FLLRQEGRPM VNLIRNRTKV RVMERDWQNT DKAVNWLRKE AINYTEPFVI
201 YLGLNLPHPY PSPSSGENFG SSTFHTSLYW LEKVSHDAIK IPKWSPLSEM
251 HPVDYYSSYT KNCTGRFTKK EIKNIRAFYY AMCAETDAML GEIILALHQL
301 DLLQKTIVIY SSDHGELAME HRQFYKMSMY EASAHVPLLM MGPGIKAGLQ
351 VSNVVSLVDI YPTMLDIAGI PLPQNLSGYS LLPLSSETFK NEHKVKNLHP
401 PWILSEFHGC NVNASTYMLR TNHWKYIAYS DGASILPQLF DLSSDPDELT
451 NVAVKFPEIT YSLDQKLHSI INYPKVSASV HQYNKEQFIK WKQSIGQNYS
501 NVIANLRWHQ DWQKEPRKYE NAIDQWLKTH MNPRAVRGSH HHHHH
```

FIGURE 3. **Purification, arylsulfatase activity, and identification of ARSK.** *A*, ARSK-His_6-expressing HEK293 cells were grown under 1% FCS conditions. 1.5 liter of conditioned medium, after ammonium sulfate precipitation and dialysis, was loaded onto a 1-ml HisTrap column (*L*, load). Unbound protein was collected (*FT*). After a washing step (*W*), ARSK eluted in a linear imidazole gradient (20–500 mM) mainly in fractions 7–10 (1 ml each), as detected by Coomassie staining (*arrow*) and by Western blotting using the anti-RGS-His_6 antibody (*bottom panel*). *B*, the ARSK-containing HisTrap fractions were pooled and loaded onto a 1-ml HiTrap SP column for a second purification step. ARSK was mainly eluted in fractions 7–9 of the applied NaCl gradient (20–1000 mM). The 68-kDa band detected by Coomassie staining upon SDS-PAGE analysis of these fractions (*arrow*) corresponded to the Western blot signal (*bottom panel*). MALDI mass fingerprint analysis of the Coomassie-stained band verified that the 68-kDa band consisted of ARSK (*D*). *C*, arylsulfatase activity of the indicated fractions from HiTrap SP chromatography (*B*) was measured at pH 4.6 using 10 mM pNCS as substrate. Activity was detected only in those fractions containing ARSK. *D*, the sequence of the ARSK precursor protein is shown with its N-terminal signal peptide (in *italics*), removed in mature ARSK, and the C-terminal RGS-His_6 tag. The sequence of the 22 tryptic peptides identified by MALDI mass fingerprint analysis of the 68-kDa band (*B*) is shown with *shading* (54% coverage of the mature sequence, Mascot score 1907). Predicted *N*-glycosylation sites are *underlined*, and the peptide carrying the FGly modification (at cysteine 80) is *boxed*.

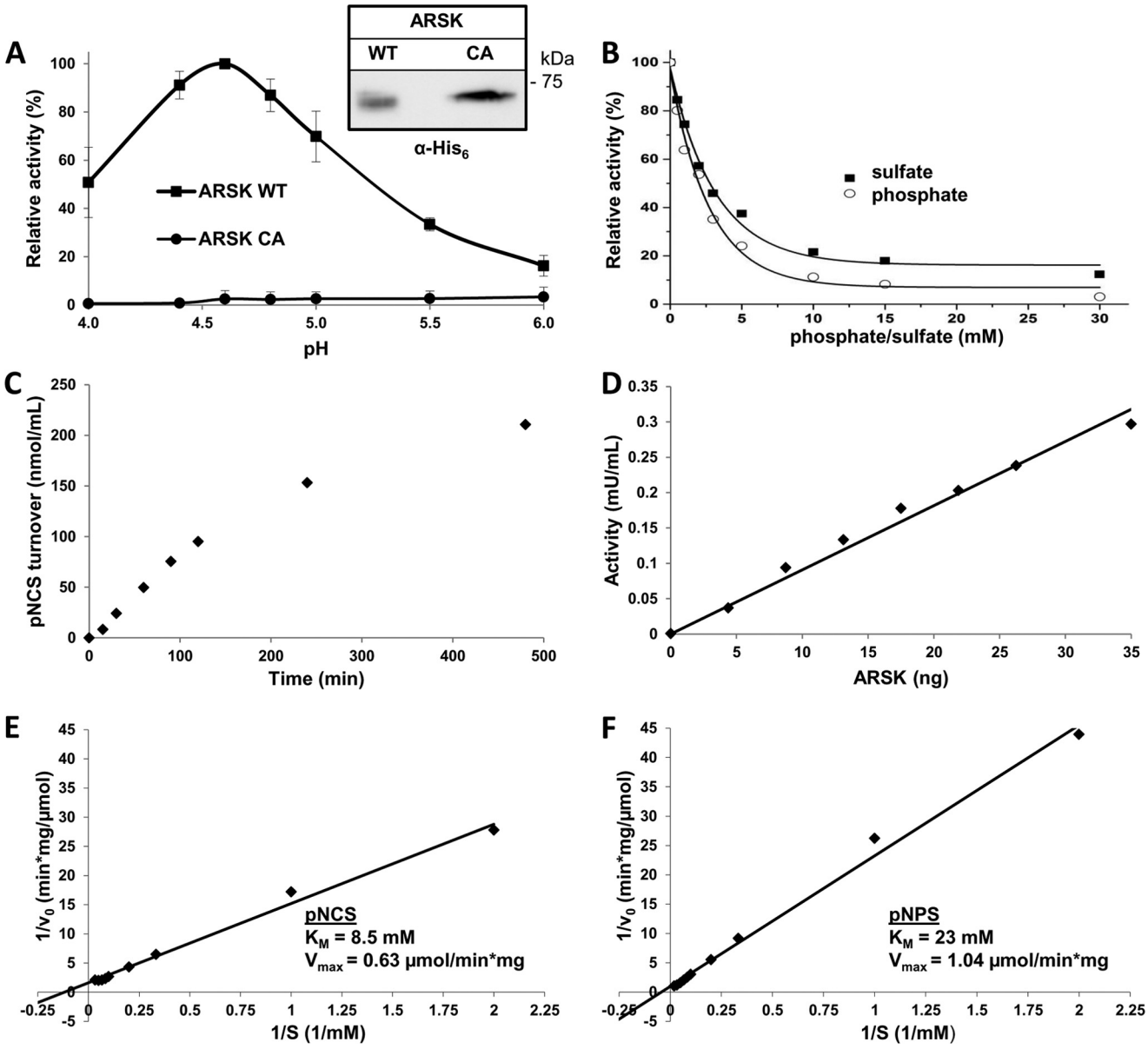


FIGURE 4. **Kinetic analysis of ARSK.** *A*, to determine the pH optimum of enzymatic activity, purified ARSK (Fig. 3*B*) was incubated for 3 h at 37 °C with 10 mM pNCS at various pH values between 4 and 6, as indicated. Similar amounts of the inactive ARSK-C/A (*CA*) mutant, purified under the same conditions (see Western blot analysis in the *inset*) were assayed in parallel. Mean values of two independent experiments ± S.D. are shown. *B*, ARSK activity was inhibited by sulfate and phosphate, as tested in the concentration range from 0.5–30 mM (at 10 mM pNCS). In two independent experiments, IC_{50} values of 2.9 ± 0.2 mM (sulfate) and 2.4 ± 0.2 mM (phosphate) were determined. *C*, the time dependence of pNCS turnover by the same ARSK preparation (35 ng) was measured for up to 8 h at 37 °C and pH 4.6. *D*, for measuring the dose dependence, different amounts (0–35 ng) of ARSK were incubated with 10 mM pNCS for 4 h at 37 °C and pH 4.6. *E* and *F*, the dependence of pNCS and pNPS turnover by 20–30 ng of ARSK on the substrate concentration was analyzed at pH 4.6 and 37 °C. The results were transformed into double-reciprocal Lineweaver-Burk plots using data points from 0.5–30 mM pNCS (*E*) and 0.5–50 mM pNPS (*F*). The kinetic constants extrapolated from these plots are given in the figure.

was ~20-fold higher as compared with ARSK-C/A (Fig. 4*A*). In fact, the background activity in the ARSK-C/A preparation was at the detection limit and, most probably, because of other contaminating sulfatases.

Characterization of ARSK Arylsulfatase Activity—Next we analyzed the enzymatic properties of ARSK and its activity toward arylsulfate pseudosubstrates. To discriminate ARSK-associated sulfatase activity from that of potentially copurified sulfatases, we measured enzymatic activity of ARSK in comparison with ARSK-C/A prepared according to the same purification protocol (see above). ARSK cleaved the small aromatic pseudosubstrates pNCS and pNPS (Fig. 4) but not the commonly used pseudosubstrate 4-methylumbelliferyl sulfate (not shown). The apparent pH optimum for ARSK was found to be at an acidic pH of about 4.6 for the pseudosubstrates pNCS (Fig. 4*A*) and pNPS (not shown), thus strongly suggesting a lysosomal localization of ARSK. Under the applied assay conditions (pH 4.6, 37 °C, 10 mM pNCS, 35 ng ARSK), substrate turnover was linear with time for about 120 min (Fig. 4*C*). Calculated activities (initial velocities) showed a direct correlation to the amount of ARSK present in the assay (Fig. 4*D*). Similar to other sulfatases, ARSK activity was inhibited by the presence of the reaction product sulfate or its analog phosphate (17, 29). For ARSK, a moderate sensitivity with

IC_{50} values of 2.9 ± 0.2 mM (sulfate) and 2.4 ± 0.2 mM (phosphate) was observed (Fig. 4*B*).

Substrate saturation curves for pNCS and pNPS were determined at the pH optimum using 20–30 ng of enzyme/assay. ARSK showed hyperbolic substrate dependence with saturation observed at 15–20 mM for pNCS and 30–40 mM for pNPS (not shown). K_m and V_{max} values were determined using Lineweaver-Burk plots. From two independent experiments, we calculated a K_m of 10.9 ± 3.3 mM for pNCS and 20.6 ± 3.6 mM for pNPS (Fig. 4, *E* and *F*, one of the two experiments shown). The maximum specific activity V_{max} was very similar for both substrates, pNCS (0.84 ± 0.29 units/mg, Fig. 4*E*) and pNPS (0.93 ± 0.16 units/mg, *F*). In comparison to most other arylsulfatases, these values are much lower than the typically observed activities of 5–100 units/mg. Instead, they are similar to the rates of those six sulfatases to which the arylsulfatase nomenclature has not been applied (3).

It should be noted that a relatively low degree of FGly modification of ARSK contributes to the low specific activity determined. FGly quantification was performed by nanoLC MALDI-MS analysis of tryptic peptides obtained by in-gel digestion of ARSK. Both the Cys-80 and the FGly-80 versions of the sulfatase signature tryptic peptide GTSFLNAYTNSPI**C-CPSR** could be clearly detected (m/z = 1969.9 and 2044.9, respectively, after carbamidomethylation). The FGly content of ARSK, however, was ~3-fold lower than that of arylsulfatase A, which we have shown to be FGly-modified by ≥ 90% (30) and which served as a control in this FGly analysis of ARSK. Of note, FGly quantification in case of ARSK was impeded by the fact that the two neighboring cysteines in the relevant peptide led to heterogenous carbamidomethylation products (data not shown). Taken together, these data suggest that ARSK is a lysosomal sulfatase with low activity and low to moderate affinity toward pseudosubstrates that, in the case of other lysosomal sulfatases, was found to correspond to a high specificity toward their natural substrates (see "Discussion").

Subcellular Localization of ARSK—The acidic pH optimum suggested a lysosomal localization of ARSK. Most soluble lysosomal enzymes are transported toward the lysosome by the mannose 6-phosphate receptors MPR46 and MPR300, which recognize an M6P-containing *N*-glycan. ARSK from conditioned medium of stably expressing HEK293 cells was partially purified by nickel-Sepharose chromatography and loaded onto a column with immobilized MPR46 and MPR300. After removal of unspecifically bound proteins with 5 mM glucose 6-phosphate, specifically bound proteins were eluted with 5 mM mannose 6-phosphate, and the fractions were analyzed by immunoblotting (Fig. 5*A*, *upper panel*). The Western blot revealed that ~70% of loaded ARSK was recovered in the mannose 6-phosphate elution fractions. As a control, recombinantly expressed murine Scpep1, another lysosomal protein (26), was analyzed on this MPR affinity column. Scpep1 bound and eluted with similar efficiency (about 60%, Fig. 5*A*, *lower panel*). In addition, the presence of M6P residues in ARSK-His_6 was confirmed on a Western blot probed with a M6P-specific antibody (25). A clear signal, even stronger than for the positive control Scpep1-His_6, was detected, whereas for the negative control FGE-His_6, only the His_6 tag but no M6P could be recognized (Fig. 5*B*).

To further verify the lysosomal localization of ARSK, we performed indirect immunofluorescence studies using stably or transiently ARSK-expressing HT1080 cells. Because of overexpression, a staining of the ER was predominant, suggesting misfolding and improper sorting (not shown). To overcome this problem, we exploited the MPR/M6P-dependent uptake and subsequent transport of many lysosomal enzymes toward the lysosomes. After incubating mouse embryonic fibroblasts for 2 h with medium to which partially purified ARSK-His_6 (~1 μg) was added, the cells were analyzed by indirect immunofluorescence using the ARSK-specific antiserum. The internalized ARSK was detectable in vesicular structures that were also positive for the commonly used lysosomal marker protein LAMP1 (Fig. 5*C*). In summary, these results indicate that ARSK is a soluble lysosomal protein that is transported to the lysosome in a MPR-dependent manner.

DISCUSSION

In 2005, four novel putative sulfatases (termed arylsulfatase H, I, J, and K) were identified bioinformatically in humans by a genome-wide screen using the sulfatase-specific signature sequence (2). Arylsulfatase I and arylsulfatase J can be considered paralogs of arylsulfatase B because of their high sequence identity (45% at the protein level). In contrast, arylsulfatase K shows low sequence identity (18–22%) with other known sulfatases (2). Despite this divergence from other sulfatases, ARSK itself is quite strongly conserved, *e.g.* human ARSK shows 76% sequence identity to chicken, 62% to zebrafish, 54% to amphioxus, and 52% to acorn worm. This conservation strengthens the prediction that ARSK has an important and conserved function. Here we demonstrate that human ARSK is a ubiquitously expressed glycoprotein that resides in the lysosome and cleaves artificial arylsulfate pseudosubstrates.

ARSK was stably expressed in human cell lines as a His-tagged derivative and exhibited an apparent molecular mass of ~68 kDa in its intracellular form and a slightly higher molecular mass of ~70 kDa when secreted into medium. Deglycosylation assays using endoglycosidases PNGaseF and EndoH clearly demonstrated that both intracellular and extracellular ARSK carry multiple complex-type as well as mannose-rich-type asparagine-linked glycans. The reduction in size of ~10 kDa after PNGaseF treatment suggests occupation of four to five of the seven predicted *N*-glycosylation sites. This agrees with our mass spectrometric analysis detecting two of the predicted glycopeptides in unglycosylated form (Fig. 3*D*).

ARSK was purified as a secreted enzyme, *i.e.* after passing intracellular quality control. Arylsulfatase activity measured in this preparation was due to recombinant ARSK because activity correlated with purified ARSK protein, as detected by mass fingerprint analysis and quantified by Western blotting or Coomassie staining. Moreover, activity was dependent on FGly modification of ARSK because the ARSK-C/A mutant, purified in parallel under identical conditions, showed no significant activity. Kinetic analysis of ARSK revealed a relatively low affinity toward artificial arylsubstrates as well as a low specific turnover of these pseudosubstrates. Similar enzymatic properties as

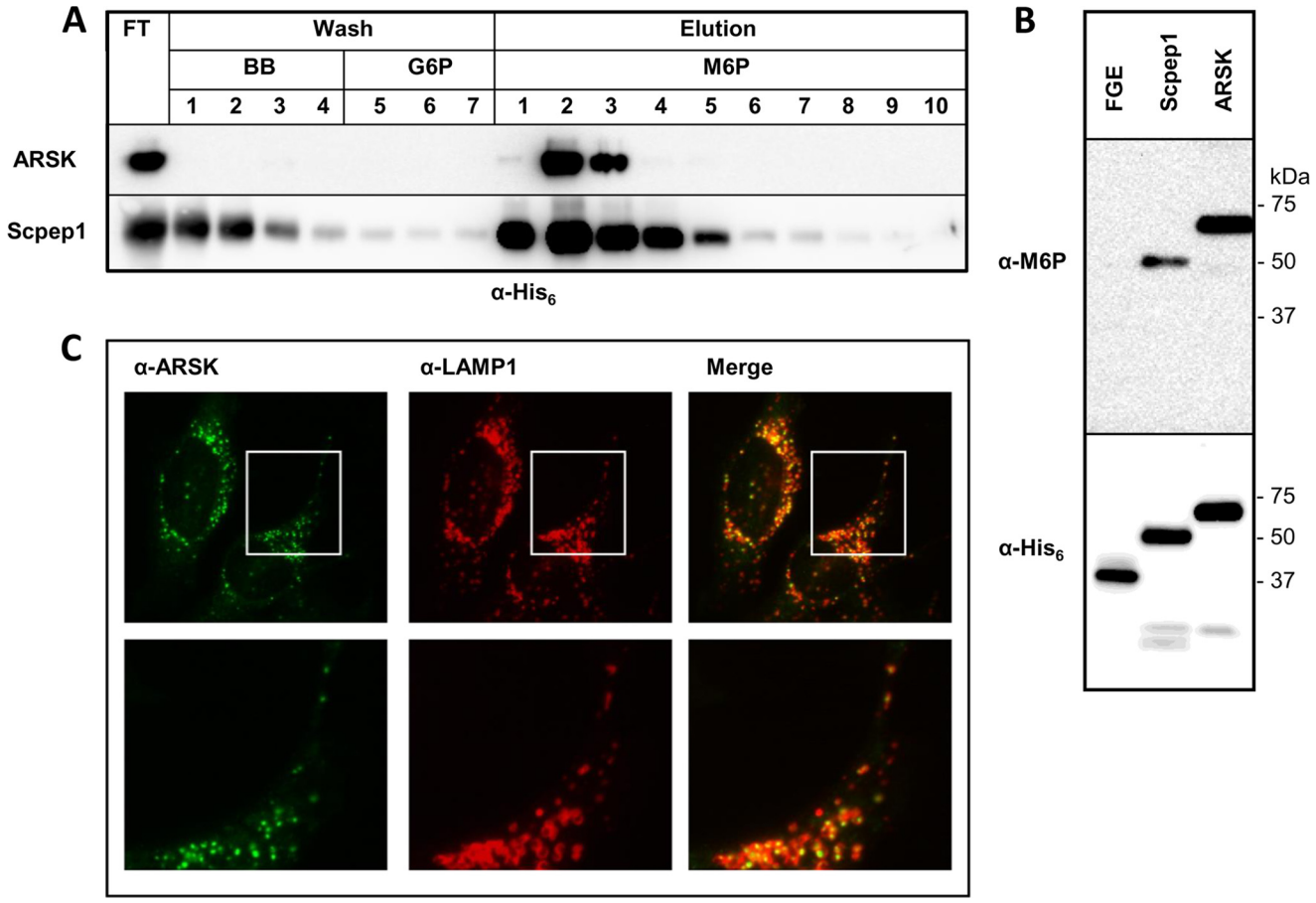


FIGURE 5. **Subcellular localization of ARSK and binding to an MPR affinity column.** *A*, HisTrap-purified ARSK (1 μg) was loaded on a matrix with immobilized MPRs and incubated overnight. After collecting the flow-through (*FT*), the column matrix was washed four times with binding buffer (*BB*) (fractions W1-W4) and three times with 5 mM glucose 6-phosphate (*G6P*) (fractions W5-W7). Bound ARSK was eluted with 5 mM M6P in 10 fractions (E1-E10). All fractions were analyzed by Western blotting using the anti-RGS-His_6 antibody (*upper panel*). The *lower panel* shows the results obtained for the established lysosomal protein Scpep1, purified as well via its RGS-His_6-tag, which was subjected to the same MPR affinity chromatography protocol. *B*, ARSK, enriched by HisTrap chromatography (Fig. 3*A*), as well as purified recombinant mouse Scpep1 (100 ng) (26) and purified recombinant FGE (40 ng) (24), both produced by HT1080 cells, were analyzed by Western blotting using the scFv M6P single-chain antibody fragment (*upper panel*) and the anti-RGS-His_6 antibody (*lower panel*), respectively. All three proteins carried the same RGS-His_6 tag. *C*, immortalized mouse embryonic fibroblasts were grown for 24 h on coverslips to 70% confluence. Then, 1 μg ARSK-His_6 was added to the cells and incubated for 2 h prior to fixation and detection of ARSK with a polyclonal ARSK antibody and detection of LAMP1 with a monoclonal LAMP1 antibody. Detection of ARSK (*green*) is shown on the *left*, detection of LAMP1 (*red*) is shown in the center, and the merged signals are shown on the *right*. The *boxed areas* are shown below at higher magnification.

reported here for ARSK, *i.e.* affinities for arylsulfates in the millimolar range (K_m 4–12 mM) and specific activities $\leq$ 1 units/mg, have been described for four other lysosomal sulfatases that show high specificity and affinity toward their natural substrates, namely iduronate 2-sulfatase, glucosamine 6-sulfatase, galactosamine 6-sulfatase, and sulfamidase (an overview is given in Ref. 3). These four sulfatases catalyze the removal of specific sulfate moieties from the sulfated glycosaminoglycans heparan, chondroitin/dermatan, or keratan sulfate, suggesting that ARSK also acts during the lysosomal degradation of sulfated glycosaminoglycans. Possible substrates include the 2-*O*-sulfate groups of glucuronic acids and the more rare 3-*O*-sulfate groups of glucosamine (in its free amine form), for which no desulfating enzyme has been identified so far.

Using the pseudosubstrates, we determined an apparent pH optimum of 4.6 for ARSK activity, which strongly suggested a lysosomal localization. This was confirmed by immunofluorescence studies demonstrating colocalization of ARSK with the lysosomal integral membrane protein LAMP-1 upon uptake of partially purified ARSK supplemented to the cell culture medium. Most lysosomal hydrolases are sorted toward the lysosome by the M6P receptor system (31), which also mediates uptake of mistargeted M6P-containing proteins from the extracellular space. Accordingly, ARSK was shown to bind efficiently to immobilized MPR in an M6P-dependent manner, and, moreover, a strong M6P-signal was detected for ARSK in Western blot analyses using a M6P-specific antibody. Taken together, these findings demonstrate a lysosomal localization of ARSK. Interestingly, and in line with our observations, ARSK had already been identified previously in studies of the lysosomal subproteome when analyzing the mannose 6-phosphate glycoproteomes from humans, mouse, and rat (32–34 and reviewed in Ref. 23). In their study, Sleat *et al.* (34) pinpointed the M6P site to asparagines Asn-498 and Asn-499 in human and mouse ARSK, respectively.

Lysosomal hydrolases are often synthesized as inactive precursors that undergo limited proteolysis during maturation into their active lysosomal forms (35), as applies also to several sulfatases, *e.g.* arylsulfatase B (*N*-acetylgalactosamine-4-sulfatase) (36, 37). In the case of ARSK, we obtained evidence for

processing of the 68-kDa precursor during 24-h pulse-chase experiments because a stable 23-kDa fragment could be immunoprecipitated by anti-ARSK antibodies from a 2-h chase onwards. An equivalent His-tagged, *i.e.* C-terminal, ARSK-derived 23-kDa fragment could be detected in Western blot analyses of ARSK enriched from conditioned medium of producer cells. Corresponding N-terminal fragment(s) could not be detected. They might have escaped our analyses on the basis of antibody recognition because of incompatible epitopes after processing. Further studies on this issue will require expression of larger amounts of ARSK and/or availability of other ARSK-specific antibodies.

ARSK is expressed in all tissues examined in this study and was also identified in eight tissues from rat in M6P glycoproteome analyses (33). Its ubiquitous expression pattern may suggest a common and widespread sulfated substrate and indicates that ARSK deficiency probably leads to a lysosomal storage disorder, as shown for all other lysosomal sulfatases. Currently, we are generating an ARSK-deficient mouse model that should pave the way to identify the physiological substrate of this sulfatase and its overall pathophysiological relevance. Finally, the mouse model could enable us to draw conclusions on ARSK-deficient human patients who so far escaped diagnosis and might be accessible for enzyme replacement therapy. The presence of M6P on ARSK qualifies this sulfatase for such a therapy, which has proven useful for treatment of numerous other lysosomal storage disorders.

Acknowledgments—We thank Bernhard Schmidt and Olaf Bernhard for mass spectrometry; Nicole Tasch, Annegret Schneemann, Britta Dreier, Martina Balleininger (all from Göttingen), William C. Lamanna, Jaqueline Alonso Lunar, Kerstin Böker, and Claudia Prange for technical assistance; Markus Damme for initial analysis of subcellular localization; and Jeffrey Esko (San Diego) for critically reading the manuscript. We also thank Kurt von Figura for support during the initial phase of this project.

REFERENCES

1. Hopwood, J. J., Ballabio A. (2001) in *The Molecular and Metabolic Bases of Inherited Diseases* (Scriver, C., Beaudet, A.L., Sly, W.S., and Valle D., eds), 8th Ed., pp 3725–3732, McGraw-Hill, New York
2. Sardiello, M., Annunziata, I., Roma, G., and Ballabio, A. (2005) Sulfatases and sulfatase modifying factors. An exclusive and promiscuous relationship. *Hum. Mol. Genet.* **14,** 3203–3217
3. Hanson, S. R., Best, M. D., and Wong, C. H. (2004) Sulfatases. Structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew. Chem. Int. Ed. Engl.* **43,** 5736–5763
4. Diez-Roux, G., and Ballabio, A. (2005) Sulfatases and human disease. *Annu. Rev. Genomics Hum. Genet.* **6,** 355–379
5. Dhoot, G. K., Gustafsson, M. K., Ai, X., Sun, W., Standiford, D. M., and Emerson, C. P., Jr. (2001) Regulation of Wnt signaling and embryo patterning by an extracellular sulfatase. *Science* **293,** 1663–1666
6. Lamanna, W. C., Baldwin, R. J., Padva, M., Kalus, I., Ten Dam, G., van Kuppevelt, T. H., Gallagher, J. T., von Figura, K., Dierks, T., and Merry, C. L. (2006) Heparan sulfate 6-O-endosulfatases. Discrete *in vivo* activities and functional co-operativity. *Biochem. J.* **400,** 63–73
7. Morimoto-Tomita, M., Uchimura, K., Werb, Z., Hemmerich, S., and Rosen, S. D. (2002) Cloning and characterization of two extracellular heparin-degrading endosulfatases in mice and humans. *J. Biol. Chem.* **277,** 49175–49185
8. Lamanna, W. C., Kalus, I., Padva, M., Baldwin, R. J., Merry, C. L., and Dierks, T. (2007) The heparanome. The enigma of encoding and decoding heparan sulfate sulfation. *J. Biotechnol.* **129,** 290–307
9. Schmidt, B., Selmer, T., Ingendoh, A., and von Figura, K. (1995) A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. *Cell* **82,** 271–278
10. von Bülow, R., Schmidt, B., Dierks, T., von Figura, K., and Usón, I. (2001) Crystal structure of an enzyme-substrate complex provides insight into the interaction between human arylsulfatase A and its substrates during catalysis. *J. Mol. Biol.* **305,** 269–277
11. Dierks, T., Lecca, M. R., Schlotterhose, P., Schmidt, B., and von Figura, K. (1999) Sequence determinants directing conversion of cysteine to formylglycine in eukaryotic sulfatases. *EMBO J.* **18,** 2084–2091
12. Dierks, T., Schmidt, B., and von Figura, K. (1997) Conversion of cysteine to formylglycine. A protein modification in the endoplasmic reticulum. *Proc. Natl. Acad. Sci. U.S.A.* **94,** 11963–11968
13. Dierks, T., Dickmanns, A., Preusser-Kunze, A., Schmidt, B., Mariappan, M., von Figura, K., Ficner, R., and Rudolph, M. G. (2005) Molecular basis for multiple sulfatase deficiency and mechanism for formylglycine generation of the human formylglycine-generating enzyme. *Cell* **121,** 541–552
14. Dierks, T., Schmidt, B., Borissenko, L. V., Peng, J., Preusser, A., Mariappan, M., and von Figura, K. (2003) Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(α)-formylglycine generating enzyme. *Cell* **113,** 435–444
15. Dierks, T., Schlotawa, L., Frese, M. A., Radhakrishnan, K., von Figura, K., and Schmidt, B. (2009) Molecular basis of multiple sulfatase deficiency, mucolipidosis II/III and Niemann-Pick C1 disease. Lysosomal storage disorders caused by defects of non-lysosomal proteins. *Biochim. Biophys. Acta* **1793,** 710–725
16. Cosma, M. P., Pepe, S., Annunziata, I., Newbold, R. F., Grompe, M., Parenti, G., and Ballabio, A. (2003) The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases. *Cell* **113,** 445–456
17. Frese, M. A., Schulz, S., and Dierks, T. (2008) Arylsulfatase G, a novel lysosomal sulfatase. *J. Biol. Chem.* **283,** 11388–11395
18. Koppe, G., Marinković-Ilsen, A., Rijken, Y., De Groot, W. P., and Jöbsis, A. C. (1978) X-linked ichthyosis. A sulphatase deficiency. *Arch. Dis. Child.* **53,** 803–806
19. 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
20. Ballabio, A., and Gieselmann, V. (2009) Lysosomal disorders. From storage to cellular damage. *Biochim. Biophys. Acta* **1793,** 684–696
21. Kowalewski, B., Lamanna, W. C., Lawrence, R., Damme, M., Stroobants, S., Padva, M., Kalus, I., Frese, M. A., Lübke, T., Lüllmann-Rauch, R., D'Hooge, R., Esko, J. D., and Dierks, T. (2012) Arylsulfatase G inactivation causes loss of heparan sulfate 3-O-sulfatase activity and mucopolysaccharidosis in mice. *Proc. Natl. Acad. Sci. U.S.A.* **109,** 10310–10315
22. Obaya, A. J. (2006) Molecular cloning and initial characterization of three novel human sulfatases. *Gene* **372,** 110–117
23. Lübke, T., Lobel, P., and Sleat, D. E. (2009) Proteomics of the lysosome. *Biochim. Biophys. Acta* **1793,** 625–635
24. Preusser-Kunze, A., Mariappan, M., Schmidt, B., Gande, S. L., Mutenda, K., Wenzel, D., von Figura, K., and Dierks, T. (2005) Molecular characterization of the human Cα-formylglycine-generating enzyme. *J. Biol. Chem.* **280,** 14900–14910
25. Müller-Loennies, S., Galliciotti, G., Kollmann, K., Glatzel, M., and Braulke, T. (2010) A novel single-chain antibody fragment for detection of mannose 6-phosphate-containing proteins. Application in mucolipidosis type II patients and mice. *Am. J. Pathol.* **177,** 240–247
26. Kollmann, K., Damme, M., Deuschl, F., Kahle, J., D'Hooge, R., Lüllmann-Rauch, R., and Lübke, T. (2009) Molecular characterization and gene disruption of mouse lysosomal putative serine carboxypeptidase 1. *FEBS J.* **276,** 1356–1369
27. Kollmann, K., Mutenda, K. E., Balleininger, M., Eckermann, E., von Figura, K., Schmidt, B., and Lübke, T. (2005) Identification of novel lysosomal matrix proteins by proteome analysis. *Proteomics* **5,** 3966–3978

28. Gieselmann, V., Pohlmann, R., Hasilik, A., and Von Figura, K. (1983) Biosynthesis and transport of cathepsin D in cultured human fibroblasts. *J. Cell Biol.* **97,** 1–5
29. Chruszcz, M., Laidler, P., Monkiewicz, M., Ortlund, E., Lebioda, L., and Lewinski, K. (2003) Crystal structure of a covalent intermediate of endogenous human arylsulfatase A. *J. Inorg. Biochem.* **96,** 386–392
30. Waldow, A., Schmidt, B., Dierks, T., von Bülow, R., and von Figura, K. (1999) Amino acid residues forming the active site of arylsulfatase A. Role in catalytic activity and substrate binding. *J. Biol. Chem.* **274,** 12284–12288
31. Braulke, T., and Bonifacino, J. S. (2009) Sorting of lysosomal proteins. *Biochim. Biophys. Acta* **1793,** 605–614
32. Qian, M., Sleat, D. E., Zheng, H., Moore, D., and Lobel, P. (2008) Proteomics analysis of serum from mutant mice reveals lysosomal proteins selectively transported by each of the two mannose 6-phosphate receptors. *Mol. Cell Proteomics* **7,** 58–70
33. Sleat, D. E., Della Valle, M. C., Zheng, H., Moore, D. F., and Lobel, P. (2008) The mannose 6-phosphate glycoprotein proteome. *J. Proteome Res.* **7,** 3010–3021
34. Sleat, D. E., Zheng, H., Qian, M., and Lobel, P. (2006) Identification of sites of mannose 6-phosphorylation on lysosomal proteins. *Mol. Cell Proteomics* **5,** 686–701
35. Hasilik, A. (1992) The early and late processing of lysosomal enzymes. Proteolysis and compartmentation. *Experientia* **48,** 130–151
36. Kobayashi, T., Honke, K., Jin, T., Gasa, S., Miyazaki, T., and Makita, A. (1992) Components and proteolytic processing sites of arylsulfatase B from human placenta. *Biochim. Biophys. Acta* **1159,** 243–247
37. 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,** 277–289