THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 280, No. 15, Issue of April 15, pp. 15173–15179, 2005
Printed in U.S.A.

Expression, Localization, Structural, and Functional Characterization of pFGE, the Paralog of the Cα-Formylglycine-generating Enzyme*

Received for publication, December 6, 2004, and in revised form, February 10, 2005
Published, JBC Papers in Press, February 11, 2005, DOI 10.1074/jbc.M413698200

Malaiyalam Mariappan, Andrea Preusser-Kunze, Martina Balleininger, Nicole Eiselt, Bernhard Schmidt, Santosh Lakshmi Gande, Dirk Wenzel‡, Thomas Dierks§, and Kurt von Figura¶

From the Institut für Biochemie und Molekulare Zellbiologie, Abteilung Biochemie II, Universität Göttingen, Heinrich-Düker-Weg 12, 37073 Göttingen and the ‡Abteilung Neurobiologie, Max-Planck-Institut für Biophysikalische Chemie, Am Fassberg, 37077 Göttingen, Germany

pFGE is the paralog of the formylglycine-generating enzyme (FGE), which catalyzes the oxidation of a specific cysteine to Cα-formylglycine, the catalytic residue in the active site of sulfatases. The enzymatic activity of sulfatases depends on this posttranslational modification, and the genetic defect of FGE causes multiple sulfatase deficiency. The structural and functional properties of pFGE were analyzed. The comparison with FGE demonstrates that both share a tissue-specific expression pattern and the localization in the lumen of the endoplasmic reticulum. Both are retained in the endoplasmic reticulum by a saturable mechanism. Limited proteolytic cleavage at similar sites indicates that both also share a similar three-dimensional structure. pFGE, however, is lacking the formylglycine-generating activity of FGE. Although overexpression of FGE stimulates the generation of catalytically active sulfatases, overexpression of pFGE has an inhibitory effect. *In vitro* pFGE interacts with sulfatase-derived peptides but not with FGE. The inhibitory effect of pFGE on the generation of active sulfatases may therefore be caused by a competition of pFGE and FGE for newly synthesized sulfatase polypeptides.

The catalytic site of pro- and eukaryotic sulfatases contains a unique amino acid, Cα-formylglycine (FGly)[1] (1–9). The FGly participates as an aldehyde hydrate in the catalytic cleavage of sulfate esters and thus is essential for the activity of sulfatases (10–12). The FGly residue is posttranslationally generated from a specific cysteine (pro- and eukaryotes) or serine (prokaryotes) residue. In humans, the FGly-generating enzyme (FGE) is encoded by the *SUMF1* gene (13, 14). Mutations in the *SUMF1* gene are the cause of multiple sulfatase deficiency, a rare autosomal recessive disorder, in which all sulfatases are synthesized as catalytically inactive polypeptides lacking FGly in their catalytic sites (1, 2, 13–15). The *SUMF1* gene is highly conserved among pro- and eukaryotes (16).

In the genomes of deuterostomia, including vertebrates and echinodermata, a paralog of the *SUMF1* gene, designated as *SUMF2*, is found (13–15).[2] Gene duplication of a common *SUMF* ancestor gene obviously has occurred at the level of a single exon gene after the evolution of insects and before that of deuterostomia. The function of the *SUMF2*-encoded paralog of FGE (pFGE) is unknown. Cotransfection of *SUMF2* with sulfatase cDNAs such as those of arylsulfatase A, C, or E enhanced moderately, but consistently, the catalytic activity of the sulfatases, suggesting that pFGE has some FGly-generating activity (13). So far, no mutations in the *SUMF2* gene have been found in multiple sulfatase deficiency patients. In the present study, we have analyzed the expression of *SUMF2* and the structural and functional properties of recombinant and endogenous pFGE and examined the interaction with and modifying activity of pFGE on sulfatases.

EXPERIMENTAL PROCEDURES

Cloning of Human pFGE Encoding cDNA (SUMF2)—Total RNA, prepared from human fibroblasts using the RNeasy Mini kit (Qiagen), was reverse-transcribed using the Omniscript reverse transcriptase kit (Qiagen), and either an oligo(dT) primer or the *SUMF2*-specific primer 1072nc (5′-GGAAGGGGAAGTTCTTTCC-3′). The first strand cDNA was amplified by PCR using the forward primer 1c (5′-ATGCG-CAGCGGGGCCGTGG-3′) and, as reverse primer, either 1072nc or 1049nc (5′-GCTGAAGCTCTCGAGCTTGG-3′). The PCR products were cloned directly into the pCR4-TOPO vector (Invitrogen). By sequencing multiples of the cloned PCR products, which had been obtained from various individuals and from independent reverse transcriptase and PCR reactions, the cDNA sequence coding for pFGE (906 bp from start to stop codon) was established. The sequence corresponds to GenBank™ entry NM_015411 but deviates in one position (87C>T). This change is a silent mutation in the third position of the serine 29 codon and may reflect a polymorphism.

Construction of Expression Plasmids—The pFGE encoding cDNA was equipped with a 5′ BamHI site and a 3′ RGS-His_6 tag sequence followed by a stop codon and a XbaI site, by add-on PCR using *Pfu* Ultra polymerase (Stratagene) and primers 5′-CGGGATCCACCATGGCCCG-GCATGGGTTAC-3′ (BamHI) and 5′-GCTCTAGATTAGTGATGGT-GATGGTGATGCGATCCTCTCAGCTCCCCTGGCGGC-3′ (RGS-His_6-XbaI). The resulting PCR product was cloned as a BamHI/XbaI fragment downstream of the cytomegalovirus promotor of the expression vector pSB4.7pA, which was provided by Transkaryotic Therapies Inc. (Cambridge, MA).

For coexpression of FGE or pFGE with steroid sulfatase, the pBI vector (BD Biosciences) was used, which allows expression of two cDNAs from a bidirectional tet-responsive promotor. The FGE- (14) and

* This work was supported by the Deutsche Forschungsgemeinschaft, the Fonds der Chemischen Industrie and Transkaryotic Therapies, Inc. (Cambridge, MA). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

§ Present address: Biochemie I, Universität Bielefeld, Universitätstrasse 25, 33615 Bielefeld, Germany.

¶ To whom correspondence is addressed. Tel.: 49-551-395948; Fax: 49-551-395979; E-mail: kfigura@gwdg.de.

[1] The abbreviations used are: FGly, C_α-formylglycine; ASA, arylsulfatase A; Bpa, *p*-benzoyl phenylalanine; FGE, FGly-generating enzyme; pFGE, paralog of FGE; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; Ni-NTA, nickel-nitrilotriacetic acid; HA, hemagglutinin; PNGase, peptide-*N*-glycosidase.

[2] T. Pringle, personal communication.

pFGE-encoding cDNAs were equipped with a 5′ EcoRV site and a 3′ HA tag sequence followed by a stop codon and a NotI site, by add-on PCR using *Pfu* Ultra polymerase (Stratagene) and the following primers: 5′-GGATATCGGGACAACATGGCTGCG-3′ (EcoRV-FGE); 5′-GGATATCACCATGGCCCGGCATGGGTTAC-3′ (EcoRV-pFGE); 5′-ATAGTTTAGCGGCCGCTTATGCGTAGTCAGGCACATCATACGGATAGTCCATGGTGGGCAGGC-3′ (FGE-HA-NotI); 5′-ATAGTTTAGCGGCCGCTTATGCGTAGTCAGGCACATCATACGGATACAGCTCCCCTGGCGGC-3′ (pFGE-HA-NotI).

The EcoRV/NotI-digested PCR products were inserted into the multiple cloning site II of the pBI vector that had been opened by PstI, blunted with T4-polymerase/dNTPs, and cut with NotI. It should be noted that after ligation, the initial EcoRV and PstI sites were destroyed. Into the multiple cloning site I of the obtained constructs and of the empty pBI vector, the steroid sulfatase cDNA was cloned as an NheI/EcoRV fragment. This fragment was generated from a PCR product that was obtained using pBEH-STS (17) as template, *Pfu* Ultra polymerase, and add-on primers 5′-CTAGCTAGCCACCATGCCTTTAAGGAAGATGAAG-3′ (NheI) and 5′-GGATATCAGCGGCTCAGTCTCTTTATCC-3′ (EcoRV).

Northern Blot Analysis—A human multitissue Northern blot (BD Biosciences) was hybridized with ^{32}P-labeled cDNA probes covering the entire coding regions of FGE or pFGE, respectively, and a β-actin cDNA probe as a control for RNA loading. Basically the same result, as shown below (see Fig. 1), was obtained when using probes hybridizing in the 3′-untranslated regions of *SUMF1* or *SUMF2* mRNA, in which the two sequences show no similarity (not shown). It was verified before each hybridization round that no radioactivity had remained from the previous hybridization. For technical details, see Ref. 18.

Cell Culture and Transfections—HT-1080 cells and human skin fibroblasts were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum and 1% penicillin/streptomycin (Invitrogen) under 5% CO_2 at 37 °C. Transient transfection was performed with Lipofectamine 2000 following the protocol recommended by Invitrogen. Typically, 4 μg of plasmid DNA were used per 6-cm dish. For coexpression experiments, the cells were transfected with a mix of 2 μg of pBI vector containing pFGE- (or FGE-) and steroid sulfatase-encoding cDNAs plus 2 μg of rtTA plasmid (reverse tetracycline controlled transactivator, BD Biosciences). For triple expression experiments, the cells were transfected with a mixture of three plasmids: 1 μg of pBI vector containing only pFGE cDNA, 1 μg of pBI vector containing FGE, and steroid sulfatase cDNAs and 2 μg of rtTA vector. After 6 h, the medium was replaced with medium containing 1 μg/ml doxycyline (BD Biosciences), and 30 h later, the cells were harvested for further analysis. For stable transfection, HT-1080 cells were transfected with pSB-pFGE-His (see above). Clones were selected and screened according to Ref. 19. HT-1080 cells stably expressing galactose 6-sulfatase were kindly provided by Transkaryotic Therapies Inc. (Cambridge, MA) and cotransfected with pSB-pFGE-His plus the puromycin resistance vector pSV-pac (20) (1:10 ratio). Transfectants were selected in normal growth medium containing puromycin (Invitrogen) with an increasing concentration from 0.2 to 0.8 μg/ml and screened by Western blotting.

Purification of Recombinant pFGE-His—HT-1080 cells stably overexpressing pFGE-His were grown to near confluence in normal growth medium. Then medium was collected every 48 h, cleared by spinning, and subjected to ammonium sulfate precipitation as described (19). The precipitate was reconstituted in and dialyzed against lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, and 10 mM imidazole, pH 8.0). The dialyzed material was cleared (142,000 × *g*, Ti45 rotor, 30 min, 4 °C) and subjected to Ni-NTA affinity chromatography under the conditions recommended by the manufacturer (Qiagen). Peak fractions, analyzed by Western blotting, were pooled and dialyzed overnight against Buffer A (10 mM Tris-HCl, pH 8.0, 100 mM NaCl, 2.5 mM dithiothreitol). The dialyzed pFGE-His was applied to a MonoQ HR10/10 anion exchange column (Amersham Biosciences) equilibrated with Buffer A. The pFGE-His protein was recovered in the flow-through and concentrated in an Ultra Thimble UH 100/25 (Schleicher & Schuell).

Indirect Immunofluorescence and Immunoelectron Microscopy—To detect endogenous expression of pFGE, human skin fibroblasts were grown on coverslips for 2 days. For recombinant expression of pFGE, the HT-1080 cells were grown on coverslips and cotransfected with pBI-pFGE-HA and rtTA plasmids, as described above. After 36 h of transfection, the cells were analyzed by indirect immunofluorescence as described (19). For colocalization studies, rabbit pFGE-His antiserum or monoclonal antibodies against the HA tag (Covance Inc., Princeton, NJ) were used as primary antibodies. The endoplasmic reticulum marker protein, protein disulfide isomerase, was detected with a monoclonal antibody of IgG_{2A} subtype (StressGen Biotechnologies). The primary antibodies were detected with isotype-specific goat secondary antibodies coupled to Cy2 or Cy3 (Molecular Probes), respectively. Confocal images were taken on a Leica TCS SP2 AOBS laser scan microscope.

For immunoelectron microscopy, ultrathin cryosections were prepared as described (19). For immunolabeling, sections were incubated with a polyclonal antibody against pFGE (1:25) for 30 min followed by a 20-min incubation with protein A-gold (10 nm). After washing steps, sections were fixed for 5 min with 1% glutaraldehyde. Sections were then quenched with glycine and blocked with bovine serum albumin before they were incubated with monoclonal antibodies against protein disulfide isomerase (1:30, StressGen Biotechnologies) or LAMP I (1:25, Hybridoma Bank) followed by a bridge antibody for 30 min (rabbit anti-mouse) and protein A-gold (5 nm) for 20 min. Sections were contrasted with uranyl acetate/methyl-cellulose for 10 min on ice, embedded in the same solution, and examined with a Philips CM120 electron microscope.

Endoglucosaminidase H/PNGase Treatment and Immunoprecipitation—HT-1080 cells stably expressing pFGE-His were grown to confluence. After 48 h in fresh medium, cells and medium were collected for partial purification of pFGE-His by Ni-NTA chromatography and endoglucosaminidase H or PNGase treatment as described (19).

Limited Proteolysis and Protein Sequence Analysis—pFGE-His purified from the secretion of HT-1080 cells was treated with thermolysin or elastase in phosphate-buffered saline at a protein/protease ratio of 25:1 (w/w) for up to 24 h at ambient temperature. For further analysis of the digests by SDS-PAGE and N-terminal sequencing of the proteolytic fragments, see Ref. 19.

Reductive Carbamidomethylation and Tryptic Digestion, MALDI-TOF Mass Spectrometry, and Edman Sequencing—Reductive carbamidomethylation and tryptic digestion of purified pFGE-His were done, as described (19). To keep disulfide bridges intact, reduction with dithiothreitol has been omitted in one experiment. Also MALDI-TOF mass spectrometry and Edman sequencing have been described (19).

Metabolic Labeling and Immunoprecipitation—Normal human skin fibroblasts were grown to near confluence in 3-cm dishes, starved in 2 ml of methionine and cysteine-free Dulbecco's modified Eagle's medium for 1 h, and pulsed for 6 h with 2 ml of medium containing 50 μCi of ^{35}S-labeled methionine and cysteine supplemented with 10% dialyzed fetal calf serum. Cells were harvested either directly or after washing with cold phosphate-buffered saline and culturing in 2 ml of chase medium (Dulbecco's modified Eagle's medium containing 10% fetal calf serum) for another 16 h (chase). Medium was collected, and cells were harvested by scraping. Immunoprecipitation of pFGE from cells and medium was performed with a pFGE anti-serum according to a protocol described earlier (21). Pellets were finally solubilized and subjected to SDS-PAGE and phosphorimaging (BAS1000, Raytest). Densitometric quantification of pFGE was done using the MacBAS software (Fuji).

Activity Assays for Sulfatases and FGE—Activity assays of FGE, steroid sulfatase, and galactose 6-sulfatase were performed as described earlier (14, 22, 23). For Western blot analysis, steroid sulfatase antiserum, galactose 6-sulfatase monoclonal antibodies (provided by Transkaryotic Therapies Inc., Cambridge, MA), and HA tag monoclonal antibodies (see above) were used as primary antibodies, which were detected with secondary horseradish peroxidase-conjugated goat anti-rabbit or anti-mouse antibodies. The Western blot signals were quantified by using the AIDA 2.1 software package (Raytest).

Photocross-linking—For photoaffinity labeling, a benzophenone-labeled peptide was synthesized corresponding to arylsulfatase A (residues 60–80) with an additional ϵ-biotinylated lysine residue at the N terminus, an additional serine amide at the C terminus, and a *p*-benzoyl phenylalanine replacing leucine 77. The reaction mixture, containing 50 μM benzophenone-labeled peptide, 1 μM pFGE-His, bovine serum albumin, or FGE-His was photocross-linked and analyzed by Western blotting as described (19).

Yeast Two-hybrid Analysis—The human pFGE cDNA (encoding residues 26–301) was inserted into the pGBKT7 bait vector (BD Biosciences) in-frame with the GAL4 DNA binding domain-encoding sequence. The human FGE cDNA (encoding amino acids 87–374) or ASA cDNA (encoding residues 19–114) was cloned into the pGADT7 prey vector (BD Biosciences) containing the GAL4 DNA-activating domain. Cotransformation of yeast cells (AH109), selection of cotransformants, and analysis for expression of *HIS3, ADE2*, and *MEL1* reporters was performed, as described (19).

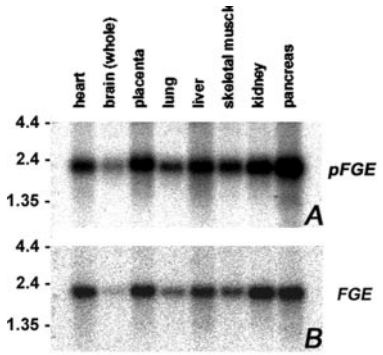

FIG. 1. **Expression of pFGE (*A*) and FGE (*B*) in human tissues.** Northern blot analysis of poly(A)+ RNA from human tissues. For β-actin abundance of the same blot, see Fig. 4 in Ref. 14. Very similar expression patterns were obtained when using probes hybridizing either in the coding region, as shown here, or in the non-conserved 3′-untranslated region (not shown).

RESULTS

Expression of SUMF2—A single *SUMF2* transcript of 2.0–2.1 kb is detectable by Northern blot analysis of poly(A)+ RNA from heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas (Fig. 1*A*) and of total RNA from skin fibroblasts (not shown). The expression pattern compares well with that of *SUMF1* (Fig. 1*B*), the relative ratio of *SUMF2* over *SUMF1* varying between 0.5 in heart and 1.4 in brain and pancreas. Relative to β-actin RNA (not shown), the abundance of *SUMF2* RNA varies by one order of magnitude, being highest in pancreas and kidney and lowest in brain.

Characterization of Recombinant pFGE—pFGE with a C-terminal HA tag was transiently expressed in the human HT-1080 sarcoma cell line. pFGE colocalized with protein disulfide isomerase, a lumenal protein of the endoplasmic reticulum (Fig. 2, *A–C*). In some cells, such as those shown in Fig. 2, *A–C*, a fraction of pFGE was associated with perinuclear structures devoid of protein disulfide isomerase. This structure is enriched in the Golgi marker protein GM130 (not shown). This indicates that recombinant pFGE is localized in the endoplasmic reticulum, but in some cells, can also accumulate in detectable amounts in the Golgi apparatus. Immunoelectronmicroscopy confirmed the localization of pFGE in the endoplasmic reticulum, where it colocalizes with protein disulfide isomerase (Fig. 3*A*) and in the Golgi stack (Fig. 3*B*). A sparse labeling for pFGE was observed in structures enriched in the endosomal/lysosomal marker LAMP I (Fig. 3*B*).

Overexpressed pFGE-His is only transiently retained in the endoplasmic reticulum. In HT-1080 cells, most of the pFGE-His that was synthesized over a period of 2 days (93% of total) was secreted. During that period, the cellular pFGE was kept at a rather constant level (Fig. 4).

pFGE contains a single potential *N*-glycosylation site at Asn-191. The intracellular 32-kDa pFGE-His was sensitive to endoglucosaminidase H and PNGaseF. Both endoglycosidases decreased the apparent molecular size of the intracellular pFGE by 2.5 kDa (Fig. 5), indicating that intracellular pFGE contains a high mannose (or hybrid) type oligosaccharide. The secreted pFGE is represented by a 31.5- and 32.5-kDa form. Both are converted into a 29.5-kDa polypeptide by PNGaseF, indicating that the two forms arise from heterogeneous *N*-glycosylation. The 31.5-kDa form is sensitive to endoglucosaminidase H, whereas the 32.5-kDa form is resistant (Fig. 5). This indicates that the 31.5-kDa form contains a high mannose type oligosaccharide, whereas the 32.5-kDa form contains a complex type oligosaccharide. MALDI-TOF analysis of the *N*-glycosylated tryptic peptide 179–192 of secreted pFGE-His confirmed its heterogeneous glycosylation at Asn-191. The oligosaccharide mixture was composed of 11 different species. They all belonged to the hybrid or complex type with up to five *N*-acetylhexosamine, six hexose, one sialic acid, and one fucose residue/

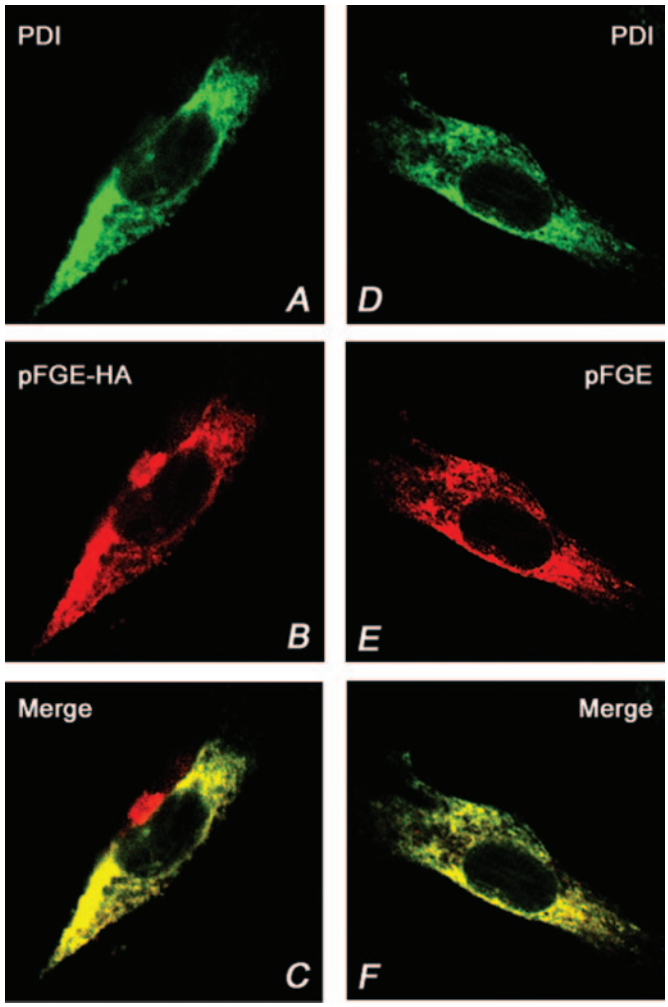

FIG. 2. **Subcellular localization of pFGE-HA in HT-1080 cells and endogenous pFGE in human skin fibroblasts.** After fixation and permeabilization of pFGE-HA overexpressing HT-1080 cells (*A–C*) or human skin fibroblasts (*D–F*), protein disulfide isomerase (*PDI*) was detected with a monoclonal IgG_{2A} antibody (*green*, *A* and *D*), the HA tag of pFGE (*red*, *B*) with a monoclonal IgG_1 antibody, and pFGE (*red*, *E*) with a rabbit antiserum. The merge reveals colocalization of protein disulfide isomerase and pFGE in *yellow* (*C* and *F*).

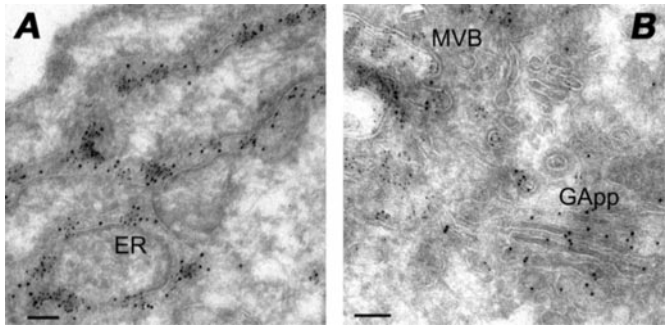

FIG. 3. **Localization of pFGE-His in HT-1080 cells by immunoelectron microscopy.** After fixation and cryosectioning, pFGE was detected with a rabbit antiserum against pFGE and protein A-gold (10 nm). *A* shows double labeling with the endoplasmic reticulum marker protein disulfide isomerase (5 nm gold), and *B* shows double labeling with the endosome/lysosome marker LAMP I (5 nm gold). *ER*, *GApp*, and *MVB* designate endoplasmic reticulum, Golgi apparatus, and multivesicular bodies, respectively. The *scale bar* corresponds to 100 nm.

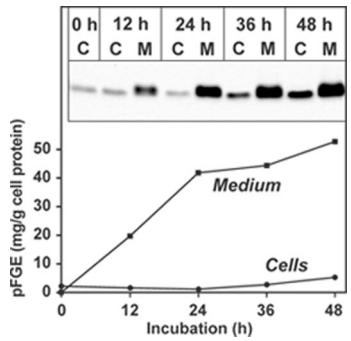

FIG. 4. **Secretion of pFGE-His by HT-1080 cells.** HT-1080 cells stably expressing pFGE-His were grown to half-confluency. At time 0, fresh medium was added, and the cells were harvested after 12, 24, 36, and 48 h of culture. pFGE was quantified in equal aliquots of cell extracts (*C*) and media (*M*) by Western blotting using a monoclonal antibody against the His_6 tag (*inset*). The amount of intracellular and secreted pFGE-His per mg of cell protein was determined by calibration of the Western blot with known amounts of purified pFGE-His and referred to total cell protein.

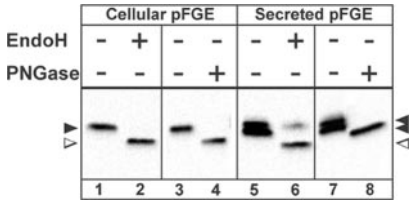

FIG. 5. ***N*-Glycosylation of intracellular and secreted pFGE-His.** Cell extracts and secretions of HT-1080 cells expressing pFGE-His were treated with endoglucosaminidase H (*EndoH*) or PNGaseF, separated by SDS-PAGE, and analyzed by Western blotting using anti-His antibodies. The *filled arrowheads* point to the glycosylated pFGE form, and the *open arrowheads* to the deglycosylated pFGE form.

oligosaccharide and masses ranging from 1768 to 2424 Da (not shown).

pFGE-His purified from the secretions of HT-1080 cells was incubated with elastase or thermolysin at a pFGE/protease ratio of 25:1. Both proteases generated two families of fragments of about 10–12 and 22–24 kDa (Fig. 6). The fragments generated by thermolysin were subjected to N-terminal amino acid sequencing. The 10–12-kDa fragments represent the amino terminus of mature pFGE starting with residue 27, whereas the 22–24-kDa fragments start with threonine 109 or with valine 115. Phylogenetic sequence conservation analysis has revealed three subdomains in pFGE (Fig. 6, *bottom*) (Ref. 16). The proteolytic cleavage sites are located in the linker region between subdomain I and II, which extends from serine 93 to leucine 116.

To examine for a possible disulfide bonding of the two cysteine residues in pFGE (at positions 156 and 290), pFGE-His purified from the secretions was first reacted with *N*-ethylmaleimide, to modify free cysteine residues, and then subjected to reductive carbamidomethylation in the presence of urea to modify disulfide-bonded cysteines. After digestion with trypsin, the peptides were separated by reversed phase high performance liquid chromatography on a C18 matrix and characterized by MALDI-TOF mass spectrometry and N-terminal amino acid sequencing. The two cysteine-containing tryptic peptides (154–159 and 290–303) were exclusively recovered as carbamidomethylated derivatives, indicating that cysteine 156 and cysteine 290 are disulfide-bonded in pFGE. In fact, when pFGE was subjected to carbamidomethylation in the presence of urea without prior reduction and then digested with trypsin, the two tryptic peptides 154–159 and 290–303 were found to be connected by a disulfide bond. Since non-reduced pFGE migrates in SDS-PAGE

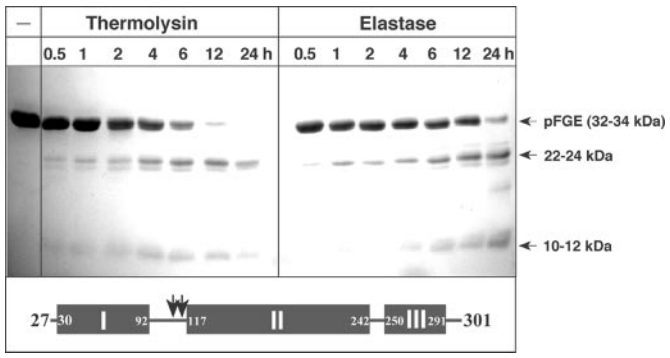

FIG. 6. **Limited proteolysis of pFGE-His by thermolysin and elastase.** pFGE-His purified from secretions of HT-1080 cells was incubated for up to 24 h with thermolysin or elastase, as indicated, at a pFGE/protease ratio of 25:1 (w/w). After separation by SDS-PAGE, the polypeptides were detected by staining with Coomassie Blue. The molecular masses of pFGE and the proteolytic fragments are indicated at the *right*. The scheme at the *bottom* shows the locations of the predicted subdomains I, II, and III together with the residue numbers defining the N and C termini of mature pFGE and its subdomains. The *arrows* indicate the two proteolytic cleavage sites of thermolysin N-terminal of threonine 109 and valine 115, as determined by Edman sequencing.

as a monomer (not shown), cysteine 156 and 290 form an intramolecular rather than an intermolecular disulfide bond.

Localization and Secretion of Endogenous pFGE—Antisera against pFGE-His, purified from the secretions of HT-1080 cells, were generated in rabbits. These antisera were suitable for indirect immunofluorescence, immunoprecipitation, and Western blots.

When fibroblasts, fixed with paraformaldehyde and permeabilized with saponin, were reacted with anti-pFGE serum, pFGE was detectable in the endoplasmic reticulum, where it colocalized with protein disulfide isomerase (Fig. 2, *D–F*). Association with the Golgi, as observed for overexpressed pFGE (Fig. 2, *B* and *C*), was not detectable for endogenous pFGE.

To follow biosynthesis and secretion of endogenous pFGE, human skin fibroblasts were metabolically labeled for 6 h with [^{35}S]methionine. After labeling, the cells were either harvested or subjected to a 16-h chase in [^{35}S]methionine-free medium. pFGE was immunoprecipitated from extracts of the cells and from media, separated by SDS-PAGE, and visualized by phosphorimaging (Fig. 7). pFGE was synthesized as a 32-kDa polypeptide that remained stable during the 16-h chase period. After the chase for 16 h, a small amount (less than 13%) was recovered in the secretions as 31.5- and 32.5-kDa forms. This indicates that differently from recombinant pFGE, only a small fraction of endogenous pFGE escapes retention in the endoplasmic reticulum. Heterogeneous processing of the *N*-linked oligosaccharide in the secretory pathway is likely to give rise to the two 31.5- and 32.5-kDa pFGE forms in the secretions, as has been shown for recombinant pFGE (cf. Fig. 5).

pFGE Lacks FGly-generating Activity—pFGE is the paralog of FGE, which has FGly-generating activity. The generation of the FGly residue is readily detectable with the help of the 23-mer peptide P23 derived from arylsulfatase A (14). This peptide contains the cysteine residue to be modified and neighboring residues, which are essential for FGly formation. The FGly-generating activity in extracts of HT-1080 cells transiently expressing pFGE-His (Fig. 8) or non-tagged pFGE (not shown) was in the range of non-transfected HT-1080 cells, whereas FGly-generating activity in extracts of cells expressing a comparable amount of FGE-His was more than 20-fold increased (Fig. 8). When purified pFGE-His was incubated with the P23 peptide under conditions that allow quantitative modification of the cysteine residue by purified FGE-His, no FGly formation was detectable. Increasing or decreasing the

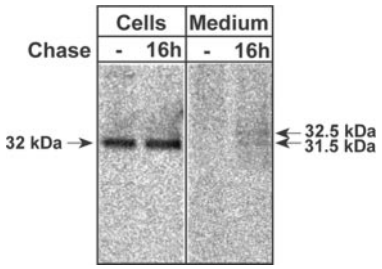

FIG. 7. **Biosynthesis and secretion of endogenous pFGE.** pFGE was immunoprecipitated from the cells and media of human skin fibroblast that had been labeled for 6 h and harvested either directly after metabolic labeling or after a chase period of 16 h. The *arrows* indicate the intracellular 32-kDa pFGE form (*left*) and the secreted 31.5- and 32.5-kDa forms (*right*).

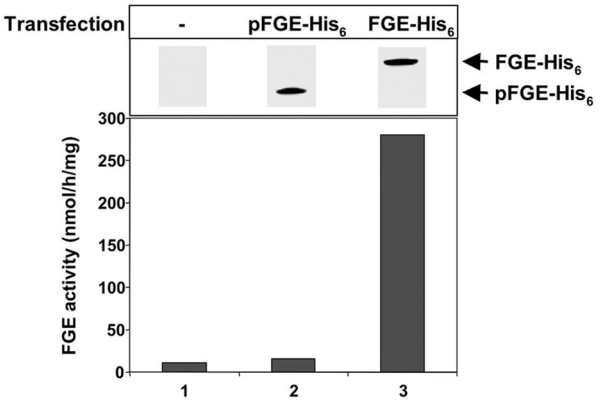

FIG. 8. **FGE activity in HT-1080 cells expressing pFGE-His or FGE-His.** The *upper panel* shows the Western blot detection of pFGE-His and FGE-His in the cell extracts with a monoclonal antibody against the His_6 tag. The *lower panel* shows the FGly-generating activity in the cell extracts determined *in vitro* with the P23 peptide as substrate. *Bar 1*, non-transfected HT-1080 cells; *bars 2* and *3*, cells transiently expressing pFGE-His or FGE-His, respectively.

amount of pFGE-His did not result in any FGly formation. In addition, we examined peptides, derived from the other 15 human sulfatases, as substrates for pFGE. Although all peptides were modified *in vitro* by FGE (19), none of them were modified by pFGE. This indicates that pFGE lacks FGly-generating activity.

Overexpression of pFGE Impairs Sulfatase Activity—In vivo, the activity of FGE can be a limiting factor for the synthesis of catalytically active sulfatases. This becomes apparent when sulfatases are overexpressed. The yield of sulfatase activity is increased by simultaneous overexpression of FGE (13, 14).

When FGE was transiently coexpressed with steroid sulfatase in HT-1080 cells, the activity of steroid sulfatase increased about 2-fold. In contrast, coexpression of pFGE reduced the sulfatase activity by about half (Fig. 9). The inhibitory effect may indicate that pFGE competes with the endogenous FGE for modification of the recombinant steroid sulfatase. In line with this assumption, we observed that coexpression of pFGE and FGE abolished the stimulatory effect of FGE on steroid sulfatase activity (Fig. 9). Control experiments verified that both tagged and non-tagged pFGE fail to activate steroid sulfatase (not shown).

A similar result was observed when pFGE-His and galactose 6-sulfatase were stably coexpressed in HT-1080 cells. Coexpression of pFGE-His reduced the specific activity of the galactose 6-sulfatase by 20% (not shown), whereas coexpression of FGE increased the specific activity of galactose-6-sulfatase more than 100-fold.[3] The different efficiencies of FGE-mediated

[3] M. Mariappan, unpublished information.

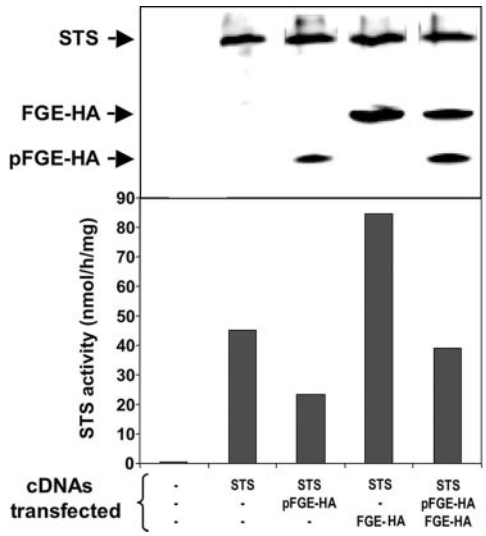

FIG. 9. **pFGE-HA reduces and FGE-HA enhances the activity of steroid sulfatase.** In HT-1080 cells, steroid sulfatase (*STS*), pFGE-HA, and FGE-HA were transiently coexpressed in the combinations indicated below the lanes. In the cell extracts, the expression of steroid sulfatase, pFGE-HA, and FGE-HA was monitored by Western blotting using a rabbit antiserum against steroid sulfatase and a monoclonal antibody against the HA tag (*top*). The activity of steroid sulfatase, referred to total cell protein, is shown at the *bottom*. It should be noted that the transfected cDNAs were located on two vectors, one vector coexpressing steroid sulfatase and, where indicated, FGE-HA from a bidirectional promotor, and another vector expressing pFGE-HA only (see "Experimental Procedures").

activation and pFGE-mediated inhibition in the two experiments are explained by the much higher expression of galactose 6-sulfatase, as compared with that of steroid sulfatase. This by far exceeded the FGly-generating capacity of endogenous FGE, thus leaving large amounts of unmodified galactose 6-sulfatase as a substrate for the recombinant FGE. Accordingly, the lower inhibitory effect of coexpressed pFGE on galactose 6-sulfatase may be due to the high expression level of this sulfatase, clearly exceeding that of pFGE. Taken together, these data support the view that pFGE has no FGly-generating activity toward the three sulfatases tested (arylsulfatase A, steroid sulfatase, and galactose-6-sulfatase). The inhibitory effect of pFGE *in vivo* on the formation of active sulfatases suggests that pFGE can interfere with the activity of FGE on sulfatases. A number of mechanisms are conceivable to explain the inhibitory effect of pFGE on the generation of catalytically active sulfatases. Among them are binding of pFGE to FGE or/and to sulfatases and thereby interfering with the modification of nascent sulfatase polypeptides by FGE.

Binding of pFGE to Sulfatases—Binding of FGE to sulfatases has been demonstrated in several experimental systems. When extracts of HT-1080 cells coexpressing FGE-His and galactose 6-sulfatase are subjected to gel permeation chromatography, complexes of FGE and galactose 6-sulfatase coelute.[3] In cells coexpressing pFGE-His and galactose 6-sulfatase, we failed to observe complexes. pFGE-His and galactose-6-sulfatase eluted as monomers of 20 and 55 kDa, respectively (not shown).

A peptide representing arylsulfatase A residues 60–80, but carrying a *p*-benzoyl phenylalanine (Bpa) at position 77 and a biotinylated lysine residue at the N terminus, can be photocross-linked to FGE (Fig. 10, *lane 5*). Replacing FGE by pFGE, we also observed cross-linking of the peptide to pFGE, albeit at only one-tenth of the efficiency observed for FGE-His (Fig. 10, *lane 1*). Cross-linking was specific as indicated by the inhibitory effect of a 2-fold excess of the arylsulfatase A 60–80 peptide P23 (Fig. 10, *lane 2*) and by the lack of cross-link

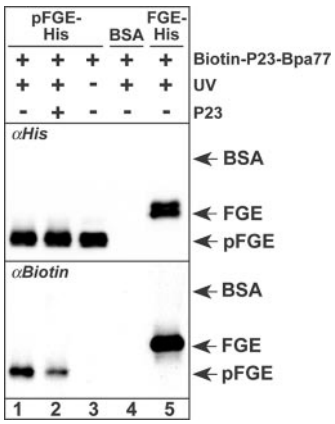

FIG. 10. **Photocross-linking of pFGE-His to an arylsulfatase A peptide.** pFGE-His (*lanes 1–3*), bovine serum albumin (*BSA*, *lane 4*) or FGE-His (*lane 5*), 1 μM each, were incubated with 50 μM biotin-P23-Bpa77, a peptide corresponding to arylsulfatase A residues 60–80 with an additional ϵ-biotinylated lysine residue at the N terminus, a C-terminal serine amide residue and a Bpa replacing leucine 77. In *lane 2*, a 2-fold excess of peptide P23, corresponding to arylsulfatase A residues 60–80 with an additional *N*-acetylated methionine and a C-terminal serine amide residue, was added. After irradiation with UV light for 30 min on ice, equal aliquots of the samples were subjected to SDS-PAGE, transferred onto a nitrocellulose membrane, and probed with antibodies against the His tag (*upper panel*) or biotin (*lower panel*). The positions of bovine serum albumin, FGE, and pFGE are indicated by *arrows*.

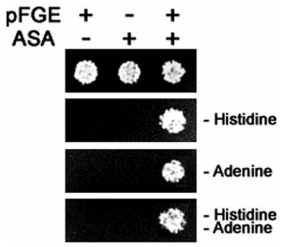

FIG. 11. **Yeast two-hybrid interaction of pFGE and arylsulfatase A.** Yeast cells were cotransformed with a prey vector containing an ASA insert (codons 19–114) and a bait vector containing a pFGE insert (codons 26–301); −, vector without insert. Cotransformants were selected on medium lacking tryptophan and leucine. Two-hybrid interaction leads to reporter activation allowing growth on selection medium that lacks, as indicated, histidine, adenine, or both, in addition to tryptophan and leucine.

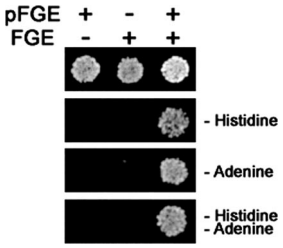

FIG. 12. **Yeast two-hybrid interaction of pFGE and FGE.** The inserts of the bait (pFGE codons 26–301) and the prey (FGE codons 87–374) vector used for cotransformation of yeast cells are indicated (−, vector without insert). For further detail, see the legend for Fig. 11.

products when bovine serum albumin (Fig. 10, *lane 4*), ovalbumin, or ribonuclease were used as acceptor proteins. When biotinylated ASA peptides with Bpa residues in position 62 or 66 were used, the cross-linking efficiency was further lowered and substitution of the critical cysteine 69 by alanine reduced cross-linking to pFGE 10-fold (not shown).

An interaction of pFGE with arylsulfatase A was observed in a yeast two-hybrid system. Using pFGE as bait and an arylsulfatase A fragment comprising residues 19–114 as prey, a two-hybrid interaction was observed for three reporter genes (*HIS3, ADE2*, and *MEL1*) expressed under the control of three distinct GAL4 responsive elements (Fig. 11).

Interaction of pFGE and FGE—The interaction between pFGE and FGE was also examined with the help of a yeast two-hybrid system. Coexpression of pFGE (residues 26–301) and FGE (residues 87–374) induced the expression of three reporter genes (*HIS3, ADE2*, and *MEL1*) (Fig. 12). Reporter gene expression was independent of whether FGE or pFGE was expressed as bait or prey hybrid proteins.

All attempts failed to verify a direct interaction of pFGE and FGE by demonstration of co-purification during gel permeation chromatography, Ni-NTA-agarose pull-down or chemical cross-linking of the two proteins employing extracts of pFGE-His and FGE coexpressing cell lines (data not shown). We also examined whether the addition of pFGE to a mixture of FGE and its peptidic substrate P23 would inhibit the FGly formation. pFGE added in an up to 300-fold molar excess over FGE had no effect on FGly formation, neither at a low P23 concentration close to its K_m (13 nM) nor at a saturating P23 concentration (200 nM). Thus, *in vitro* pFGE has no inhibitory effect on FGE activity toward peptidic substrates.

DISCUSSION

The present study demonstrates that pFGE and FGE share many structural and topological properties but differ functionally. FGE catalyzes in sulfatases the oxidation of a cysteine residue to the active site FGly residue, whereas pFGE lacks such an activity both *in vitro* and *in vivo*.

The nucleotide sequence of pFGE predicts an N-terminal signal peptide directing translocation of the nascent polypeptide into the lumen of the endoplasmic reticulum. pFGE was indeed found to be a soluble glycoprotein of the endoplasmic reticulum. Although only a small fraction of endogenous pFGE escapes from the endoplasmic reticulum into secretions, overexpression of pFGE results in its massive secretion. pFGE is therefore likely to be retained in the endoplasmic reticulum by a saturable mechanism. It should be noted that pFGE lacks an endoplasmic reticulum retention signal of the KDEL type, suggesting that it is retained by binding to some other resident component of the endoplasmic reticulum. During secretion, the single high mannose type oligosaccharide of pFGE becomes processed to hybrid and complex type structures containing fucose and sialic acid residues. Since for intracellular pFGE only the high mannose type oligosaccharide was found, the residence in the Golgi, as detected in some overexpressing cells (Fig. 2*C*), is negligible, and exposure to Golgi-dependent modification reactions is restricted to the fraction of pFGE that escapes endoplasmic reticulum retention and becomes secreted.

Phylogenetic sequence conservation analysis has revealed that pFGE and FGE contain three highly conserved regions. In pFGE, these subdomains make up more than 85% of the molecule. Digestion with elastase, a serine proteinase, or thermolysin, a zinc-metalloproteinase, generated two stable fragments by cleavage within the short linker sequence connecting the first and the second subdomain. This may indicate that the three conserved regions are part of two folding domains.

FGE contains three fully conserved cysteine residues in the highly conserved C-terminal subdomain III. Two of them have been suspected to be critical for catalytic activity (19). They both are lacking in pFGE. pFGE contains only two cysteine residues, one located in subdomain II (Cys-156) and another one located in subdomain III (Cys-290). They form an intramolecular disulfide bond linking subdomains II and III. This disulfide bond is conserved in FGE, in which two intersubdomain disulfide bridges link subdomains II and III (19). Therefore, in

pFGE as well as in FGE, subdomains II and III are tightly connected and together form a large protease-resistant domain.

The localization, carbohydrate processing, and secretion upon overexpression as well as the protease sensitivity of the linker region between the first two subdomains are properties that pFGE shares with its paralog FGE (19). Moreover, the relative abundance of their RNAs in different tissues is rather similar. All these observations suggest that pFGE and FGE fulfil similar functions. A known and important functional property of FGE is its FGly-generating activity, an essential protein modification for sulfatases that renders them catalytically active. The critical role of FGE for the FGly formation is evident from the loss of catalytically active, FGly-containing sulfatases in patients that carry mutations in the *SUMF1* gene encoding FGE (13–15). Cells of these patients accumulate catalytically inactive sulfatases that retain the cysteine residue, which normally is oxidized by FGE to FGly (2). So far, no mutations in the pFGE-encoding *SUMF2* gene have been found in multiple sulfatase deficiency patients.

Under conditions appropriate for the FGly-generating activity of FGE toward peptidic substrates, pFGE was inactive. This was true for the peptides derived from the 16 sulfatases known or predicted from the human genome sequence. This indicates that under *in vitro* conditions, pFGE lacks FGly-generating activity. The increase of sulfatase activity upon coexpression of pFGE with sulfatases had suggested that *in vivo* pFGE has FGly-generating activity or is stimulating the activity of endogenous FGE (13). We failed to observe an increase of sulfatase activity upon transient or stable coexpression of pFGE. In HT-1080 cells, the coexpression of pFGE decreased rather than increased the activity of coexpressed sulfatases. In an attempt to understand this inhibitory effect, we examined the interaction of pFGE with sulfatases and FGE. We could demonstrate the interaction of an ASA fragment and short ASA-derived peptides with pFGE in a yeast two-hybrid approach and by photocross-linking, respectively. In a yeast two-hybrid assay, we could also observe an interaction between pFGE and FGE. This, however, could not be confirmed in approaches monitoring a direct physical interaction of pFGE and FGE. Taken together, our results support the view that pFGE lacks FGly-generating activity. *In vivo* the overexpression of pFGE interferes with the FGly formation in sulfatases indicated by the decrease of sulfatase activity without concomitant decrease in sulfatase polypeptide. This effect may be due to binding of sulfatases to pFGE, thereby sequestrating them from FGE. However, the effects of pFGE on FGE are also conceivable either in an indirect manner, *e.g.* by dislocating FGE from the endoplasmic reticulum due to competition for a common retention mechanism, or directly by heterodimerization. The latter possibility agrees with the crystal structure of pFGE and is discussed by Dickmanns *et al.* (24).

This study underlines the similarities of the non-catalytic properties of pFGE and FGE and provides evidence that both contact newly synthesized sulfatases in the endoplasmic reticulum. However, the biological role of pFGE remains to be defined. Detailed analysis of the molecular interaction of pFGE and FGE and nascent sulfatase polypeptides as well as ablation of pFGE may help elucidating its function.

Acknowledgments—We thank Achim Dickmanns, Ralf Ficner, Gabriele Fischer von Mollard, Stefan Höning, Olaf Jahn, Klaus Neifer, Jianhe Peng, Thomas Pringle, Markus Rudolph, Petra Schlotterhose, Annegret Schneemann, Peter Schu, and Nicole Tasch for excellent technical assistance and fruitful discussions.

REFERENCES

1. Hopwood, J. J., and Ballabio, A. (2001) in *The Metabolic and Molecular Bases of Inherited Disease* (Scriver, C. R., Beaudet, A. L., Valle, D., and Sly, W. S., eds) pp. 3725–3732, McGraw-Hill Professional, New York
2. Schmidt, B., Selmer, T., Ingendoh, A., and von Figura, K. (1995) *Cell* **82,** 271–278
3. Selmer, T., Hallmann, A., Schmidt, B., Sumper, M., and von Figura, K. (1996) *Eur. J. Biochem.* **238,** 341–345
4. Miech, C., Dierks, T., Selmer, T., von Figura, K., and Schmidt, B. (1998) *J. Biol. Chem.* **273,** 4835–4837
5. Dierks, T., Schmidt, B., and von Figura, K. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 11963–11968
6. Dierks, T., Miech, C., Hummerjohann, J., Schmidt, B., Kertesz, M. A., and von Figura, K. (1998) *J. Biol. Chem.* **273,** 25560–25564
7. von Figura, K., Schmidt, B., Selmer, T., and Dierks, T. (1998) *BioEssays* **20,** 505–510
8. Dierks, T., Lecca, M. R., Schlotterhose, P., Schmidt, B., and von Figura, K. (1999) *EMBO J.* **18,** 2084–2091
9. von Figura, K., Borissenko, L. V., Fey, J., Peng, J., Schmidt, B., and Dierks, T. (2004) in *Lysosomal Disorders of Brain* (Platt, F. M., and Walkley, S. U., eds) pp. 131–140, Oxford University Press, Oxford
10. Boltes, I., Czapinski, H., Kahnert, A., von Bülow, R., Dierks, T., Schmidt, B., von Figura, K., Kertesz, M. A., and Usón, I. (2001) *Structure* **9,** 483–491
11. Lukatela, G., Krauss, N., Theis, K., Selmer, T., Gieselmann, V., von Figura, K., and Saenger, W. (1998) *Biochemistry* **37,** 3654–3664
12. von Bülow, R., Schmidt, B., Dierks, T., von Figura, K., and Usón, I. (2001) *J. Mol. Biol.* **305,** 269–277
13. Cosma, M. P., Pepe, S., Annunziata, I., Trott, D. A., Parenti, G., and Ballabio, A. (2003) *Cell* **113,** 445–456
14. Dierks, T., Schmidt, B., Borissenko, L. V., Peng, J., Preusser, A., Mariappan, M., and von Figura, K. (2003) *Cell* **113,** 435–444
15. Cosma, M. P., Pepe, S., Parenti, G., Settembre, C., Annunziata, I., Wade-Martins, R., Di Domenico, C., Di Natale, P., Mankad, A., Cox, B., Uziel, G., Mancini, G. M., Zammarchi, E., Donati, M. A., Kleijer, W. J., Filocamo, M., Carrozzo, R., Carella, M., and Ballabio, A. (2004) *Hum. Mutat.* **23,** 576–581
16. Landgrebe, J., Dierks, T., Schmidt, B., and von Figura, K. (2003) *Gene* (*Amst.*) **316,** 47–56
17. Stein, C., Hille, A., Seidel, J., Rijnbout, S., Waheed, A., Schmidt, B., Geuze, H., and von Figura, K. (1989) *J. Biol. Chem.* **264,** 13865–13872
18. Lübke, T., Marquart, T., Etzioni, A., Hartmann, E., von Figura, K., and Körner, C. (2001) *Nat. Genet.* **28,** 73–76
19. Preusser-Kunze, A., Mariappan, M., Schmidt, B., Gande, S. L., Mutenda, K., Wenzel, D., von Figura, K., and Dierks, T. (2005) *J. Biol. Chem.* ***280,*** 14900–14910
20. Vara, J. A., Portela, A., Ortin, J., and Jimenez, A. (1986) *Nucleic Acids Res.* **14,** 4617–4624
21. Gieselmann, V., Schmidt, B., and von Figura, K. (1992) *J. Biol. Chem.* **267,** 13262–13266
22. Rommerskirch, W., and von Figura, K. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 2561–2565
23. Glössl, J., and Kresse, H. (1978) *Clin. Chim. Acta* **88,** 111–119
24. Dickmanns, A., Schmidt, B., Rudolph, M. G., Mariappan, M., Dierks, T., von Figura, K., and Ficner, R. (2005) *J. Biol. Chem.* **280,** 15180–15187