Molecular Characterization of pFGE, the Paralog of the Cα-Formylglycine-generating Enzyme

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg-August-Universität Göttingen

vorgelegt von
Malaiyalam Mariappan
aus Mamsapuram (Indien)

Göttingen 2005

D7
Referent: **Prof. Dr. h.c. Kurt von Figura**
Koreferent: **Prof. Dr. Hans-Joachim Fritz**

Tag der mündlichen Prüfung:01.11.2005

Dedicated to my Father

Contents

Abbreviations

Amp	Ampicilline
APS	Ammonium peroxide sulfate
ASA	Arylsulfatase A
ASB	Arylsulfatase B
Arg	Arginine
ATP	Adenodin-5'-triphosphate
BHK21	baby hamster kidney cells
bp	Base pairs
BPa	p-benzoyl phenylalanine
BPB	Brom phenol blue
BSA	Bovine serum albumin
^{o}C	Grade Celsius
cDNA	Complementary DNA
CHO	Chinese Hamster Ovary
Ci	Curie
CMV	cyclomegalovirus
ConA	Concanavalin A
cpm	Impulse per minute
C69	ASA 65-80 peptide
C69S	ASA 65-80 C69S peptide
Da	Dalton
ddH2O	double distilled H2O
DMSO	Dimethylsulfoxide
DNA	Desoxyribonucleic acid
DNPH	Dinitrophenylhydrasine
DSP	dithiobis(succinimidylpropionate)
DTT	Dithiothreitol
dNTPs	2'-Desoxyribonucleoside-5'-phosphate
E	Extinction

E.coli	Escherichia coli
EDTA	Ethylendiamintetraacetate
eq	equivalents
ER	Endoplasmic reticulum
EST	Expressed sequence tag
et al	et alteri (and others)
FGly-residue	Formylglycine residue
FGE	C-αFormylglycine generating enzyme
g	gram
GApp	Golgi apparatus
Gly	Glycine
Gal-6-S	Galactose-6-Sulfatase
GM130	Golgi matrix protein 130
hr	hour
HPLC	High perfomance liquid chromatography
His	Histidine
HT1080	Human fibrosarcoma cells
Ig	Immunoglobulin
kb	kilobase
K.pneumoniae	Klebsiella pneumoniae
l	liter
LAMP	Lysosome associated matrix protein
LB-medium	Luria-Broth medium
m	meter
M	molar
MALDI TOF-MS	Matrix Assisted Laser Desorption Ionization Time-of-flight Mass Spectrometry
μ	micro, -(x10^{-6})
mA	milliamper
min	minute
MgCl$_2$	Magnesium chloride
MSD	Multiple sulfatase deficiency
mRNA	messenger RNA
MVB	Multi vesicular bodies
n	nano, -(x10^{-9})
NEM	N-ethyl maleimide
NIH/3T3	Mouse embryo fibroblast
Ni-NTA	Nickel-nitrilotriacetic acid

OD	Optical density
PAGE	Polyacrylamide gel electrophoresis
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDI	Protein disulphide isomerase
pFGE	paralog of C-αFormylglycine generating enzyme
pNCS	para-nitrocatecholsulfate
P.aeruginosa	Pseudomonas aeruginosa
P23	ASA 60-80
PT67	NIH/3T3-based packaging line
RNA	Ribonucleic acid
RNAse	Ribonuclease
RP	Reversed phase
rpm	Rotation per minute
RT	Room temperature
rRNA	Ribosomal RNA
s	second
^{35}S	Sulphur 35
SDS	Sodium dodecyl sulfate
SUMF1	Sulfatase modifying factor 1
SUMF2	Sulfatase modifying factor 2
t	time
Taq	Thermophilius aquatics
TAE	Tris-acetate-EDTA-buffer
TBS	Tris-buffered solution
TCA	Trichlor-acetic acid
TGN	trans-Golgi-network
TE	Tris-EDTA
TEMED	N,N,N,N-tetramethylethylendiamine
Tris	Tris-(hydroxymethyl)-aminomethane
UTR	Untranslated region
UV	Ultraviolet
V	Volt
v/v	Volume pro volume
wt	Wild type
w/v	Weight pro volume

Chapter 1

Introduction

Sulfate esters are participating in a wide spectrum of biological reactions. They are found in a number of organic compounds of the cell, like glycosaminoglycans (dermatan sulfate, chondroitin sulfate, keratan sulfate, heparan sulfate) and sulfated glycolipids (cerebroside sulfate), tyrosine-, serine-, and threonine-sulfate carrying proteins[1] and sulfated hydroxysteroids (cholesterol sulfate, dehydroepiandrosterone sulfate). Sulfate esters are synthesized by a class of enzymes known as sulfotransferases. Sulfotransferases catalyze the transfer of a sulfate (i.e, sulfonate, SO_3^-) group from an activated donor onto a hydroxyl or less frequently an aminogroup of the acceptor molecule. The nucleotide analogue 3' phosphoadenosine 5' phosphosulfate (PAPS) invariably serves as the sulfate donor [2]. The sulfate esters are hydrolyzed by enzymes belonging to the family of sulfate ester sulfohydrolases or sulfatases. In this thesis, we take a comprehensive look at the sulfatases and their activating enzyme(s).

Sulfatases cleave sulfate esters such as N and O-sulfate esters in biological systems (Eq.1 and 2). It plays a key role in regulating the sulfation states that determine the function of many physiological molecules.

$R - SO_3^- + H_2O \Rightarrow R - OH + H^+ + SO_4^{2-} (1)$

$R - NHSO_3^- + H_2O \Rightarrow R - NH_3^+ + SO_4^{2-} (2)$
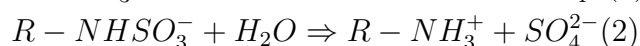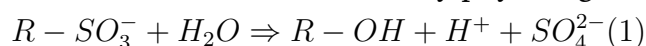
Sulfatase substrates range from small cytosolic steroids, such as estrogen sulfate, to complex cell surface carbohydrates, such as the glycosaminoglycans. The transformation of these molecules has been linked with important cellular functions including hormone regulation, cellular degradation and modulation of signaling pathways. Sulfatases have also been implicated in the onset of various pathophysiological conditions including hormone-dependent cancers, lysosomal storage disorders, developmental abnormalities and bacterial pathogenesis.

1.1   Sulfatase family and Subcellular localization

The sulfatase family comprises a class of enzymes that is highly conserved sequentially, structurally and mechanistically across eukaryotic and prokaryotic species. The remarkable similarities shared by sulfatases include: 1) 20-60 % sequence homology over the entire protein length, 2) a highly conserved N-terminal region containing consensus sulfatase motifs [3, 4](Fig1.1). The extent of sequence similarity and structural likeness suggests that they emerged long ago from a commmon ancestral gene [5, 6]

Human sulfatases (from genome):													Length (residues)	Signal peptide
Arylsulfatase A	C	T	P	S	R	A	A	L	L	T	G	R	(Pos.69-80)	507	+
Arylsulfatase B	C	T	P	S	R	S	Q	L	L	T	G	R	(Pos.91-102)	533	+
Arylsulfatase C	C	T	P	S	R	A	A	F	M	T	G	R	(Pos.83-94)	583	+
Arylsulfatase D	C	T	P	S	R	A	A	F	L	T	G	R	(Pos.89-100)	593	+
Arylsulfatase E	C	T	P	S	R	A	A	F	L	T	G	R	(Pos.86-97)	589	+
Arylsulfatase F	C	S	P	S	R	S	A	F	L	T	G	R	(Pos.79-90)	591	+
Arylsulfatase G	C	T	P	S	R	A	A	F	L	T	G	R	(Pos.181-192)	688	?
N-Acetylgalactosamine 6-sulfatase	C	S	P	S	R	A	A	L	L	T	G	R	(Pos.79-90)	522	+
Nacetylglycosamine 6-sulfatase	C	C	P	S	R	A	S	I	L	T	G	K	(Pos.91-102)	552	+
Iduronate sulfatase	C	A	P	S	R	V	S	F	L	T	G	R	(Pos.84-95)	550	+
Sulfamidase	C	S	P	S	R	A	S	L	L	T	G	L	(Pos.70-81)	502	+
Sulf 6	C	S	P	S	R	A	S	L	L	T	G	R	(Pos.84-95)	525	+
Sulf 1	C	C	P	S	R	S	S	M	L	T	G	K	(Pos.87-98)	871	+
Sulf 2	C	C	P	S	R	S	S	I	L	T	G	K	(Pos.88-99)	870	+
Sulf 3	C	C	P	S	R	A	A	M	W	S	G	L	(Pos.70-81)	526	+
Sulf 4	C	T	P	S	R	S	Q	F	I	T	G	K	(Pos.96-107)	573	+
Sulf 5	C	T	P	S	R	S	Q	L	L	T	G	R	(Pos.93-104)	569	+
Lower eukaryotic sulfatases:														
Coturnix coturnix(Qsulf1)	C	C	P	S	R	S	S	M	L	T	G	K	(Pos.87-98)	869	+
Hemicentrotus pulcherrimus	C	T	P	S	R	S	A	I	M	T	G	R	(Pos.100-111)	551	+
Strongylocentrotus purpuratis	C	T	P	S	R	S	A	I	V	T	G	R	(Pos.115-126)	567	+
Heliocidaris erythrogramma	C	T	P	S	R	S	A	I	M	T	G	R	(Pos.106-117)	559	+
Volvox carteri	C	C	P	S	R	T	N	L	W	R	G	Q	(Pos.72-86)	649	+
Chlamydomonas reinhardtii	C	C	P	S	R	T	N	L	W	R	G	Q	(Pos.73-84)	646	+
Neurospora crassa	C	C	P	A	R	V	S	L	W	T	G	K	(Pos.89-100)	639	+

Figure 1.1. Partial alignment of sulfatase protein family. This consensus sequence is important for directing the first amino acid residue to the catalytically active FGly. Highly conserved residues are shown in gray.

The human genome predicts 17 sulfatases of which 11 have been characterized and found to show distinct sulfatase activities and subcellular localization (1.1). A subgroup of seven human sulfatases were termed arylsulfatases due to their ability to cleave synthetic chromogenic or fluorogenic substrates that are applied in the enzyme assays and histochemistry. At least eight human sulfatases were localized in the lysosomes, where they catalyse the degradation of glycosaminoglycans and sulfolipids. The arylsulfatase C (Steroid Sulfatase) is an integral membrane protein of the endoplasmic reticulum and the plasma membrane. It is involved in cholesterol and steroid hormone metabolism. Arylsulfatase D, E, F and G are non lysosomal proteins that have been localized to the endoplasmic reticulum(arylsulfatase D,F and G) and Golgi apparatus(arylsulfatase E) [4, 7].The physiological roles of arylsulfatase D,E,F and G are still to be determined, although there are indications that arylsulfatase E is possibly involved in Vitamin K metabolism [8].

The newly discovered sulfatases Sulf1 and Sulf2 are localized at the cell surface[7].Sulf1 has been involved in the modification of the sulfation of glycosaminoglycans (GAGs) which dictate developmental cell signaling and patterning processes. QSulf1 (Quail Sulfatase1) regulates the Wnt signaling by cleaving the 6-O sulfate from heparan sulfate proteoglycan which releases Wnt. The "Catch and Present" model was suggested for the regulation of Wnt signaling by QSulf1[9]. In this model, desulfation by QSulf1 is proposed to stimulate a shift from a high affinity heparan sulfate-Wnt complex (the catch phase) to a low affinity heparan sulfate-Wnt complex, which allows interaction of Wnt with frizzled receptors (the present phase) to initate the Wnt signaling cascade. Desulfation of heparan sulfate proteoglycans by HSulf1 influences many other signaling pathways including Epidermal Growth Factor (EGF) and Fibroblast Growth Factor (FGF) mediated signaling [10]. In addition, Sulfs have been implicated in pathophysiological conditions including tumor onset and progression [11].

1.2   Single sulfatase deficiencies

The biological importance of human sulfatases is highlighted by the manifestation of eight known inherited metabolic disorders. They are associated with impaired desulfation of specific substrate metabolites and with their lysosomal storage (Table1.2). For example, arylsulfatase A is required to catalyze the first step in the degradation pathway of galactosyl-3-sulfate ceramide (Sulfatide), a

Table 1.1. Substrates and subcellular location of human sulfatases

Sulfatase	Abbr	Physiological substrate	Subcellular location
arylsulfatase A	ARSA	sulfatide	lysosome
arylsulfatase B	ARSB	dermatan sulfate chondroitin sulfate	lysosome
arylsulfatase C	ARSC	steroid sulfates	ER
arylsulfatase D	ARSD	unknown	ER
arylsulfatase E	ARSE	unknown	Golgi network
arylsulfatase F	ARSF	unknown	ER
arylsulfatase G	ARSG	unknown	ER
galactosamine-6-sulfatase	GalN6S	chondroitin sulfate keratan sulfate	lysosome
glucosamine-3-sulfatase	GlcN3S	heparan sulfate	lysosome
glucosamine-6-sulfatase	GlcN6S	heparan sulfate keratan sulfate	lysosome
glucuronate-2-sulfatase	GlcA2S	heparan sulfate	lysosome
heparan-N-sulfatase	GlcNS	heparan sulfate	lysosome
iduronate-2-sulfatase	IdoA2S IdoA2S	heparan sulfate dermatan sulfate	lysosome
endo sulfatase 1	Sulf1	heparan sulfate	cell surface
endo sulfatase 2	Sulf2	heparan sulfate	cell surface

major sphingolipid of myelin[12]. Deficiency of this sulfatase is the cause of Metachromatic Leukodystophy (MLD), a lysosomal storage disorder that leads to severe neurological symptoms because of extensive demyelination in the central and peripheral nervous system. Arylsulfatase B is involved in degradative pathways of dermatan sulfate and chondroitin 4-sulfate [5]. These glycosaminoglycans are constituents of proteoglycans existing in the extracellular matrix. Deficiency of arylsulfatase B leads to Mucopolysaccaridosis VI, a lysosomal storage disorder with a variety of clinical phenotypes such as dwarfism, skeletal abnormalities, cloudy cornea and heart failure [5]. Deficiency of arylsulfatase C leads to X-linked ichthyosis, a relatively mild disease of the skin with increased levels of cholesterol sulfate in the *stratum corneum* [13]. These distinct single sulfatase deficiencies emphasise the high substrate specificity of the sulfatases and their limited functional redundancy. The resulting phenotypes range from extremely severe and life threatening with early death to reduced life expectancy. The deficiency of sulfatases in case of each syndrome is caused by different genetic defects. For instance, one of the most common mutations associated with a mild

form of MLD is a substitution of proline 426 by leucine in arylsulfatase A (ASA). The resulting deficiency of ASA is caused by its instability in the lysosomes. It was shown that the instability of the mutated enzyme, caused by its defective oligomerization in the lysosmes, drastically increases the vulnerability of the enzyme to degradation by lysosomal cysteine proteases [14]. Point mutations of the corresponding sulfatase genes are also the cause of many other sulfatase deficiencies, among them various mucopolysaccharidoses (MPS).

Table 1.2. Inherited disorders resulting from deficiency in a single sulfatase

Subcellular location	Sulfatase	Genetic disorder
lysosome	ARSA	metachromatic leukodystrophy (MLD)
	IdoAS	Hunter's disease (MPS II)
	GlcNS	Sanfilippo A (MPS IIIA)
	GlcN6S	Sanfilippo (MPS IIID)
	GalN6S	Morquio A (MPS IVA)
	ARSB	Maroteaux-Lamy (MPS VI)
ER	ARSC	X-linked ichthyosis (XLI)
Golgi	ARSE	chondrodysplasia punctata (CDPX)

1.3   Multiple sulfatase deficiency and Cα-formylglycine

In addition to single sulfatase deficiencies, a rare autosomal recessive disorder called Multilpe Sulfatase Deficiency (MSD) is known, in which the activities of all known sulfatases are severely decreased. Austin described the first case of MSD in two siblings in 1963 and 1965. Since then a number of MSD patients have been identified [12]. The clinical phenotype of MSD combines features characteristic of single sulfatase deficiencies, such as rapid neurologic deterioration and developmental delay.

Single sulfatase deficiency syndromes are caused by mutations in the corresponding sulfatase genes. In case of each syndrome the activity of only one sulfatase is decreased, all the other sulfatases are active within the normal range. On the contrary, the activity of all the sulfatases are dramatically reduced in MSD patients, this finding has led to the proposal that there must be a common reason for the deficiency of all sulfatases. Studies in cultured cells from MSD patients have

shown that synthesis of sulfatase polypeptides is normal but their catalytic activity is severely diminished and for some of the sulfatase polypeptides their stability as well [15, 16]. Expression of sulfatase cDNAs in MSD fibroblasts yielded sulfatase polypeptides with a severely reduced activity [17]. Therefore it was proposed that sulfatases require co- or post- translational modification that is missing in case of MSD. By Mass spectrometric analysis of the sulfatase polypeptides (ASA and ASB) synthesized in MSD fibroblasts revealed that they lack a Cα-formylglycine (FGly; 2-amino-3-oxopropionic acid) residue and contain a cysteine instead, as predicted by the cDNA [18]. The FGly residue is found at a position where the cDNA sequences of all known eukaryotic sulfatases predict a cysteine within a highly conserved sequence. This sequence of 12 amino acids is shown in Fig1.1. Later the presence of FGly was shown also for lower eukaryotic [19] and prokaryotic [20, 21] sulfatases. This demonstrates that the FGly residue is conserved in prokaryotic and eukaryotic members of the sulfatase family and can be generated by oxidation of either a cysteine (eukaryotes and prokaryotes) or, a serine residue (prokaryotes only) as will be explained below.

1.3.1 Role of Formylglycine in sulfate ester hydrolysis

The crystal structures of human sulfatases ASA and and arylsulfatase B (ASB) have been solved at 2.1 Åand 2.5 Å, respectively. The FGly residue was shown to be located in the active site cavity (Fig1.2) representing part of metal binding site with an octahedrally coordinated metal ion [22, 23]. The 3-D structure of the bacterial sulfatase from *Pseudomonas aeruginosa* (PAS), resolved at 1.3 Å, established the metal ion in the active site as a Ca^{2+}[24]. The structures of pro- and eukaryotic active sites of sulfatases are identical within the error limits and show topographic similarity to that of alkaline phosphatase. The side chain of FGly is superimposable to that of the catalytically essential serine 102 in the alkaline phosphatase active center [22].

Structural and enzymatic studies of ASA mutants, in which the cysteine was replaced by a serine or alanine [25, 14] and the recent solution of the X-ray structure of PAS [24] provided insight into the actual mechanism of sulfate ester hydrolysis. The electron density map of the FGly side chain showed the presence of two hydroxyl groups bonded to Cβ in FGly 51, the key catalytic site chain in PAS, corresponding to FGly 69 in ASA and FGly 91 in ASB. Thus, the FGly residue in the active enzyme is present as a FGly-hydrate (Fig1.3). One of the geminal hydroxyl groups of the aldehyde hydrate serves as an acceptor for sulfate leading
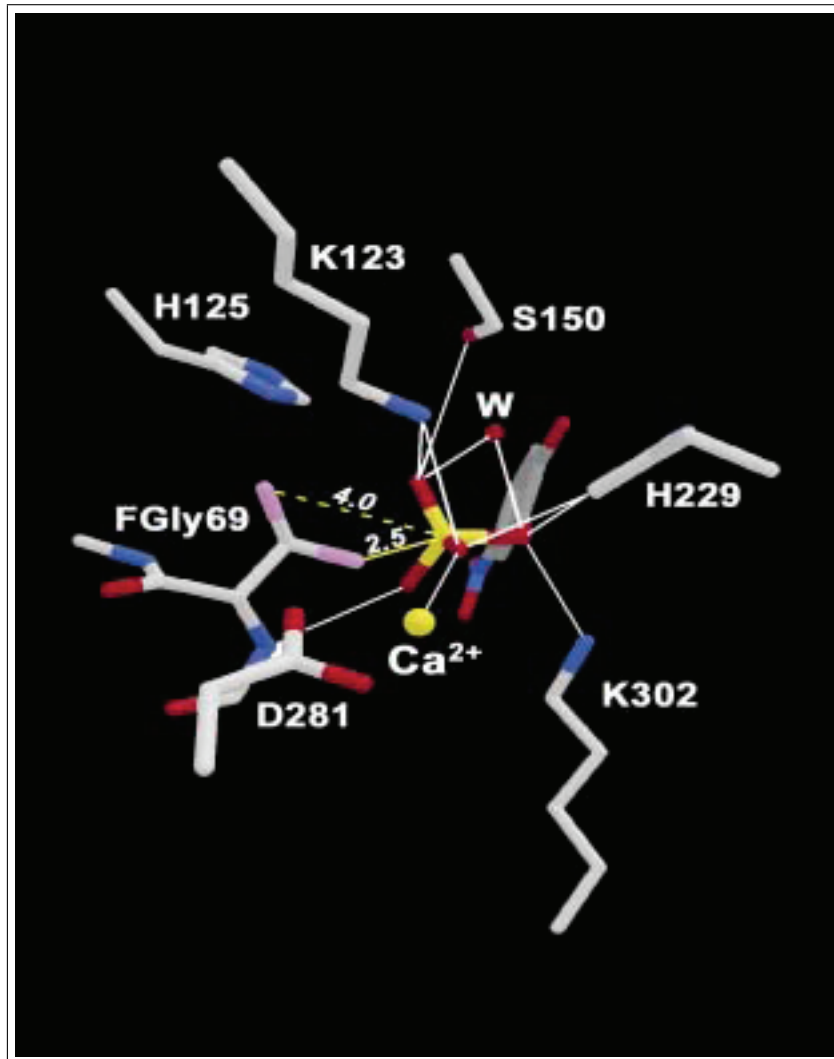
Figure 1.2. Active centre of human arylsulfatase A [14, 24]. The sulfate group of the substrate is positioned in the active center and coordinated by lysine 123, serine150, histidine 229, lysine 302, a water molecule (W) and a Ca^{2+} ion. The nucleophilic attack on the sulfur atom by one of the hydroxyl groups of FGly69-hydrate (distance 2.5 A) is facilitated by deprotonation of this hydroxyl group by aspartate.

to the formation of a covalently sulfated enzyme intermediate. The function of this hydroxyl group is similar to that of Ser 102 in alkaline phosphatase. The second hydroxyl group of the aldehyde hydrate is proposed to induce an intramolecular rearrangement leading to elimination of the sulfate and regeneration of the aldehyde group (Fig1.3).

   After substrate desulfation, the alcohol is released and diffuses out of the catalytic pocket. The sulfate is now covalently bound to the enzyme and additionally coordinated to the calcium atom through two of its oxygen atoms. The C-O bond in the FGly-sulfate ester is polarized through all the contacts of the sulfate
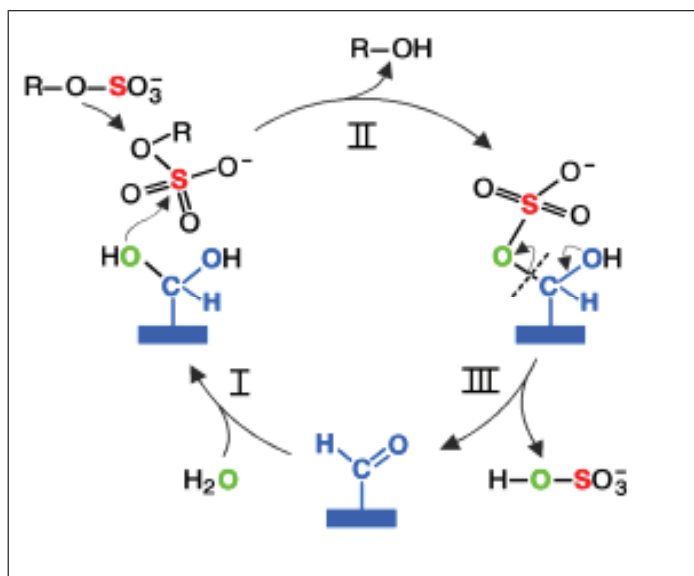
Figure 1.3. Scheme of the proposed catalytic mechanism for sulfate ester hydrolysis [24]. Sulfate ester cleavage is initiated by formation of an aldehyde hydrate at the catalytic FGly residue (I). Neucleophilic attack of the sulfate group by one of the hydroxyls of the aldehyde hydrate leads to substitution of the alcohol (II). The sulfate group is then released from the sulfated enzyme intermediate by an elimination reaction cleaving a C-O bond and regenerating the aldehyde (III).

group to positively charged residues, favoring its elimination. The final step of the reaction is the elimination of sulfate and regeneration of the aldehyde. The aldehyde hydrate is regenerated by addition of a water molecule and is stabilized by hydrogen bonds to histidine, arginine and the calcium atom. The mechanism, as described, is based on the structural analyses of sulfatases and their substrate complexes [23, 22, 14] and was also confirmed by kinetic studies of numerous active site mutants [25, 26, 27]. The ASA-C69S and the corresponding ASB-C91S mutants allowed to trap the sulfated FGly intemediate (Fig.1.4C). The sulfate could no longer be eliminated from this intermediate, since the second hydroxyl group was absent in these mutants [25]. The stuctural similarity of sulfatases and in particular of their active site regions, strongly suggests that the proposed catalytic mechanism is common for all sulfatases. The key function of the FGly residue in this mechanism explains the critical role of the posttranslational generation of this residue in the biogenesis of enzymatically active sulfatases.

1.3.2   Formylglycine modification motif

In experiments with *in vitro* synthesized ASA fragments it was shown that a sequence of 16 amino acid residues, encompassing positions -4 to +11 with respect

to the cysteine to be modified (Fig.5), is essential and sufficient for the formation of FGly [28], . Formation of FGly was observed with the same efficiency as that of the control after transferring these 16 residues to a heterologous protein [29]. Analysis of different deletion- and substitution- mutants of ASA could show that a sequence motif consisting of 12 consecutive residues starting with the cysteine to be modified is nesessary to reach optimal FGly formation [29]. This sequence consists of an essential core motif CxPxR and an auxiliary motif xxxL/MTGR/K/L. Apart from cysteine, the key residues of the core motif are proline and arginine in positions +2 and +4, respectively, which can be found in all known and putative members of the sulfatase family. The auxiliary motif (position +5 to +11 after cysteine that is to be modified) is playing a supporting but not an essential role in FGly formation. Even complete substitution of the highly conserved amino acid residues LTGR (Fig.1.5) by an AAAA-tetrapeptide led only to 50% reduction of FGly formation [29]. It is proposed that this sequence of seven amino acid residues facilitates presentation of the core motif to the modifying enzyme.

Inhibition of the FGly formation (%) after substitution by Ala/Gly

7	0	28	0	100	0	94	33	95	33	0	0	49	0	26	17
P	V	S	L	C 69	T	P	S	R	A	A	L	L	T	G	R
P	V	S	L	C 69	T	P	S	R	A	A	L	L	T	G	R
13	13	38	50	100	50	100	94	100	44	56	44	56	88	94	69

Conservation (%) in eukaryotic sulfatases
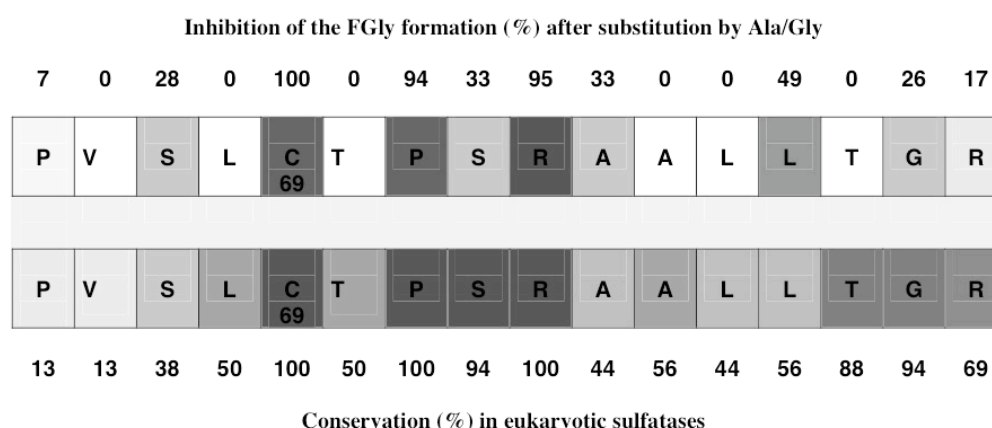
Figure 1.4. A short linear sequence determines the FGly formation [29]. The 16 mer sequence of the human Arylsulfatase A is essential and sufficient for the modification of Cys69. Inhibition of FGly formation after substitution of each single amino acid residue by alanine or glycine is given in the upper row in numbers and indicated by different intensity of grey tones. In the lower row, intensity of grey tones and percent values indicate the conservation level of the amino acid residues among eukaryotic sulfatases. Comparison of the two rows demonstrate that all three residues, C69, P71, R73, which are essential for FGly formation are 100% conserved among all eukaryotic sulfatases.

1.3.3   Formation of formylglycine in the endoplasmic reticulum

The cysteine residue encoded in the sulfatase gene, is incorporated into the nascent sulfatase polypeptide chain during translation [28]. The cysteine is converted to FGly during or shortly after protein translocation into the endoplasmic reticulum (Fig1.3), as could be shown in an *in vitro* translation/translocation system comprising import competent dog pancreas microsomes [28, 30]. It was shown that, in an arrested translocation intermediate, the cysteine 69 of ASA was accessible inside the microsomes. However, FGly formation was observed only after releasing the nascent chain from the ribosome by puromycin [28]. Thus, FGly formation occurs after or at a later stage of cotranslational protein translocation.
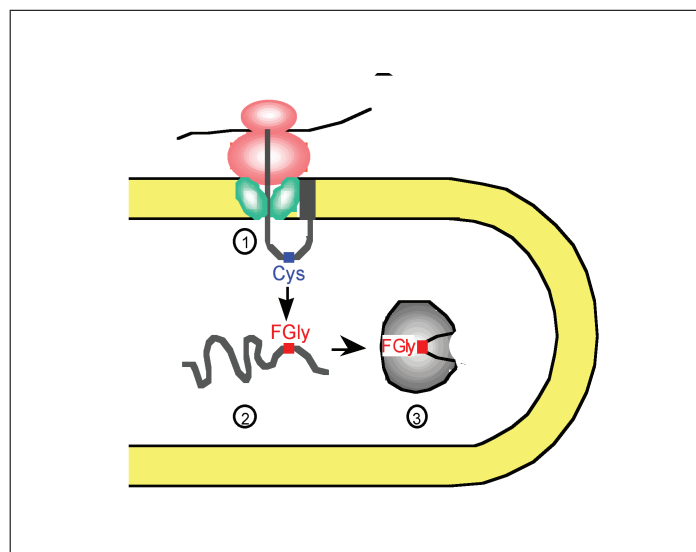
Figure 1.5. Conversion of the cysteine to a FGly residue in sulfatases occurs in the endoplasmic reticulum. Early translocation intermediate (Stage 1) retaining the signal peptide (open box) still carry the conserved cysteine residue. Completion of translocation (Stage 2) is associated with cleavage of the signal peptide and conversion of the conserved cysteine (red) to FGly (blue). The modified sulfatase polypeptides are supposed to complete folding in the endoplasmic reticulum (Stage 3)[28].

In a recent study of *in vitro* FGly formation [31] it was investigated whether the components of the FGly-generating machinery are part of the ER membrane or the ER lumen or both. The dog pancreas microsomes were treated with increasing detergent concentrations and separated into supernatant and pellet fractions by centrifugation. Increasing concentrations of the detergent solubilized increasing amounts of FGly-generating activity that was recovered in the supernatant fraction. Under conditions that selectively extract the luminal components but leave mem-

brane components in the pellet fraction, almost 100 percentage of FGly generating
activity was recovered in the supernatant and virtualy no activity remained in
the pellet fraction. Thus, the FGly modifying machinery is part of the soluble
components of the ER lumen referred to as reticuloplasm. In the same study it
was shown that *in vitro* formation of FGly by reticuloplasm does not depend on the
presence of the signal peptide sequence in the sulfatase substrate. Furthermore, the
FGly-generating enzyme was characterised with respect to some basic biochemical
properties such as pH optimum, which was unexpectedly high (pH 10-10.5). Kinet-
ically controlled conditions were established for assaying FGly formation *in vitro*.
The FGly forming activity showed typical enzymatic properties with a characteristic
temperature dependence (activation energy 61 kJ/mol) and sensitivity to inhibitory
peptides [31]. Reticuloplasm as a source for the FGly-generating enzyme and
the *in vitro* assay for detection of its activity are the basis for the functional and
chromatographic characterisation of the so far unknown modifying protein on the
way to its eventual identification.

1.4 FGly generating machinery in prokaryotes

Mammalian sulfatases are involved in the turnover of endogenous sulfated sub-
strates [32]. On the other hand, sulfatases of lower eukaryotes, like algae or fungi
and of bacteria, are expressed under conditions of sulfur starvation and function in
sulfate scavenging from exogenous substrates[33]. Due to conservation of the FGly
modification motif, most of the sulfatases encoded in various bacterial and archaea
bacterial genomes are predicted also to undergo FGly modification by oxidation
of cysteine or serine. This was shown experimentally for the arylsulfatase from
Pseudomonas aeruginosa (PAS) and *Methanosarcina mazei sulfatase* (MMS), two
members of the cysteine-type sulfatases. Even after strong overexpression in *E. coli*
PAS was quantitatively converted to the active FGly-bearing enzyme [21]. Thus,
the *E. coli* cytosol must contain the FGly modifying machinery. This machinery
is expressed even under excessive inorganic sulfate supply conditions and thus is
independent of the sulfur status of the cell, in contrast to expression of the sulfatase
structural genes, as shown for *P.aeruginosa* and *K.pneumoniae* [34, 35, 33]

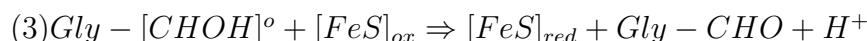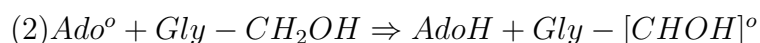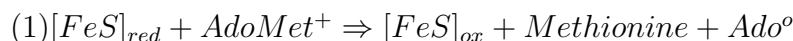The other well characterised bacterial sulfatase, the arysulfatase of K.pneumoniae,
is a serine- type sulfatase that carries FGly residue generated by oxidation of a serine
rather than a cysteine [20]. In contrast to the cytosolic cysteine-type sulfatases,
serine-type sulfatases are located in the periplasm [35, 36]. The key FGly motif

(SXPXR) and also the auxiliary downstream element (LTG) are conserved in serine-type sulfatases [29].

Despite these similarities, bacteria have two different pathways for FGly generation either from cysteine or serine. This is indicated by two observations: first, substitution of the cysteine to be modified in *Pseudomonas* sulfatase by serine totally blocks FGly formation [30]; second, expression of an active, FGly containing *Klebsiella* sulfatase in *E. coli* essentially requires coexpression of the *Klebsiella* atsB gene [37], while the genomic background of *E. coli* is sufficient for expressing an active and modified *Pseudomonas* sulfatase. It was shown recently, that AtsB, which is a iron-sulfur protein, is located in the cytosol and is interacting directly with the serine-type FGly modification motif, which allows to consider the possibility that AtsB itself is the oxidising enzyme converting serine to FGly [38]. Some iron sulfur proteins related to AtsB were shown to have enzymatic redox functions by generating radical species by reductive cleavage of S-adenosylmethionine (SAM) through an unusual FeS center that is also present in AtsB [39, 40]. A possible reaction sequence for AtsB-mediated FGly formation is outlined in Scheme.1 Transfer of an electron from the reduced FeS center to S-adenosylmethionine leads to its reductive cleavage (step 1). The generated deoxyadenosyl radical abstracts a hydrogen atom from the substrate, *i.e.* the Ser72 side chain, under formation of deoxyadenosine and a substrate radical (step 2). The single electron of this radical is then accepted by the FeS center, leading to its re-reduction, under formation of FGly (step 3).

Scheme 1: Proposed mechanism for AtsB-mediated FGly-formation [38]

$(1) [FeS]_{red} + AdoMet^+ \Rightarrow [FeS]_{ox} + Methionine + Ado^o$

$(2) Ado^o + Gly - CH_2OH \Rightarrow AdoH + Gly - [CHOH]^o$

$(3) Gly - [CHOH]^o + [FeS]_{ox} \Rightarrow [FeS]_{red} + Gly - CHO + H^+$

where, $AdoMet^+$ = S-adenosylmethionine

Ado^o = 5'-deoxyadenosyl radical

AdoH = 5'-deoxyadenosine

Gly-CH_2OH = Serine

Gly-CHO = FGly

As FGly formation from a serine is most likely a single enzymatic reaction, AtsB may suffice for FGly formation in serine-type sulfatases. On the contrary, the cysteine-modifying enzyme could not be identified to date in bacteria. A transposon-mutagenesis approach failed (in *E. coli*), indicating that the cysteine-modifying system is either redundant or essential, i.e. required for other functions apart from FGly- modification of sulfatases.

1.5 Purification and identification of the mammalian Cα-formylglycine generating enzyme

1.5.1 *In vitro* assay for FGly formation

Based on the information that translation of a 16 mer ASA sequence in a heterologous polypeptide backround and its translocation into the endoplasmic reticulum allows the generation of FGly, an *in vitro* assay for the FGly formation was established. In this assay a detergent solubilized extract from bovine pancreas microsomes served as a source of the modifying enzyme. As substrate (^{35}S) methionine labeled sulfatase polypeptide was used that was translated *in vitro* by an mRNA programmed reticulocyte lysate. The polypeptides were added to the assay mixture as ribosome-associated nascent chain complexes. The quantification of the product included tryptic digestion, separation of the peptides by RP-HPLC, and identification and quantification of the ^{35}S-labeled FGly containing tryptic peptide by a combination of chemical derivatization to hydrazones, RP-HPLC separation , and liquid scintillation counting [31]. For monitoring the enzyme activity during purification, this procedure was tedious to carry out. A new assay based on a peptidic substrate and matrix-assisted laser desorption/ionisation time-of-flight (MALDI-TOF) mass spectrometry was developed.

Previously it was shown that the 16-mer peptide corresponding to ASA residues 65-80 and containing the sequence motif required for FGly formation inhibited the FGE activity in the *in vitro* assay [31]. This suggested that peptides such as ASA 65-80 may serve as substrate for FGE. Thus a 23-mer peptide p23 which corresponds to ASA residues 60-80 was synthesized. Incubation of P23 with extracts from microsomes of bovine pancreas or bovine testis converted up to 95% of the peptide into the FGly containing derivative (Fig1.6). Under standard conditions, the FGly formation was proportional to the amount of enzyme and time of incubation as long as less than 50% of the substrate was consumed and the incubation period did not exceed 24hr. The K$_M$ for P23 was 13nM[41].

1.5.2 Purification of FGE

FGE was purified from the soluble fraction of bovine testis microsomes by a four-step chromatography. The first two steps were chromatography on a MonoQ anion exchanger and on concanavalin A-Sepharose. At pH8, the FGE activity bound to MonoQ and was eluted at 50-165mM NaCl. When this fraction was mixed
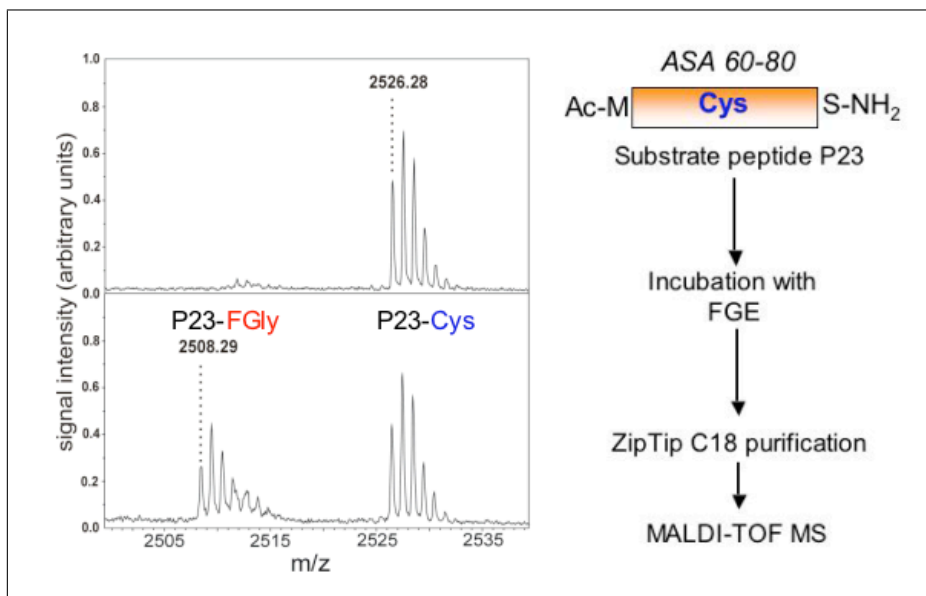
Figure 1.6. FGly Modification of P23. 6 pmol of P23 were incubated under standard conditions for 10min at 37^oC in the absence (top) or presence (bottom) of $1\mu l$ of microsomal extract. The samples were prepared for MALDI-TOF mass spectrometry (see Materials and methods). The monoisotopic masses MH^+ of P23 (2526.28) and its FGly derivative (2508.29) are indicated.

with concanavalin A-Sepharose, FGE was bound and eluted with 0.5 M α-methyl mannoside. The two final purification steps were chromatography on affinity matrices derivatized with 16-mer peptides. The first affinity matrix was Affigel 10 substituted with a variant of the ASA65-80 peptide, in which three residues critical for FGly formation, Cys-69, Pro-71, and Arg-73, were scrambled (scrambled peptide PVSL**PTRSC**AALLTGR). This peptide did not inhibit FGE activity when added at 10 μM concentration to the *in vitro* assay and when immobilized to Affigel 10, did not retain FGE activity. Chromatography on the scrambled peptide affinity matrix removed peptide binding proteins including chaperones of the endoplasmic reticulum. The second affinity matrix was Affigel 10 substituted with a variant of the ASA 65-80, in which the Cys-69 was replaced by a serine (Ser-69 peptide PVSL**S**TPSRAALLTGR). The FGE activity could be eluted with either 2 M KSCN or 25 μM Ser-69 peptide. Prior to activity determination, the KSCN or Ser-69 peptide had to be removed by dialysis. The substitution of Cys-69 by serine was crucial for the elution of active FGE. Affigel 10 substituted with the wild-type ASA65-80 peptide bound FGE efficiently. However, nearly no activity could be recovered in eluates with chaotropic salts (KSCN, $MgCl_2$), peptides (ASA65-80, or Ser-69 peptide), or buffers with low or high pH. In Figure 1.6 the polypeptide pattern of the starting material and of the active fractions obtained after the four chromatographic steps of
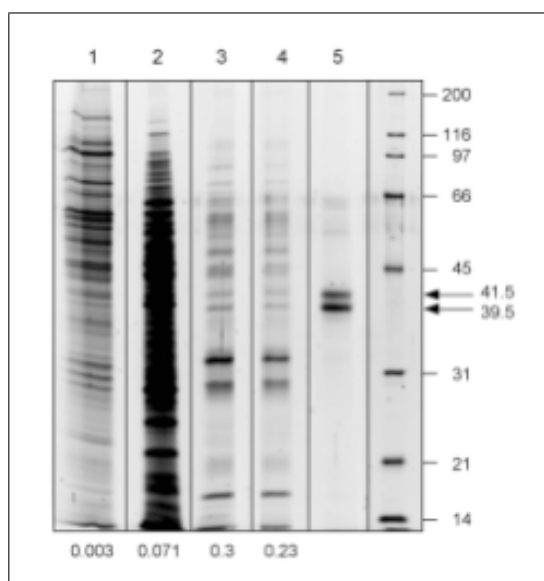
Figure 1.7. Purification of FGE from Bovine Testis (Dierks et al., 2003)
Aliquots of the soluble extract from microsomes (lane 1), of the pooled fractions after chromatography on MonoQ (lane 2), concanavalin A-Sepharose (lane 3), and scrambled peptide-Affigel 10 (lane 4) were separated by SDS-PAGE. The entire material eluted from Ser-69 peptide-Affigel 10 was concentrated and loaded in lane 5. Molecular weight standards are shown

a typical purification is shown. In the final fraction, 5% of the starting FGE activity and 0.0006% of the starting protein were recovered (8333-fold purification).

1.5.3   Identification of FGE gene

Mass finger print analysis revealed that the 39.5 and 41.5 kDa polypeptides (see Figure 1.7, lane 5) were encoded by a single gene called *SUMF1*. The *SUMF1* gene encoding human FGE is located on chromosome 3p26. Orthologous genes were found in mouse (87% identity), rat (94% identity), *Drosophila melanogaster* (48% identity), *Anopheles gambiae* (47% identity), and *Fugu rubripes* (63% identity). Orthologous EST sequences were found for a number of other species too[42]. The genomes of *S. cerevisiae* and *C. elegans* lack *SUMF1* homologs. In prokaryotes, 13 orthologs of the *SUMF1* gene were found. Sequence comparison analysis of human FGE suggested that FGE comprises 374 aminoacids including a cleavable endoplasmic reticulum(ER) signal sequence. It is composed of three distinct subdomains [42], the N-terminal subdomain (residues 91-154) has a potential N-glycosylation site at 141 which is conserved in other orthologs. The middle part of FGE (residues 179-308) is represented by a tryptophan rich subdomain. The C-terminal subdomain (residues 327-366) has the most conserved sequence within the

FGE family. Among 40 residues of the subdomain four residues are fully conserved. Three of these cysteines are also conserved in the prokaryotic FGE orthologs.

1.5.4   Expression and subcellular localization of FGE

The *SUMF1* expression level is highest in pancreas and kidney and lowest in brain. Transient expression of tagged and un-tagged FGE increased the FGE activity 1.6- to 3.9-fold. Stable expression of FGE in PT67 cells increased the activity of FGE about 100 fold [41]. Indirect immunofluoresence showed the colocalization of the variously tagged forms of FGE with protein disulfide isomerase, a lumenal protein of the endoplasmic reticulum in BHK21, CHO and HT1080 cells [41].

1.5.5   Mutation in *SUMF1* causes MSD

MSD is caused by a deficiency to generate FGly residues in sulfatases [18]. It was shown that mutations in the *SUMF1* gene are the cause of multiple sulfatase deficiency. So far 22 mutations of *SUMF1* gene were described in 20 different patients [41, 43]. Interestingly, seven mutations appear to be clustered in the C-terminal subdomain of the protein, suggesting that this subdomain is critical for *SUMF1* activity. Five other mutaions are located in the tryptophan-rich subdomain and one muataion is in the N-terminal subdomain. In addition, two mutations of the first methionine and one located within the signal peptide were also found [43]. For the three MSD-cell lines, it was shown that transduction of the FGE encoding cDNA partially or fully restores the sulfatase activities. On the contrary, transduction of the FGE encoding cDNA carrying one of the mutations observed in MSD patients did not restore sulfatase activities [41]. MSD is both clinically and biochemically heterogenous. Biochemically it is characteristic that a residual activity of sulfatases is detected, which in cultured skin fibroblasts is generally below 10% of controls [44, 45]. However, in some MSD cell lines, the activity of selected sulfatases can reach the normal range [46]. Furthermore the residual activity is subject to variations, depending on the cell culture conditions and unknown factors [47, 48, 49, 50]. Biochemically MSD has been classified into two groups [50, 15]. In group I, the residual activity of sulfatases is below 15%, including that of ASB. In group II, the residual activity of sulfatases is higher and particularly that of ASB may reach values of up to 50-100% of control. The phenotypic heterogeneity suggests that the different mutations in MSD patients are associated with different residual activities of FGE [41].

1.6 Aim

1.6.1 Characterization of pFGE, the paralog of FGE

The *SUMF1* gene is conserved among pro- and eukaryotes [42]. In the genome of deuterostomia, including vertebrates and echinodermata, a paralog of the *SUMF1* gene, designated as *SUMF2* is found [41, 51, 43]. 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. A high degree of sequence similarity was observed along the entire length of *SUMF1* and *SUMF2* protein sequences, with the exception of the hydrophobic N-terminal region. Therefore, *SUMF2* encoded protein was termed paralog of FGE (pFGE). Among the family of *SUMF2* encoded pFGE, the identity is much higher than to the *SUMF1* encoded FGE of the same species. For example, the identity between human and mouse pFGE is 86%, while the identity with the corresponding FGE is 47%-49%. pFGE shares the subdomain structure with FGE. In subdomain III, however, they lack two of the three cysteines conserved among pro- and eukaryotic FGEs [41]. The human *SUMF2* gene is located on chromosome 7q11. The human pFGE shares 47.1% sequence identity and 62.1% similarity with human FGE (Fig1.8). The function of the paralog FGE is unknown. Cotransfection of *SUMF2* with sulfatase cDNAs such as those of
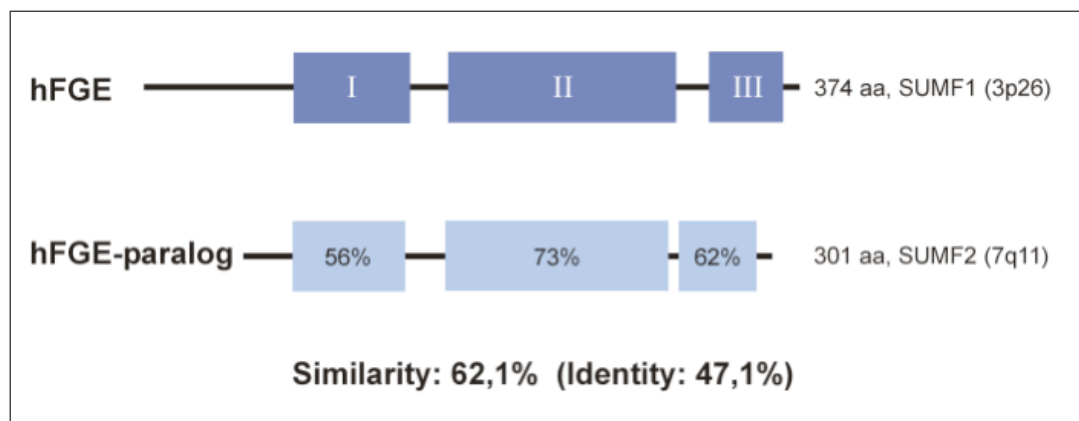
Figure 1.8. Human FGE and paralog FGE

arylsulfatase A, C, or E enhanced moderately, but consistently, the catalytic activity of the sulfatases, suggesting that pFGE has some FGly -generating activity[51]. So far, no mutations in the the *SUMF2* gene have been found in multiple sulfatase deficiency patients. In this thesis we have began to analyze the expression of *SUMF2*, structural and functional properties of recombinant and endogenous pFGE.

We have also examined the interaction of pFGE with sulfatases and determined the modifying activity of pFGE on sulfatases.

1.6.2   Effect of FGE on sulfatases activity

Normally, the necessary amount of FGE required to modify sulfatases is constitutively expressed in the cell. However, FGE is limiting to modify the recombinant sulfatase which is excessively expressed[52, 18]. Furthermore, the recombinant catalytically active sulfatases are needed for enzyme replacement therapy of patients suffering from the deficiency of lysosomal sulfatase, e.g., of arylsulfatase B in Maroteaux-Lamy disease, Iduronate sulfatase in Hunter disease, and galactose 6-/N-acetylgalactosamine 6-sulfatase in Morquio A disease (for review see Neufeld and Muenzer, 2001). In the present study, we have examined the utility of FGE for expression of functional sulfatases. And we also investigated the interaction between FGE and sulfatases.

Chapter 2

Materials and Methods

2.1   Materials

2.1.1   Laboratory equipment

Analytic balances type 1602 MP and 1265 MP	Sartorius, Goettingen
Analytical and preparative HPLC:	
SMART-system with the following columns:	Amersham Biosciences
Gel filtration columns:	
Fast Desalting PC 3.2/10 (3.2 x 100 mm)	Amersham Biosciences
Superdex 200 PC 3.2/30 (2.4ml)	Amersham Biosciences
Anion exchange columns:	
MonoQ PC 1.6/5 (0.1ml)	Amersham Biosciences
MonoQ HR 10/10 (8ml)	Amersham Biosciences
Reversed Phase columns:	
Aquapore RP-300 (C8, 2.1 x 220 mm)	
μPeak C2/C18 PC 3.2/3 (C2/C18, 2.1 x 30 mm)	Amersham Biosciences
UV-detectors for SMART-system:	
μPeak Detector	Amersham Biosciences
UV-Detector for Vision Workstation:	
FLUOR-305	PerSeptive Biosystems
Intelligent Dark Box II, Las-1000+	Fuji, Japan
Ice machine	Ziegra, Isernhagen
Centrifuges:	
Eppendorf centrifuge Type 5415C and 5402	Eppendorf, Hamburg
Table ultracentrifuge TL-100	Beckmann, München
Ulracentrifuge L8-70M	Beckmann, München

Cold Centrifuge J-21C and J2-MC   Beckmann, München
Labofuge GL   Heraeus Sepatech
 Rotors:

JA10, JA 20   Beckmann, München
Ti 45, Ti 60, Ti 70   Beckmann, München
TLA 45,   Beckmann, München
TLA-100.3   Beckmann, München
Electrophoresis chambers for agarose gels   Workshop of the Institute
Electrophoresis chambers for polyacrilamide gels   Workshop of the Institute
Liquid scintillation counter 1900TR   Packard, Frankfurt/Main
Gel dryer   Bio-Rad, Hilden
Magnetic mixer   IKA, Works, INC.
MALDI-TOF Mass Spectrometer, REFLEX III   Bruker Daltonics, Bremen
Microwave oven   Siemens, München
pH-Meter   Beckmann, München
Photometer, UV 160 A   Shimadzu, Kioto/Japan
UV-hand lamp (365/254nm), Type 5415 and 5402   Eppendorf, Hamburg
Vacuum concentrator model 100H   Bachhofer, Reutlingen
Vortex-Genie   Scientific Industries, USA.
DNA-Sequencer Type 310   ABI, PE Biosystems
Electroporator 1000 (used for bacteria)   Stratagene, USA
Confocal Laser Scanning Microscope   Leica, Bensheim
Leica TCS SP2 AOBS
(Ar: 488, 514 nm; He/Ne: 543 nm; 63x Oil Objective)
Incubators Innova 4230 and 4330   New Brunswick Scintific
Phosphoimager Fujix BAS1000   Fuji, Japan
Ultra turrax T8   IKA Labortechnik, Staufenv
Supersignal Chemiluminiscent Substrate   Pierce, Illinois
Thermocycler GeneAmp PCR system 9600   Perkin-Elmer Cetus

2.1.2   Chemicals, plasticware and membranes

Chemicals   Boehringer/Roche, Mannheim,
                                   Merck, Darmstadt,
                                   Roth, Karlsruhe,
                                   Serva, Heidelberg,
                                   Sigma, Deisenhofen
Cell culture plasticware   Greiner, Frickenhausen

	Nalge Nunc International, Denmark
Nitrocellulose membrane	Schleich and Schüll, Dassel
PVDF membrane, 0.2μM	Schleich and Schüll, Dassel
Hybond-N Nylon membrane	Amersham Biosciences, UK
Whatman GB002 paper	Schleich and Schüll, Dassel
Whatman GB003 paper extra thick	Schleich and Schüll, Dassel

2.1.3   Kits, spin columns and reagents

DNA, RNA:

HiSpeed Plasmid Midi kit	Qiagen
Omniscript Reverse Transcription kit	Qiagen
PCR purification kit	Qiagen
QIAprep Spin Miniprep kit	Qiagen
QIAquick Gel Extraction kit	Qiagen
Rapid-hyb hybridisation buffer	Amersham Pharmacia Biotech
rediprimeTMII labeling system	Amersham Pharmacia Biotech
RNAlater RNA Stabilisation Reagent	Qiagen
RNase inhibitor RNase Out	Invitrogen
RNeasy Mini and Midi Kit	Qiagen
TA Cloning Kit	Invitrogen
Effectene Transfection kit	Qiagen
Fugene Transfection reagent	Roche
Lipofectamine 2000	Invitrogen

Protein:

Bio-Rad Protein Assay	Bio-Rad
DAKO fluorescent mounting medium	DakoCytomation, USA
ECL Plus	Amersham Biosciences
ConcanavalinA-Sepharose 4B	Amersham Biosciences
Ni-agarose	Qiagen
Protease Inhibitor Cocktail	Sigma
PD-10 Sephadex G-25 M column	Amersham Pharmacia Biotech
Sep-Pak C18 Cartridge	Waters, USA
Stimune Adjuvant	Cedi Diagnostics B. V., Netherlands
Supersignal Chemiluminescence Kit	Pierce, USA
Roti-blue Colloidal Coomassie Brilliant Blue	Roth

Titermax Gold Adjuvant   Sigma
Vivaspin   Vivascience, Goettingen
Bovine Serum Albimin (BSA)   Serva
Iodoacetamide, Iodoacetic acid   Sigma
PANSORBIN cells   Calbiochem
Protein A sepharose   Sigma
Protein G agarose   Sigma
Prestained Marker   Biorad

2.1.4   Vectors and DNA standards

pBI   BD CLONTECH
pSB 4.7 pA   Transkaryotic Therapies Inc, Cambridge, MA
pSV-pac   gift from Prof. Stefan Höning
pGK Hygro   Invitrogen life technologies
pUB/Bsd   Invitrogen life technologies
1-kb DNA ladder   Gibco BRL

2.1.5   Antibiotics and drugs

Ampicillin   Serva
Blasticidin-S Hydrochloride   Invitrogen life technologies
Hygromycin   Calbiochem
Neomycin (Gentamycin sulfate or G418)   Gibco
Penicillin/Streptomycin   Gibco
(100 x =10,000 U/ml)

2.1.6   Radioactive substances

$[^{35}S]$-Methionine, 10 mCi/ml   Amersham Pharmacia Biotech

2.1.7   Enzymes, substrates and nucleotides

Restriction endonucleases   New England Biolabs
Klenow DNA polymerase   New England Biolabs
DNA ligase   New England Biolabs
Taq DNA polymerase   Amersham Pharmacia Biotech
Alkaline phosphatase   Boehringer

Ultrapure dNTP Set	Amersham Pharmacia Biotech
Adenosine 5'-triphosphate (ATP)	Sigma
Endoglycosidase H, PNGase	Roche Diagnostics

2.1.8 Primary antibodies

antigen	type	western blot	reference
RGS-His$_6$ tag	mouse mAb	1:2000	Qiagen
HA tag	mouse mAb	1:2000	Covance Inc., Princeton
pFGE	rabbit pAb	1:2500	this study
pFGE	mouse mAb	1:2500	this study
FGE	rabbit pAb	1:2500	this study
FGE	mouse mAb	1:2500	this study
Galactose-6-sulfatase	mouse mAb	1:1000	Transkaryotic Therapies, Inc, USA
Steroid sulfatase	rabbit pAb	1:5000	
Arylsulfatase A	rabbit pAb	1:10000	

2.1.9 Secondary antibodies

Goat anti-rabbit Horseradish peroxidase conjugate
Goat anti-mouse Horseradish peroxidase conjugate
Goat anti-rabbit Cy3 conjugate
Goat anti-mouse Cy2 conjugate
Goat anti-mouse Cy3 conjugate
Goat anti-mouse Cy2 conjugate
All secondary antibodies were purchased from Dianova, Hamburg.

2.1.10 Stock solutions and buffers

1 M Sodiumphosphate buffer
1 M solution of sodium di-hydrogen phosphate was slowly added to 1 M di-sodium hydrogenphosphate solution with constant mixing on a magnetic stirrer till the pH came down to 7.4.
10 x PBS
100 mM sodium phosphate pH 7.4
9 % sodium chloride
Dissolved in 800 ml water and pH was adjusted to 7.4 with HCl, volume was made

up to 1000 ml and autoclaved. Stored at room temperature.

1 x TBS

10 mM Tris/ HCl pH 7.4

150 mM Sodium chloride

1 x TAE 0.04 M Tris-acetate

1mM EDTA (pH 8.0)

50x TAE

242 g Tris base

57.1 g glacial acetic acid

100 ml of 0.5 M EDTA (pH 8.0)

Dissolved in water and the final volume was made upto one litre.

TE Buffer

10 mM Tris/ HCl pH 7.5

1 mM EDTA

2.2 Molecular Biology Methods

2.2.1 Cultivation of E.coli

Luria Bertani (LB) medium

10 g Bacto-Tryptone

5 g Bacto-yeast extract

5 g NaCl

Dissolved in 900 ml distilled water, pH adjusted to 7.0 with 10 N NaOH, made up the volume to one liter, sterilized by autoclaving and stored at room temperature.

LB-Ampicillin Agar Plates

1.5% of Agar was added to the LB medium and autoclaved. After autoclaving, the medium was let to cool down to $55°C$ and ampicillin was added to a final concentration of $100\mu g/ml$. This medium was poured into 10 cm petriplates in the hood and left undisturbed for about 30 min to solidify. LB-Agar plates were stored in the cold room.

Preparation of competent E. coli cells

Buffers

TFB I	30 mM Calcium acetate, pH 5.8
	100 mM Rubidium chloride

	10 mM Calcium chloride
	50 mM Manganese chloride
	15% Glycerol (w/v)
TFB II	10 mM MOPS, pH 6.5
	75 mM Calcium chloride
	10 mM Rubidium chloride
	15% (w/v) Glycerol

2-3 ml bacterial pre-culture was grown in LB medium at 37°C overnight. 1 ml of the pre-culture was then inoculated into 99 ml of LB medium and grown at 37°C to an O.D 600 of 0.4-0.6. Cells were pelleted at 3,000 rpm for 5 min at 4°C, resuspended cell pellet in 1/3 of the original volume ice-cold buffer TFB I and incubated on ice for 5 min. Cells were centrifuged at 3000 rpm for 5 min at 4°C, the pellet was resuspended in about 1/25 of the original volume ice-cold buffer TFB II and incubated on ice for 30 min. Aliquots of 50μl were frozen in liquid nitrogen and stored at -80°C.

2.2.2 Transformation of E.coli competent cells

50-100 ng of DNA was added to each 50μl aliquot of competent cells and incubated on ice for 30 min. Cells were subjected to heat shock by incubating at 42°C for 2 min and incubated on ice for 1-2 min. 0.9 ml of LB medium was added to the cells followed by incubation at 37°C in the shaker for 1 hour. Cells were plated on LB-agar plates containing appropriate antibiotic. For blue-white selection, cells were plated on plates containing IPTG (Isopropyl-β-Dthiogalactopyrahoside; 40 μl of 100 mM solution in water) and X-Gal (40μl of 4% solution in dimethylformamide).

2.2.3 Preparation of electrocompetent DH5α cells

10% (v/v) Glycerol : 1 ml (1.26 g) of glycerol in 10 ml sterile water
A single E. coli colony was inoculated into 5 ml of LB media and allowed to grow overnight at 37°C in a shaker incubator. 2.5 ml of this pre-culture was inoculated into 500 ml LB media and allowed to grow to an OD 600 of 0.5-0.7 at 37°C. Cells were pre-chilled on ice for 15 min and then pelleted at 5000 rpm for 15 min at 4°C. Pellet was resuspended in 500 ml of ice cold water and centrifuged as described above. This washing was repeated one more time. To the pellet an equal volume of 10% glycerol was added and resuspended, aliquots of 100μl were stored at -80ºC.

2.2.4   Transformation of the electrocompetent cells

SOC medium

0.5% Yeast extract

2% Bacto-tryptone

10 mM Sodium chloride

2,5 mM Potassium chloride

10 mM Magnesium sulfate

10 mM Magnesium chloride

20 mM Glucose

For each electroporation 40μl of the electrocompetent cells were used. 0.5ng of DNA was added to the cells and the contents were transferred into a pre-chilled electroporation cuvette. The cuvette was placed in the electroporater and pulse was applied. SOC medium was added to the cells and they were allowed to recover in sterile tubes for 30 min at 37°C in a shaker incubator. Cells were plated on LB plates containing appropriate antibiotic.

2.2.5   Glycerol stocks of bacterial strains

Bacterial cultures were grown overnight at 37°C in a shaker-incubator. 0.3 ml of sterile 100% glycerol was taken in freezing vials to which 700μl of the overnight culture was added. The contents were gently mixed, shock frozen on dry ice and stored at -80°CC.

2.2.6   Mini preparation of plasmid DNA

Plasmid DNA was isolated from E.coli cultures using the Qiagen. Isolation was done according to the instructions of the manufacturer.

Buffer P1	50 mM Tris/ HCl pH 8,0
	10 mM EDTA
	100 μg/ml RNase A
Buffer P2	0.2 M NaOH
	1% SDS
Buffer P3	3 M Potassium acetate pH 5.5

Single E. coli colony was inoculated into 3 ml of LB medium containing 100μg/ml of antibiotic and grown overnight at 37°C in a shaker incubator. Cells were pelleted in a table-top centrifuge at 3,000 rpm for 5 min. Cell pellet was

resuspended in 250 μl of buffer P1 and 250μl of buffer P2 was added, mixed gently by inverting the tube 4-6 times. To this, 350 μl of buffer P3 was added and gently mixed and centrifuged for 10 min at 13,000 rpm in a table-top eppendorf centrifuge. The supernatant was applied onto a QIAprep spin column and centrifuged for 1 min at 13,000 rpm. Flow through was discarded, the column was washed with 0.75 ml of buffer PE and centrifuged again for 1 min. Flow through was discarded and the column was centrifuged for an additional 1 min to remove any residual wash buffer. The column was placed in a clean eppendorf tube and 50 μl of double distilled water was added directly to the centre of the column. The column was let to stand for 1 min and DNA was eluted by centrifuging at 13,000 rpm for 1 min.

2.2.7 Determining the concentration of DNA

DNA concentration was determined using a spectrophotometer. DNA was diluted in water and the absorbance was measured at 260 nm. Absorbance or optical density (OD) of 1 at 260 nm corresponds to $\sim$50μg/ml of double stranded DNA or $\sim$40μg/ml of single stranded DNA and RNA or $\sim$20μg/ml of oligonucleotides. The ratio between the readings at 260 nm and 280 nm (OD260/ OD280) provides an estimate of the purity of the nucleic acid. Pure preparations of DNA and RNA have OD-260/OD-280 values of 1.8 and 2.0, respectively. Any contamination with proteins or phenol would yield values less than mentioned above.

2.2.8 Restriction endonuclease digestion of DNA

The activity of restriction enzymes is measured in terms of 'Units' (U).
One unit of restriction enzyme is the amount of enzyme required to completely digest 1μg substrate DNA in 1hr.

Plasmid DNA	0.5 - 1μg
10 x buffer	2μl
Restriction enzyme	1-2 U
BSA	1μg/μl (added where essential)

Water was added to a final volume of 20μl. Reaction mix was incubated at 37°C* for 2 hours.
*Incubation temperatures were set as recommended by the manufacturer (New England Bio Labs Beverly U.S.A.) which varies from enzyme to enzyme. Reaction mix was analyzed on agarose gel (0.8 to 2% depending on the size of the DNA of interest) The buffers and enzymes used were all from New England Biolabs. The

composition of the buffers is as follows:

NEB 1: 10 mM Bis Tris Propane-HCl (pH 7.0), 10 mM $MgCl_2$, 1 mM DTT

NEB 2: 10 mM Tris-HCl (pH 7.9), 10 mM $MgCl_2$, 50 mM NaCl, 1 mM DTT

NEB 3: 50 mM Tris-HCl (pH 7.9), 10 mM $MgCl_2$, 100 mM NaCl, 1 mM DTT

NEB 4: 20 mM Tris-acetate (pH 7.9), 10 mM Magnesium acetate, 1 mM DTT, 50 mM Potassium acetate

2.2.9   Agarose gel electrophoresis of DNA

The size and purity of DNA is analyzed by agarose gel electrophoresis. Concentration of agarose used for analysis is inversely proportional to the size of the DNA of interest, that is, the larger the DNA the lower the concentration of agarose.

Agarose concentration (%)	DNA size (kb)
0.7	20 - 1
0.9	7 - 0.5
1.2	6 - 0.4
1.5	4 - 0.2
2.0	3 - 0.1

Gel loading buffer (10x)   0.25% (w/v) Bromophenol blue
40% Saccharose in 1x TAE

Agarose was weighed and dissolved in 1 x TAE by boiling in microwave oven. The agarose solution was cooled to 60°C and ethidium bromide was added to a final concentration of 0.5μg/ml. This was poured into the agarose gel cassette and allowed to polymerize completely. The sample DNA was mixed with gel loading buffer and loaded onto the gel. The gel electrophoresis was carried out at 100 V. Ethidium bromide is a fluorescent dye which contains a planar group that intercalates between the stacked bases of the DNA. The fixed position of this group and its close proximity to the bases cause the dye to bound to DNA. It results in an increased fluorescent yield compared to that of the dye in free solution. Ultraviolet radiation at 254 nm is absorbed by the DNA and transmitted to the dye; radiation at 302 nm and 366 nm is absorbed by the bound dye itself. In both cases, the energy is re-emitted at 590 nm in the red orange region of the visible spectrum. Hence DNA can be visualized under a UV transilluminator. The gel was photographed using a gel documentation system.

2.2.10 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'-ATGCGCAGCGGGGCCGTGG-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 entry 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.

2.2.11 Construction of expression plasmids

The pFGE encoding cDNA was equipped with a 5' BamHI site and a 3' RGS-His6 tag sequence followed by a stop codon and a XbaI site, by add-on PCR using Pfu Ultra polymerase (Stratagene) and primers 5'-CGGGATCCACCATGGCCCGGCATGGGTTAC-3' (BamHI) and 5'-GCTCTAGATTAGTGATGGTGATGGTGATGCGATCCTCTCAGCTCC CCTGGCGGC-3' (RGS-His6-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- and 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'-ATAGTTTAGCGGCCGCTTATGCGTAGTCAGGCACATCATACGGATAGTCCATGGTGGGC AGGC-3' (FGE-HA-NotI); 5'-ATAGTTTAGCGGCCGCTTATGCGTAGTCAGGCACATCATA CGGATACAGCTCCCCTGGCGGC-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 as template, Pfu Ultra polymerase, and add-on primers 5'-CTAGCTAGCCACCATGCCTTTAAGGAAGATGAAG-3' (NheI) and 5'-GGATATCAGCGGCTCAGTCTCTTATCC-3' (EcoRV).

2.2.12   Northern blot analysis

A human multi tissue Northern blot (BD Biosciences) was hybridized with 32P-labelled 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 in Fig.3.2.A, was obtained when using probes hybridizing in the 3' UTR regions of SUMF1 or SUMF2 mRNA, where 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 refer[53].

2.3   Cell culture and transfections

2.3.1   Basics

Normal growth medium:	DMEM, 10% fetal calf serum, Pen-Strep
Trypsin-EDTA solution:	0.5 g/l Trypsin, 0.2 g/l EDTA
phosphate-buffered saline (PBS):	150 mM NaCl, 120 mM KCl,
	10 mM Na_2HPO_2/KH_2PO_2 pH 7.4
Freezing medium:	10% DMSO in normal growth medium

HT1080 cells and Human skin fibroblasts cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum (FCS) and 1% penicillin/streptomycin (Invitrogen) under 5% CO_2 at 37°C.

Passaging by trypsinisation

Cell line was passaged every 2-7 days depending on the growth rate. After two washes with PBS, 0.5 ml trypsin solution per 25 cm^2 was added to the cell layer. The cells were incubated at 37 °C for 5 min, then singled by tapping to the dish and

resuspended in medium. To concentrate, wash or harvest cells, centrifugation was always done at 500xg for 5 min.

Freezing cells for stock maintenance

Usually, cells from a confluent 35 cm^2 flask were frozen as 2-4 vials. Cells were pelleted after trypsinisation, resuspended in 1 ml freezing medium per vial, immediately placed on ice and stored at -80 °C overnight. The vials were transferred to the liquid nitrogen container the following day.

Thawing cells

The frozen vial was taken out of liquid nitrogen, the cap was carefully loosened to release pressure and the vial was thawed in 70% ethanol at 37 °C, until the suspension had almost completely melted. Under the hood, the cells were then transferred to a tube containing 5 ml chilled cell culture medium. The cells were pelleted, resuspended in prewarmed culture medium and transferred to a culture flask. Medium was replaced on the next day.

2.3.2 Transfections

Transient transfection

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 hr, the medium was replaced with medium containing 1 μg/ml doxycyline (BD Biosciences), and 30 hr later, the cells were harvested for further analysis.

Stable transfection

HT1080 cells in 6cm dishes were transfected with pSB-pFGE-His or FGE-His using Lipofectamine 2000 as mentioned above. After 36h, cells were transferred to two 100-mm dishes at a 1:20 dilution and cultured for 2 weeks in normal growth

medium containing genticin (G418 sulfate; Invitrogen) with an increasing concentration from 200 to 800μg/ml. Drug resistant clones were picked and expanded in growth medium containing genticin (800μg/ml). Clones were screened by western blot analysis for expression of the appropriate protein. HT1080 cells stably expressing galactose-6-sulfatase (kindly gifted by Transkaryotic Therapies Inc, Cambridge, MA) were transfected with pSB-pFGE-His or FGE-His plus puromycin resistance vector (10:1 ratio) using Lipofectamine 2000. After 36hr, cells were passaged to two 100 mm dishes and cultured for two to three weeks in normal growth medium containing puromycin (Invitrogen) with an increasing concentration from 0.2 to 0.8μg/ml. The drug resistant clones were expanded and screened by western blotting.

2.4 Biochemical Methods

2.4.1 Analysis of Protein

Protein estimation by BIORAD reagent

Bovine Serum Albumin (BSA) stock solution 1 mg/ml
Concentration range : 2 -16μg/μl
A standard curve was made using BSA in the range of 2 -16μg. 10μl of the sample was used for the protein estimation. The sample volume was made up to 800μl with water. 200μl of the BIORAD reagent was added to each samples, votexed and incubated for 3 - 5 min at room temperature, 200μl from each tube was pipetted into ELISA strips and optical density was measured at 595 nm in the ELISA reader.

Solubilisation of proteins

2 x Laemmli buffer
125 mM Tris-Cl, pH 6.8
4% SDS
0.004%Bromophenol blue
20% (w/v) Glycerol
10% β-mercaptoethanol
Protein samples after speed vac were resuspended in 1x Laemmli buffer. The resuspended proteins were boiled at 95°C for 5 min, cooled on ice for 1 min, centrifuged at 13,000 rpm for 2 min and resolved by SDS-PAGE.

2.4.2   SDS-Polyacrylamide Gel Electrophoresis (Laemmli et al., 1970)

The elecrophoretic separation of the proteins was performed through high-tris discontinual SDS polyacrylamide gel electrophoresis. A system with vertical oriented glass plates (16 x 16 cm; 1 mm spacer) was used. The gels were prepared as follows: first the resolving gel (12.5-17.5% acrylamide) was put between glass plates and covered with a layer of butanol. After polymerisation of the gel was completed, butanol was removed and the space between glass plates was dried with Watmann paper. Then the concentrating gel was put on top of the polymerised resolving gel and the sample combs were introduced immediately. After approximately 15 min polymerisation was completed and the sample combs were removed. The glass plates with the gel between them were fixed inside the electrophoresis chamber and covered with electrophoresis buffer. Protein samples were mixed with Laemmli loading buffer 1:1, denatured by 95 °C and centrifuged with 14000 rpm. The supernatant was introduced into the stacking gel pockets. Electrophoresis was performed for 2-3 hours with the constant current of 35 mA.

Electrophoresis buffer:

SDS (w/v) – 10g/10 litre

Glycine (w/v)– 144,27g/10 litre

Tris (w/v) – 60,53g/10 litre

20% Ammoniumperoxiddisulfate (APS) in ddH2O

Stacking gel(1X)

Ingredients	5%
30% Acrylamide solution: (ml)	0.4
1% bisacrylamide solution (ml)	0.275
ddH2O (ml)	1.175
0,5 M Tris-HCl, pH 6.8 (ml)	0.625
(10%) SDS (μl)	25
TEMED (μl)	2.5
(20%) APS (μl)	10

Separating gel (1X)

Ingredients	15%
30% Acrylamide solution: (ml)	8.75

1% bisacrylamide solution (ml)	2.74
ddH2O (ml)	1.4
1.5 M Tris-HCl, pH 8.8 (ml)	4.3
(10%) SDS (μl)	175
TEMED (μl)	14.58
(20%) APS (μl)	58.33

2.4.3   Detection of proteins in polyacrylamide gels

Staining with Roti-Blue Colloidal Coomassie

Colloidal Coomassie staining is one of the most sensitive staining protocols. Due
to the colloidal properties of the dye, it binds with high specificity to proteins and
only minimal to the gel matrix. This allows visualisation of proteins separated by
SDS-PAGE with sensitivity as high as 30ng of protein. Roti-Blue collodial Coomassie
was purchased from Roth and staining was performed according to the protocol as
follows. Immediately after completion of electrophoresis, gels were incubated in
the fixing solution for 60 min, shaking. The staining solution was applied on to
the gels for 60 min. In special cases the gels were stained for several days, which
improved the sensitivity of the staining. After staining, gels were incubated in the
washing solution for 5-10 min and then kept in the stabilising solution or prepared
for drying in the drying solution.

Fixing solution:	20% (v/v) methanol
	8.5% (v/v) o-phosphoric acid
Staining solution:	20% (v/v) methanol
	20% (v/v) 5x concentrate Roti-Blue colloidal Coomassie
Washing solution:	25% (v/v) methanol
Stabilising solution:	20% (w/v) ammonium sulphate
Drying solution:	10% (v/v) glycerol
	20% (v/v) ethanol

2.4.4   Western blot (semi-dry)

Cathode buffer	40 mM -aminocaproic acid, 20 mM Tris-Cl, 20% (v/v) methanol, pH 9.0 (pH was adjusted with free Tris base before addition of methanol)
Anode buffer	75 mM Tris-Cl, 20% (v/v) methanol, pH 7.4
Phosphate-buffered saline (PBS)	140 mM NaCl, 2.7 mM KCl, 10mM-

	Na_2HPO4, 1.8 mM KH_2PO_2, pH 7.3
PBST	0.1% (v/v) Tween-20 in PBS
Blocking buffer	5% milk powder in PBST
Glycine stripping solution	0.2 M glycine pH 2.8, 0.5 M NaCl
SDS stripping solution	16 mM Tris-Cl pH 6.8, 2% SDS,
	0.1 M -mercaptoethanol

Six pieces of 3 mm Whatman paper and one piece of nitrocellulose membrane were cut to the size of the SDS gel. Gel and membrane were equilibrated for 5-15 min in cathode buffer. The blot was assembled without air bubbles according to the following scheme:

cathode (-)
3x paper in cathode buffer
gel
membrane
3x paper in anode buffer
anode (+)

For transfer, the current was set to 1 mA/cm^2 gel size for 45-60 min. The membrane was then briefly washed with PBST and incubated in blocking buffer for one hour at room temperature. Decoration with the primary antibody diluted in blocking buffer occurred overnight at 4 °C. After three 30-45 min washes with PBST, the membrane was incubated with horse radish peroxidase (HRP)-coupled secondary antibody, diluted 1:10 000 in blocking buffer, for one hour at room temperature. The blot was washed three times 30-45 min with PBST and incubated with chemiluminescence substrate solution. Intelligent Dark Box II Camera was used for signal detection.

Stripping of nitrocellulose membranes

Removal of antibodies from a blot was done under mild conditions, if it should serve to reduce the background for incubation with another primary antibody, either from a different species or for a protein of clearly distinct size than in the first decoration. After washing the membrane in PBST, it was incubated 5-20 min in glycine stripping solution. The solution was neutralised with 1 M Tris-Cl pH 8.5, followed by several washes in PBST. If it was crucial to remove antibodies completely, the blot was incubated in SDS stripping solution for 30 min, tightly closed, on a wheel at 50 °C, followed by several washes in PBST. Decoration of the membrane was done as described above, starting from the blocking step again.

2.4.5   Detection of radioactively labelled polypeptides

For detection of radioactively labelled polypeptides polyacrylamide gels were fixed in destaining solution for 15 min and then washed with water for 30 min. Then the gels were dried between cellophane layers in the gel-dryer for 2 hours. For visualisation of the incorporated radioactive label the gels were exposed on to the Imaging plate (Fuji) overnight or longer period at RT. The image on the plate was analysed by the program Image reader (Fuji-Film, Vers. 1.4 E) with the help of the BAS 1000 phosphoimager (Ray-test).

2.4.6   Staining with silver (Schevchenko et al., 1996)

Staining of proteins separated on SDS-gel with silver allows quick and effective visualization of even small amounts of protein. After completion of electrophoresis gels were incubated with shaking in fixing solution for 1 hour. After that gels were washed 2 times for 20 min in 30% ethanol and equilibrated for 20 min in water. Equilibrated gels were washed for 1 min in 0.03% Na_2S_2O3, washed with water and incubated in staining solution for 20 min at 4°C. Then the gels were washed in water twice for 30 seconds and developed in the developing solution until desired intensity of the bands was obtained. Developing was stopped by transferring the gels into a new chamber with 5% acetic acid. Gels were kept in 1% acetic acid or washed with water to prepare them for drying.

Fixing solution:	10% (v/v) acetic acid
	40% (v/v) ethanol
Staining solution:	0.2% (w/v) silver nitrate
	0,008% (v/v) formic aldehyde
Developing solution:	3% (w/v) sodium bicarbonate
	0,018% (v/v) formic aldehyde

2.4.7   In vitro assay for FGE

For monitoring the activity of FGE, the N-acetylated and C-amidated 23-mer peptide P23 (MTDFYVPVSLCTPSRAALLTGRS) was used as substrate. The conversion of the cysteine residue in position 11 to FGly was monitored by MALDI-TOF mass spectrometry. A 6 μM stock solution of P23 in 30% acetonitrile and 0.1% trifluoroacetic acid (TFA) was prepared. Under standard conditions, 6 pmol P23 were incubated at 37 °C with up to 10 μl enzyme in a final volume of 30 μl 50 mM Tris/HCl (pH 9.0), containing 67 mM NaCl, 15 M $CaCl_2$, 2 mM DTT, and 0.33 mg/ml bovine serum

albumin. To stop the enzyme reaction, 2 μl 10% TFA was added. P23 then was bound to ZipTip C18 (Millipore), washed with 0.1% TFA, and eluted in 3 μl 50% acetonitrile, 0.1% TFA. 0.5 μl of the eluate was mixed with 0.5 μl of matrix solution (5 mg/ml -cyano-4-hydroxy-cinnamic acid (Bruker Daltonics) in 50% acetonitrile, 0.1% TFA) on a stainless steel target. MALDI-TOF mass spectrometry was done with a Reflex III (Bruker Daltonics) using reflectron mode and laser energy just above the desorption/ionization threshold. All spectra were averages of 200-300 shots from several spots on the target. The mass axis was calibrated using peptides of molecular masses ranging from 1000-3000 Da as external standards. Monoisotopic MH of P23 is 2526.28 and of the FGly containing product 2508.29. Activity (pmol product/hr) was calculated on the basis of the peak height of the product divided by the sum of the peak heights of P23 and the product.

2.4.8 Determination of FGE activity in cells

The trypsinized cells were washed with phosphate-buffered saline (PBS) and resuspended in 10 mM Tris (pH 8.0), containing 2.5 mM DTT, the proteinase inhibitor cocktail (Sigma). The cells were lysed by ultrasonication and cleared by centrifugation at 125,000 Xg for 30min. The supernatant was subjected to chromatography on a MonoQ PC 1.6/5 column using a 1 ml gradient of 0 to 375 mM NaCl in the Tris buffer described above and a flow rate of 0.1 ml/min. Fractions eluting at 50-200mM NaCl were pooled, lyophilized, and reconstituted in one-tenth of the original pool volume prior to determination of FGE activity with peptide P23.

2.4.9 Sulfatase assays

Activity of ASA, STS, and galactose-6-sulfatase were determined as described (Rommerskirch and von Figura, 1992; Glössl and Kresse, 1978).

2.4.10 Purification of recombinant pFGE-His

HT1080 cells stably overexpressing pFGE-His were grown to near confluence in normal growth medium. After 48h, the medium was collected and fresh medium was added. Medium can be collected up to 8 times for purification of pFGE-His. The medium was spun down at 1100 X g (JA10 rotor) for 30 minutes and the supernatant was precipitated by 50% ammonium sulfate. The precipitated medium was centrifuged at 11300 Xg (JA10) for 30 minutes and the pellet was reconstituted in lysis buffer (50mM NaH_2PO_4, 300mM NaCl and 10mM Imidazole).

The suspension was dialyzed against lysis buffer for 16h. The dialysed material was cleared (142414 X g, Ti45 rotor, 30min, 4°C) and subjected to Ni^{2+}-NTA affinity chromatography under the conditions recommended by the manufacturer (Qiagen). Peak fractions, analysed by western blotting were pooled and dialysed against buffer A (10mM Tris-Hcl pH8.0, 100mM NaCl, 2.5mM DTT) for overnight. The dialysed pFGE-His was applied to a MonoQHR10/10 anion exchange column (Amersham Biosciences) equilibrated with Buffer A. The unbound pFGE-His was recovered in the flow through and other contaminants were eluted with a linear gradient to 400 mM NaCl in buffer A. The flow through fractions were pooled and concentrated by Ultrathimble (Schleicher and Schuell). 50 μg of pFGE-His applied to a Superdex 200 gel-filteration column (Amersham Biosciences) that had been equilibrated in PBS and eluted isocratically with PBS. Peak fractions were run on a 15% SDS polyacrylamide gel and followed by staining with Coomassie Brilliant Blue.

2.4.11 Immunisation of rabbits and serum preparation

Rabbits were housed by the animal facility of the medical faculty.

Preparation of antigen emulsions for immunisation:

300 μg and 200 μg of purified pFGE-His from secretion was emulsified for the first injection and subsequent booster injections, respectively. Stimune or Titermax Gold served as adjuvant. Pre-immune serum was collected before immunization to use as a negative control in western blots and immunoprecipitations. Booster injections were given once in two weeks and bleeds were collected two weeks after booster injections and analyzed by Western blot. The rabbit was sacrificed two weeks after the fourth booster injection.

Preparation of serum from blood samples:

Blood samples were incubated for 4 h at room temperature or overnight at 4°C to allow for clot formation and centrifuged for 15 min at 13,000 Xg and 4°C. The supernatant was carefully transferred into new tubes. Small aliquots were either frozen at -20 °C or at -80 °C for long term storage.

2.4.12 Gel filtration on a Superdex-200

HT1080 cells stably expressing galactose-6-sulfatase and FGE-His or pFGE-His were grown to near confluence in 100-mm dishes. Cells were harvested by scrapping and

lysed in 2ml of PBS by ultrasonication. The cell lysate was cleared by centrifugation at 125,000 Xg for 30 min at 4 °C. 50 μl of supernatent was subjected to Superdex-200 PC 3.2 (2.4 ml) which had been equilibrated with PBS. The column was eluted with the same buffer at flow rate of 0.04 ml/min. Fractions, each 0.08 ml were collected and run on 15% SDS-polyacrylamide gel. After transferring to nitrocellulose membrane, probed with a mixture of monoclonal antibodies against galactose-6-sulfatase and His-tag.

For coelution of purified pFGE-His and FGE-His, both were mixted in a 2:1 molar ratio and incubated at room temperature for 30 min. The reaction was carried out in PBS at pH 7.4. Then the mixture was passed through Superdex-200 as described above. Protein fractions were analysed by Western blot using His-tag monoclonal antibody.

2.4.13 *In Vivo* and *in vitro* protein-protein cross-linking

In vivo– HT1080 cells trasiently expressing pFGE and FGE were grown on 60-mm plates. Cells were washed two times with phosphate-buffered saline (PBS) at room temperature and mixed with 5 ml of PBS, and dithiobis (succinimidyl propionate) (DSP) dissolved in Me_2SO was added to a final concentration of 2 mM. The plates were incubated on ice for 30 min, and the reaction was quenched by adding 5 ml of TE buffer (50 mM Tris-HCl, pH 7.6). Cells were further washed twice with PBS and harvested by scrapping. The cell pellet was resuspended in PBS and lysed by ultrasonication. For western blot analysis, equal volume of cell extracts were applied for SDS-PAGE under reducing and non-reducing conditions (with or without 50 mM DTT), transferred to nitrocellulose membrane, and probed with FGE or pFGE antiserum and horse radish peroxidase-conjugated goat anti-rabbit antibody, as primary and secondary antibody, respectively.

In vitro– Purified pFGE-His and FGE-His were mixed in 2:1 molar ratio and incubated at 37 °C for 15 min. The reaction was carried out in PBS. After 15 min, the various concentrations (0.5, 0.25, 0.125 and 0.062 mM) of DSP was added to the reaction mixture. The reactions were incubated at room temperature for 60 min and quenched by adding 5 ml of 50 mM Tris-HCl, pH 7.6. Samples were electrophoresed on a 15% SDS-PAGE, transfered onto nitrocellulose membrane and detected with FGE or pFGE antiserum.

2.4.14   Treatment with N-ethylmaleimide (NEM)

HT1080 cells were transiently transfected with pFGE and FGE or pFGE. After 24h of transfection, cells were washed with PBS and incubated with 10 mM NEM (Fluka)/PBS for 10 min at 37 °C. Controls were incubated with PBS without NEM. After trypsinization, cells were harvested by centrifugation, washed once with PBS, and lysed in PBS containing 10 mM NEM. Protein concentration of cell extracts was determined by Bradford's reagent (Bio-Rad). For Western blot analysis, equal amounts of proteins were applied for SDS-PAGE under reducing and non-reducing conditions (with and without 5% mercaptoethanol), transferred to nitrocellulose, and probed with a mixture of pFGE and FGE antiserum and horseradish peroxidase-conjugated goat anti-rabbit antibody as primary and secondary antibody, respectively.

2.4.15   Indirect immunofluorescence

Cells were seeded on round coverslips in a 24-well plate 1-2 days in advance and grown to 50-70% confluency. All the incubation steps given were performed by filling the wells with 0.2 ml of the respective solution unless indicated otherwise. Solutions were changed carefully for the cells should neither be washed away nor become dry. The lid of the cell culture plate was closed during prolonged incubations and upon application of the secondary, fluorescent antibody the samples were protected from light.

Fixation with para-formaldehyde:

PFA reacts with nitrogen-containing groups and in effect crosslinks the proteins in a cell.
3.7% para-formaldehyde (PFA):
The required amount of PFA was suspended in 1/5 of the final volume, mixed with 50 μl of 1 M NaOH/g PFA and dissolved by incubation at 60 °C for about 1 h. The solution was stored for 1-2 weeks at 4 °C.
0.5% saponin-PBS:
10% stock solution, stored at -20 °C

Protocol

Cells were washed once with PBS and fixed with 3.7% PFA for 20 min at room temperature. After two washes with PBS, aldehyde groups were blocked by 50 mM NH$_4$Cl for 10 min at room temperature. The cells were washed twice with PBS and permeabilised for 10 min at room temperature with 0.5% saponin in PBS. After this step, all PBS solutions should contain 0.1% saponin. After permeabilization, cells were wased once with PBS/saponin and incubated with primary antibody for 1hr/37 °C. Then cells were washed twice with PBS/saponin and incubated with 10% goat serum in PBS. After incubation, cells were washed once with PBS/saponin and incubated with secondary antibody coupled with flurochrome for 30 min at 37 °C. Then cells were washed twice with PBS and left them in water. The coverslips were carefully taken out with a forceps, dried from excess water and adhered upside-down to a drop of DAKO mounting medium on a glass slide. The slides were protected from light to dry overnight and sealed with nail-polish the next day.

Indirect immunofluoresence of fibroblasts and HT1080 cells

To detect endogenous expression of pFGE, the human skin fibroblast cells were grown on cover slips for two days. For recombinant expression of pFGE, the HT1080 cells were grown on cover slips and cotransfected with pBI-pFGE.HA and rtTA using Lipofectamine 2000 (Invitrogen) as described above. After 36hr of Transfection, the cells were analysed for indirect immunofluoresence as described above. For the colocalization studies, rabbit pFGE-His antiserum and monoclonal antibodies against HA tag (Covance, Princeton) were used as primary antibodies. The endoplasmic reticulum marker protein, Protein Disulfide Isomerase (PDI) was detected with a monoclonal antibody of different subtypes (IgG2A, Stressgen). The primary antibodies were detected with isotype specific goat secondary antibodies coupled to Cy2 or Cy3 (Molecular Probes), respectively. Confocal images were taken on Leica TCS SP2 AOBS laser scan microscope.

2.4.16   Immunoelectron microscopy

For immunoelectron microscopy, ultrathin cryosections were prepared as described [54]. 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.

2.4.17 EndoH / PNGase treatment

Stably expressing pFGE.His HT1080 cells were plated in 6 cm dishes at 50% confluence. Cells were grown till the confluence and medium was changed with fresh medium. After 48 h in fresh medium, cells and medium were collected for partial purification of FGE-His by chromatography on Ni-NTA agarose under native conditions according to the manufacturer's protocol (Qiagen). Instead of using imidazole, protein was eluted by boiling the Ni-agarose beads in 0.1% SDS/0.1 M -mercaptoethanol at 95 °C for 5 min. The eluate was used for PNGase (with addition of Triton X-100) and EndoH treatment as recommended by the manufacturer (Roche Applied Science).

2.4.18 Limited proteolysis and protein sequence analysis

pFGE-His purified from the secretion of HT1080 cells was treated with thermolysin or elastase in phosphate-buffered saline at a protein to protease ratio of 25:1 (w/w) for up to 24 h at room temperature. The reaction was stopped by adding SDS-PAGE loading buffer and heating at 95 °C. The digests were separated by SDS-PAGE, followed by Coomassie Blue staining. For N-terminal sequencing, the proteolytic fragments were resolved by SDS-PAGE, blotted onto a polyvinylidene difluoride membrane (Schleicher and Schuell, Germany), stained with Coomassie Blue, excised from the membrane and subjected to Edman degradation on a Procise cLC sequenator (Applied Biosystems).

2.4.19 Reductive carbamidomethylation and tryptic digestion

Purified pFGE-His (20-200 μg) was dried, solubilized in 45 μl of 8 M urea, 0.4 M ammonium bicarbonate, pH 8.6 (buffer U), and reduced with 5 μl of 100 mM DTT in buffer U at 56 °C for 1 h. Cysteines were carbamidomethylated by incubation with 5 μl of 300 mM iodoacetamide in buffer U in the dark at 25 °C for 30 min. Reduction and carbamidomethylation were repeated (adding 7.5 μl of 100 mM DTT and 5 μl of 300 mM iodoacetamide, respectively) and the excess of iodoacetamide quenched

with 2.5 μl of 100 mM DTT. The pH was adjusted to about 2 by adding 130 μl2% trifluoroacetic acid/water, and the enzyme was purified on a C4 reversed-phase column (Aquapore BU-300 7-micron 20 x 2.1 mm) applying a 0-90% acetonitrile gradient in 0.1% trifluoroacetic acid (1.5% per min). The pFGE-His containing fraction was lyophilized, solubilized in 25 μl buffer U, diluted to 100 μl with water, and digested with trypsin (2% w/w) at 37 °C overnight. For the identification of disulfide bridges, reduction and carbamidomethylation were omitted, and native pFGE was directly digested with trypsin (at 10% w/w). After digestion the peptides were purified in aliquots, corresponding to 20 μg of pFGE, on a microbore C18 reverse phase column (C18 PepMap100 1 x 150 mm, Dionex) applying a 0-63% acetonitrile gradient in 0.1% trifluoroacetic acid (1.0% per min). All reverse phase purifications were performed on a SMART system (Amersham Biosciences).

2.4.20 Mass spectrometry and Edman sequencing

For the analysis of cysteine modifications and the oligosaccharide sidechain structure, tryptic peptides of pFGE were subjected to matrix-assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF MS) using a Reflex III (Bruker Daltonik, Germany) and dihydroxybenzoic acid as matrix. In order to prove the identity of cysteine or oligosaccharide containing peptides selected fractions were subjected to edman sequencing (Procise cLC, Applied Biosystems).

2.4.21 Metabolic labelling and immunoprecipitation

Normal human skin fibroblast cells were grown to near confluence in 30 mm dishes, cells were starved in 2 ml of Methionine and cysteine-free DMEM for 1h, pulsed for 6 hours with 2ml of medium containing 50 μci of 35S-labelled Methionine and Cysteine supplemented with 10% dialysed fetal calf serum and harvested or washed with cold PBS and cultured in 2ml of chase medium (DMEM containing Methionine, Cysteine and 10% FCS) for 16h. Medium was collected from all the plates and cells were harvested by scrapping. Immunoprecipitation of pFGE from cells and medium was performed with pFGE anti-serum as described earlier [55]. Pellets were finally solubilized and subjected to SDS-PAGE and fluorography. Densitometric quantification of pFGE was done using a Bio-Imaging Analyser, FUJI.

2.4.22   Photocrosslinking

For photoaffinity labelling, a benzophenone labelled peptide was synthesised corresponding to Arysulfatase A (residues 59-81) with an additional ϵ-biotinylated lysine residue at the N-terminus and a benzoyl phenylalanine at position 77 instead of leucine. The reaction mixture, containing 50μM of benzophenone labelled peptide, 1μM of each pFGE-His and bovine serum albumin or FGE-His was irradiated for 30 minutes on ice with an Ultra tech 400W halogen metal vapor lamp (Osram) [54]. The photoadducts were analysed by western blotting using the His tag or biotin monoclonal antibody as primary antibody and HRP conjugated goat anti- mouse antibody as secondary antibody.

2.4.23   Co-immunoprecipitation

Transfected cells were removed from dishes by scrapping, pelleted by centrifugation and lysed by ultrasonication in PBS containing protease inhibitors (protein inhibitor cocktail, Sigma). The cell lysates were spun down at 125,000 Xg for 30 min. The supernatent was incubated with respective antibody for overnight at 4 °C. To the suspension prewashed 10% StaphA (Pansorbin) was added, and the sample was incubated for 1 hr at 4 °C. After centrifugation of the StaphA, antibody-antigen complex was washed five times with PBS, boiled in 1x laemmli buffer, and separated by 15% SDS-PAGE under reducing conditions. For immunoblotting, gels were transfered onto nitrocellulose membranes, incubated with primary antibodies followed by peroxidase-conjugated anti-rabbit or anti-mouse antibody, and developed by chemiluminescence.

Chapter 3

Results

3.1 Molecular characterization of pFGE, the paralog of FGE

Sequence comparisons led to the discovery of a paralogue of *SUMF1* in the vertebrate genomes, named *SUMF2*. *SUMF2* encoded pFGE shares 47% sequence identity and 62% similarity with *SUMF1* encoded FGE (Fig 3.1). The region between

```
FGE     89   SKMVPIPAGVFTMGTDDPQIKQDGEAPARRVTIDAFYMDAYEVSNTEFEKFVNSTGYLTE   148
             + MV +  G F MGT+ P   +DGE P R  T+  F +D + V+N +F   FV    Y TE
pFGE    28   TSMVQLQGGRFLMGTNSPD-SRDGEGPVREATVKPFAIDIFPVTNKDFRDFVREKKYRTE    86


FGE     149  AEKFGDSFVFEGMLSEQVKTNIQQAVAAAPWWLPVKGANWRHPEGPDSTILHRPDHPVLH   208
             AE FG SFVFE +S++++   Q + +  WWLPV+ A WR P GP S I  R +HPVLH
pFGE    87   AEMFGWSFVFEDFVSDELRNKATQPMKSVLWWLPVEKAFWRQPAGPGSGIRERLEHPVLH   146


FGE     209  VSWNDAVAYCTWAGKRLPTEAEWEYSCRGGLHNRLFPWGNKLQPKGQHYANIWQGEFPVT   268
             VSWNDA AYC W GKRLPTE EWE++ RGGL  +++PWGN  QP   + N+WQG+FP
pFGE    147  VSWNDARAYCAWRGKRLPTEEEWEFAARGGLKGQVYPWGNWFQP---NRTNLWQGKFPKG   203


FGE     269  NTGEDGFQGTAPVDAFP-PNGYGLYNIVGNAWEWTSDWWTVHHSVEETLNPKGPPSGKDR   327
             +  EDGF G +PV+AFP  N YGLY+++GN WEWT+    + + + E+ +          R
pFGE    204  DKAEDGFHGVSPVNAFPAQNNYGLYDLLGNVWEWTA---SPYQAAEQDM----------R   250


FGE     328  VKKGGSYM--CHRSYCYRYRCAARSQNTPDSSASNLGFRCAAD                   368
             V +G S++      S  +R R   R  NTPDS++ NLGFRCAAD
pFGE    251  VLRGASWIDTADGSANHRARVTTRMGNTPDSASDNLGFRCAAD                   293
```

Figure 3.1. Sequence alignment of human-FGE and human-pFGE
The letters in between sequences indicates, identities in FGE and pFGE. The plus symbol indicates, similarities in FGE and pFGE.

aminoacids 303 and 351 of the human FGE differs from the corresponding region of

human pFGE. *SUMF1* is conserved from bacteria to man, whereas *SUMF2* evolved later and is present only in present in deuterostomia. [41, 51].

3.1.1 Cloning of human pFGE encoding cDNA (*SUMF2*)

Total RNA was prepared from human fibroblasts and reverse transcribed to cDNA by reverse transcriptase using an oligo (dT) primer or a *SUMF2*-specific primer (see Methods). The first strand cDNA was amplified by PCR using the *SUMF2* specific forward and reverse primers. The PCR products were cloned directly into the pCR4-TOPO vector. 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^{TM} 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.

3.1.2 Expression of SUMF2
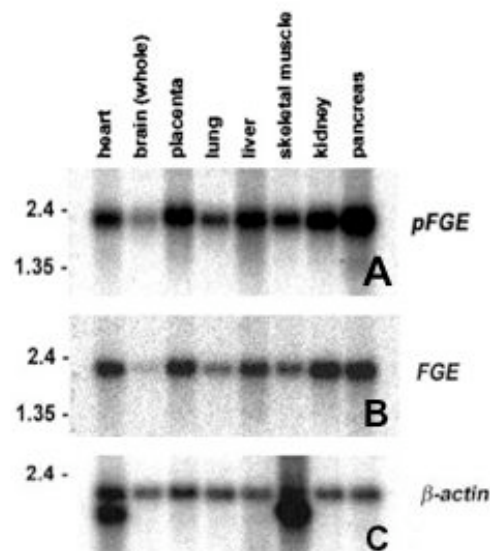
Figure 3.2. Expression of pFGE (A), FGE (B) and β-actin (C) in human tissues Northern blot analysis of polyA+ RNA from human tissues. Very similar expression patterns were obtained when using probes hybridizing either in the coding region, as shown here, or in the non-conserved 3' untranslated region (not shown).

A human multi tissue Northern blot was hybridized with ^{32}P-labelled cDNA

probes covering the entire coding regions of pFGE or FGE, respectively, and a β-actin cDNA probe as a control for RNA loading. The results show, a single transcript of 2.0-2.1 kb is detectable by Northern blot analysis of polyA+ RNA from heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas (Fig 3.2.A). The expression pattern compares well to that of *SUMF1* (Fig 3.2.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 (Fig 3.2.C) the abundance of *SUMF2* RNA varies by one order of magnitude, being highest in pancreas and kidney and lowest in brain. Basically the same result, as shown in Fig.1, was obtained when using probes hybridizing in the 3' UTR regions of *SUMF1* or *SUMF2* mRNA, where the two sequences show no similarity (not shown).

3.1.3 Subcellular localization of pFGE
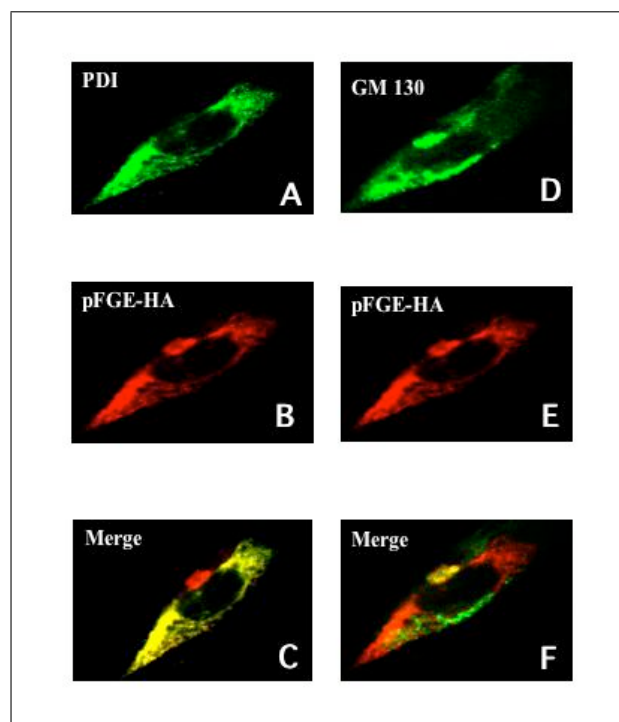
Figure 3.3. Subcellular localization of pFGE-HA in HT-1080 cells.
After fixation and permeabilization the proteindisulfide isomerase was detected with a monoclonal IgG$_{2A}$ antibody (green, A) and GM130 (green, D) with a rabbit antiserum. The HA-tag of pFGE (red, B and E) with a monoclonal IgG1 antibody. The merge reveals co localization of PDI or GM130 and pFGE in yellow (C, F)

pFGE with a C-terminal HA- or His-tag was transiently expressed in the human

HT-1080 sarcoma cell line. pFGE colocalized with proteindisulfide isomerase, a lumenal protein of the endoplasmic reticulum (Fig 3.3, A-C). In some cells as the one shown in Fig 3.3, a fraction of pFGE was associated with perinuclear structures devoid of proteindisulfide isomerase. This structure is enriched in the Golgi marker protein GM130 (Fig 3.3), D-F). This indicates that recombinant pFGE is localized in the endoplasmic reticulum. In some cells, however it accumulates in detectable amounts also in the Golgi apparatus. Immunoelectronmicroscopy confirmed the localization of pFGE in the endoplasmic reticulum where it colocalizes with protein disulfide isomerase and in the Golgi stack (Fig 3.4). A sparse labelling for pFGE was observed in structures enriched in the endosomal / lysosomal marker LAMP I.
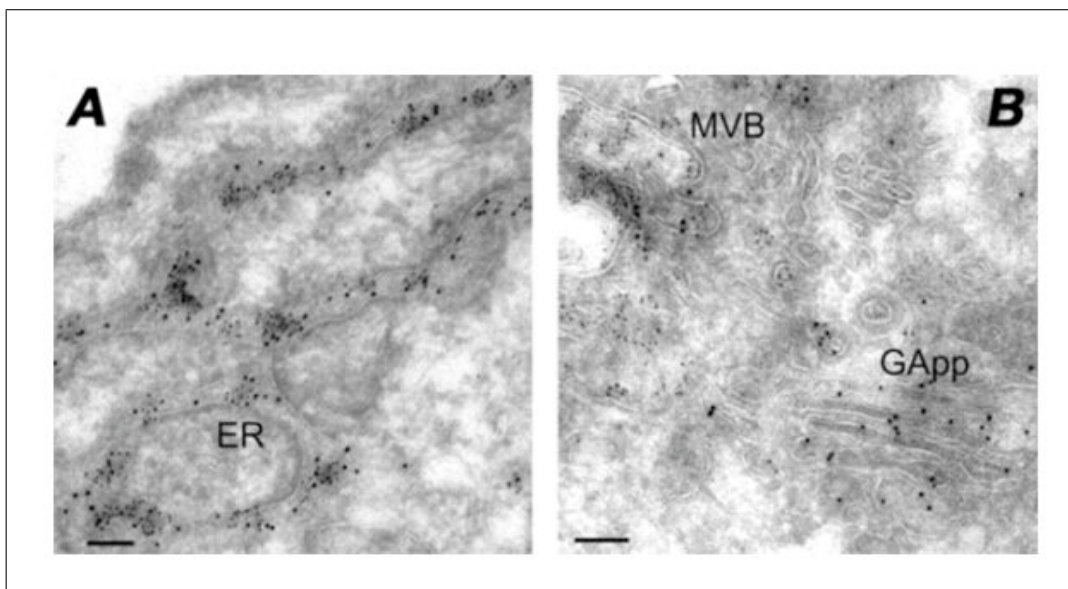
Figure 3.4. Localization of pFGE-His in HT-1080 cells by immunoelectron microscopy After fixation and cryosectioning, pFGE was detected with a rabbit antiserum against pFGE and protein A-gold (10 nm). A shows double labeling with the endoplasmic reticulum marker protein disulfide isomerase (5 nm gold), 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.

3.1.4 Secretion of pFGE-His by HT-1080 cells

Overexpressed pFGE-His, is only transient by being 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. Cellular pFGE was kept at a rather constant level (Fig 3.5).
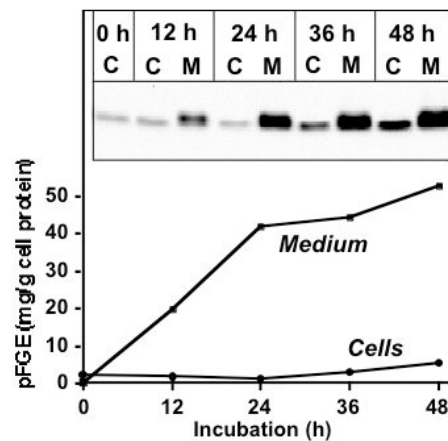
Figure 3.5. Secretion of pFGE-His in HT-1080 cells
HT-1080 cells stably expressing pFGE-His were grown to half confluency. At time zero fresh medium was added and the cells were harvested 0, 12, 24, 36 and 48 h of culture. pFGE was quantified in equal aliquots of cell extracts (C) and media(M) by Western blot using a monoclonal antibody against the His-tag (inset). The amount of intracellular and secreted pFGE-His per mg cell protein was determined by calibration of the Western blot with known amounts of purified pFGE-His and refered to total cell protein.

3.1.5 Analysis of N-glycosylation

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 3.6), indicating that intracellular pFGE contains a high-mannose (or hybride) 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 heterogenous N-glycosylation. The 31.5 kDa form is sensitive to endoglucosamidase H, while the 32.5 kDa form is resistant (Fig 3.6). This indicates that the 31.5kDa form contains a high mannose type oligosaccharide, while the 32.5kDa form contains a complex type oligosaccharide. MALDI-TOF analysis of the N-glycosylated tryptic peptide 179-192 of secreted pFGE-His

confirmed the heterogenous glycosylation at Asn 191. The oligosaccharide mixture was composed of 11 different species (Fig 3.8). They all belonged to the hybride or complex type with up to 5 N-acetylhexosamine, six hexose, one sialic acid and one fucose residue per oligosaccharide and masses ranging from 1768 to 2424 Da (Fig 3.7).
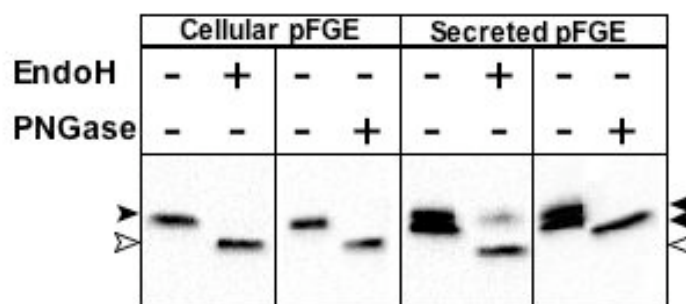
Figure 3.6. N-Glycosylation of intracellular and secreted pFGE-His
Cell extracts and secretions of HT-1080 cells expressing pFGE were treated with endoglucosaminidase H (Endo H) or N-glycosidase I (PNGase), separated by SDS-PAGE and detected by Western blot using anti-His antibodies. The filled arrows point to the glycosylated, the open arrows to the deglycosylated pFGE forms

Figure 3.7. Complex type structure of the oligosaccharide side chain at aspargine 191 of secreted pFGE, containing 5 N-acetylhexosamine, six hexose, one sialic acid and one fucose residue per oligosaccharide

3.1.6 Large scale purification of pFGE-His

For purification of pFGE-His, the medium collected from HT1080 cells stably expressing pFGE-His was used as a starting material. The cells were grown to near confluency in normal growth medium. The medium was collected every 48h, cleared by spinning, and subjected to ammonium sulfate precipitation (see
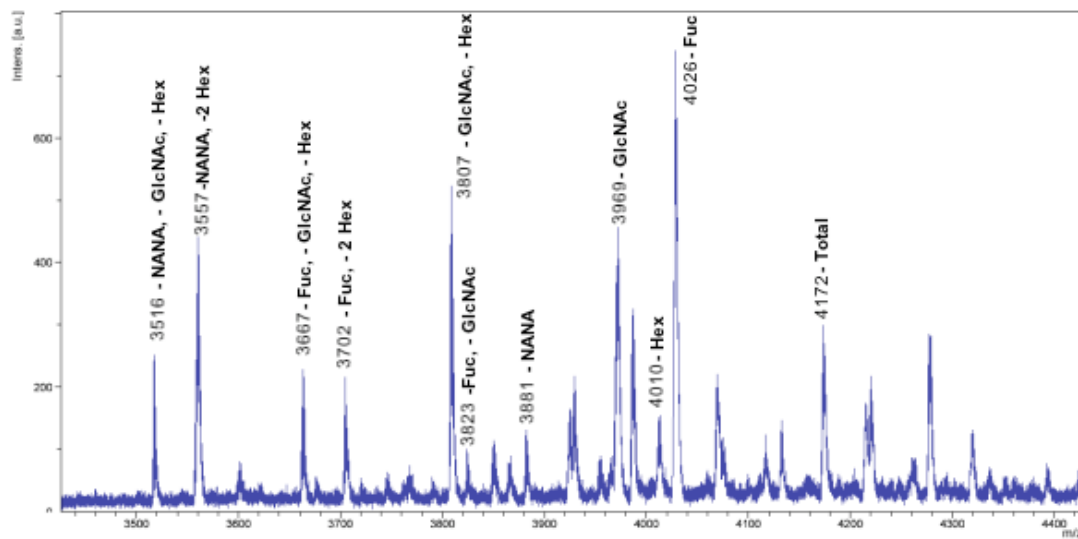
Figure 3.8. N-Glycosylation secreted pFGE-His

N-glycosylated tryptic peptide 179-192 of secreted pFGE was purified by reverse-phase chromatography and analyzed by MALDI-TOF MS. A series of 11 different mass signals was found corresponding to peptide 179-192. The total mass of glycopeptide is 4172 Da.

Methods). The precipitate was reconstituted and dialyzed in Ni-NTA lysis buffer. The dialysed material was cleared by spinning and applied to Ni-NTA affinity chromatography (Fig 3.9).

Peak fractions were pooled and dialyzed against buffer A (see Methods). The dialyzed protein was loaded to MonoQ anion exchange chromatography in contrast to FGE, the pFGE-His protein was not bound to the column at pH 8. It was recovered in the flow-through and other contaminants were bound to the column (Fig 3.10). Flow through fractions were pooled, concentrated and dialyzed against Buffer A. Analytical size-exclusion chromatography of the purified pFGE-His protein revealed its homogeneity (Fig 3.11). From 1 liter of medium 8mg of pFGE-His could be purified. On SDS gels pFGE-His protein appeared in two molecular forms (Fig 3.10). The larger 32.5-KDa and smaller 31.5-KDa forms corresponded to full length pFGE with alanine 27 at its N terminus, as analyzed by Edman Sequencing. However both forms are distinguished by different glycosylation pattern (Fig 3.6).
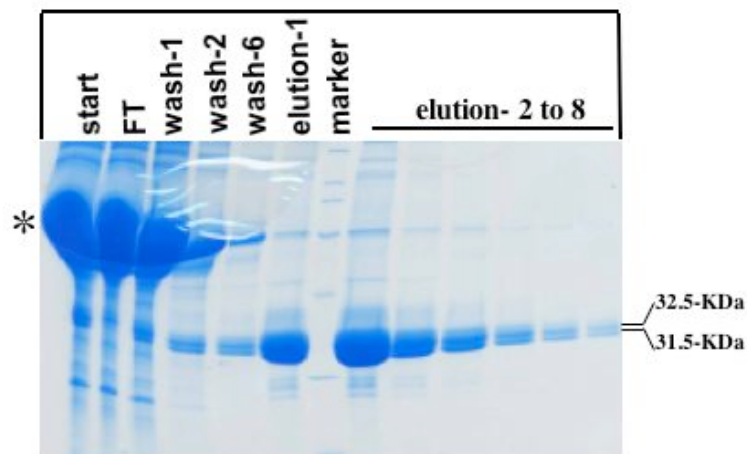
Figure 3.9. Ni-NTA purified pFGE-His

Ammonium sulfate precipitated pFGE-His medium was passed through Ni-NTA -agarose and the fractions were run on a 15% SDS polyacrylamide gel. Staining was with Coomassie Brillant Blue. The two different size of 32.5 and 31.5-KDa point to the differentially glycosylated forms of pFGE. ∗ indicates bovine serum albumin from growth medium.
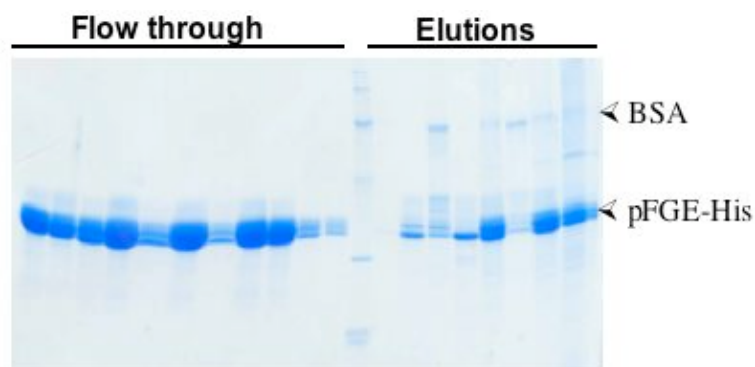
Figure 3.10. SDS-PAGE of fractions collected from MonoQ chromatography

Material eluted from Ni-NTA further purified by MonoQ chromatography . Aliquots of the fractions were separated by 15% SDS polyacrylamide gel. Separated proteins were visualized with Coomassie Brilliant Blue staining.
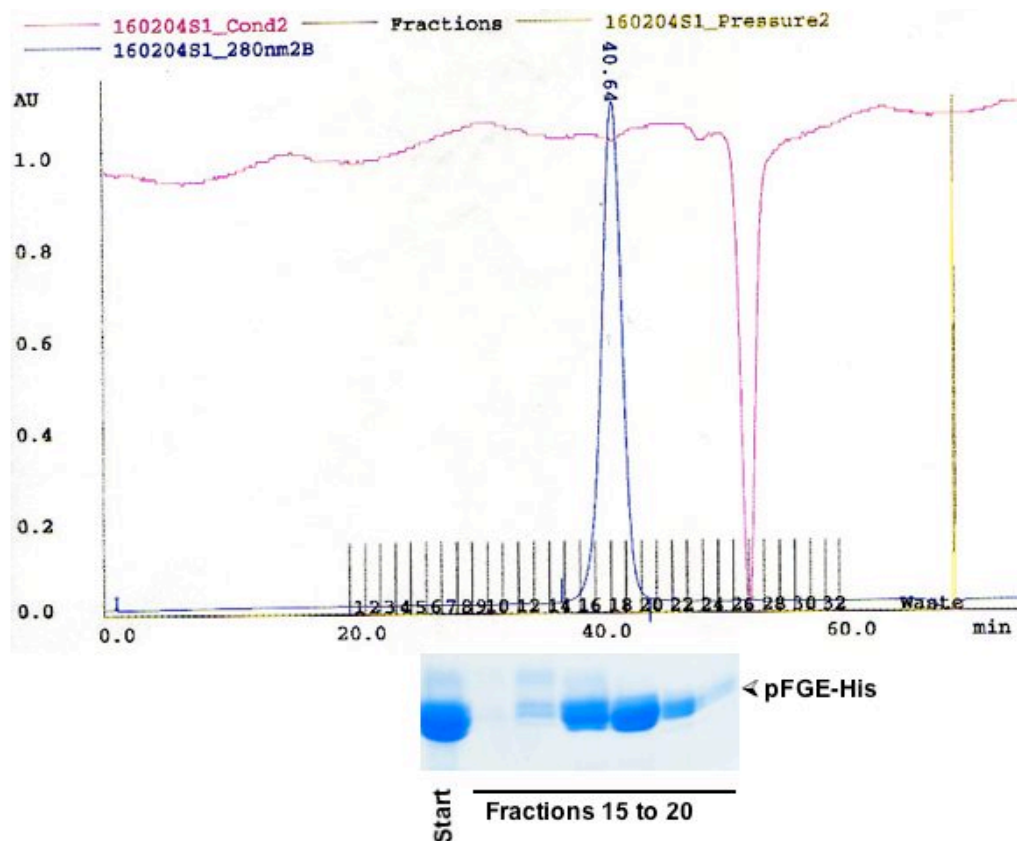
Figure 3.11. Gel filtration chromatography

The flow-through fractions from MonoQ chromatography were pooled, concentrated and dialyzed (see Methods). An aliquots of purified pFGE-His applied to a Superdex 200 gel-filtration column that had been equilibrated in PBS and eluted isocratically with PBS. pFGE-His eluted as a single peak and the fractions were run on a 15% SDS polyacrylamide gel and followed by staining with Coomassie Brilliant Blue.

3.1.7   Domain structure of pFGE-His

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 kDa and 22-24 kDa (Fig 3.12). The fragments
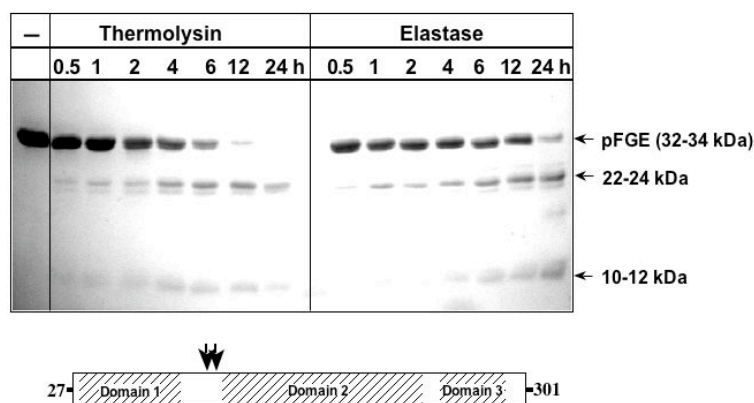
Figure 3.12. Limited proteolysis of pFGE-His by thermolysin and elastase
pFGE-His purified from secretions of HT-1080 cells was incubated for up to 24 h with thermolysin or elastase at a FGE-protease ratio of 25:1(w/w). After separation by SDS-PAGE the polypeptides were detected by staining with Coomassie blue. The molecular weights of pFGE and the proteolytic fragments are indicated at the right. The scheme at the bottom shows the locations of the sub domains 1, 2 and 3 together with the residue numbers defining the N- and C- termini of mature pFGE and its sub domains. The arrows indicate the two proteolytic cleavage sites of thermolysin N-terminal of threonine 109 and valine 115, as determined by Edman sequencing.

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, while the 22-24kDa fragments start with threonine 109 or with valine 115. Sequence comparison has revealed three subdomains in pFGE [42] (Fig 3.12 bottom). The thermolysin cleavage sites are located in the linker region between subdomain 1 and 2, which extends from serine 93 to leucine 116.

3.1.8   Disulfide bridges of pFGE-His

To examine for a possible disulfide bonding of the two existing cysteine residues in pFGE (at positions 156 and 290), pFGE-His purified from the secretions was first reacted with NEM 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 HPLC on a C18 matrix and characterized by MALDI-TOF 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. As non-reduced pFGE migrates in SDS-PAGE as a monomer (not shown) cysteine 156 and 290 form an intramolecular and not an intermolecular disulfide bond.

3.1.9 Generation and characterization of antisera against pFGE-His

To allow for detection of endogenous pFGE, we wanted to generate antisera against pFGE-His which was purified from secretions. A rabbit was immunized with purified pFGE-His using stimune adjuvant (Cedi Diagnostics B.V). To screen high titer antibodies, the rabbit antisera was regularly monitored by immunoprecipitation of purified pFGE-His. The degree of immunoprecipitation of purified pFGE-His was not changed between bleed 3 and bleed 4 antiserum (Fig 3.13). Hence, the rabbit
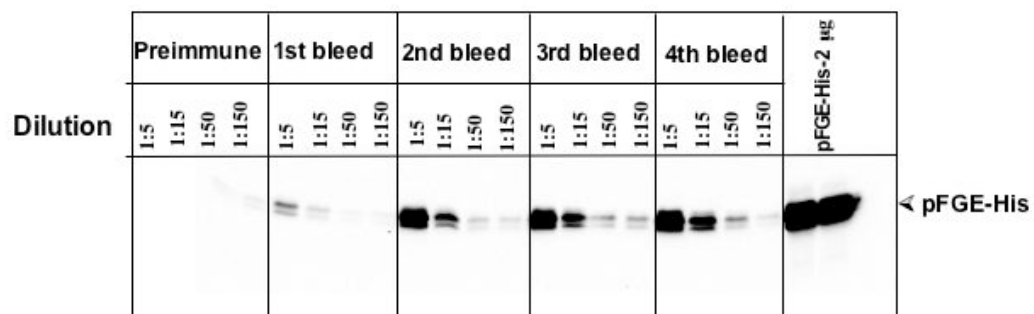
Figure 3.13. Titration of pFGE antisera by Immunoprecipitation
The rabbit sera were diluted to 1:5, 1:15, 1:50, and 1:150 in PBS and incubated with purified pFGE-His. The mixture was precipitated by Pansorbin (see Methods). The precipitant was resolved on SDS-PAGE and transfered to nitrocellulose and followed by probed with monoclonal antibody against His-tag.

was sacrificed to collect blood after 3 booster injection (see Methods). pFGE is highly similar to FGE, therefore it was essential to check whether pFGE antisera could precipitate recombinant FGE. The results show that pFGE antiserum could precipitate recombinant pFGE but not recombinant FGE (Fig 3.25 and 3.26). Next we analyzed the sensitivity and cross-reactivity of pFGE antiserum towards purified
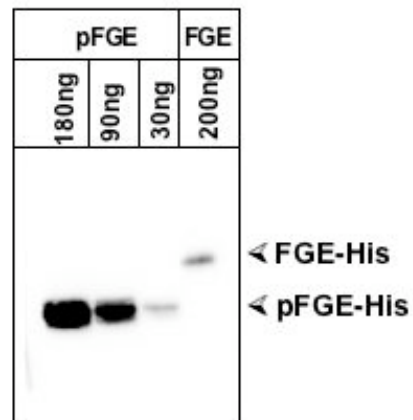
Figure 3.14. Sensitivity and cross-rectivity of pFGE antiserum

Various amounts of purified pFGE-His and purified FGE-His were run on a SDS-PAGE and followed by immunodetection with pFGE antiserum (1:2500 dilution).
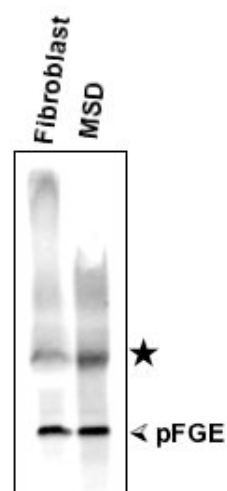
Figure 3.15. Immunoprecipitation of endogenous pFGE

The cell extracts of human skin fibroblast or MSD fibroblast were immunoprecipitated with pFGE antiserum. The presence of endogenous pFGE was revealed by immunoblotting with pFGE antiserum. ⋆ indicates the heavy chain of pFGE antibody.

pFGE-His and FGE-His, respectively. This experiments reveal that pFGE antiserum could detect 30 ng of purified pFGE-His. Nevertheless it weakly cross-reacted with purified FGE (Fig 3.14). Above experiment was repeated with FGE antiserum where in FGE antiserum exclusively reacts with FGE.

In an attempt to check whether pFGE antiserum was able to precipitate endogenous pFGE, the extracts of human skin fibroblast cells or MSD fibroblasts were subjected to immunoprecipitation with pFGE antiserum. The presence of endogenous pFGE was detected by immunobloting with pFGE antiserum (Fig 3.15). However, to detect efficiently enogenous pFGE we had to start with ∼2mg of cell extracts. In a control experiment, preimmmune serum replaced pFGE antiserum which failed to precipitate specifically endogenous pFGE (data not shown).

pFGE antiserum allowed us to precipitate the endogenous pFGE, therefore we were interested to know whether this antiserum is suitable for immunofluoresence microscopy. The experiments demonstrate that pFGE antiserum was able to detect endogenous pFGE (Fig 3.18, B and E).

3.1.10   Generation and characterization of monoclonal antibodies against pFGE-His

We have learnt from our earlier experiments that pFGE antiserum could be used for immunofluoresence and immunoprecipitation studies to detect specifically pFGE. Albeit, it weakly cross-reacts with FGE on western blot. To circumvent this problem again we raised polyclonal antibodies against purified pFGE-His in rabbit. Invariably, we found that this antiserum also cross-reacts with FGE in western blot.
In order to have specific antibody against pFGE, we intended to generate monoclonal antibody against purified pFGE-His. For this we had sent our purified pFGE-His to Transkaryotic Therapies, Inc (cambridge, MA) to make hybridoma. Supernatants from hybridomas were screened in our lab by western blot, immunoprecipitation and immunofluoresence studies. It should be noted that all the screening experiments were carried out with purified pFGE-His or cells expressing His-tagged pFGE. Highly reactive best hybridoma clones were chosen in among 100 different hybridoma clones. For these clones, ascites were produced to have concentrated monoclonal pFGE antibodies. Finally we investigated all these monoclonal antibodies for many purposes. Surprisingly we noticed that all these monoclonal antibodies recognize only His-tagged intracellular and purified pFGE, but not untagged intracellular pFGE. Moreover they cross-react with intracellular

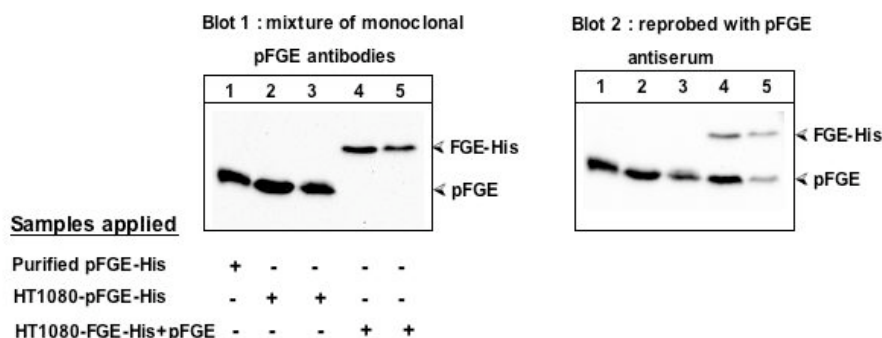His-tagged FGE but not purified His-tagged FGE (Fig 3.16,3.17).

Figure 3.16. Cross-reactivity of pFGE monoclonal antibodies
Extracts from HT1080 cells expressing pFGE-His or FGE-His+pFGE were loaded on SDS-PAGE including purified pFGE-His as a control. After transferring to nitrocellulose membrane, the blot was probed with mixture of three monoclonal antibodies of pFGE-His (Blot.1). In lane 4 and 5 untagged pFGE was not detected by monoclonal antibodies but they recognized His-tagged pFGE (lane1-3) and intracellular His-tagged FGE (lane 4 and 5). The blot.1 was re-probed with pFGE antiserum in which untagged pFGE was detected by pFGE antiserum.

To verify this observation, we performed immunofluoresence using these an-tibodies. This also suggested that pFGE monoclonal antibodies recognize His-tagged pFGE but not untagged pFGE. And they were not sensitive enough to detect endogenous pFGE in human skin fibroblasts either by immunofluoresence or immunoprecipitation studies. Taken together all these data support that monoclonal antibodies were likely raised against epitopes that include that His-tag of pFGE-His. This is likely to explain their cross reactivity with FGE-His.
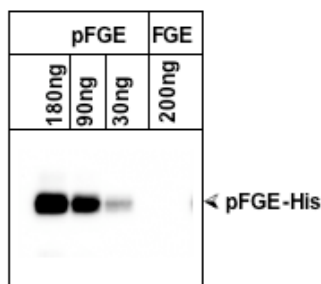
Figure 3.17. Sensitivity and cross-rectivity of monoclonal pFGE antibodies
Various amounts of purified pFGE-His and purified FGE-His were run on SDS-PAGE and followed by immunodetection with pFGE monoclonal antibody.

3.1.11   Localization and secretion of endogenous pFGE

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 3.18, A-C). Association with the Golgi, as observed for overexpressed pFGE (Fig 3.3, D-F), was
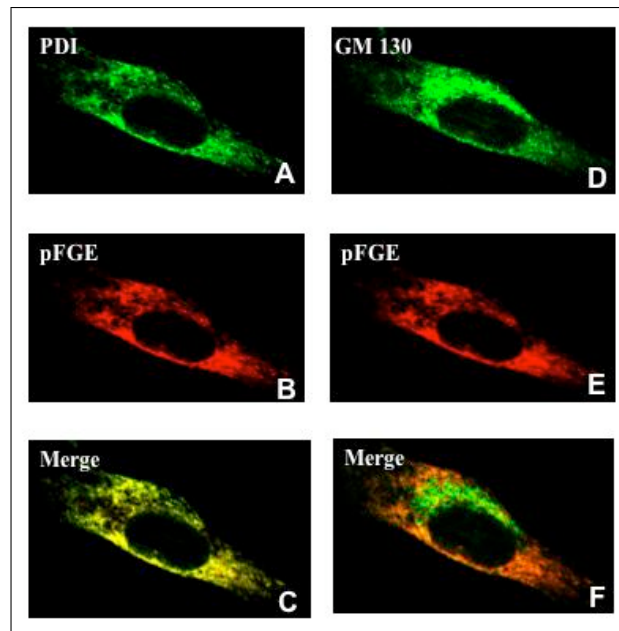
Figure 3.18. Subcellular localization of endogenous pFGE in human skin fibroblasts. After fixation and permeabilization the proteindisulfide isomerase was detected with a monoclonal IgG$_{2A}$ antibody (green, A) and GM130 (green, D) with a monoclonal IgG1 antibody. The endogenous pFGE (red, B and E) was detected with a rabbit antiserum. The merge reveals colocalization of PDI and pFGE in yellow (C).There is no colocalization of endogenous pFGE with the Golgi marker GM130 (F).

not detectable for endogenous pFGE (Fig 3.18, D-F).

To follow biosynthesis and possible secretion of endogenous pFGE, human skin fibroblasts were metabolically labelled for 6 h with ^{35}S- methionine. After labelling, the cells were either harvested or subjected to a 16 h chase in ^{35}S-methionine-free medium. pFGE was immuno-precipitated from extracts of the cells and media, separated by SDS-PAGE and visualized by phosphoimaging (Fig 3.19). 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 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
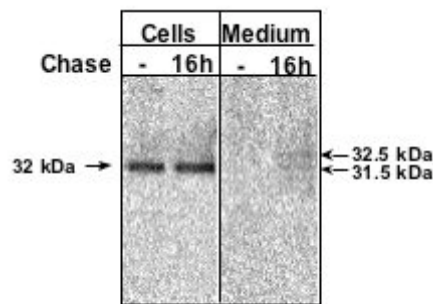
Figure 3.19. Biosynthesis and secretion of endogenous pFGE
pFGE was immunoprecipitated from the cells and media of human skin fibroblast that had been labelled for 6 h and harvested either directly after metabolic labelling or after a chase for 16 h. The arrows indicate the intracellular 32 kDa pFGE form (left) and the secreted 31.5 and 32.5 kDa forms (right).

rise to the two 31.5 and 32.5 KDa pFGE forms in the secretions, as has been shown for recombinant pFGE (Fig 3.6).

3.1.12   pFGE lacks FGly-generating activity

pFGE is the paralog of FGE, which has FGly generating activity. The generation of FGly residue is readily detectable with the help of the 23mer peptide P23. This peptide corresponds to arylsulfatase A 60-80, with an additional N-acetylated methionine and a C-amidated serine residue to protect its N- and C-terminus, respectively[41]. It contains the cysteine residue to be modified and neighbouring residues, which are essential for the FGly formation. The FGE activity in extracts of HT-1080 cells transiently expressing pFGE-His was in the range of non-transfected HT-1080 cells, whereas FGE activity in extracts of cells expressing a comparable amount of FGE-His was more than 20-fold increased (Fig 3.20). When purified pFGE-His was incubated with P23 the peptide under conditions that allow quantitative modification of the cysteine residue by purified FGE-His, no FGly formation was detectable. Increasing or decreasing the amount of pFGE-His did not result in FGly-formation.

The human genome predicts 17 sulfatases of which 16 have been shown to be expressed in various tissues. To test whether pFGE can modify any of them, we synthesized peptides derived from all 16 sulfatases as N-acetylated 23-mers comprising residues -10 to +11, with the critical cysteine in position 0 and a C-terminal serine amide. These peptides were used as pFGE or FGE substrates at 200nM concentration, this is about 15-fold above the half-saturating concentration
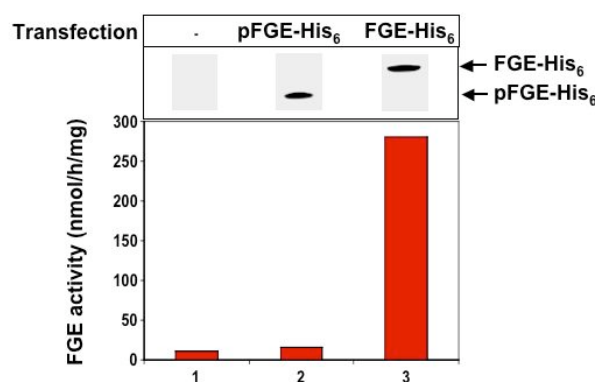
Figure 3.20. FGE activity in HT-1080 cells expressing pFGE-His or FGE-His
The upper panel shows the Western blot detection of pFGE-His and FGE-His in the cell extracts with a monoclonal antibody against the His tag. The lower panel shows the FGly-generating activity in the cell extracts determined *in 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.

determined for P23 [41]. As shown in Table 3.1, FGE was able to generate FGly in all tested peptides. But none of them were modified by pFGE. This indicates that pFGE lacks FGly-generating activity.

3.1.13 Recombinant 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 [41, 51]. 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 3.21). 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 3.21). 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 % , whereas coexpression of FGE increased the specific activity of galactose-6-sulfatase more than 100-fold (see in the next section Fig 3.28). The different efficiencies of FGE-mediated activation and pFGE-mediated inhibition in

Table 3.1. FGE-mediated turnover of peptides derived from human sulfatases
The indicated synthetic peptides are derived from the 16 human sulfatases predicted by the human genome, with the homologous amino acid positions given in the peptide name. All peptides are N-terminally acetylated and carry a serine amide in the C-terminal position as a protective group against peptidases. Two peptides (P23 and I2S) carry a point mutation in the first position (F59M and F74L, respectively) in order to avoid coincidence of their masses with those of others. The core FGly modification motif [29] of all peptides is underlined. The cysteine to be modified and its 1 neighbor are boldface. The peptides are grouped according to the nature of this 1 residue (threonine, serine, cysteine, or alanine). 6 pmol of each peptide were incubated for 30 min under standard assay conditions (see "Materials and Methods") with 1ng of FGE, purified from HT1080 secretions. The FGly modification of the peptides by FGE is given as absolute (pmol/h) and relative activity (percentage of P23 turnover).

Peptide	Sequence	FGly modification	
		pmol/h	*%*
ASA-F59M-(59–80)[a] (P23)	MTDFYVPVSL_**C**TPSR_SRAALLTGRS	6.0	100
ASB-(81–102)	LLDNYYTQPL_**C**TPSR_SQLLTGRS	6.8	113
ASC-(73–94)	LTQHLAASPL_**C**TPSR_AAFMTGRS	4.0	66
ASD-(79–100)	LTQHLAAAPL_**C**TPSR_/AAFLTGRS	2.5	41
ASE-(76–97)	LTQHISAASL_**C**TPSR_AAFLTGRS	8.7	145
Sulf4-(86–107)	KLENYVVQPI_**C**TPSR_SQFITGKS	5.8	96
Sulf5-(83–104)	KLENYYIQPI_**C**TPSR_SQLLTGRS	4.6	76
ASF-(69–90)	LTQHISAASL_**C**SPSR_SAFLTGRS	3.0	50
ASG-(74–95)	FVDFHAAAST_**C**SPSR_ASLLTGRS	3.6	60
Sulfamidase-(60–81)	FRNAFTSVSS_**C**SPSR_ASLLTGLS	0.5	8
NAcGalN6S-(69–90)[b]	FPNFYSANPL_**C**SPSR_AALLTGRS	0.6	10
NAcGlcN6S-(81–102)[c]	FSSAYVPSAL_**CC**PSR_ASILTGKS	2.8	46
Sulf1-(74–95)	FINAFVTTPM_**CC**PSR_SSMLTGKS	2.0	34
Sulf2-(77–98)	FINAFVTTPM_**CC**PSR_SSILTGKS	3.1	52
Sulf3-(70–91)	FLNAYTNSPI_**CC**PSR_AAMWSGLS	2.2	36
I2S-F74L-(74–95)[d]	LQNAFAQQAV_**C**APSR_VSFLTGRS	1.8	30

[a] ASA-ASG, arylsulfatase A-G.
[b] NAcGalN6S, N-acetylgalactosamine 6-sulfatase.
[c] NAcGlcN6S, N-acetylglucosamine 6-sulfatase.
[d] I2S, iduronate 2-sulfatase.

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 these sulfatases, 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
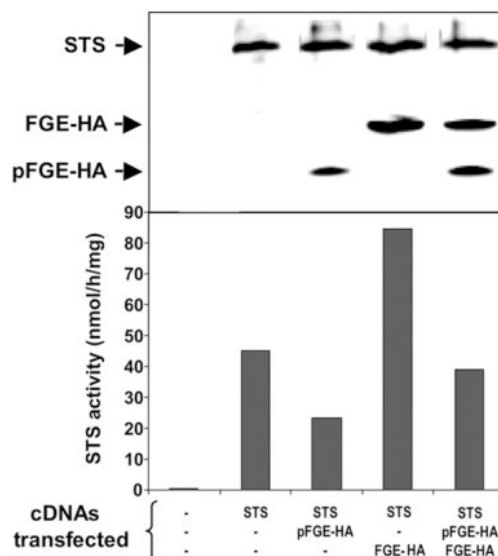
Figure 3.21. pFGE-HA reduces and FGE-HA enhances the activity of steroid sulfatase. In HT-1080 cells, steroid sulfatase (STS), pFGE-HA, and FGE-HA were transiently coexpressed in the combinations indicated below the lanes. In the cell extracts, the expression of steroid sulfatase, pFGE-HA, and FGE-HA 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 "Materials and Methods").

sulfatase polypeptides by FGE.

3.1.14   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(see in the next section Fig 3.26). 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 (Fig 3.22).

Photocrosslinking of pFGE-His to an arylsulfatase A peptide

(in cooperation with Andrea Preusser-Kunze)

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 3.23, lane 5). Replacing FGE by

Coelution of Gal-6-S and pFGE
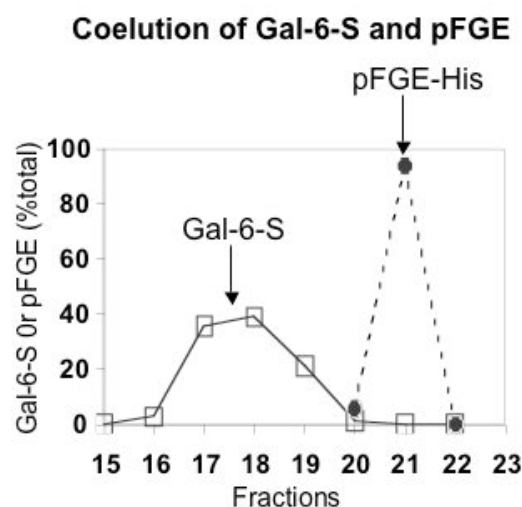
Figure 3.22. Galactose-6-sulfatase and pFGE-His do not coelute
The extracts of HT1080 cells expressing pFGE-His and galactose-6-sulfatse (Gal-6-S) were chromatographed by Superdex-200 as described under "Materials and Methods". The fractions were analyzed by immunobloting using a mixture of galactose-6-sulfatase and His-tag monoclonal antibodies. Percentage of pFGE or Gal-6-S were calculated by quantifying the western blot signals using densitometry.

pFGE, we also observed cross-linking of the peptide to pFGE, albeit at only one-tenth of the efficiency observed for FGE-His (Fig 3.23, 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 3.23, lane 2) and by the lack of cross-link products when bovine serum albumin (Fig 3.23, 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.

3.1.15   Interaction of pFGE and FGE

While the results clearly demonstrate that pFGE and FGE have much in common with respect to their expression pattern, localization and structural features. Therefore we were interested to check whether pFGE interacts with FGE.

Biochemical approaches to study pFGE and FGE interaction

To determine whether pFGE forms intermolecular disulfide bridge with FGE, cells expressing pFGE and FGE were treated N-ethyl maleimide to irreversibly modify
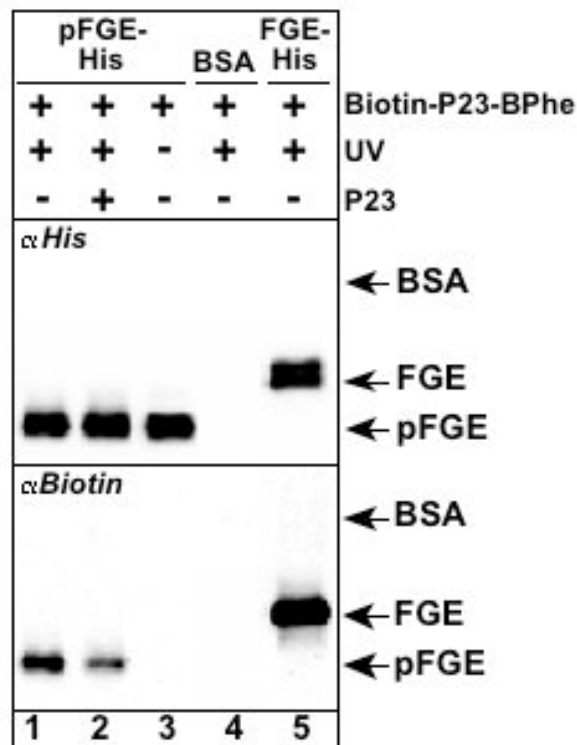
Figure 3.23. Photocross-linking of pFGE-His to an arylsulfatase A peptide.
pFGE-His (lane 1-3), bovine serum albumin (lane4) or FGE-His (lane5), 1μM each, were incubated with 50μM of a peptide corresponding to arylsulfatase A residues 59-81 with an additional ϵ- biotinylated lysine residue at the N-terminus and a benzyl phenylalanine at position 77 instead of leucine 77. In lane 2 a twofold excess of P23, corresponding to arylsulfatase residues 60-80 with additional N-acetylated methionine and a C- amidated serine residues at the N- and C- terminus, was added. After irradiation with UV- light for 30min on ice, equal aliquots of the samples were subjected to SDS-PAGE, transfer onto nitocellulose membrane and probed with antibodies against biotin (upper panel) the His-tag (lower panel). The position of bovine serum albumin, FGE and pFGE are indicated by arrows.

thiol groups. The cells were then lysed and extracts applied on 15% SDS polyacrylamide gels in reducing and non reducing conditions. After transfer to nitrocellulose membrane, pFGE and FGE were identified by immunodetction with a mixture of FGE and pFGE antiserum. The data show that under non reducing condition there was no cross reacting material at the size expected for FGE-pFGE complexes.

We next addressed whether pFGE and FGE form non-covalent complexes, Ni-agarose pull-down assay was performed using the extracts of HT1080 cells coexpressing pFGE-His and FGE. His-tagged pFGE was precipitated from the cell extracts with Ni-NTA agarose. The precipitant was examined for the presence of FGE by
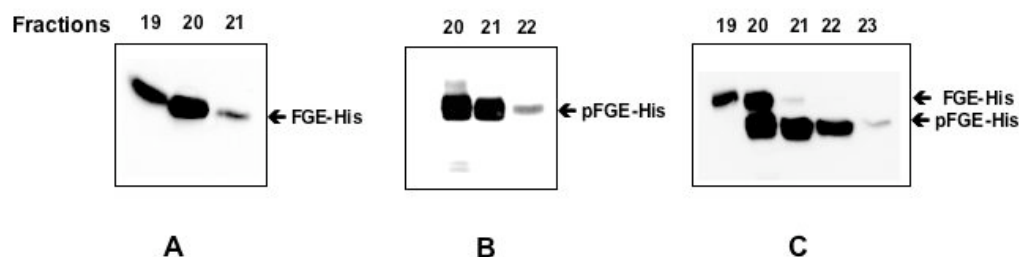
Figure 3.24. Purified pFGE-His and FGE-His on gel permeation chromatography
Fig A, B: Purified FGE-His or pFGE-His was chromatographed on Superdex-200. Protein fractions were analyzed by Western blot using a monoclonal antibody against His-tag. Fig C: Purified pFGE-His and FGE-His were mixed in a 2:1 molar ratio and incubated at room temperature for 30 minutes. The reaction was carried out in PBS at pH 7.4. Then the mixture was passed through Superdex-200. Protein fractions were analyzed by Western blot using His-tag monoclonal antibody.

Western blotting with FGE antiserum. Ni-NTA agarose pulled down pFGE-His and did not coprecipitate FGE along with pFGE-His(not shown).

We failed to detect complexes of pFGE and FGE by demonstration of co-elution during gel permeation chromatography, employing extracts of FGE-His and FGE coexpressing cell lines (not shown). This observation is substantiated by the same experiment was repeated with a mixture of purified FGE-His and pFGE-His. Whereas pFGE-His and FGE-His get eluted as monomers of 25 and 45 kDa, respectively (Fig 3.24).

Further we studied whether pFGE interacts with FGE in *in vivo* condition. To demonstrate this, the cells transiently coexpressing FGE-His and pFGE-HA were treated with the chemical cross-linker dithiobis (succinimidyl propionate) (DSP), which has a fixed spacer arm of 12 $\mathring{A}$ and a disulfide bond that can be cleaved with reducing agents. Cells were then lysed and the cell extracts were treated with or without DTT to cleave the DSP cross-linker. The extracts were resolved by SDS-PAGE and pFGE-HA and FGE-His were identified by immunoblotting with monoclonal antibodies of HA- or His-tag. The results show that there was no cross-linked adduct of pFGE and FGE, however, we observed FGE homodimer. In parallel *in vitro* cross-linking was performed using a mixture of purified pFGE-His and FGE-His. The cross-linked products were treated with or without DTT and
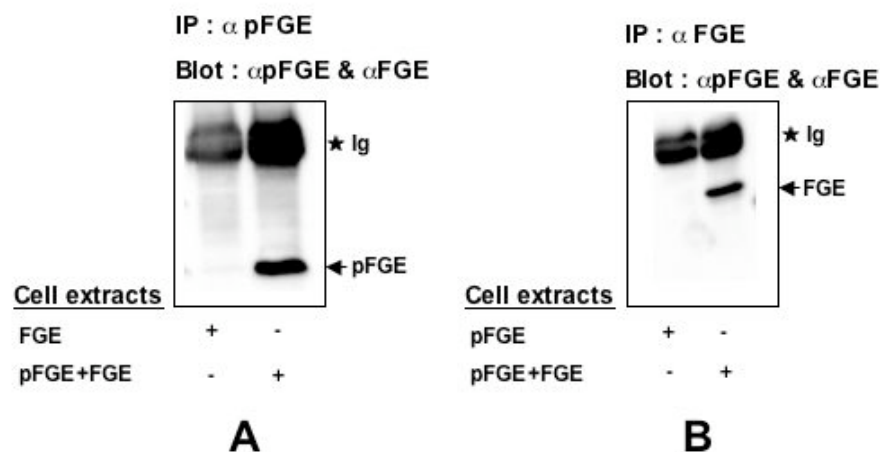
Figure 3.25. Coimmunoprecipitation of pFGE and FGE

Fig A: HT1080 cells were transiently transfected with FGE or FGE and pFGE. The extracts from these cells were immunoprecipiated with pFGE antiserum and immunoblotted with a mixture of pFGE and FGE antiserum. It should be noted that the transfected cDNAs were located on two vectors, one vector coexpressing FGE and pFGE from a bidirectional promoter, and another vector expressing either pFGE or FGE only (see "Materials and Methods"). Fig B: HT1080 cells were transfected with pFGE or FGE and pFGE. The extracts from these cells were immunoprecipiated with pFGE antiserum and immunoblotted with a mixture of pFGE and FGE antiserum

separated on SDS-PAGE and immunodetected with monoclonal antibody of His-tag. Again, no crosslinked products were obtained except FGE homodimer.

So far most of the biochemical approaches were done with using C-terminally His or HA tagged pFGE or FGE. In order to exclude the possibility of C-terminal His or HA tag interfering the interaction of pFGE and FGE, we performed coimmunoprecipitation experiments using cells expressing untagged pFGE and FGE. The protein extracts from these cells were immunoprecipitated by either pFGE or FGE antiserum. The precipitant was resolved on 15% SDS-PAGE and immunoblotted with a mixture of pFGE and FGE antiserum. The results demonstrate, neither pFGE antiserum nor FGE antiserum immunoprecipitate the complex of pFGE and FGE. However, pFGE antiserum or FGE antiserum could immunoprecipitate pFGE or FGE alone, respectively (Fig 3.25). The above coimmunoprecipitation was repeated with a mixture of purified pFGE-His and FGE-His. Invariably, there was no interaction between pFGE and FGE observed.

3.1.16 Interaction of pFGE and FGE in presence of Sulfatase substrate

We could not observe direct interaction of pFGE and FGE through biochemical approaches. Nevertheless we cannot exclude that pFGE and FGE interaction could be substrate-dependent. To check this, we 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 Km (13 nM) nor at a saturating P23 concentration (200 nM). Thus, in vitro pFGE has no inhibitory effect on FGE activity toward peptidic substrates.

Binding of pFGE to either peptidic substrate or FGE may not inhibit the FGly formation. To rule out this premise, we incubated pFGE and FGE in presence of peptidic substrate P23. Then the mixture was subjected to gel permeation chromatography. The elution profile reveals, pFGE do not co-purify with FGE in presence of substrate peptide. It is possible that FGE might interact with pFGE upon binding with substrate peptide. The transition state of the enzyme substrate complex would last for a short time therefore we may not observe the interaction of pFGE and FGE. The transition state can be stabilized by using mutant peptidic substrate C69S-ASA (positions 65-80 with the critical cysteine 69 mutated to serine) which can not be modified by FGE [41]. The above experiments was repeated with this mutant peptidic substrate. This data also supported that apparently there was no coelution of pFGE and FGE in presence of mutant peptidic substrate (not shown).

In parallel in vitro cross-linking was performed with a mixture of pFGE, FGE and peptidic substrate or mutant peptidic substrate. Invariably there was no crosslinked adducts were obtained except FGE homodimer. Taken together, all these in vitro experiments confirmed that pFGE does not interact with FGE under variable conditions.

To verify the in vivo interaction of pFGE and FGE in presence of sulfatase substrate. We co-transfected HT1080 stably expressing arylsulfatase A with FGE and pFGE. The protein extracts from these cells were immunoprecipitated with the pFGE or FGE antiserum and immunoblotted with a mixture of pFGE and FGE antiserum. Consistently, we found neither pFGE nor FGE antiserum precipitate the complex of pFGE and FGE (Fig 3.26). However, pFGE or FGE antiserum could precipitate pFGE or FGE alone, respectively.
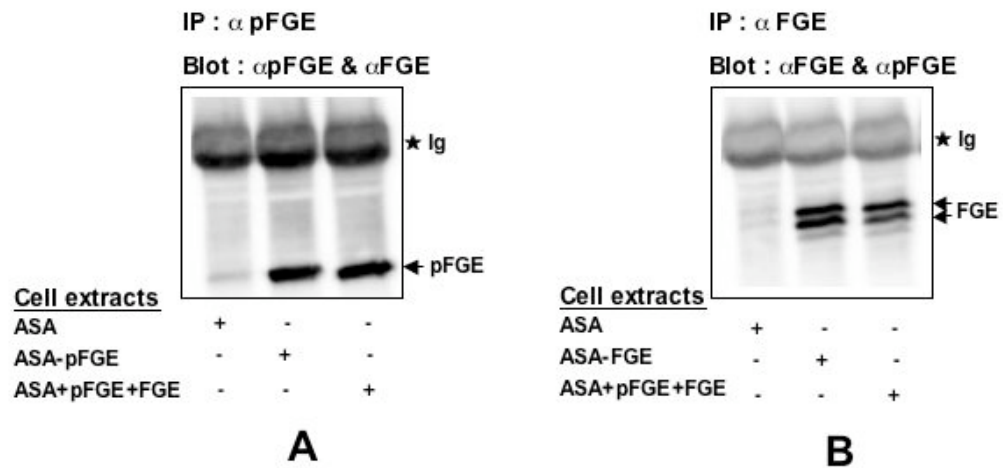
Figure 3.26. Coimmunoprecipitation of pFGE and FGE in presence of arylsulfatase A
Fig A: HT1080 cells stably overexpressing arylsulfatase A were transiently transfected with pFGE or FGE and pFGE. The extracts from these cells were immunoprecipiated with pFGE antiserum and immunoblotted with a mixture of pFGE and FGE antiserum. Fig B: HT1080 cells stably overexpressing arylsulfatase A were transfected with FGE or FGE and pFGE. The extracts from these cells were immunoprecipiated with pFGE antiserum and immunoblotted with a mixture of pFGE and FGE antiserum

3.2 Effect of FGE on sulfatases activitiy

Catalytically active recombinant sulfatases are needed for enzyme replacement therapy of patients suffering from the deficiency of specific lysosomal sulfatases. It was shown that endogenous FGE is limiting the synthesis of active recombinant sulfatases[52, 18]. Therefore we studied the effect of over-expression of FGE on sulfatase activity.

3.2.1 Establishing cell line and Transfection system

We examined steroid sulfatase activity (STS) in four different cell lines (CHO, BHK, Cos-7 and HT1080) cotransfected with FGE-HA and STS. There was no significant difference in STS activity between cells expressing STS together with or without FGE-HA. The failure to detect a stimulatory effect of FGE-HA on STS activity may be due to a low expression of STS which could be modified by endogenous FGE. It is also possible that FGE-HA and STS do not efficiently reach the same cells. Therefore we were seeking for a different experimental set up that allows efficient coexpression of FGE-HA and STS in the same cells. This was achieved by using tet-responsive expression vector pBI (Clontech), which allows regulated coexpression of two cDNAs (STS and FGE-HA) from a bidirectional CMV promoter.

3.2.2 Coexpression of FGE-HA and steroid sulfatase in HT1080 cells

Using the above set up, we transiently transfected HT1080 cells with STS or STS plus FGE-HA. The results show that FGE coexpression enhanced the activity of STS about three fold. The activity of FGE was increased up to 22 fold (Fig 3.27). In addition, we determined the activity of endogenous arylsulfatase A (ASA). But there was no enhancing effect of FGE on ASA. These observations emphasize that the endogenous FGE is sufficient to modify most of the recombinant STS. Thus, the stimulatory effect of FGE on STS is rather low upto 3 fold, however it is consistent.

3.2.3 Expression of FGE-HA or pFGE-HA in HT1080-galactose-6-sulfatase cells

To gain insight into the stimulatory effect of FGE on different sulfatase activity, we transiently transfected FGE-HA or pFGE-HA into HT1080 cells stably expressing
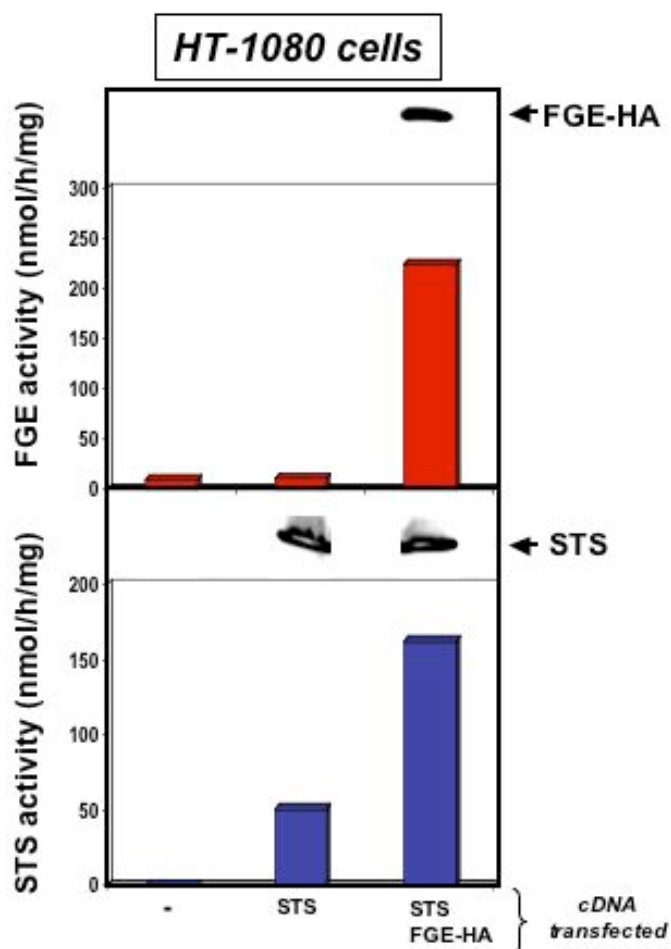
Figure 3.27. Coexpression of FGE-HA and steroid sulfatase
In HT1080 cells, steroid sulfatase (STS), FGE were transiently transfected in the combinations
indicated below the lanes. In the cell extracts, the expression of STS, FGE-HA was monitored
by western blot using a rabbit antiserum against STS and monoclonal antibody against the HA
tag. The activity of STS refered to total cell protein (bottom). The FGE activity was measured
after partial purification of cell extracts by MonoQ [41] and referred to total cell protein (top).

galactose-6-sulfatase. It should be noted that the surplus expression of galactose-6-sulfatase in this stable cell line by far exceeded the FGly generating capacity of endogenous FGE, thus leaving large amounts of unmodified galactose-6-sulfatase as a substrate for the recombinat FGE-HA. To our surprise, the result shows that FGE-HA expression enhanced the activity of galactose-6-sulfatase only 2.5 fold . There was no effect of pFGE-HA on galactose-6-sulfatase (Table3.2). It is conceivable that the expression of galactose-6-sulfatase and FGE-HA are not concomitant. Because galactose-6-sulfatase expression is stable and FGE-HA expression is transient. Therefore we were not able to see high stimulatory effect of FGE on galactose-6-

sulfatase activity.

Table 3.2. Transfection of FGE-HA or pFGE-HA in HT1080 galactose-6-sulfatase

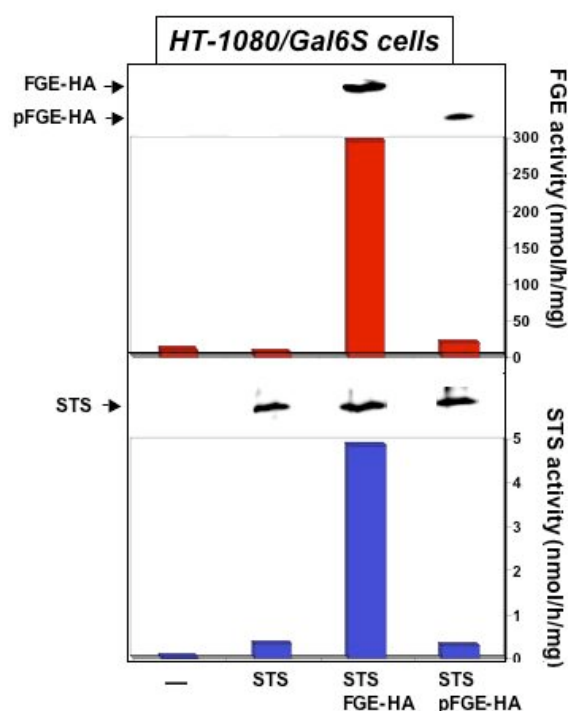| Expression	Galactose-6-Sulfatase nmol/h*mg	Steroid sulfatase nmol/h*mg	Arylsulfatase A mu/mg
-	6.2	0.01	0.02
FGE-HA	15.3	0.01	0.02
pFGE-HA	4.0	0.01	0.02

Figure 3.28.
HT-1080 cells stably expressing galactose-6-sulfatase were transfected with pFGE-HA, FGE-HA and STS in the combinations indicated below the chart. 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. The activity of steroid sulfatase and FGE, referred to total cell protein.

3.2.4 Coexpression of FGE-HA or pFGE-HA and STS in HT1080-galactose-6-sulfatase cells

Our aim was to establish the cell line in which sulfatase stimulation by FGE could be studied in more detail. For this we exploited a cell line stably overexpressing

galactose-6-sulfatase. Using this cells, we transfected STS or STS plus FGE-HA and STS plus pFGE-HA. Finally, we found FGE coexpression markedly enhanced STS activity about 12 fold. As expected, pFGE-HA coexpression did not increase the activity of STS rather it has very low inhibitory effect (Fig 3.28). Furthermore, we determined activity of endogenous ASA activity in these cells. But, there was no effect on ASA activity neither by FGE-HA or pFGE-HA.

3.2.5 Stable expression of FGE-His and galactose-6-sulfatase in HT1080 cells

In order to achieve high stimulatory effect of FGE on galactose-6-sulfatase, we stably transfected FGE-His into HT1080 galactose-6-sulfatase cells. In these cells, specific activity of galactose-6-sulfatase increased more than 100 fold, whereas coexpression of pFGE-His reduced specific activity of the galactose-6-sulfatase by 20% (as described in the earlier section)(Fig 3.29)
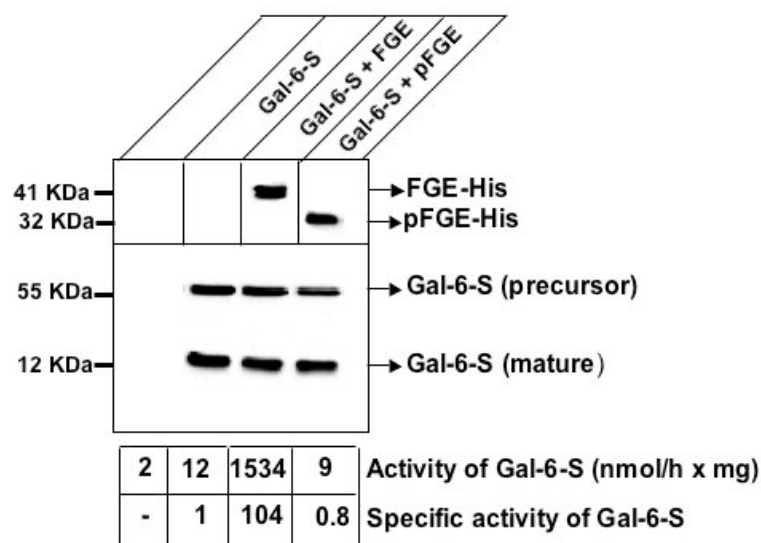
Figure 3.29.
In HT-1080 cells, galactose-6-sulfatase (Gal-6-S) and FGE-His or pFGE-His were stably coexpressed. In the cell extracts, the expression of galactose-6-sulfatase, FGE-His and pFGE-His was monitored by Western blotting using monoclonal antibodies against galactose-6-sulfatase (lower panel) and the His tag (upper panel). The activity and specific activity of galactose-6-sulfatase, referred to total cell protein and western blot signal of Gal-6-S, respectively are shown at the bottom.

3.2.6 Coelution of FGE-His and galactose-6-sulfatase

Through photocrosslinking experiment, we could demonstrate *in vitro* cross-linking of purified FGE-His and photo labeled peptidic substrate (Fig 3.23, lane 5). Therefore we were interested to check whether FGE and galactose-6-sulfatase form a complex *in vivo* . This was done by passing the extracts of HT1080 cells coexpressing FGE-His and galactose-6-sulfatase through gel filtration column. The result shows that the complexes (270KDa) of FGE-His and galactose-6-sulfatase elute in fractions 13 to 16 (Fig 3.30. open square). The monomers of FGE-His (35KDa) elutes in
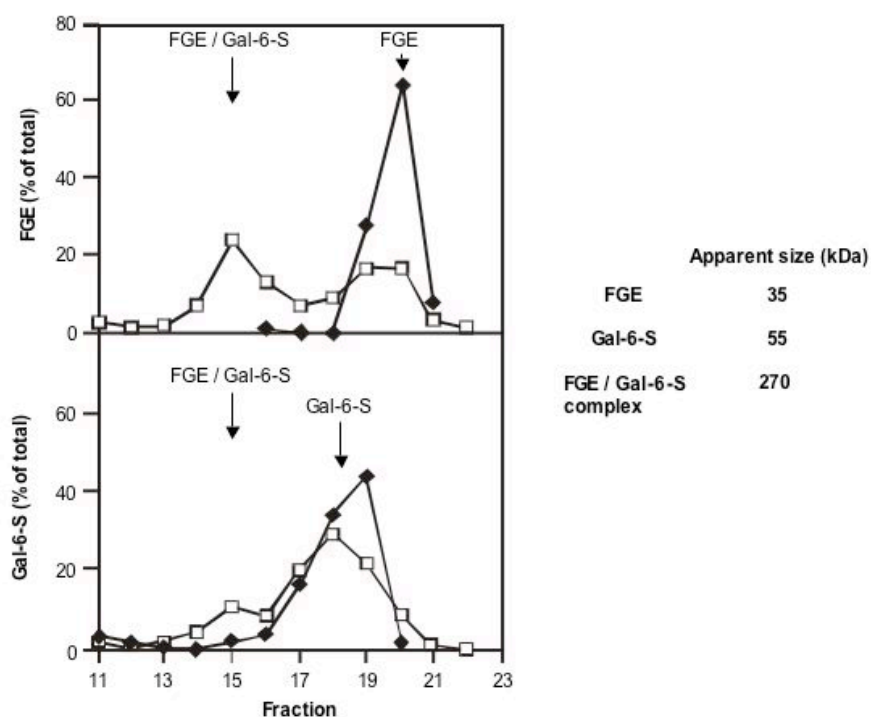
Figure 3.30. Coelution profile of galactose-6-sulfatase (Gal-6-S) and FGE-His
The extracts of HT1080 cells expressing FGE-His and galactose-6-sulfatase were fractionated by Superdex-200 as described under "Materials and Methods". The fractions were analyzed by immunobloting using a mixture of galactose-6-sulfatase and His-tag monoclonal antibodies. Percentage of FGE-His or Gal-6-S were calculated by quantifying the western blot signals using densitometry. In the graph, open square indicates that elution profile from cell expressing galactose-6-sulfatase and FGE-His and closed square points, elution profile from cells expressing either FGE-His or galactose-6-sulfatase alone.

fractions 18 to 21 and galactose-6-sulfatase elutes in fractions 16 to 21. In control experiments, extracts from FGE-His or galactose-6sulfatase expressing HT1080 cells were subjected to gel filtration chromatography. As can be seen from elution pattern, galactose-6-sulfatase or FGE-His exclusively eluted as monomers in the

same fractions where non-complexed fractions of FGE and Galactose-6-sulfatase elute (Fig 3.30. closed square).

3.2.7 Co-immunoprecipitation of FGE-His and galactose-6-sulfatase

The occurence of the FGE-His+galactose-6-sulfatase complex *in vivo* was evaluated by immunoprecipitation of recombinant proteins. We observed specific co-immunoprecipiation of FGE-His with galactose-6-sulfatase from lysates of HT1080 cells stably expressing FGE-His and galactose-6-sulfatase using FGE antiserum (Fig 3.31). The extracts of cells expressing only galactose-6-sulfatase was weakly precipitated by FGE antiserum, indicating that galactose-6-sulfatase forms complex with endogenous FGE (Fig 3.31). However, we cannot exclude the possibility of complex formation after extracting the cells.
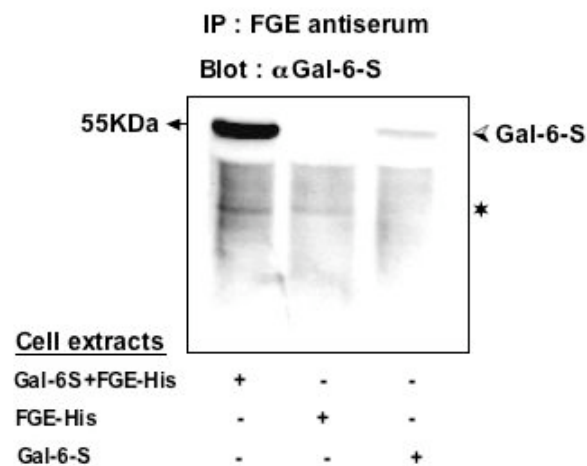
Figure 3.31. Interaction of FGE-His and galactose-6-sulfatase
The HT1080 cell extracts of galactose-6-sulfatase and FGE-His or FGE-His or galactose-6-sulfatase was subjected to immunoprecipitation with FGE antiserum. The presence of co-precipitated galactose-6-sulfatase was detected by immunobloting with a monoclonal of antibody against galactose-6-sulfatase. ∗ indicates the non specific detection from *Staphylococcus aureus* cells which was used for immunoprecipitation

3.2.8 *In vivo* interaction between FGE-His and galactose-6-sulfatase

Further, we investigated whether FGE-His forms complex with galactose-6-sulfatase *in vivo*. To demonstrate this, cells stably expressing FGE-His and galactose-6-

sulfatase were treated with the thio-cleavable chemical crosslinker dithiobis (succin-imidyl propionate) (DSP). Cells were then lysed and the cell extracts were treated with or without DTT to cleave DSP crosslinker. The extracts were resolved by SDS-PAGE and FGE-His and galactose-6-sulfatse were identified by immunoblotting with monoclonal antibodies against His-tag and galactose-6-sulfatase. The result reveals that high molecular weight cross linked adduct (above 250 KDa) was detectable (Fig 3.32). This observation agrees with the 270KDa complex of FGE-His and galactose-6-sulfatase which was shown by coelution experiment.
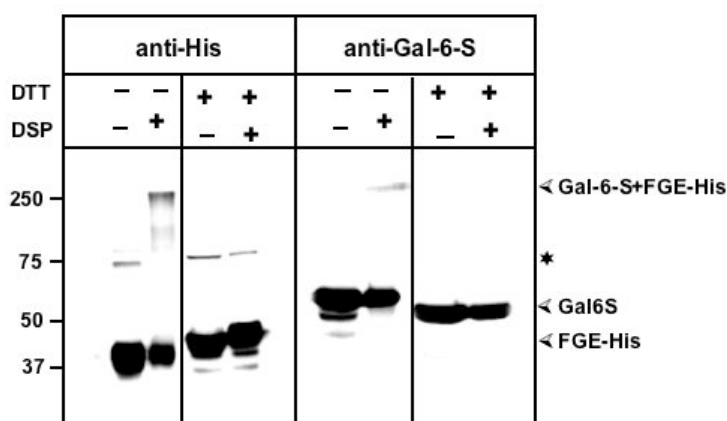
Figure 3.32. Chemical crosslinking of FGE-His and galactose-6-sulfatase
The HT1080 cells stably expressing FGE-His and galactose-6-sulfatase were treated with 2mM DSP. Crosslinked proteins were treated with or without 50mM DTT, resolved on 15% SDS-PAGE. After transferring into the nitrocellulose membrane, probed with either monoclonal antibodies against His or galactose-6-sulfatase.

3.2.9   Retention of FGE in sulfatase expressing cells

FGE is localized in the endoplasmic reticulum. However it does not contain an ER-retention signal of KDEL type. Its retention in the edoplasmic reticulum may therefore be mediated by the interaction with other ER proteins. Upon stable overexpression FGE escapes from the ER and is secreted. The time course of intracellular accumulation and secretion of FGE-His by HT1080 cells showed that a constant level of FGE is maintained in the cells and that the excess is secreted (Fig 3.33, A). An intriguing result was observed when FGE-His and galactose-6-sulfatase were stably coexpressed in HT1080 cells. Coexpression of galactose-6-sulfatase dramatically reduced the secretion of FGE-His (Fig 3.33, B). Based on these observations, we attempted to check whether overexpression of sulfatase
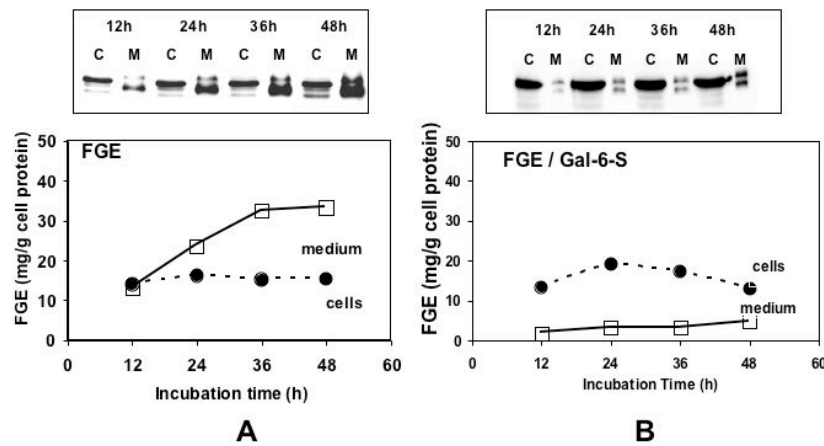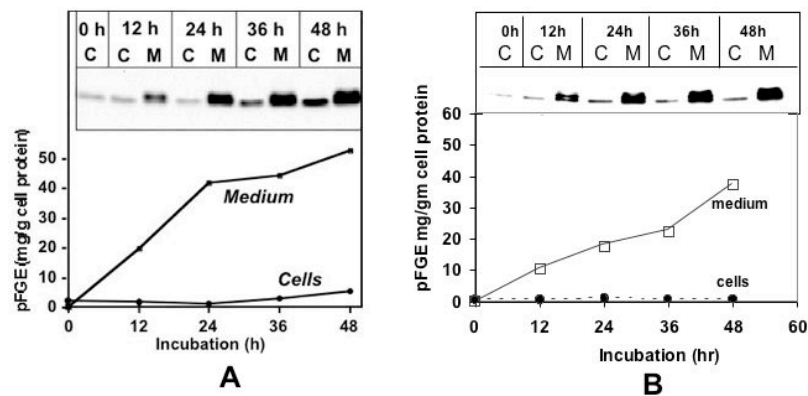
Figure 3.33. Secretion of FGE-His

Fig.A, HT1080-cells stably expressing FGE-His were grown to half-confluency. At time 0, fresh medium was added, and the cells were harvested after 12, 24, 36, and 48h of culture. FGE was quantified in equal aliquots of cell extracts and media by western blotting using a monoclonal antibody against His tag. The amount of intracellular and secreted FGE-His per mg of cell protein was determined by calibration of the Western blot with known amounts of purified FGE-His and referred to total cell protein. Fig.B, As described above, the secretion of FGE-His in HT1080 cells overexpressing galactose-6-sulfatase and FGE-His was performed.

Figure 3.34. Secretion of pFGE-His

Fig.A, HT1080-cells stably expressing pFGE-His were grown to half-confluency. At time 0, fresh medium was added, and the cells were harvested after 0, 12, 24, 36, and 48h of culture. pFGE was quantified in equal aliquots of cell extracts (C) and media (M) by western blotting using a monoclonal antibody against His tag. 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.B, As described above, the secretion of pFGE-His in HT1080 cells overexpressing galactose-6-sulfatase and pFGE-His was performed

impinge on the retention of pFGE. Coexpression of galactose-6-sulfatase did not affect the secretion profile of pFGE-His (Fig 3.34). This result supports the view that the substrate mediated retention is specific for FGE and not shared by its paralogue.

Chapter 4

Discussion

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

4.1 Molecular and cell biological characterization of pFGE

Sequence searches revealed that *SUMF2* encoded pFGE is found only in deuterostomia, including vertebrates and echinodermata, and it shares 47% amino acid identity and 62% similarity with *SUMF1* encoded FGE [41]. Northern blot analysis shows that pFGE is expressed ubiquitously in different tissues. Interestingly, we found a comparable amount of SUMF2 messenger RNA with respect to the SUMF1 transcript in all the tissues analyzed with the highest level in pancreas and kidney and the lowest in brain. The similar levels of SUMF1 and SUMF2 expression, point to a function of pFGE that is related to FGE activity. Using real time PCR analysis, Zito (2005) showed that SUMF2 expression in MSD cell lines was significantly lower with respect to the controls, suggesting that the transcription of SUMF2 is dependent on SUMF1.

The nucleotide sequence of pFGE predicts a N-terminal signal peptide, directing translocation of the nascent polypeptide into the lumen of the enoplasmic reticulum. HT1080 cells transiently expressing pFGE-HA were analyzed by immunofluorescence, where pFGE colocalized with PDI, an endoplasmic reticulum marker protein. Interestingly in some over-expressing cells, pFGE was detected in the Golgi (Fig

3.3, E). In contrast to recombinant pFGE, the endogenous pFGE was exclusively localized in the endoplasmic reticulum. In addition, HT1080 cells stably over-expressing pFGE-His were analyzed by immuno electron microscopy, where colocalization of pFGE-His with PDI or Glogi marker GM130 was detectable. pFGE-His was also found in structures enriched in the endosomal /lysosomal marker LAMP1 (Fig 3.4). These data have to be further verified for endogenous pFGE in human skin fibroblasts.

pFGE is localized in the lumen of the endoplasmic reticulum and it is found to be a soluble glycoprotein. Nevertheless, a small fraction of endogenous protein escapes from the endoplasmic reticulum into the secretions. Overexpression of pFGE-His in HT1080 cells, results in its massive secretion. However, the cellular pFGE-His is maintained at a constant level (Fig 3.5). pFGE is therefore likely to be retained in the endoplasmic reticulum by a saturable mechanism. It should be noted that pFGE lacks a canonical endoplasmic reticulum retention signal of the KDEL type. It remains unclear how soluble resident proteins lacking this or a related motif are retained in the endoplasmic reticulum. One possibility could be that pFGE interacts with a KDEL type containing protein. Fishing of partners interacting with immobilized pFGE was unsuccessful so far. It was recently demonstrated that the C-terminal domain (57 residues) of Dictyostelium PDI lacking the HDEL retrieval signal is necessary for endoplasmic reticulum retention [56]. Therefore it is interesting to check whether the C-terminal region of pFGE confers the endoplasmic reticulum retention. To our surprise, the secretion of FGE-His is markedly decreased from cells coexpressing galactose-6-sulfatase and FGE-His (Fig 3.33), suggesting that FGE retention may be mediated by their substrate sulfatases. Conversely, coexpression of galactose-6-sulfatase has no impinge on secretion of pFGE-His (Fig 3.34). This result suggests that pFGE may use a retention machinery which is not shared by FGE.

pFGE has a single potential N-glycoyslation site at Asn 191. Full length pFGE from cells and secretions differed in electrophoretic mobility by 0.5KDa (Fig 3.6). The endoglycosidases experiment revealed that intracellular pFGE is sensitive to Endo-H and PNGase F. Whereas, in two forms of secreted pFGE, the smaller form is sensitive to Endo-H and the larger form is resistant. Both forms are sensitive to PNGase F. MALDI-TOF analysis of tryptic glycopeptides of pFGE confirmed heterogenous glycosylation at Asn 191. All these data suggest that 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 3.3, D-F), is negligible. It is evident that Golgi-dependent modification reactions are restricted to the fraction of pFGE that escapes endoplasmic reticulum retention and becomes secreted. A similar result was observed when intra- and extracellular FGE were digested with Endo-H or PNGase [54].

pFGE-His was purified from the secretions in a two step protocol on Ni-NTA agarose followed by MonoQ anion exchange chromatography. On SDS gels and Western blot, purified pFGE-His was detected in two molecular forms, the larger with an apparent mass of 32.5 KDa and the smaller of 31.5 KDa. However, both forms corresponded to full length pFGE with alanine 27 at its N-terminus. The obvious explanation for this heterogeneity in size of secreted pFGE-His is its different glycosylation pattern (Fig 3.6).

In order to detect endogenous and recombinant pFGE, we raised polyclonal antibodies against purified pFGE-His in rabbit. Through immunoprecipation experiments, the polyclonal antibodies were shown to precipitate recombinant as well as endogenous pFGE. In addition, it could be used to specifically detect pFGE in immunofluoresence studies. Nevertheless, on Western blots pFGE antiserum weakly cross-reacted with FGE.

After two attempts, we failed to produce polyclonal antiserum which could be used to detect exclusively pFGE on Western blots. One explanation for this problem could be that pFGE is too similar to FGE. On Western blots, FGE polypeptides are largely denatured thereby all the epitopes were accessible to pFGE antiserum. To avoid this problem, we raised monoclonal antibodies against purified pFGE-His in mouse. All the hybridoma clones were screened by immunoprecipitation, immunoblotting and immunofluoresence experiments. Among many different clones, the highly reactive clones were selected to produce ascites. Surprisingly, all these monoclonal antibodies recognized His-tagged intracellular and secreted pFGE but not untagged intracellular pFGE. Furthermore, they cross-reacted with His tagged intracellular FGE and not with untagged intracellular and secreted FGE-His. These observations were substantiated by immunofluoresence experiments. Taken together all these data support the conclusion that monoclonal antibodies were likely raised against epitopes that includes the His-tag of pFGE-His.

Phylogenetic sequence comparison revealed that pFGE and FGE contain three highly conserved regions. In pFGE, these subdomains make up more than 85% of the molecule. Digestion with either elastase, a serine protease, or thermolysin, a zinc-metalloproteinase, generated two stable fragments by proteolytic cleavage within the short linker sequence connecting the first and the second subdomain. This suggested that the three conserved regions correspond to folding domains.

In collaboration with Achim Dickmanns, the crystal structure of pFGE was determined , which represents the first three dimensional structure of a DUF323 domain [57]. In contrast to the domain mapping studies, the pFGE structure revealed a single domain of a novel fold with a strikingly low degree of secondary structure. pFGE structure was used to determine the crystal structure of FGE, revealing a high structural similarity [58](Fig 4.1)). FGE contains three highly conserved
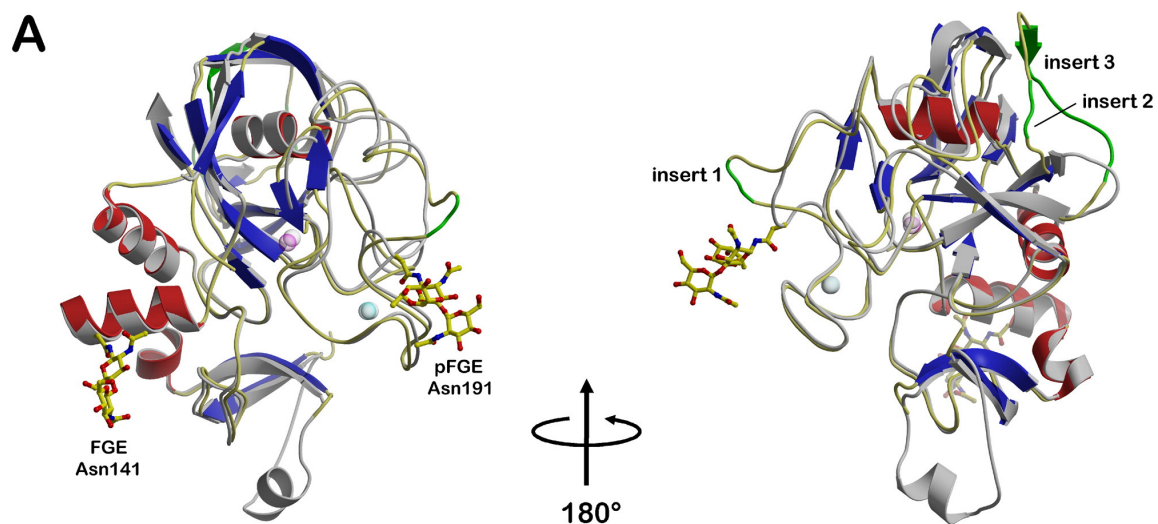

Figure 4.1. Comparison of FGE and pFGE (Dierks et., al 2005)
Superposition of FGE and pFGE. pFGE is shown as a *gray* ribbon representation, whereas for FGE α helices are colored *red*, β strands *blue*, and loop regions *yellow*. The two Ca^{2+} ions in pFGE are drawn in *gray*, and the Ca^{2+} ions in FGE are shown as transparent spheres in *cyan* (site 1) and *magenta* (site 2). The three insert regions in FGE relative to pFGE are labeled and shown in *green*. The carbohydrate moieties are located on opposite sides of the molecules and are drawn as stick models.

cysteine residues in the highly conserved C-terminal region (subdomain III [54]). Two of them have been shown to be critical for its catalytic activity [58]. Both

these cysteines are missing in pFGE. pFGE contains two cysteine residues located in subdomain II (Cys156) and another in subdomain III (C290). They form an intra- molecular disulfide bond linking subdomains II and III. This disulfide bond is conserved in FGE, in which two inter subdomain disulfide bridges link subdomains II and III [54]. Therefore, in pFGE as well as in FGE, subdomains II and III are tightly connected and together form a large protease-resistant domain.

4.1.1   Functional properties of pFGE

The localization, carbohydrate processing and secretion upon overexpression as well as protease sensitivity of the linker region between the first two subdomains are properties that FGE shares with its paralog FGE [54]. Moreover, the relative abundance of their mRNAs in different tissues is rather similar. All these observations suggest that pFGE and FGE fulfill 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 MSD patients that carry mutations in the gene encoding FGE [51, 41]. Cells of these patients accumulate catalytically inactive sulfatases that retain the cysteine residue, which normally is oxidized by FGE to FGly [18]. So far, no mutations in the pFGE-encoding *SUMF2* gene have been found in MSD patients.

Under conditions appropriate for the FGly-generating activity of FGE towards peptidic substrates, pFGE was inactive (Fig 3.20,Table 3.1). This was true for the peptides derived from the 16 sulfatases known or predicted from the human genome sequence. Thus, under *in vitro* conditions, pFGE lacks FGly-generating activity. This result agrees with crystallographic studies of pFGE and FGE showing that pFGE lacks the redox-active Cys 336 of FGE [58]. The increase of sulfatase activity upon coexpression of pFGE with sulfatases as observed by Cosma et al., (2003), had suggested that *in vivo* pFGE has FGly-generating activity or is stimulating the activity of endogenous FGE . We failed to observe an increase of sulfatase activity upon transient or stable coexpression of pFGE.

Surprisingly, in HT-1080 cells, the coexpression of pFGE decreased rather than increasing the activity of coexpressed sulfatases (Fig 3.21). 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 photocrosslinking, respectively. The possibility of interaction between pFGE and ASA derived peptides could be explained at molecular level by crystallographic studies. Many of the residues (Ala149, Ser155, Trp299, Ser356, Asn360, and Leu361) lining the substrate binding groove in FGE are conserved in pFGE. This points to the possibility that pFGE competes with FGE for substrates [58]. Pro-182 of FGE, which was shown to chemically cross-link with the bound substrate related peptide, is structurally conserved in pFGE (Pro-120). These prolines border one end of a cleft thought to be involved in peptide binding [57](Fig4.2).
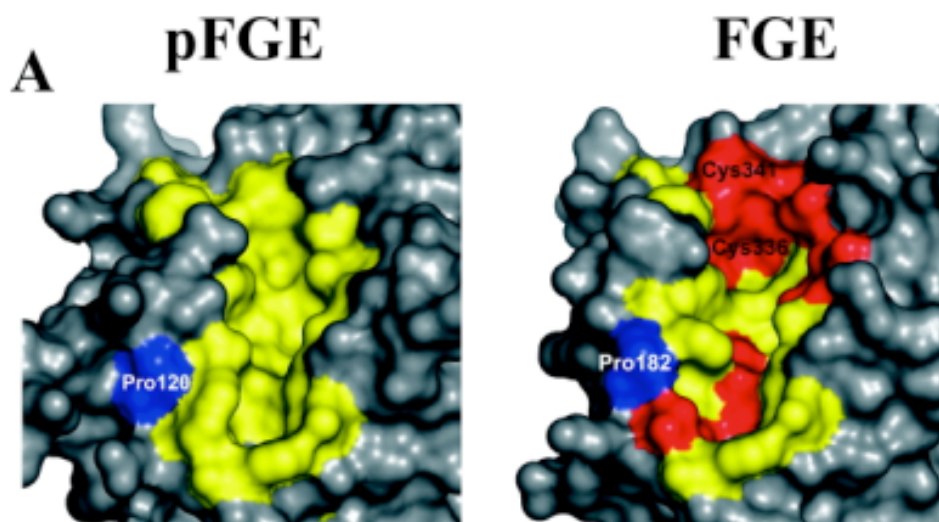
Figure 4.2. Comparison of the putative peptide binding clefts of pFGE and FGE (Achim et., al 2005)
A, left, pFGE in *gray* with the cleft (*yellow*) most likely to be involved in the binding of the sulfatase polypeptide chain colored in *yellow*. Pro-120, colored in *blue*, corresponds to Pro-182 of FGE that has been shown to cross-link to a substrate peptide. Right, the FGE cleft is shown in the same color code used for pFGE. The amino acid changes are highlighted in *red* and are predominantly located in the region of the putative reaction center.

In yeast two hybrid assay, we could also observe an interaction between pFGE and FGE. It was shown that the first two cysteine residues in FGE are involved in the formation of homodimer [54]. Therefore, we were keen to check whether pFGE makes intermolecular disulfide with FGE. In cells expressing FGE and pFGE treated with NEM prior to harvesting, no heterodimer of pFGE and FGE was detected on Western blots of non reducing SDS-gels. Furthermore, we were unable to detect a thiol mediated homodimer of pFGE from cells over-expressing pFGE. Nevertheless, pFGE was crystallized as a homodimer in the asymmetric unit [57]. This observation

suggests that homodimer formation of pFGE might implicate to its function. Further, we have to verify the possibility of homodimer formation of pFGE through non covalent interactions by using differentially tagged pFGE molecules.

We also analysed whether pFGE forms a direct physical interaction with FGE. Ni-agarose pull down assay, gel permeation chromatography and cross-linking were performed, employing extracts of pFGE-His and FGE coexpressing cells. Yet, we failed to detect complexes of pFGE and FGE. In solution, no heterodimers of purified pFGE and FGE were observed by gel permeation chromatography, suggesting that the affinity of pFGE for FGE, if present, is low. pFGE/FGE dimers are sterically possible, albeit with a small buried surface area and low surface complementarity [57]. In order to rule out the possibility of His tag interfering the interaction of pFGE and FGE, we made coimmunoprecipitation experiments using cells expressing untagged pFGE and FGE. The results clearly state that no complexes of pFGE and FGE were precipitated (Fig 3.25). However, Zito et al., (2005) described, the complexes of pFGE and FGE detected by coimmunoprecipiation experiments [59]. The contradictory results between our two groups remain to be resolved and theoretically can be due to the cell lines which were used for co-expression of pFGE and FGE.

Apparently, the interaction of pFGE and FGE as observed on the yeast two hybrid experiments could not be confirmed by biochemical approaches. Albeit, we cannot exclude that pFGE and FGE interaction could be substrate-dependent. In solution, heterodimers of pFGE and FGE in the presence of peptidic substrates were not detected by gel permeation chromatography. This observation was substantiated by cross linking experiments with a mixture of pFGE, FGE and substrate peptide. Further, we checked this interaction by using cells expressing pFGE, FGE and arylsulfatase A. Consistently, we found that neither pFGE nor FGE antiserum could precipitate complexes of pFGE and FGE.

4.1.2 Which is the biological role of pFGE?

Taken together, our results support the view that pFGE lacks FGly-generating activity. Thus, we may conclude that pFGE is not responsible for the residual activity of sulfatases in MSD patients [41]. Nevertheless the function of pFGE is likely to be related to FGE funtion for seven reasons.

1. The SUMF2 expression levels are substantial though absolutely variable and the relative pFGE/FGE ratios are approximately equal in all tissue types studied.

2. In MSD cells, *SUMF2* expression was significantly lower with respect to the controls, suggesting that the transcription of *SUMF2* is dependent on *SUMF1*[59].

3. pFGE and FGE are colocalized in the endoplasmic reticulum.

4. Recently pFGE/FGE heterodimerization has been observed by coimmunoprecipitation of the proteins from transfected Cos7 cells [59]. In addition, this study reported a trimolecular complex of pFGE, FGE, and sulfatases.

5. The compromising effect of pFGE on FGE mediated generation of catalytically active sulfatases in coexpression studies

6. pFGE has been shown to bind peptides bearing the recognition motif required for the generation of FGly in arylsulfatase A *in vitro*.

7. The three dimensional structures of pFGE and FGE are closely related and share residues bearing the peptide binding groove [57, 58].

A number of proposals can be suggested to explain the role of pFGE. *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 sequestering them from FGE. *In vitro*, addition of pFGE to a mixture of FGE and its peptidic substrate P23 has no effect on the FGly formation. While this result suggests an effect of pFGE on FGE activity *in vivo*, it can be expected that another unidentified protein with a function that links pFGE to FGE activity may be induced. The effects of pFGE on FGE are also conceivable either in a indirect manner, e.g., by dislocating FGE from the endoplasmic reticulum due to competition for a common retention mechanism, or directly by heterodimerization. The earlier possibility is not likely to happen because the retention of FGE appears to be mediated by the sulfatase substrates, which is not true for pFGE. The latter possibility is so far not confirmed by biochemical approaches but agrees with the crystal structure of pFGE which shares structural similarity with FGE. Moreover, the formation of a heterodimer appears to be stereochemically possible [57]. It is plausible, that pFGE may act as a modulator of FGE activity. Thus, pFGE may control the amount of the FGly activated sulfatases to be synthesised. Further, it will be fascinating to determine whether pFGE regulation is sulfatase specific and/ or tissue specific. To identify the precise biological role of pFGE, the detailed analysis of the molecular interaction of pFGE and FGE and nascent sulfatase polypeptides as well as ablation of pFGE are required. Soon pFGE knock-out mice will be available, which hopefully will help to elucidate the precise role that pFGE plays on sulfatase activation.

4.2   FGE as a tool to improve the production of recombinant sulfatases

Inactivity of FGE causes multiple sulfatase deficiency (MSD), an autosomal recessive lysosomal storage disease. In sulfatases, FGE posttranslationally converts a cysteine residue to FGly, which is part of the catalytic site and is essential for sulfatases activity. Overexpression of sulfatases can be associated with a decrease of the activity of endogenous sulfatases [52]. Furthermore, the endogenous FGE is limiting the synthesis of active recombinant sulfatases. The recent preclinical studies from Matzner et al., (2005) in a mouse model of metachromatic leukodystrophy (MLD) revealed that intravenous injection of the human recombinant ASA can reduce excess sulfatide in peripheral nervous system and kidney by up to 70%. They also demonstrate the passage of lysosomal enzymes through the blood brain barrier and their delivery to the brain. These studies encourage the enzyme replacement therapy for single sulfatases deficiency. In order to produce catalytically active sulfatases, we began to examine the effect of FGE on sulfatases. Our findings show that FGE is absolutely necessary for the expression of catalytically active recombinant sulfatases. Thus, these studies have profound implications for the mass production of active sulfatases to be utilized for enzyme replacement therapy.

4.2.1   Concurrent expression of FGE and sulfatase is essential to produce highly active sulfatase

When FGE-HA and steroid sulfatase (STS) were coexpressed in HT1080 or HT1080-galactose-6-sulfatase cells, STS activity increased in both the cells, but the increase was more pronounced in HT1080-galactose-6-sulfatase. The variable results between the two cell lines could be explained. In HT1080 cells, the endogenous FGE is able to modify recombinant STS to some extent, thus, we could not see a distinct effect of FGE on STS. But, in HT1080-galactose-6-sulfatse cells, the superfluous expression of galactose-6-sulfatase by far exceeded the FGly generating capacity of endogenous FGE, thus leaving large amounts of unmodified galactose-6-sulfatase and STS as substrates for recombinant FGE.

Strikingly, the specific activity of galactose-6-sulfatase increased more than 100 fold in HT1080 cells stably coexpressing FGE-His and galactose-6-sulfatase. Conversely, we failed to notice a significant difference of Gal-6-S activity when we transiently transfected FGE-His into HT1080-galactose-6-sulfatase cells. The dis-

crepancy between the two experiments can be explained. In transient transfection FGE-His does not reach all cells and so concomitant. But, in the stable cell line, the synthesis of FGE-His and galactose-6-sulfatase is concomitant, hence FGE-His could modify almost all newly synthesized galactose-6-sulfatase. However, it is unclear whether all the galactose-6-sulfatases are catalytically active in these cells. One potential way to address this question is to quantify the FGly content in galactose-6-sulfatase. Takakusaki et al., (2005) also examined the validity of FGE coexpression for the expression of a catalytically active ASA enzyme, as a step in the development of gene therapy for metachromatic leukodystrophy (MLD). In addition this study reported that FGE coexpression did not affect ASA mRNA or protein levels and half life of ASA. Synergistic activation was clearly demonstrated in their in vivo experiments. Simultaneous injection of ASA cDNA with FGE cDNA by a hydrodynamic procedure significantly increased ASA activity in the liver and serum of ASA deficient (MLD) mice. However, they could not show the transfer of ASA protein from blood to the brain.

4.2.2   Formation of FGE-sulfatase complexes in the ER: retention mechanism or non-productive complexes?

FGE-His can be crosslinked with photolabelled substrate (Fig 3.23, lane 1). A yeast two hybrid assay revealed that ASA fragment harboring FGly modification motif interacts with FGE [54]. These data tempted us to check whether FGE and sulfatases form a complex *in vivo* . When extract of HT1080 cells coexpressing FGE-His and galactose-6-sulfatase were subjected to gel permeation chromatography, complexes of FGE and galactose-6-sulfatase coeluted (Fig 3.30). The apparent size of the FGE-His and galactose-6-sulfatase complex was 270 KDa. However, FGE-His and galactose-6-sulfatase partially eluted as monomers of 35 KDa and 55 KDa, respectively. The complex formation must be specific because pFGE-His and galactose-6-sulfatase did not colelute when passing the cell extracts of pFGE-His and galactose-6-sulfatase on gel permeation chromatography (Fig 3.24). The stoichiometry of this complex is yet to be characterized.

Further, we verified the interaction of FGE-His and galactose-6-sulfatase by coimmunoprecipitation. Again, we found that FGE-His coprecipitated with galactose-6-sulfatase (Fig 3.31). However, we cannot exclude that the complex formation might occur after extracting the cells. To exclude the later possibility, we have done *in vivo* chemical cross-linking of FGE-His and galactose-6-sulfatase. In agreement with the coelution experiment, we could observe the high molecular weight cross

linked adduct (above 250 KDa). To substantiate these data, we should repeat these studies with different sulfatases coexpressed with FGE.

It is unclear whether the complex comprises only FGE-His and galactose-6-sulfatase or some unknown factor. This can be achieved by purifying the complex and analysing by MALDI-TOF. Further, it will be interesting to know whether FGE forms complexes with catalytically active or inactive galactose-6-sulfatase or whether inactive complexes are eventually converted into active complexes allowing for quantitative FGly-modification. The later is suggested by the observation that the specific activity of galactose-6-sulfatase increased up-to 100 fold in these cells. It has been shown that FGE can not modify completely folded and inactive sulfatases because the FGly motif is buried inside the sulfatases. To exclude whether the complex formation occurs through non specific interactions, one has to coexpress FGE-His with mutated galactose-6-sulfatase (Cys79Ser) and to analyze for the complex formation. Further it will be intriguing to understand the physiological role of FGE-sulfatase complexes and what is limiting the release of non-complexed sulfatase.

Recombinant and endogenous FGE are localized in the endoplasmic reticulum [41, 54]. But, FGE by-itself does not contain a canonical endoplasmic reticulum retention signal. The retention mechanism of FGE in the endoplasmic reticulum remains elusive. Upon stable overexpression, FGE escapes from the endoplasmic reticulum and is secreted (Fig 3.33, A). However, the intracellular FGE was kept at a constant level. When we stably coexpressed FGE-His and galactose-6-sulfatase in HT1080 cells, the secretion of FGE-His was dramatically reduced (Fig 3.33, B). In parallel, we repeated this experiment with cells coexpressing pFGE-His and galactose-6-sulfatase, where coexpression of galactose-6-sulfatase did not have an influence on secretion of pFGE-His (Fig 3.34). This result supports that sulfatase-mediated retention of FGE is specific for FGE and not shared by its paralogue. In order to certify these observations, we should repeat this experiment with different sulfatases coexpressing with FGE. These findings lead us to propose that FGE is constantly interacting and modifying all the newly synthesized nascent sulfatase polypeptides and thus it is retained in the endoplasmic reticulum. In order to confirm this observation, we should repeat these secretion studies with different sulfatases coexpressed with FGE.

Summary

Cα-formylglycine (FGly) is the catalytic residue in the active site of sulfatases. In eukaryotes, it is generated in the endoplasmic reticulum by post-translational modification of a conserved cysteine residue. The enzymatic activity of sulfatases depends on this posttranslational modification which is mediated by a recently identified unique enzyme named as formylglycine generating enzyme (FGE). The genetic defect of FGE causes multiple sulfatase deficiency (MSD), a lysosomal storage disorder. In the genomes of deuterostomia, including vertebrates and echinodermata, a paralog of FGE, designated as pFGE, is exist. The human pFGE shares 47.1% sequence identity and 62.1% similarity with human FGE. The function of the pFGE is unknown. The goal of this study was to identify and analyze the structural and functional properties of pFGE. In comparison with FGE, pFGE shares a tissue specific expression pattern and the localization in the lumen of the endoplasmic reticulum. Upon overexpression, both FGE and pFGE escape from the endoplasmic reticulum and get secreted. However a constant level of pFGE or FGE is maintained in the cells. Although both lack endoplasmic reticulum retention signal, they are still retained by an unknown saturable mechanism. pFGE was purified to homogeneity from the secretions in a two-step protocol on Ni-NTA agarose followed by anion exchange chromatography and polyclonal and monoclonal antibodies were raised. 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 towards peptidic substrates. This observation holds true for peptidic substrates derived from all 16 sulfatases known or predicted from the human genome. 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.

In an independent study, we examined the utility of FGE for overexpression of catalytically active sulfatases in a way to improve the efficacy of sulfatase replace-

ment therapy. Our data demonstrating that FGE coexpression is indispensable for efficient synthesis of functional sulfatases. HT1080 cells were stably transfected with galactose-6-sulfatase and FGE. The specific activity of galactose-6-sulfatase increased up to 100 fold in cell lysates by coexpression of FGE. Interestingly, stimulation of sulfatase activity is more pronounced when FGE and sulfatase are concomitantly expressed in the HT1080 cells. Moreover, stable FGE-sulfatase complexes are formed in the endoplasmic reticulum but its physiological significance remains elusive. However, this complex formation may aid in the retention of FGE as a substrate-mediated phenomenon; substantiated by the observation that the secretion of FGE is reduced markedly when coexpressed with sulfatase.

Bibliography

[1] KF. Medzihradszky, Z. Darula, E. Perlson, M. Fainzilber, RJ. Chalkley, H. Ball, D. Greenbaum, M. Bogyo, DR. Tyson, RA. Bradshaw, and AL. Burlingame. O-sulfonation of serine and threonine: mass spectrometric detection and characterization of a new posttranslational modification in diverse proteins throughout the eukaryotes. *Mol Cell Proteomics.*, 3(5):429–40, 2004. 1

[2] CD. Klaassen and JW. Boles. Sulfation and sulfotransferases 5: the importance of 3'-phosphoadenosine 5'-phosphosulfate (PAPS) in the regulation of sulfation. *FASEB J*, 11(6):404–18, 1997. 1

[3] B. Franco, G. Meroni, G. Parenti, J. Levilliers, L. Bernard, M. Gebbia, L. Cox, P. Maroteaux, L. Sheffield, and GA. Rappold. A cluster of sulfatase genes on Xp22.3: mutations in chondrodysplasia punctata (CDPX) and implications for warfarin embryopathy. *Cell*, 81(1):15–25, 1995. 2

[4] G. Parenti, G. Meroni, and A. Ballabio. The sulfatase gene family. *Curr Opin Genet Dev*, 7(3):386–91, 1997. 2, 3

[5] E.F. Neufeld and J. Muenzer. The Metabolic and Molecular Bases of Inherited disease. *McGraw-Hill, Newyork*, III, 8th ed (Ed.: C. R. Scriver):2465–2494, 1999. 2, 4

[6] C. Peters, B. Schmidt, W. Rommerskirch, K. Rupp, M. Zuhlsdorf, M. Vingron, HE. Meyer, R. Pohlmann, and K. von Figura. Phylogenetic conservation of arylsulfatases cDNA cloning and expression of human arylsulfatase B . *J Biol Chem*, 265(6):3374–81, 1990. 2

[7] Wong CH Hanson SR, Best MD. Sulfatases: structure, mechanism, biological activity, inhibition, and synthetic utility. *Angew Chem Int Ed Engl*, 43(43):5736–63, 2004. 3

[8] A. Daniele, G. Parenti, M d'Addio, G. Andria, A. Ballabio, and G. Meroni. Biochemical characterization of arylsulfatase E and functional analysis of

mutations found in patients with X-linked chondrodysplasia punctata. *Am J Hum Genet*, 62(3):562–72, 1998. 3

[9] X. Ai, AT. Do, O. Lozynska, M. Kusche-Gullberg, U. Lindahl, and Jr. Emerson CP. QSulf1 remodels the 6-O sulfation states of cell surface heparan sulfate proteoglycans to promote Wnt signaling. *J Cell Biol*, 162(2):341–51, 2003. 3

[10] J. Lai, J. Chien, J Staub, R. Avula, EL. Greene, TA. Matthews, DI. Smith, SH. Kaufmann, LR. Roberts, and V. Shridhar. Loss of HSulf-1 up-regulates heparin-binding growth factor signaling in cancer. *J Biol Chem*, 278(25):23107–17, 2003. 3

[11] R. Sasisekharan, Z. Shriver, G. Venkataraman, and U. Narayanasami. Roles of heparan-sulphate glycosaminoglycans in cancer. *Nat Rev Cancer*, 2(7):521–8, 2002. 3

[12] E.H. Kolodny and A.L. Fluharty. Metachromatic leucodystrophy and multiple sulfatase deficiency: Sulfatide lipidosis. In: Scriver, C.R., Beaudet, A.L., Sly, W.S., Valle, D. (eds.),. *The Metabolic and Molecular Bases of Inherited disease. McGraw-Hill, Newyork*, pages 2693–2741, 1995. 4, 5

[13] A. Ballabio and L. Shapiro. STS deficiency and X-linked ichtyosis; in: The Metabolic and Molecular Bases of Inherited Disease; publ. Scriver C.R., Beaudet. *McGraw-Hill, New York*, pages 2999–3022, 1995. 4

[14] R. von Bulow, B. Schmidt, T. Dierks, K. von Figura, and I. Uson. 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, 2001. 5, 6, 7, 8

[15] F. Steckel, A. Hasilik, and K. von Figura. Synthesis and stability of arylsulfatase A and B in fibroblasts from multiple sulfatase deficiency. *Eur J Biochem*, 151(1):141–5, 1985. 6, 16

[16] Y. Eto, I. Gomibuchi, F. Umezawa, and T. Tsuda. Pathochemistry, pathogenesis and enzyme replacement in multiple-sulfatase deficiency. *Enzyme*, 38(1-4):273–9. Review, 1987. 6

[17] W. Rommerskirch and K. von Figura. Multiple sulfatase deficiency: catalytically inactive sulfatases are expressed from retrovirally introduced sulfatase cDNAs. *Proc Natl Acad Sci U S A*, 89(7):2561–5, 1992. 6

[18] B. Schmidt, T. Selmer, A. Ingendoh, and K. von Figura. A novel amino acid modification in sulfatases that is defective in multiple sulfatase deficiency. *Cell*, 82(2):271–8, 1995. 6, 16, 18, 70, 83

[19] T. Selmer, A. Hallmann, B. Schmidt, M. Sumper, and K. von Figura. The evolutionary conservation of a novel protein modification, the conversion of cysteine to serinesemialdehyde in arylsulfatase from Volvox carteri. *Eur J Biochem*, 238(2):341–5, 1996. 6

[20] C. Miech, T. Dierks, T. Selmer, K. von Figura, and Schmidt B. Arylsulfatase from Klebsiella pneumoniae carries a formylglycine generated from a serine. *J Biol Chem*, 273(9):4835–7, 1998. 6, 11

[21] T. Dierks, C. Miech, J. Hummerjohann, B. Schmidt, MA. Kertesz, and K. von Figura. Posttranslational formation of formylglycine in prokaryotic sulfatases by modification of either cysteine or serine. *J Biol Chem*, 273(40):25560–4, 1998. 6, 11

[22] G. Lukatela, N. Krauss, K. Theis, T. Selmer, V. Gieselmann, K. von Figura, and W. Saenger. Crystal structure of human arylsulfatase A: the aldehyde function and the metal ion at the active site suggest a novel mechanism for sulfate ester hydrolysis. *Biochemistry,* 37(11):3654–64, 1998. 6, 8

[23] CS. Bond, PR. Clements, SJ. Ashby, CA. Collyer, SJ. Harrop, JJ. Hopwood, and JM. Guss. Structure of a human lysosomal sulfatase. *Structure*, 5(2):277–89, 1997. 6, 8

[24] I. Boltes, H. Czapinska, A. Kahnert, R. von Bulow, T. Dierks, B. Schmidt, K. von Figura, MA. Kertesz, and I. Uson. 1.3 A structure of arylsulfatase from Pseudomonas aeruginosa establishes the catalytic mechanism of sulfate ester cleavage in the sulfatase family. *Structure,* 9(6):483–91, 2001. 6, 7, 8

[25] M. Recksiek, T. Selmer, T. Dierks, B. Schmidt, and K. von Figura. Sulfatases, trapping of the sulfated enzyme intermediate by substituting the active site formylglycine. *J Biol Chem*, 273(11):6096–103, 1998. 6, 8

[26] A. Waldow, B. Schmidt, T. Dierks, R. von Bulow, and K. von Figura. Amino acid residues forming the active site of arylsulfatase A. Role in catalytic activity and substrate binding. *J Biol Chem,* 274(18):12284–8, 1999. 8

[27] A. Knaust, B. Schmidt, T. Dierks, R. von Bulow, and K. von Figura. Residues critical for formylglycine formation and/or catalytic activity of arylsulfatase A. *Biochemistry,* 37(40):13941–6, 1998. 8

[28] T. Dierks, B. Schmidt, and K. von Figura. Conversion of cysteine to formylglycine: a protein modification in the endoplasmic reticulum. *Proc Natl Acad Sci U S A*, 94(22):11963–8, 1997. 9, 10

[29] T. Dierks, MR. Lecca, P. Schlotterhose, B. Schmidt, and K. von Figura. Sequence determinants directing conversion of cysteine to formylglycine in eukaryotic sulfatases. *EMBO J*, 18(8):2084–91, 1999. 9, 12, 62

[30] T. Dierks, MR. Lecca, B. Schmidt, and K. von Figura. Conversion of cysteine to formylglycine in eukaryotic sulfatases occurs by a common mechanism in the endoplasmic reticulum. *FEBS Lett*, 423(1):61–5, 1998. 10, 12

[31] J. Fey, M. Balleininger, LV. Borissenko, B. Schmidt, K. von Figura, and T. Dierks. Characterization of posttranslational formylglycine formation by luminal components of the endoplasmic reticulum. *J Biol Chem*, 276(50):47021–8, 2001. 10, 11, 13

[32] K. von Figura, B. Schmidt, T. Selmer, and T. Dierks. A novel protein modification generating an aldehyde group in sulfatases: its role in catalysis and disease. *Bioessays*, 20(6):505–10. Review, 1998. 11

[33] K.S. Dodgson, G.F. White, and J.W. Fitzgerald. Sulfatases of microbial origin. CRC Press, Inc, Boca Raton, USA, 1982. 11

[34] S. Beil, H. Kehrli, P. James, W. Staudenmann, AM. Cook, T. Leisinger, and MA. Kertesz. Purification and characterization of the arylsulfatase synthesized by Pseudomonas aeruginosa PAO during growth in sulfate-free medium and cloning of the arylsulfatase gene (atsA). *Eur J Biochem*, 229(2):385–94, 1995. 11

[35] Y. Murooka, K. Ishibashi, M. Yasumoto, M. Sasaki, H. Sugino, H. Azakami, and M. Yamashita. A sulfur- and tyramine-regulated Klebsiella aerogenes operon containing the arylsulfatase (atsA) gene and the atsB gene. *J Bacteriol*, 172(4):2131–40, 1990. 11

[36] A. Schirmer and R. Kolter. Computational analysis of bacterial sulfatases and their modifying enzymes. *Chem Biol*, Review:R181–6, 1998. 11

[37] C. Szameit, C.and Miech, M. Balleininger, B. Schmidt, K. von Figura, and T. Dierks. The iron sulfur protein AtsB is required for posttranslational formation of formylglycine in the Klebsiella sulfatase. *J Biol Chem*, 274(22):15375–81, 1999. 12

[38] C. Marquordt, Q. Fang, E. Will, J. Peng, K. von Figura, and T Dierks. Posttranslational modification of serine to formylglycine in bacterial sulfatases. Recognition of the modification motif by the iron-sulfur protein AtsB. *J Biol Chem*, 278(4):2212–8, 2003. 12

[39] HJ. Sofia, G. Chen, BG. Hetzler, JF. Reyes-Spindola, and NE. Miller. Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: functional characterization using new analysis and information visualization methods. *Nucleic Acids Res*, 29(5):1097–106, 2001. 12

[40] J. Cheek and JB. Broderick. Adenosylmethionine-dependent iron-sulfur enzymes: versatile clusters in a radical new role. *2001*, Review:209–26, J Biol Inorg Chem. 12

[41] T. Dierks, B. Schmidt, LV. Borissenko, J. Peng, A. Preusser, M. Mariappan, and K. von Figura. Multiple sulfatase deficiency is caused by mutations in the gene encoding the human C(alpha)-formylglycine generating enzyme. *Cell*, 113(4):435–44, 2003. 13, 16, 17, 46, 60, 61, 68, 71, 79, 83, 85, 89

[42] J. Landgrebe, T. Dierks, B. Schmidt, and K. von Figura. The human SUMF1 gene, required for posttranslational sulfatase modification, defines a new gene family which is conserved from pro- to eukaryotes. *Gene*, 316:47–56, 2003. 15, 17, 54

[43] MP. Cosma, S. Pepe, G. Parenti, C. Settembre, I. Annunziata, R. Wade-Martins, C. Di Domenico, A. Di Natale, P. Mankad, B. Cox, G. Uziel, GM. Mancini, E. Zammarchi, MA. Donati, WJ. Kleijer, M. Filocamo, R. Carrozzo, M. Carella, and A. Ballabio. Molecular and functional analysis of SUMF1 mutations in multiple sulfatase deficiency. *Hum Mutat*, 23(6):576–81, 2004. 16, 17

[44] R. Basner, K. von Figura, J. Glossl, U. Klein, H. Kresse, and W. Mlekusch. Multiple deficiency of mucopolysaccharide sulfatases in mucosulfatidosis. *Pediatr Res.*, 13(12):1316–8, 1979. 16

[45] M. Burch, AH. Fensom, M. Jackson, T. Pitts-Tucker, and PJ. Congdon. Multiple sulphatase deficiency presenting at birth. *Clin Genet.*, 30(5):409–15, 1986. 16

[46] T. Yutaka, S. Okada, T. Kato, K. Inui, and H. Yabuuchi. Properties of sulfatases in cultured skin fibroblasts of multiple sulfatase deficient patients. *Clin Genet.*, 20(4):296–303, 1981. 16

[47] AL. Fluharty, RL. Stevens, LL. Davis, LJ. Shapiro, and H. Kihara. Presence of arylsulfatase A (ARS A) in multiple sulfatase deficiency disorder fibroblasts. *Am J Hum Genet.*, 30(3):249–55, 1978. 16

[48] AL. Fluharty, RL. Stevens, SD. de la Flor, LJ. Shapiro, and H. Kihara. Arysulfatase A modulation with pH in multiple sulfatase deficiency disorder fibroblasts. *Am J Hum Genet.*, 31(5):574–80, 1979. 16

[49] Kresse H. and D. Holtfrerich. Thiosulfate-mediated increase of arylsulfatase activities in multiple sulfatase deficiency disorder fibroblasts. *Biochem Biophys Res Commun.*, 97(1):41–8, 1980. 16

[50] PL. Chang and RG. Davidson. Pseudo arylsulfatase-A deficiency in healthy individuals: genetic and biochemical relationship to metachromatic leukodystrophy. *Proc Natl Acad Sci U S A.*, 80(23):7323–7, 1983. 16

[51] MP. Cosma, S. Pepe, I. Annunziata, RF. Newbold, M. Grompe, G. Parenti, and A. Ballabio. The multiple sulfatase deficiency gene encodes an essential and limiting factor for the activity of sulfatases. *Cell*, 113(4):445–56, 2003. 17, 46, 61, 83

[52] DS. Anson, V. Muller, J. Bielicki, GS. Harper, and JJ. Hopwood. Overexpression of N-acetylgalactosamine-4-sulphatase induces a multiple sulphatase deficiency in mucopolysaccharidosis-type-VI fibroblasts. *Biochem J*, 15;294:657–62, 1993. 18, 70, 87

[53] T. Lübke, T. Marquardt, A. Etzioni, E. Hartmann, K. von Figura, and C. Korner. Complementation cloning identifies CDG-IIc, a new type of congenital disorders of glycosylation, as a GDP-fucose transporter deficiency. *2001*, 28(1):73–6., Nat Genet. 30

[54] A. Preusser-Kunze, M. Mariappan, B. Schmidt, SL. Gande, K. Mutenda, D. Wenzel, K. von Figura, and T. Dierks. Molecular characterization of the human Calpha-formylglycine-generating enzyme. *J Biol Chem*, 280(15):14900–10, 2005. 41, 44, 81, 82, 83, 84, 88, 89

[55] V. Gieselmann, B. Schmidt, and K. von Figura. In vitro mutagenesis of potential N-glycosylation sites of arylsulfatase A. Effects on glycosylation, phosphorylation, and intracellular sorting. *J Biol Chem.*, 267(19):13262–6, 1992. 43

[56] J. Monnat, EM. Neuhaus, MS. Pop, DM. Ferrari, B Kramer, and T. Soldati. Identification of a novel saturable endoplasmic reticulum localization mechanism mediated by the C-terminus of a Dictyostelium protein disulfide isomerase. *Mol Biol Cell.*, 11(10):3469–84, 2000. 80

[57] A. Dickmanns, B. Schmidt, M.G. Rudolph, M. Mariappan, T. Dierks, K. von Figura, and R. Ficner. Crystal structure of human pFGE, the paralog of the Calpha-formylglycine-generating enzyme. *J Biol Chem*, 280(15):15180–7, 2005. 82, 84, 85, 86

[58] T. Dierks, A. Dickmanns, A. Preusser-Kunze, B. Schmidt, M. Mariappan, K. von Figura, R. Ficner, and M.G. Rudolph. Molecular basis for multiple sulfatase deficiency and mechanism for formylglycine generation of the human formylglycine-generating enzyme. *Cell*, 121(4):541–52, 2005. 82, 83, 84, 86

[59] E. Zito, A. Fraldi, S. Pepe, I. Annunziata, G. Kobinger, P. Di Natale, A. Ballabio, and MP. Cosma. Sulphatase activities are regulated by the interaction of sulphatase-modifying factor 1 with SUMF2. *EMBO Rep*, 6(7):655–60, 2005. 85, 86

[60] U. Matzner, E. Herbst, KK. Hedayati, R. Lullmann-Rauch, C. Wessig, S. Schroder, C. Eistrup, C. Moller, J. Fogh, and V. Gieselmann. Enzyme replacement improves nervous system pathology and function in a mouse model for metachromatic leukodystrophy. *Hum Mol Genet.*, 1;14(9):1139–52, 2005.

[61] Y. Takakusaki, S. Hisayasu, Y. Hirai, and T. Shimada. Coexpression of formylglycine-generating enzyme is essential for synthesis and secretion of functional arylsulfatase A in a mouse model of metachromatic leukodystrophy. *Hum Gene Ther*, 16(8):929–36, 2005.

[62] JT. Conary, A. Hasilik, and K. von Figura. Synthesis and stability of steroid sulfatase in fibroblasts from multiple sulfatase deficiency. *Biol Chem Hoppe Seyler.*, 369(4):297–302, 1988.

Acknowledgements

Publications

1. 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 Cα-formylglycine generating enzyme. *Cell* 113, 435-444.

2. Preusser-Kunze, A*., **Mariappan, M***., Schmidt, B., Gande, S. L., Mutenda, K., Wenzel, D., von Figura, K., and Dierks, T. Molecular characterization of the human Cα-formylglycine generating enzyme. (2005) J. Biol. Chem. 280, 14900-14910.
* *The firs two authors contributed equally to this work.*

3. Dickmanns, A., Schmidt, B., Rudolph, M. G., **Mariappan, M**., Dierks, T., von Figura, K., and Ficner, R. Crystal structure of human pFGE, the paralog of the Cα-formylglycine-generating enzyme.(2005) *J. Biol. Chem*. 280, 15180-15187.

4. **Mariappan, M**., Preusser-Kunze, A., Balleininger, M,, Eiselt, N., Schmidt, B., Gande, S. L., Wenzel, D., Dierks, T., von Figura, K. (2005) *J. Biol. Chem*. Expression, localization, structural, and functional characterization of pFGE, the paralog of the Cα-formylglycine-generating enzyme. 280, 15173-15179.

5. Dierks, T., Dickmanns, A., Preusser-Kunze, A., Schmidt, B., **Mariappan, M**., von Figura, K., Ficner, R., and Rudolph, M. G. Molecular basis for multiple sulfatase deficiency and mechanism for formylglycine generation of the human formylglycine-generating enzyme.*Cell*. (2005) 20;121(4):541-52.

Curriculum Vitae

Name: Malaiyalam Mariappan.

Date of Birth: 06.04.1979.

Father's Name: Mr. Mariappan Ramaiah.

Mother's Name: Ms. Kaliyammal Mariappan.

Place of Birth: Mamsapuram, India.

1994 - 1996: High School, Mamsapuram.

1996 - 1999: Bachelor of Science in Biochemistry, Senguthar Arts and Science college, Tiruchengodu, India

1999 - 2001: Master of Science in Biochemistry, University of Madras, India.

M.Sc Project: Isolation and characterization of calcium oxalate binding proteins from human urine, University of Madras, India.

July 2001 - Aug 2002:   Worked as Junior Research Fellow with Prof. Raghavan-Varadharajan, Indian Institute of Science, Bangalore, India.

Sept 2002 - 2005   Doctorate student under Prof. Dr. h.c. Kurt von Figura, Biochemistry II, University of Goettingen,