Lisa Wechselberger, BSc

Identification of sequences derived from G0S2 required for ATGL inhibition

MASTER'S THESIS

to achieve the university degree of

Master of Science

Master's degree programme: Molecular Microbiology

submitted to

Graz University of Technology

Supervisor

Ass.-Prof. Mag. Dr.rer.nat. Monika Oberer

Institute of Molecular Biocsciences

University of Graz

Graz, March, 2015

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present master‘s thesis dissertation.

27/03/2015

Date

Wendelberger Lisa

Signature

Acknowledgements

First of all, I would like to thank *Monika Oberer* for giving me the possibility to work on such an interesting project and for including me as a member of her research group. Thank you for your guidance and support during my whole master thesis.

I am also very grateful to *Ines* who introduced me in the whole lab work and who has always been there for me, both professionally and personally. Thanks you Ines, I've learned a lot from you!

I want to thank as well my "office-colleagues" *Kenny*, *Krishna* and *Roland*, who made my daily life at work very convenient, and who always answered all my questions very patiently. Thank you also for your support in the lab.

A special thanks also to *Harald* for his kind help concerning the TGH-assays.

I am very grateful to all *Strubis*, it was a great experience to work with all of you.

I would also like to thank my friends, without who I would not have had such wonderful six years in Graz.

Finally, big thanks to my family, especially my mother *Astrid* and my sister *Viktoria* who have always supported me and who have always been there for me!

Enfin et surtout, Farouk, merci pour tout!

L'homme qui ne médite pas vit dans l'aveuglement, l'homme qui médite vit dans l'obscurité.

Nous n'avons que le choix du noir

Victor Hugo

Wer in sich ruht ist Herr der Welt

Stefan Zweig

Table of Contents

Abbreviations

aa	amino acids
ATGL	adipose triglyceride lipase
APS	ammonium persulfate
BAT	brown adipose tissue
BSA	bovine serum albumin
cAMP	cyclic AMP
CGI-58	comparative gene identification-58
cpm	counts per minute
ECL	enhanced chemiluminescence
EDTA	ethylendiaminetetraacetic acid
FW	fresenius water (sterile, nuclease free water)
dNTP	dinucleoside triphosphate
DG	diglyceride
DGAT2	diacylglycerol acyltransferase2
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DTT	dithiothreitol
G0S2	G0/G1 switch protein 2
HMM	hidden Markov model
HSL	hormone sensitive lipase
IPTG	isopropyl β-D-1-thiogalactopyranoside
Kan	kanamycin
kDa	kilo Dalton
KLD	kinase-ligase-DpnI enzyme mix
KPB	potassium phosphate buffer
LB	lysogeny broth
MG	monoglyceride
MGL	monoglyceride lipase
NEB	New England Biosciences
NFκB	nuclear factor κB
OD	optical density
ONC	overnight culture
PC	phosphatidylcholine
PCR	polymerase chain reaction
PVDF	polyvinylidene difluoride
PGLB	protein gel loading buffer
PI	phosphatidylinositol
PLIN1	perilipin-1
PNPLA2	patatin-like phospholipase domain containing A 2
PPARγ	peroxisome proliferator-activated receptor γ
SDM	site-directed mutagenesis
SDS	sodium dodecylsulfate
SDS-PAGE	sodium dodecylsulfatepolyacrylamide-gel electrophoresis
SOC	super optimal broth (SOB) with catabolite repression
TB	tris-borate
TBST	tris-buffered saline + Tween 20
TEMED	tetramethylethylenediamine
TG	triglyceride
TGH	triglyceride hydrolase assay
TO	triolein
TRIS	tris(hydroxymethyl) aminomethane
WAT	white adipose tissue

1. Introduction

According to the WHO, overweight and obesity are defined as abnormal or excessive fat accumulation that presents a risk to health. The cause of obesity is on the one side an increased high-fat and high-sugar nutrition and on the other side a decreased physical activity, due to altered diet and lifestyle conditions after industrialization of our societies. These fundamental changes go along with the appearance of the so-called disease of civilization in Western populations [1].

Lipids are key components of metabolism and therefore need to be synthesized, stored and processed constantly. In most organisms energy in form of fat is stored mainly in lipid droplets (LD) of white and brown adipose tissue (WAT, BAT) as neutral and inert triglycerides (TG). In response to increased energy demand or in time of nutrient deprivation these TG stores are mobilized and lead to release of FA into the circulation. FAs have multiple biological functions, they are used as energy substrates and are also integral components of membrane lipids and function as second messengers.

Lipid metabolism takes place on the surface of LDs, which are dynamic organelles involved in lipid synthesis, turnover and trafficking. LDs consist of an hydrophobic core comprising neutral lipids, which are mainly TG and sterol esters, surrounded by a monolayer of phospholipids with their hydrophobic carbon-chains on the inside and the hydrophilic head groups pointing outwards towards the aqueous cytosol [2]. This arrangement provides a particular separation of the aqueous and organic phases of the cell. The surface of LDs is coated by several types of proteins, including lipid synthesis enzymes as DGAT2 [3,4], lipases, membrane trafficking proteins and structural proteins as proteins of the perilipin family [5].

1.1 Lipolysis

The hydrolysis of TGs is defined as lipolysis and occurs in all tissues and cell types, but is most abundant in WAT and BAT. Lipolysis is mediated by lipolytic enzymes in three subsequent reactions that degrade adipose TGs and release non-esterified fatty acids (FA) into the circulation [6,7].

Adipose triglyceride lipase (ATGL) catalyses the first step of TG breakdown, which leads to the release of diglycerides (DG) and FA [8]. The second reaction is mediated by hormone sensitive lipase (HSL), which leads to the generation of monoglycerides (MG) and FA. HSL is a

tri- and a diglyceride lipase, whereas its activity on DG is ten times higher than on TG [9,10]. MG is hydrolysed by monoglyceride lipase (MGL) and results in the release of glycerol (G) and FA [11]. Figure 1: Simplified scheme of TG breakdown by lipolytic enzymes ATGL, HSL and MGL.

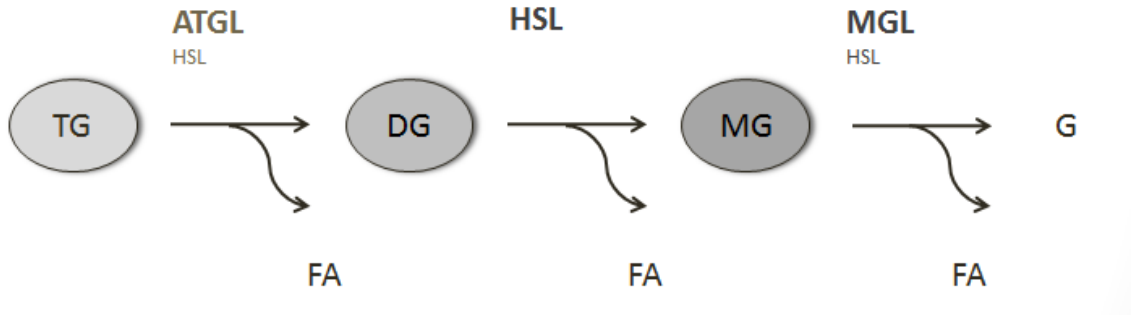


Figure 1: Simplified scheme of TG breakdown by lipolytic enzymes ATGL, HSL and MGL. ATGL hydrolyses TG to DG and FA. HSL hydrolyses DG to MG and FA. MG is hydrolysed to glycerol and FA by MGL. Fatty acids and glycerol exit the adipocyte and enter the circulation.

1.1.1 Adipose triglyceride lipase

Prior to the discovery of ATGL [8,12,13], hormone sensitive lipase (HSL) was considered to be the only enzyme to hydrolyse TG in mammalian adipose tissue and non-adipose tissue.

In 2004, three groups simultaneously reported the identification of a new cytosolic triacylglycerol lipase [8,12,13]. It was named ATGL, Desnutrin and Phospholipase A2 (now annotated as the patatin-like phospholipase domain containing protein 2 (PNPLA2)) [14].

Mouse ATGL consists out of 486 aa and it shares 55% amino acid identity with human ATGL. Characteristic for this superfamily of patatin-like phospholipases is the patatin-related region harboring an αβα sandwich (Figure 2). This domain is named after the potato tuber protein patatin, a lipid hydrolase [15]. So far, 9 proteins, PNPLA1 to PNPLA9, belonging to this protein family are encoded in the human genome, yet no 3D structures have been determined [16,17]. The hydrolytic reaction in these proteins is mediated by a catalytic serine-aspartate dyad. The serine is located within a GXSXG motif that is typically found in lipases of the α/β hydrolase fold family [17].

Very few is known about the 3D structure of ATGL, but it is predicted that it also acts through this catalytic dyad composed out of Ser47 and Asp166 [18] shown in Figure 2. This was confirmed by mutations studies, which lead to catalytically inactive proteins [7]. The C-

terminal half of the protein contains a hydrophobic stretch, which has been recognized to mediate translocation of ATGL to LD [19].

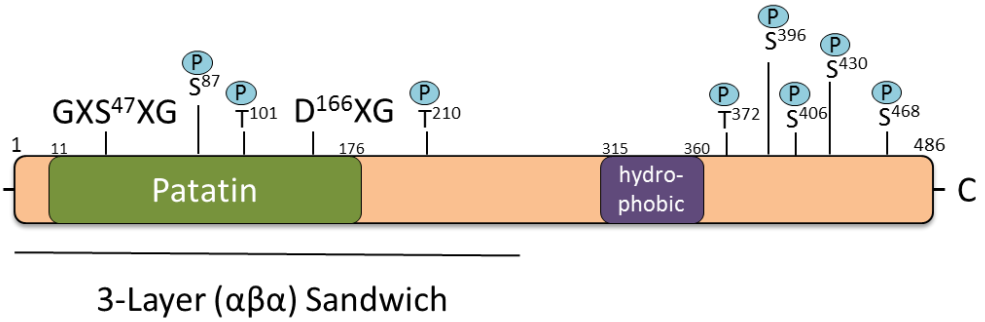


Figure 2: Conserved residues and domain organization of mouse ATGL. ATGL carries a patatin-domain (aa 11-176) in the N-terminal αβα sandwich. The catalytic dyad is formed by Ser^{47} and Asp^{166} within the consensus motifs GXSXG and DXG. The C-terminal region contains a hydrophobic stretch from aa 315-360. Phosphorylation sites at Ser^{87}, Thr^{101}, Thr^{210}, Thr^{372}, Ser^{396}, Ser^{406}, Ser^{430} and Ser^{468}.

Increased TG catabolism and FA metabolism come along with abnormalities in lipid metabolism and are directly related to metabolic diseases, including type-2 diabetes and cancer. Free circulating FA lead to deleterious effects on the organism, including ectopic lipid accumulation in liver, pancreas and muscle. This goes hand in hand with tissue inflammation and insulin resistance, summarized in the term lipotoxicity [20,21].

Complete inactivation of ATGL leads to massive TG accumulation, whereas ATGL knock out studies revealed that decreased ATGL activity exerts beneficial effects on the organism. ATGL knock out mice showed increased glucose tolerance and improved insulin sensitivity as well as protection from cancer induced cachexia [22,23].

1.1.2 Regulation of lipolysis

Lipid homeostasis in humans is given by the cycle of lipolysis and re-esterification of lipids according to energetic demands. The regulation occurs on different levels and is strongly influenced on the post-translational and post-transcriptional level by hormones, other biochemical signals, and by regulatory proteins.

1.1.2.1 Hormonal regulation

In response to the nutrient status of the organism, lipolysis rates are precisely regulated through hormonal and biochemical signals. Studies revealed that beta-adrenergic mediated signalling is required for fully hormone-activated lipolysis in WAT [8,23,24]. Thus basically, the neuroendocrine control of lipolysis is exerted positively by this catecholaminergic

stimulation and negatively by insulin-induced suppression, both of which affect c-AMP levels and hence the protein kinase A (PKA)-mediated phosphorylation of HSL and perilipin-1 (PLIN1) [25].

PLIN1 is both the most abundant protein associated with LDs and a major PKA substrate in adipocytes [26,27]. It controls access of lipases to TGs thus regulates the lipids homeostasis [28]. In the basal status, PLIN1 sequesters the comparative gene identification-58 (CGI-58) needed for ATGL activation.

Catecholamines as adrenalin and noradrenalin bind to β-adrenergic G-protein-coupled receptors at the plasma membrane and activate adenylate cyclase, followed by increased levels of cAMP. This second messenger activates the protein kinase A (PKA) [29], which directly phosphorylates PLIN1 and HSL.

The hormone-stimulated phosphorylation of PLIN1 leads to the binding of cytosolic HSL that can now translocate to the LD [30]. Activation of PLIN1 also leads to the release of CGI-58, now available for ATGL activation [31,32]. ATGL is activated 14-fold through direct binding of CGI-58 [7,23]. A simplified model of this regulation is shown in Figure 3.

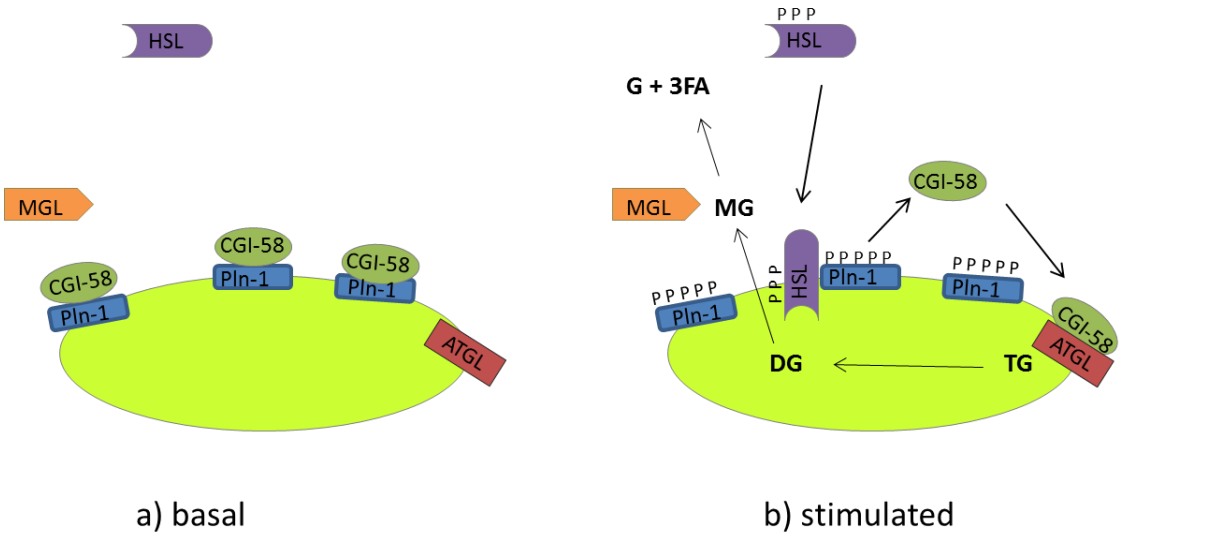


Figure 3: Simplified model of regulation of lipolysis on LD by perilipin-1 (PLIN1). a) In the basal status PLIN1 bound to LD prevents localization of HSL to the LD and also sequesters comparative gene identification-58 (CGI-58) needed for ATGL activation. b) β-adrenergic stimulation activates adenylate cyclase, followed by increased levels of cAMP, which triggers the protein kinase A (PKA)-mediated phosphorylation of HSL and PLIN1 (not shown). Phosphorylated PLIN1 changes conformation to facilitate increased lipolysis. Phosphorylated HSL docks on PLIN1 and gains access to LD. PLIN1 releases CGI-58 which then forms a complex with ATGL. The ATGL/CGI-58 complex hydrolyses TG to DG and FA. HSL hydrolyses DG to MG and FA. MG is hydrolysed to glycerol and FA by MGL. Fatty acids and glycerol exit the adipocyte and enter the circulation.

Direct interactions of ATGL with its protein inhibitor G0/G1 switch protein 2 (G0S2) results in complete ATGL inactivation [33,34,35,36,37].

ATGL is upregulated during adipose regulation [8,13] and it is a target for transcription factors PPARγ [38,39,40]. mRNA levels of ATGL are upregulated during fasting and in presence of glucocorticoids [13]. In contrast, ATGL mRNA expression is repressed by insulin, TNF-α and feeding [41,42].

ATGL is known to be a phosphoprotein, and can at least be phosphorylated on eight different sites, shown in Figure 2 [43,44]. Previous studies have identified two phosphorylation sites, Ser^{403} and Ser^{430}, in its C-terminal region and in recent studies it was shown that Ser^{406} is a direct target for PKA, and that its phosphorylation-status correlates with lipolytic activation upon β-adrenergic stimulation [44,45]. Very recent studies described Thr^{372} as a novel site which prevents LD localization of ATGL upon phosphorylation [43]. However the complete relevance of phosphorylation involved in regulation of ATGL remain incompletely defined. Figure 2 shows conserved residues and domain organisation of mATGL including phosphorylation sites.

1.1.2.2 Comparative gene identification 58

Mouse CGI-58 is a 349 aa protein and its name is derived from an proteomic approach to identify conserved genes between *C. elegans* and humans. There is no 3D structure available yet, but it is predicted to constitute a fold typical for α/β hydrolase. The α/β-hydrolase-fold family is characterized by a catalytic triad including a nucleophile, an acid, and an histidine that are distant from each other in sequence but close in 3D structure [46]. The nucleophile in this catalytic triad is usually a serine, but in CGI-58 it is replaced by $aspargine^{166}$. The acid of the 'triad' could be either $aspartate^{301}$, within the sequence GARSCIDG or the glutamic acid residue at 260, inside the motif PSGETA, whereas the latter is less conserved [47,48,49]. The third member of the inactive catalytic triad-like arrangement is $histidine^{329}$. Conserved residues are shown in Figure 4.

Activation of ATGL requires direct interaction with CGI-58 [7,50]. Mutation studies show that ATGL activation occurs upon binding of CGI-58 within the N-terminal patatin-domain of the lipase through direct protein-protein interaction [51]. However, this interaction is not sufficient for complete ATGL activation, which requires additionally the binding of CGI-58 to LD [50].

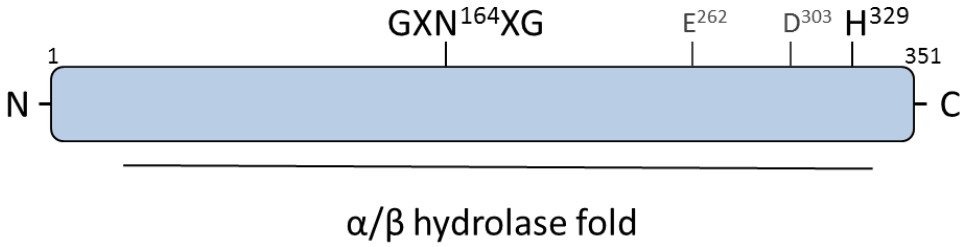


Figure 4: Conserved residues and domain organization of murine CGI-58. It is a member of the α/β-hydrolase-fold family. Residues of a functional catalytic triad would correspond to N^{164}, E^{262} or D^{303} and H^{329}.

1.1.2.3 G0/G1 switch protein 1

G0S2 is a small basic protein of 103 aa and was first identified in blood mononuclear cells following drug-induced cell cycle transition from G0 to G1 phase in the early 1990s [52]. It was only 20 years later when Yang et al described G0S2 playing a crucial role in regulation of lipid metabolism [53].

Structural knowledge on G0S2 is very limited as it does not share any sequence similarities of proteins with known 3D structures. However, based secondary structure prediction it is an alpha-helical protein. It has an hydrophobic stretch located between aa Lys27 and Met43 (Figure 5) and deletion of this hydrophobic region disables the interaction of G0S2 with ATGL [53]. Figure 5 illustrates these few predicted biochemical properties.

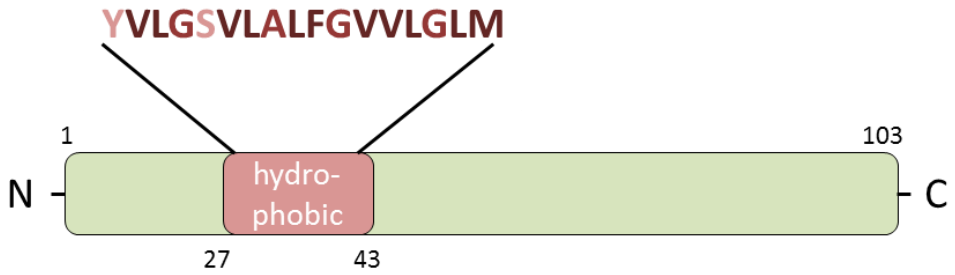


Figure 5: Organization of hG0S2: hydrophobic stretch between amino acids 27 to 43. Amino acids are colored according to level of hydrophobicity (dark: very hydrophobic, pale: less hydrophobic)

G0S2 binds directly to the patatin-domain of ATGL and the inhibition takes place even in presence of CGI-58, which suggests that the interaction does not compete with binding of the activator protein [54]. Upon stimulation of lipolysis, ATGL and G0S2 translocate to LD, which is not the case when ATGL expression is downregulated [53]. These findings suggest that G0S2 alone may not be able to bind to LDs, but requires ATGL as a binding partner.

The G0S2 promoter contains a potential PPAR-responsive element (PPRE) and was shown to be a target of peroxisome proliferator-activated receptor γ (PPARγ). PPARγ is a member of the PPAR family of nuclear receptor transcription factors in adipose tissue and functions as a master regulator of adipogenesis [55].

In adipocytes, G0S2 expression is downregulated by beta-adrenergic stimulation and upregulated by insulin [7] and by PPARγ during adipogenesis, suggesting a role for G0S2 as either preventing energy catabolism or promoting energy storage [53]. Unsurprisingly, expression levels of G0S2 in adipose tissue were found to be low during fasting but increased after feeding [56].

G0S2 also plays an important role in cancer development as a tumour suppressor gene as it is reported to be silenced in lung and neck cancer [57]. G0S2 was also identified as a mitochondrial protein [58]. Welch showed that it interacts with Bcl-2, an anti-apoptotic factor at the mitochondria, promoting apoptosis in human cancer cells.

Thus, G0S2 appears in different cellular processes and might be a link between cell cycle, cell survival and lipolysis.

G0S2 is only found in vertebrates and no homologs in plants or in lower organisms as *C. elegans* and Drosophila are known [59]. G0S2 is highly conserved between species, the mouse and the human isoform share 78 % identity on the protein level. This conservation also ranges to evolutionary very distant species, and the conservation highly occurs in the N-terminal region of the protein as shown in Figure 6.

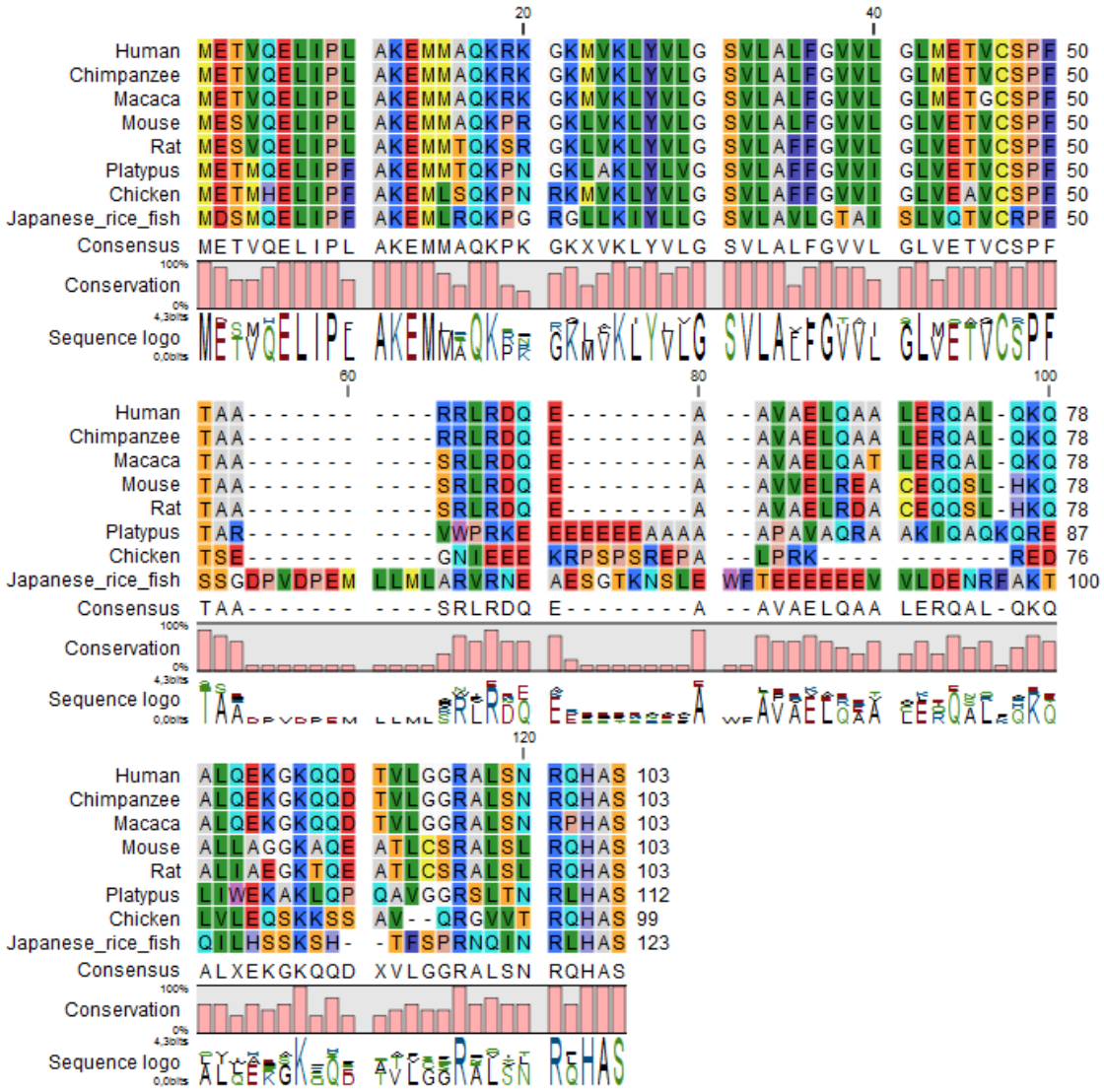


Figure 6: Sequence alignment of G0S2 from 8 different species including the level of conservation. This figure was generated using a trial version of CLC workbench. The conservation highly occurs in the N-terminal part of the protein.

2. Relevance

My master thesis is embedded in a subproject of the SFB project 'Lipotox' on the characterization of G0S2 and ATGL to fully understand the molecular basis of G0S2-ATGL interaction.

Previous studies in our group led to the identification of shortened variants of human G0S2 that inhibit ATGL's hydrolysing activity [60]. Extensive N- or C-terminal truncations of wild-type hG0S2 are still fully capable of inhibiting ATGL (Met1-Met43 and Tyr27-Ser103) [60]. These findings suggest that it is the hydrophobic stretch between amino acids Tyr27-Met43 that plays an essential role for hG0S2 activity. Interestingly, this short peptide by itself was not capable of inhibiting ATGL. This led to the conclusion that there are crucial residues for ATGL interaction N or C-terminal of this hydrophobic region. Based on these data, Cerk *et al.* identified a short stretch of 33 amino acids, the Lys20-Ala52 variant, acting as an efficient inhibitor of ATGL. A synthetic peptide corresponding to the hG0S2_Lys20-Ala52 sequence was synthesized and tested for its inhibitory capacity. Cerk *et al.* showed that this peptide completely inhibited ATGL, which indicates that this synthesized peptide adopts a tertiary structure compatible for direct interaction with ATGL.

As direct follow-up of this work, the first aim of my thesis was to narrow the N- and the C-terminal boundaries from Lys20-Ala52 towards Tyr27-Met43 to identify the shortest peptide still capable of inhibiting ATGL.

We hypothesize that the hydrophobic core peptide plays an essential role in the inhibition of ATGL, yet that there are additional hydrophilic residues required to solubilize the inhibitory core peptide in order to reach and interact physically with ATGL in the lipid-water interface. Thus my second aim was to replace conserved non-hydrophobic residues into oppositely charged or/and non-polar residues in the Lys20-Ala52 construct.

G0S2 is highly conserved between species and as a third aim I tested whether G0S2 from evolutionary more distant species are still capable of inhibiting TG hydrolysis of mATGL *in vitro.* G0S2 genes from chicken, Japanese rice fish and platypus were ordered and tested for their inhibitory capability towards ATGL.

3. Experimental procedures

3.1 Consumables and buffer protocols

Plastic consumables were ordered from Greiner Bio-One, glassware from Fisher Scientific.

Chemicals were ordered from Sigma-Aldrich, Merck and Roth.

Most of used enzymes and reaction buffers were ordered from NEB and are represented in

Agarose Gel
- TAE-Buffer (1x)
- SYBR®Safe DNA gel stain 1:10000
- 1 % agarose

SDS-Running Buffer denat. (10x)
- 144 g/l glycine
- 30 g/l Tris
- 10 g/l SDS

SDS-Gel (18%)
- 8 ml H_2O
- 36 ml 30%/0.5% acrylamide
- 15 ml Tris (pH 8.8)
- 600 µl SDS
- 60 µl TEMED
- 240 µl APS (25%)

Stacking gel
- 5 ml H_2O
- 10 ml 30%/0.5% acrylamide
- 4.8 ml Tris (pH 8.8)
- 200 µl SDS (10%)
- 20 µl TEMED
- 80 µl APS (25%)
- 100 µl bromphenol blue

Solution A (pH 7)
- 0.25 M sucrose
- 1 mM EDTA
- 1 mM DTT

PGLB (2x)
- 125 mM Tris-HCl (pH 6.8)
- 20% glycerol
- 4 % SDS
- 0.0005 % bromphenol blue
- 700 mM β-mercaptoethanol

Transfer Buffer
- 100 ml 10x Tris/Glycine
- 1 ml SDS (10%)
- 200 ml MeOH
- 700 ml H_2O

TBST (10x)
- 1 % Tween-20
- 1.5 M NaCl
- 500 mM Tris-HCl

Table 1 Enzymes and reaction-buffers used for experiments

Product	Manufacturer	Ref. Number
Albumin Standard (2.0mg/ml)	Thermo Scientific	23209
BamHI-HF (20 000 U/ml)	NEB	R3236S
DpnI (20 000 U/ml)	NEB	R0146S
XhoI (20 000 U/ml)	NEB	R0146S
CutSMART buffer (10x)	NEB	B7204S
Antarctic phosphatase (5 000 U/ml)	NEB	M0289S
Antarctic phosphatease buffer (10x)	NEB	B0289S
dNTPs (10mM each)	NEB	N0447S
Q5 HF-Polymerase (2 000 U/ml)	NEB	M0491S
Q5 High GC Enhancer (5x)	NEB	B9028A
Q5 Reaction buffer (5x)	NEB	B9027S
T4 Ligase (400 000 U/ml)	NEB	M0202S
T4 Ligase buffer (10x; 10 mM ATP)	NEB	B0202S

3.2 Cloning of recombinant proteins

Different cloning techniques were used to generate the hG0S2 variants. The primers were designed using 12-15 bases from the 5' (fwd) or 3' (rev) end of the desired sequence. For better annealing, primers with bases of guanine or cytosine at the 3' end are preferred. Sequences that are recognized by specific restriction enzymes were added.

3.2.1 Conventional Cloning

N-terminal truncations were generated by PCR using forward primers that flank the corresponding sequence of human G0S2 and contain a BamHI cleavage site. The reverse primer introduced a stop codon after Ala52 and contained an XhoI cleavage site. The primers were ordered from Invitrogen ™ Life Technologies. All constructs were cloned originally into the pET-28a(+) vector. This vector was once equipped in our lab with a smt3-tag and a TEV cleavage site. The vector map is shown in Figure 7.

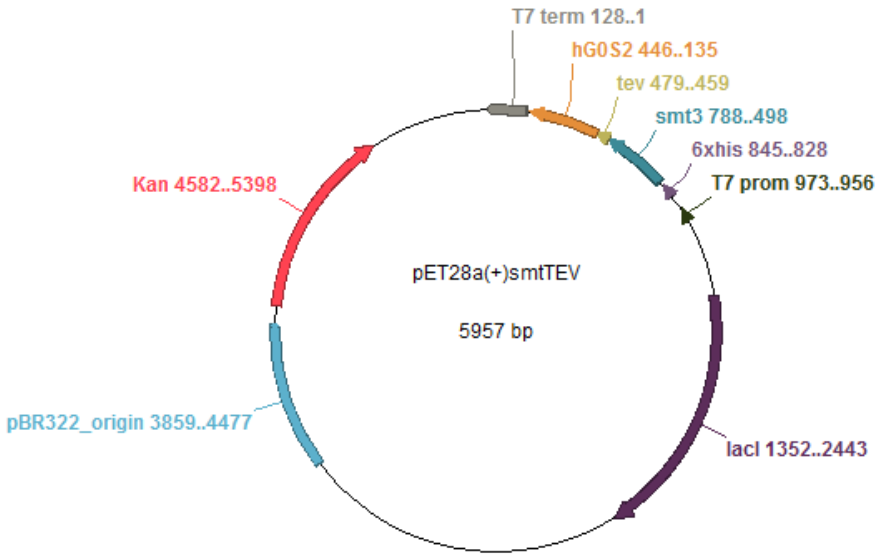


Figure 7 Vector map of the pet28a(+) vector with highlighted features and corresponding basepairs: 6xhis, smt3 tag and TEV cleavage site, Kan, T7 prom, T7 term and lacI. The map was generated using ApE.

The pet28a_SMT-tev vector carrying the full length hG0S2 was used as the template DNA for the N-terminal truncations and the pet28b_SMT_tev_hG0S2_Met1-Val46 for the Lys20-Val46 construct.

3.2.1.1 Amplification of insert DNA

The PCR mixture contained the following reagents:

template DNA 100 ng /µl	1 µl
primer forward (200 pmol)	1 µl
primer reverse (200 pmol)	1 µl
dNTPs (10mM)	1 µl
Q5 Buffer (5x)	10 µl
GC Enhancer (5x)	10 µl
Q5 Polymerase (2 000 U/ml)	0.5 µl
FW	to a total volume of 50 µl

The primer sequences are illustrated in Table 2. For a negative control the same mixture without template DNA was used. All PCR mixtures were prepared on ice in a 200 µl reaction tube.

The amplification of the DNA fragments was performed in the C1000 Thermal Cycler (Bio-Rad) using following protocol:

98°	1 min	
98°	30 sec	⎫
57°/58°	1 min	⎬ 30x
72°	1 min	⎭
72°	10 min	

Table 2 List of primers used for N-terminal and partial C-terminal truncations of hG0S2. Tm's as used in PCR reaction. BamHI cleavage sites highlighted in yellow, XhoI cleavage sites in blue and the introduced STOP-codon in red.

Desired product	Primer	Tm
hG0S2_Gly21-Ala52	fwd: GGAGTCGGATCCGGGAAGATGGTGAAGC rev: CCGCTCGAGCTAAGAGGCGTGCTGC	57°
hG0S2_Lys22-Ala52	fwd: GGCGCCGGATCCAAGATGGTGAAGCTGTACG rev: CCGCTCGAGCTAAGAGGCGTGCTGC	58°
hG0S2_Met23-Ala52	fwd: GGAGCCGGATCCATGGTGAAGCTGTACGTGC rev: CCGCTCGAGCTAAGAGGCGTGCTGC	58°
hG0S2_Val24-Ala52	fwd: GGAGTCGGATCCGTGAAGCTGTACGTGCTGG rev: CCGCTCGAGCTAAGAGGCGTGCTGC	58°
hG0S2_Lys25-Ala52	fwd: GGCGTCGGATCCAAGCTGTACGTGCTGG rev: CCGCTCGAGCTAAGAGGCGTGCTGC	57°
hG0S2_Leu26-Ala52	fwd: GGATTGGATCCCTGTACGTGCTGGGCAGC rev: CCGCTCGAGCTAAGAGGCGTGCTGC	57°
hG0S2_Lys20-Val46	fwd: GGAGCCGGATCCAAGGGGAAGATGGTGAAGC rev: CCGCTCGAGCTAAGAGGCGTGCTGC	58°

3.2.1.2 Insert PCR Clean up

To check whether the amplification was done successfully, an aliquot of 5 µl of each PCR-reaction was loaded on an agarose gel. The agarose-gel electrophoresis was done according to following protocol, using a 40 ml 1% agarose gel:

- Prepare gel-buffer by adding 1:10000 SYBR®Safe DNA gel stain (Invitrogen – Life technologies) in 40 ml TAE (1x)
- Weight out 0.4 g agarose in an Erlenmeyer flask and add buffer
- Heat to boiling in a microwave until the solution becomes clear and the agarose is completely solved
- Let the flask cool a bit
- Cast the liquid into a gel tray and remove bubbles with the comb
- Let the gel polymerise for 20 minutes
- Gel loading dye (6x) need to be added to the samples
- Fill gel-chamber with 1x TAE buffer, put gel in and apply samples
- Run at 80 V for 10 – 20 min.
- Visualize SYBR®Safe bound to nucleic acid under UV-lamp or blue light.

The PCR products were then cleaned up using the Wizard® SV Gel and PCR Clean-Up System following the manufacturer's instructions. The DNA was eluted in 30 µl ultrapure water.

3.2.1.3 Isolation of the vector DNA

100 ng DNA of the pet28a_SMT_tev_hG0S2 vector was transformed into chemically competent *E. coli* Top10 cells according to following protocol:

- Mix 200 µl T10 cells with the ligation mix
- Let stand on ice for 10 minutes
- Heat shock at 42°C for 45 seconds
- Let stand on ice for 2 minutes
- Add 800 mL of LB media
- Incubate at 37°C for 45 minutes shaking

100 µl of the mix was spread out on selective agar plates (LB + kan [40µg/µl]) for incubation at 37° overnight.

A single colony was picked to start a preculture in 10 ml LB. After 8 hours incubation at 37° and 180rpm, 300 ml LB were inoculated with 3 ml of the preculture and incubated overnight at 37°C and 180rpm.

The plasmid DNA was isolated using the Plasmid DNA purification (NucleoBond® XtraMidi) according to manufacturer's protocol.

3.2.1.4 Double Digest of vector and insert DNA

The amplified DNA and the isolated plasmid DNA were then digested by the restriction enzymes BamHI and XhoI.

The mixture was prepared according to the following protocol:

	Vector digestion:	Insert digestion:
Cut smart buffer (10x)	5 µl	5 µl
BamHI (20 000 U/ml)	2 µl	1 µl
XhoI (20 000 U/ml)	1 µl	1 µl
Plasmid DNA/Insert DNA	10 µg	~30 µl
FW	to a total volume of 50µl	

The mixtures were digested at 37° for three hours.

3.2.1.5 Vector dephosphorylation

Digested DNA typically possesses a 5' phosphate group that is required for ligation. In order to prevent self-ligation, the 5' phosphate can be removed prior to ligation to decrease the background activity of the cloning process.

The dephosphorylation mix was prepared according to the following protocol:

Vector digestion mix	50 µl
Antarctic phosphatase buffer (10x)	6 µl
Antarctic phosphatase (5000U/ml)	3 µl
FW	to a total volume of 60 µl

3.2.1.6 Gel clean-up of vector and insert DNA

The purification of the digested insert and vector DNA was carried out by agarose gel electrophoresis. Loading Dye (6x) was added to the DNA and loaded on a 1% agarose gel, which ran at 100V for 15-20 minutes.

The corresponding gel pieces were cut out of the gel and cleaned up using the Wizard® SV Gel and Clean-Up System according to the manufacturer's protocol.

Concentrations of nucleic acids were measured using NanoDrop due to their absorbance of ultra violet light.

3.2.1.7 Ligation

Ligation is used to insert the digested and purified DNA fragments into the plasmid DNA. The calculation below gives the amount of fragments to use in relation to the vector at a ratio of 3:1.

$$insert\ mass\ (ng) = 3\ \times \frac{Insert\ length\ in\ bp}{vector\ length\ in\ bp}\ \times vector\ mass\ in\ ng$$

The ligation mix consisted of following components:

Insert DNA	12ng
Plasmid DNA	100 ng
T4 Ligase (400 000 U/ml)	1 µl
T4 Ligase buffer (10x)	1.5 µl
FW	to a total volume of 15 µl

The ligation mix was incubated overnight at room temperature.

3.2.1.8 Transformation

The ligation mix was transformed into chemically competent *E.coli* Top10 cells according to the protocol described on page 21.

The samples were centrifuged at 3500 rpm for five minutes. The supernatant was discarded and the pellet disolved in a small volume of LB, approximately 100 µl, which was spread out on selective agar plates (LB + kan [40µg/µl]) for incubation at 37° overnight.

A single colony was picked to inoculate 10 mL of LB media for an ONC. The ONC's were spun down at 4000 rpm for 10 minutes and the plasmid DNA was purified using QIAprep Spin Miniprep Kit according to the manufacturer's protocol.

15 µl of 100ng/µl plasmid DNA was added with 10 µM T7 primer and was sent for sequencing to Eurofins.

3.2.2 Side directed Mutagenesis

For the generation of the hG0S2_Lys20-52 mutants and also for the Lys20-Val46 construct two different site directed mutagenesis methods were used

3.2.2.1 QuickChange® Site directed mutagenesis kit

For the hG0S2_20-52_K25D and the hG0S2_20-52_T45G mutant a side directed mutagenesis using a QuickChange® protocol was performed.

The mutagenic primers need to be designed individually according to the desired mutation, where at both of the primers (rev and fwd) need to contain the same mutation, as they anneal to the same sequence on the opposite strand of the plasmid.

The primers should be between 25 and 45 bases in length, with a Tm ≥ 78°C, which is calculated using following formula:

$$Tm = 81.5 + 0.41 \times \%GC - \frac{675}{N} - \%\, mismatch$$

N .. primer length in bases

The pet28a_SMT_tev_hG0S2_Lys20-52 construct was used as a template DNA and the PCR mixture was prepared according to following protocol:

template DNA (100 ng/µl)	1 µl
primer forward (200 pmol)	1 µl
primer reverse (200 pmol)	1 µl
dNTPs (10 mM)	1 µl
Q5 buffer (5x)	5 µl
Q5 polymerase (2 000 U/ml)	0.5µl
FW	to a total volume of 25 µl

The used primers are shown in Table 3. As a negative control the same mix was prepared without primers. The amplification was done in a C1000™ Thermal Cycler according to the following protocol:

98°	1 min	
98°	30 sec	30x
60°/64°	1 min	
72°	3 min	
72°	10 min	

Table 3 List of mutagenic primers used for Quikchange SDM. Tm's as used in PCR reaction. The desired mutation is highlighted in green.

Desired product	Primer	Tm
hG0S2_Lys20-Ala52_K25D	fwd: GGGGAAGATGGTGGACCTGTACGTGCTGGGC rev: GCCCAGCACGTACAGGTCCACCATCTTCCCC	60°C
hG0S2_Lys20-Ala52_T45G	fwd: GGAAGATGGTGAAGCTGGGCGTGCTGGGCAGC rev: GCTGCCCAGCACGCCCAGCTTCACCATCTTCC	64°C

1 µl of DpnI was added to the PCR mix and incubated for 1h at 37°, which was done twice.

The electroporation of the whole mixture into *E. coli* Top10 electro-competent cells was carried out using following protocol:

- desalt PCR-Mix on minifilter for 30min
- Sterilize and pre-cool cuvettes
- mix desalted DNA with competent cells
- transfer cells into cuvettes
- electroshock
- add 800µl LB media
- incubate for 30-45 min at 37°C

An aliquot of 100µl was plated out on selective agar plates (LB + kan [40µg/µl]) for incubation at 37° overnight

A single colony was picked to inoculate 10 mL of LB media for an ONC. The ONC's were spun down at 4000 rpm for 10 minutes and the plasmid DNA was purified using QIAprep Spin Miniprep Kit according to manufacturer's protocol.

3.2.2.2 Q5® Site directed mutagenesis kit

For the hG0S2_20-52_Y27G mutant and the hG0S2_20-45 and hG0S2_20-44 construct a side directed mutagenesis using the Q5® Site directed mutagenesis kit according to manufacturer's protocol was performed.

The primers were designed using the NEBaseChanger™ tool, which included also the calculation of the Tm shown in Table 4.

As a template DNA the pet28b_SMT_tev_hG0S2_20-52 construct was used.

Table 4 List of mutagenic primers used for Q5 SDM. Tm's as used in PCR reaction. The desired mutation is highlighted in green and the inserted stop-codons in red.

Desired product	Primer	Tm
hG0S2_Lys20Ala52_Y27G	fwd: GGTGAAGCTGGGCGTGCTGGGCA rev: ATCTTCCCCTTGGATCCC	64°C
hG0S2_Lys20_Thr45	fwd: GATGGAGACTTAGTGCAGCCCCTTCACGG rev: AGGCCGAGCACCACGCCG	72°C
hG0S2_Lys20_Glu44	fwd: CCTGATGGAGTGAGTGTGCAGCCCCTTCAC rev: CCGAGCACCACGCCGAAG	69°C

The PCR mixture was prepared according to following protocol

Q5 Hot Start High-Fidelty Master Mix (2x)	12.5 µl
forward primer (10 µM)	1.25 µl
reverse primer (10 µM)	1.25 µl
Template DNA (25 ng/µl)	1 µl
FW	9 µl

The amplification was done in a C1000™ Thermal Cycler according to the following protocol; the Tm for each construct is shown in Table 4.

98°	30 sec	
98°	10 sec	25x
64°-72°	30 sec	
72°	3 min	
72°	10 min	

After the PCR an aliquot of the sample was loaded on a 1% agarose gel, to check the amplification of the desired DNA. Only a small volume of the PCR mix was added with KLD enzyme mix and KLD buffer according to following protocol:

PCR product	1 µl
KLD reaction buffer (2x)	5 µl
KLD enzyme mix (10x)	1 µl
FW	3 µl

The mix was incubated for 5 minutes at room temperature and then transformed according to the following protocol:

- Add 5 µl of KLD mix to 50 µl of chemically-competent cells
- Incubate on ice for 30 minutes

- Heat shock at 42°C for 30 seconds
- Incubate on ice for 5 minutes
- Add 950 µl SOC, gently shake at 37° for one hour
- Spread 40-100 µl onto appropriate selection plate (LB + kan [40µg/µl]) and incubate overnight at 37°C

3.2.3 Cloning of G0S2 genes from evolutionary far distant species

G0S2 genes from chicken, Japanese rice fish and platypus with a pEX-A2 vector backbone were ordered at Eurofins and then cloned in our expression vector pet 28a_SMT_tev.

The lyophilised plasmid DNA was diluted in FW to a concentration of 100ng/µl and transformed in *E. coli* T10 chemically competent cells. 100 µl of the transformation mix was spread out on LB plates + Amp [100µg/µl| as the pEX-A2 vector has an ampicillin resistance.

A single colony was picked to start a preculture in 10 ml LB. After 8 hours incubation at 37° and 180rpm, 300 ml LB were inoculated with 3 ml of the preculture and incubated overnight at 37°C and 180rpm.

The plasmid DNA was purified using the Plasmid DNA purification (NucleoBond® XtraMidi) according to manufacturer's protocol. Double digest, ligation, transformation in *E. coli* T10 cells, and plasmid purification was done as described above.

To check if the cloning was done properly and to check if the DNAs carry the desired mutation, 15 µl of 100ng/µl plasmid DNA of each construct as added with 10 µM T7 primer and sequenced by Eurofins Scientific.

3.3 Expression of recombinant proteins

For protein expression, chemically competent *E.coli* BL21 (DE3) CodonPlus® cells were used. There are several properties of this strain that contribute to its usefulness in protein expression, which are the T7 polymerase, the lon protease deficiency and T1 phage resistance.

T7 RNA Polymerase: (*T7 gene1)* is encoded by the lambda *DE3* prophage present within the chromosome. T7 RNA polymerase is expressed from the *lacUV5* promoter, thus the protein expression is induced with IPTG.

The transformation was carried out according to the protocol at page 23 and afterwards 100 µl of the transformation mix was plated out on selective agar plates (LB + kan [40µg/µl]) and incubated at 37° overnight. To start an ONC a single colony was picked and incubated in 10ml LB medium at 37° shaking.250 ml LB medium is inoculated with 5 ml of the ONC and incubated at 180 rpm and 37° until an OD_{600} of 0.5. The expression was induced with 0.5M IPTG.Cell samples were taken before the induction (500 µl) and after three hours of expression (250 µl). The samples were spun down for five minutes and 5000 rpm, and the pellet resolved in 30 µl 2x PGLB for SDS-PAGE to check the expression level of the different constructs.

After three hours of expression at 30°C and 180 rpm the cultures were aliquoted in 15 ml falcons and spun down at 4°C at 4000 rpm. The supernatant was discarded and the pellets kept on -20°C until they were used in the TGH assay.

3.3.1 SDS PAGE

Gel electrophoresis is a biochemically method to separate proteins according to their electric charge and molecular weight. In comparison to a native gel electrophoresis, the addition of SDS leads to a negative charge of all proteins and ensures that the proteins are only separated by molecular weight and not due to different charges. The SDS-PAGE gel consists out of a stacking gel and a separation gel. The stacking gel (4%) is on the top of the latter. The percentage of polyacrylamide in the gel depends on the size of the proteins used in the experiments.

The protein samples are mixed with 2xPGLB, heated in boiling water for 10min and spun down shortly in a table-top centrifuge. 5 µl were loaded on the gel and the chamber was filled with SDS-Running buffer.

For a distinct separation of the G0S2 protein, 18% acrylamide gels were used. The SDS-Page was performed with 200V for approximately 80 minutes using Mini Protean Tetra System from Bio-Rad.

After the separation the gels were either coloured with Coomassie-Blue or transferred on a PDVF membrane (ROTH) for Western Blotting

As a standard, to determine the exact molecular size of the proteins in the sample the Unstained Protein Molecular Weight Marker was used.

When SDS-PAGE was followed by Western blotting also the NEB prestained broad range marker was used. Both markers are shown in Figure 8.

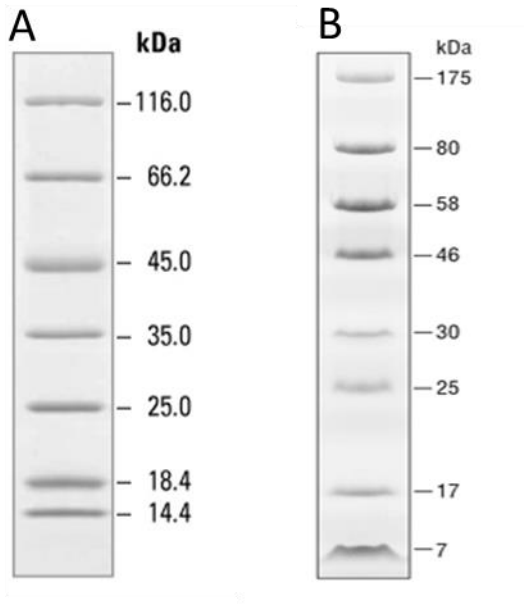


Figure 8 Proteinmarkers used in SDS-PAGE and Westernblot; A: Unstained protein marker; B: prestained protein marker broadrange

3.3.2 Coomassie Staining

For visualization of the separated protein bands, the gel was stained with Coomassie Brilliant Blue R-250, 50 % ethanol and 10 % acetic acid, for 10-30 minutes and distained with water or distain-solution. Coomassie Blue colours proteins unspecifically as it attaches to basic residues of amino acids.

3.3.3 Western Blot

Western blotting is an important technique in molecular biology to detect specific proteins in a complex mixture of cells.

It consists out of three steps:

- separation of proteins due to their size by gel electrophoresis
- transfer of the separated proteins on a membrane
- detection of the target protein by specific antibodies

After the separation on a SDS-gel the blotting sandwich was built up in following order: negative (black) plate, sponge, filter paper, gel, membrane, filter paper, sponge, positive (red) plate. As the negatively charged proteins migrate to the anode the membrane needs to be placed between the gel and the positive electrode. Before use, the membrane needs to be activated in methanol and air bubbles between gel and membrane have to be removed to ensure an efficient transfer. The whole sandwich was relocated into the blotting chamber with the black plate behind and was filled up with transfer buffer and cooled with ice.

After the transfer of 60 minutes at 200 mA the membrane was blocked in 10% milk powder in 1x TBST at room temperature for one hour or at 4°C overnight shaking.

The primary antibody mouse-anti-HIS-N-terminal was diluted 1:10000 in 1x TBST with 5% milk powder and the membrane incubated at room temperature for one hour shaking.

After the incubation period the membrane was washed 3 times, 10 minutes each wash with 1x TBST.

The secondary antibody goat-anti-mouse-HRP was diluted 1:5000 in 1x TBST with 5% milk powder and then applied for one hour shaking at room temperature.

The membrane was then washed 3 times, 10 minutes each wash with 1x TBST.

The membrane was then incubated with the chemiluminescent reagent (Amersham ECL Prime Western Blotting Detection Reagent) for 5 minutes.

The secondary antibody is coupled to horseradish peroxidase (HRP), which converts the added chemiluminescent substrate into a light emitting luminescent molecule that can be detected by exposure of the membrane to an x-ray film.

3.4 TGH Assay

The triglyceride hydrolase assay is a biochemical *in vitro* assay, which uses a radioactive-labelled substrate to directly measure the lipolytic activity in a given sample by counting the amount of released FA in a sample.

The assays performed in this master thesis are an optimized, downscaled version of the TGH assay described by Schweiger et al. [61].

For cell lysis 15 ml bacterial cell pellets were resuspended in 1 ml solution A with freshly

added protease inhibitors (1 mg/ml pepstatin, 2 mg/ml antipain, and 20 mg/ml leupeptin) and DTT (1mM) and then sonicated (Bandelin electronic UW2070 Berlin, Germany) for 1 minute and 20% power using program 5. The cell lysates were then centrifuged for 20 minutes at 14000 rpm and 4°C. The protein concentration in the supernatant was determined using Bradford protein assay.

The Bradford assay was developed by M. Bradford to detect proteins via an absorbance shift of the Coomassie-Brilliant Blue G-250 dye from 470 to 595 nm. 20 µl of protein solution (diluted 1:40 and 1:80 in solution A) were mixed with 200 µl Bio-Rad Protein Assay (Bio-Rad Laboratories Inc. Hercules, USA) in a 96 well plate. For the generation of an accurate standard curve 2 mg/ml BSA (Thermo Fisher Scientific Inc. Waltham, USA) was serially diluted and also mixed with 200 µl Bradford assay. Duplicates of each sample and BSA dilution series were measured and the absorbance at 595 or rather 620 nm was measured using a plate reader. The assays were performed with human G0S2, mouse ATLG and mouse CGI-58.

As triplicates of each sample were measured, a mastermix was prepared, whereas in each reaction a total protein amount of 25 µg of each, hG0S2 and Strep-mATGL, was mixed with 1 µg of mCGI-58 filled up to the total volume with solution A.
For 80 reactions of a 1.67 mM substrate 30 µl PC:PI (3:1) [10mg/ml], 29.5 µl TO [100mg/ml] and 25 µl hot TO [0.5 µCi] are mixed and dried off under a stream of nitrogen.
1.8 ml of KPB (0.1 M, pH 7) were added to the dried substrate and sonicated on ice three times for one minute with one minute pause in between the cycles, which was followed by four times 30 seconds with 30 seconds pause and 20% power. After the sonication 200 µl of 20 % BSA (diluted in KPB) were added.
For each reaction 25 µl of sample was mixed with 25 µl of the substrate and incubated for one hour at 37°C in a water bath shaking.

The reaction was stopped by adding 650 µl of stopping solution (methanol/chloroform/n-heptane 10:9:7) and 200 µl of 0.1 M potassium carbonate. The samples were vortexed and then centrifuged for ten minutes at 2500 rpm.
200 µl of the upper aqueous phase was mixed with 2 ml scintillation cocktail in a scintillation

vial and counted by the Beckman LS 6500 multipurpose scintillation counter (Beckman Coulter Inc.).

For the calculation of the specific substrate activity triplets of 10 µl and 20 µl substrates were also mixed with 2 ml of scintillation cocktail and counted.

Several control reactions need to be performed, which include solution A as blank, basal ATGL, ATGL* (the asterisk represents the presence of CGI-58), ATGL* mixed with the SMT protein and ATGL* mixed with the full length hG0S2 as a positive control.

The rates of ATGL's hydrolysing activity are presented as amounts of released FA per hour and mg protein referred to as counts per minute (cpm). Rates are calculated using the following equation:

$$\frac{(cpm\ sample - cpm\ blank) \times \frac{V1}{V2}}{\frac{cpm\ substrate}{nmol\ FA} \times\ mg\ protein\ \times\ 0.715\ \times\ t} = nmol\ FA\ /mg\ protein/\ h$$

V1.. is the total volume of the upper water phase (0.49 ml)

V2.. is the volume measured by liquid scintillation (0.2 ml)

t.. one hour

The measured counts per minute were divided by 0.715, because only 71.5 % of all FA were recovered into the water phase.

4 Results and discussion

4.1 N- and C-terminal truncations

Using methods as conventional cloning and site-directed mutagenesis led to the generation of either N- or C-terminally truncated hG0S2 variants. The constructs and their sizes are shown in Table 5.

Table 5 Amplified N- and C-terminal constructs and corresponding size in basepairs and kDa

Construct	bp	kDa	Numbers of residues from hG0S2 protein
Smt3-tev_hG0S2_21-52	507	18.9	32
Smt3-tev_hG0S2_22-52	504	18.8	31
Smt3-tev_hG0S2_23-52	501	18.7	30
Smt3-tev_hG0S2_24-52	498	18.6	29
Smt3-tev_hG0S2_25-52	495	18.5	28
Smt3-tev_hG0S2_26-52	492	18.3	27
Smt3-tev_hG0S2_20-44	486	18.2	25
Smt3-tev_hG0S2_20-45	489	18.3	26
Smt3-tev_hG0S2_20-46	492	18.4	27

a.) SDS-PAGE and Western blot from N- and C-terminally truncations of hG0S2

The SDS-gel (Figure 9) shows good overexpression of the hG0S2 full-length protein and of the C-terminally truncated variants 20-44, 20-45 and 20-46. N-terminal truncations do not show any visible overexpression and it is not clear which bands belong to the hG0S2 proteins.

However, on the western blot (Figure 10) the overexpression of all the constructs could be detected. Except for the variants Met23-Ala52 and Val25-Ala52, all proteins show quite a similar expression level, where at the full-length protein shows the highest overexpression.

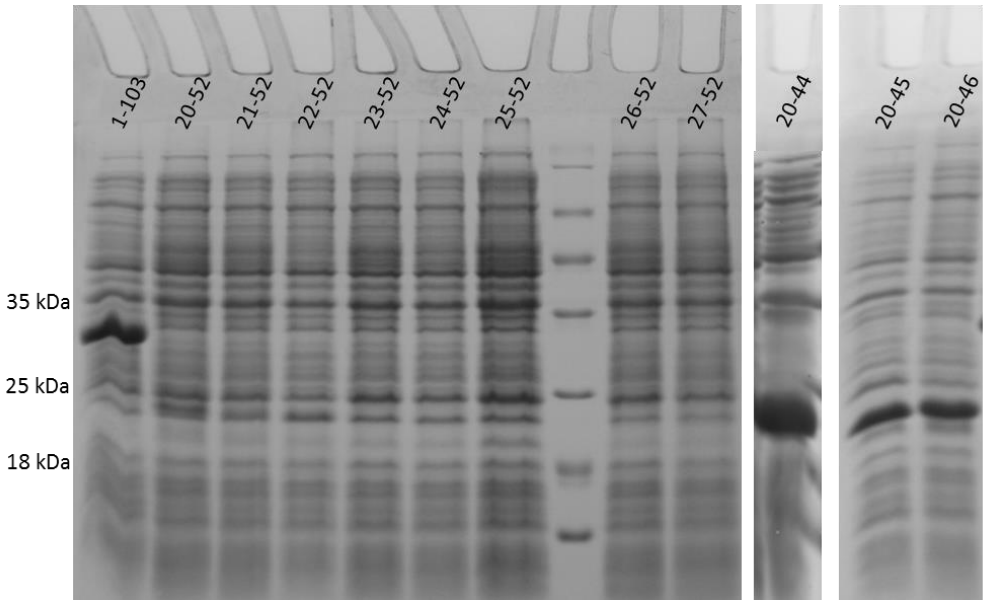


Figure 9: SDS-page of N- and C-terminally truncated G0S2 proteins after 3 hours of expression at 30°C. The cultures were induced with 0.5M IPTG. The samples were separated then on an 18% gel. Unstained protein marker with sizes in kDa.

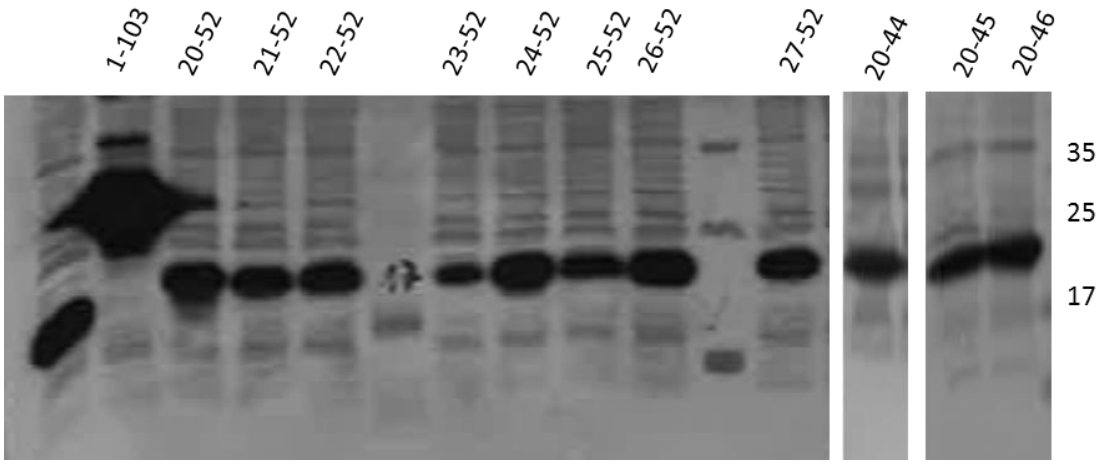


Figure 10: Western blot of N- and C-terminally truncated hG0S2 proteins after 3 hours of expression at 30°C. The cultures were induced with 0.5M IPTG. The samples were separated then on an 18% gel. Stained and unstained protein marker with sizes in kDa. Primary antibody: mouse-anti-HIS-N-terminal; secondary antibody: goat-anti-mouse-HRP

b) TGH assay of N-terminally truncated hG0S2 proteins

To test if the truncated variants are still capable of inhibiting ATGLs hydrolysing activity *in vitro*, a triglyceride hydrolase assay is used. During the assay the amounts of released FA are measured, which are seen directly proportional to ATGLs activity and indirectly proportional to G0S2 activity, respectively.

As controls basal mATGL, mATGL in presence of CGI-58 (indicated by an asterisk (*)), mATGL*G0S2-full-length as a positive- and mATGL*pSMT as an additional negative-control are used.

Each construct was transformed and expressed twice and for each expression at least tested three times for its ATGL inhibiting capability, whereas the correct number of repeats is indicated in each assay.

When testing the N- terminally truncated proteins, none of the variants from Gly21-Ala52 to Leu26-Ala52 show complete inhibitory activity towards ATGL, comparable to the full-length hG0S2 and also the Lys20-Ala52, shown in Figure 11. The newly amplified constructs are more or less half active, except of the Gly21-Ala52 variant, which always varied between 100% ATGL inactivation and 90% in all the assays done. Thus, it is difficult to state if it is able to inhibit ATGL completely and also if this 90% would also already be enough for further studies, respectively.

Notably, the variant Met23-Ala52 has markedly less ability to inhibit ATGL compared to the other variants. But when checking its expression level on the western blot, it clearly shows a lower expression compared to the other variants. Thus, different concentrations are used in the following TGH assays, to find out if it is inhibiting ability increases in a dose dependent manner (Figure 12: TGH assay: Different concentrations of the hG0S2_23-52 mutant were tested to investigate a possible dose-dependence of its inhibiting ability towards mATGL. Equal amounts of mATGL and hG0S2 were used (25µg each), additionally twofold (50µg) and threefold (75µg) amounts of hG0S2_23-52 were used. The measured FA in the samples were calculated into nmol FA/h*mg protein to obtain the percentage of mATGL hydrolysing activity. Normally, equal amounts of ATGL and G0S2 are used in the assay, which are 25µg each. In the assay in Figure 12 additional samples with 50 and 75µg of the hG0S2_Met23-Ala52 protein are tested.

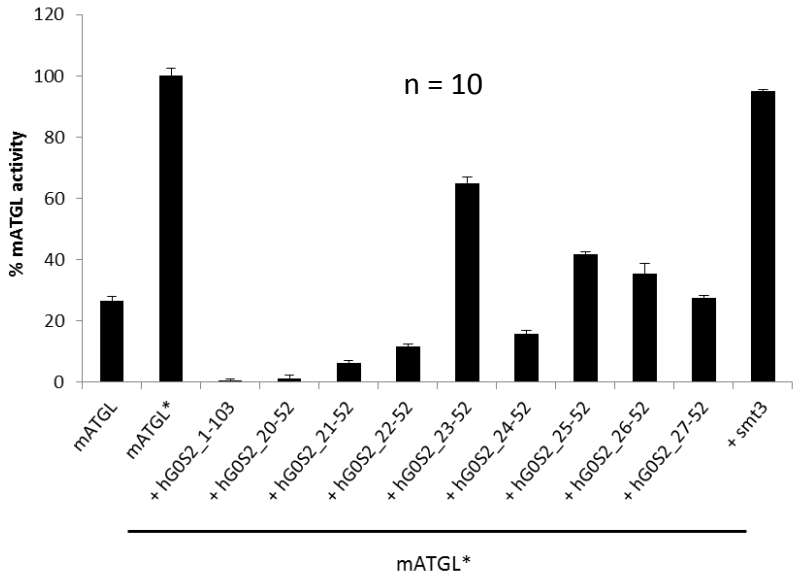


Figure 11: TGH assay: N-terminal truncated variants of hG0S2 were tested for their inhibitory activity on mATGL in vitro. Equal amounts of mATGL and hG0S2 were used, 25µg each. The measured FA in the samples were calculated into nmol FA/h*mg protein to obtain the percentage of mATGL hydrolysing activity

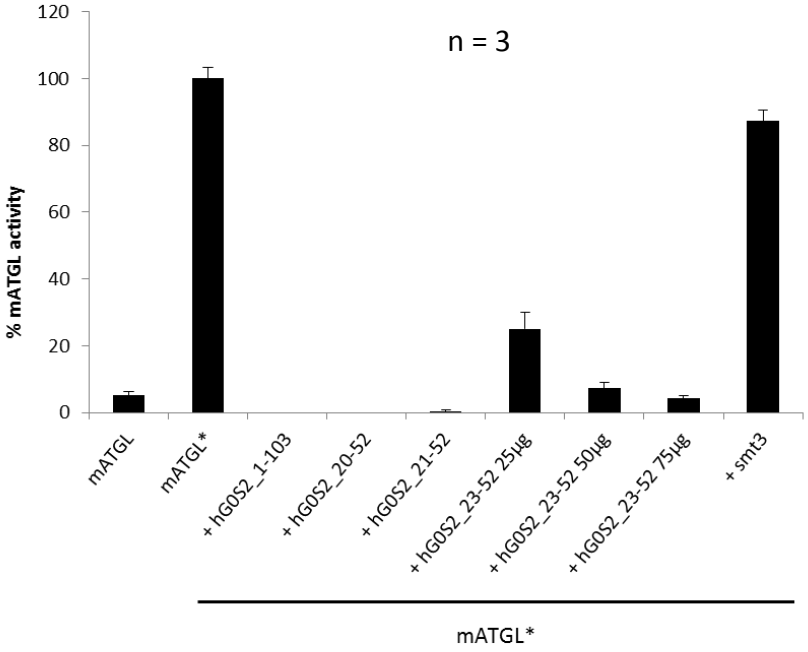


Figure 12: TGH assay: Different concentrations of the hG0S2_23-52 mutant were tested to investigate a possible dose-dependence of its inhibiting ability towards mATGL. Equal amounts of mATGL and hG0S2 were used (25µg each), additionally twofold (50µg) and threefold (75µg) amounts of hG0S2_23-52 were used. The measured FA in the samples were calculated into nmol FA/h*mg protein to obtain the percentage of mATGL hydrolysing activity

The variant Lys20-Ala52 is fully active in inhibiting ATGL's TG hydrolysing activity, but the Lys20-Met43 variant is only half-active. Consequently the C-terminal boundary has to be somewhere between Met43 and Ala52. Therefore we first generated the Lys20-Val46 variant. The resulting variant was able to fully inhibit ATGL. As a consequence we only had to produce the Lys20-Glu44 and the Lys20-Thr45 variants to determine the shortest peptide sequence, which is 25 aa when taking the Lys20-Glu44 variant.

When testing the inhibitory activity of the C-terminal truncations, one can clearly see that they have the same effect on ATGL as the full-length hG0S2 protein and the Lys20-Ala52 construct, respectively (Figure 13).

Its activities towards ATGL are totally comparable to that of the Lys20-Ala52 whereas the latter does also not show any overexpression on the SDS-PAGE. The higher over-expression of the C-terminally truncated proteins might be due to the fact that they are more stable. Stable proteins can be expressed to higher amounts, which are needed for crystallography for example or other experiments.

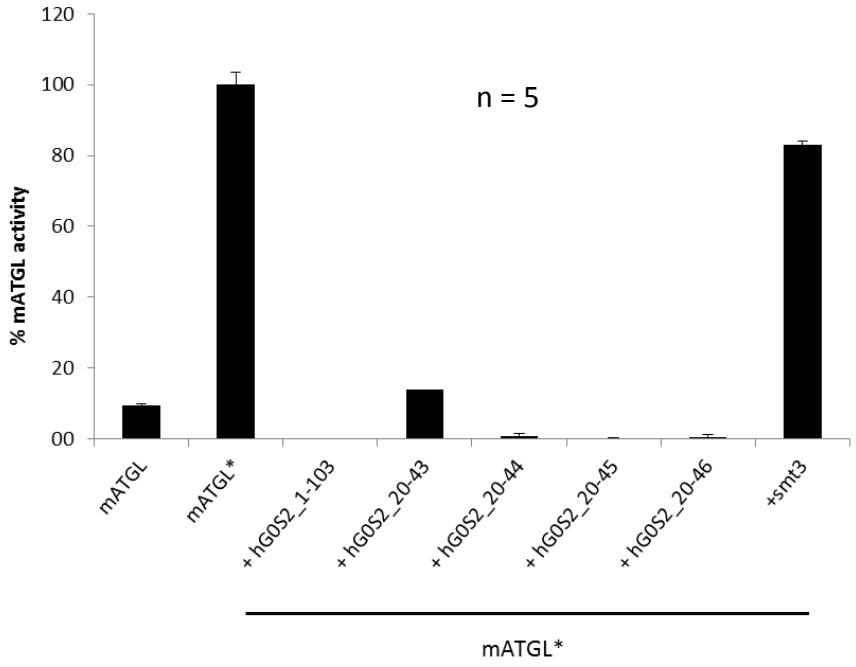


Figure 13: TGH assay: C-terminal truncated variants of hG0S2 were tested for their inhibitory activity on mATGL in vitro. Equal amounts of mATGL and hG0S2 were used, 25µg each. The measured FA in the samples were calculated into nmol FA/h*mg protein to obtain the percentage of mATGL hydrolysing activity

4.2 Mutations of hG0S2Lys20-Ala52

Due to two different site-directed mutagenesis strategies (Quickchange and Q5), three single point mutations could be generated in the hG0S2_Lys20-Ala52 construct, which are listed in Table 6.

Table 6 Amplified Lys20-Ala52 mutants and corresponding size in basepairs and kDa

Construct	bp	kDa	Numbers of residues from hG0S2 protein
Smt3-tev_hG0S2_20-52_K25D	510	19.0	33
Smt3-tev_hG0S2_20-52_Y27G	510	18.9	33
Smt3-tev_hG0S2_20-52_T45G	501	19.0	33

a.) SDS-PAGE and Western blot of hG0S2_Lys20-Ala52 mutants

The K25D and the T45G mutants do not show any clear overexpression on the gel compared to the full-length hG0S2 protein (Figure 14). Only the Y27G is slightly overexpressed. However, the western blot (Figure 15) shows a high overexpression of the Y27G and T45G mutant, whereas the K25D mutant only expressed in low amounts.

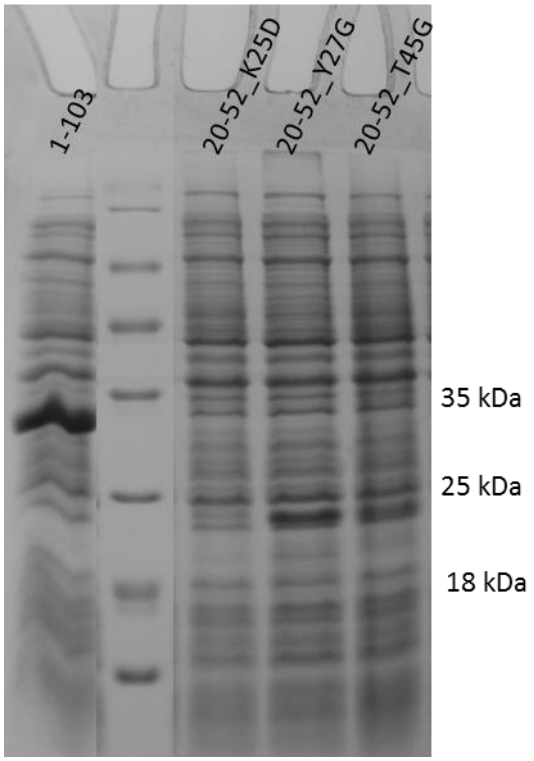


Figure 14: SDS-page of hG0S2_Lys20-Ala52 mutants after 3 hours of expression at 30°C. The cultures were induced with 0.5M IPTG. The samples were separated then on an 18% gel. Unstained protein marker with sizes in kDa.

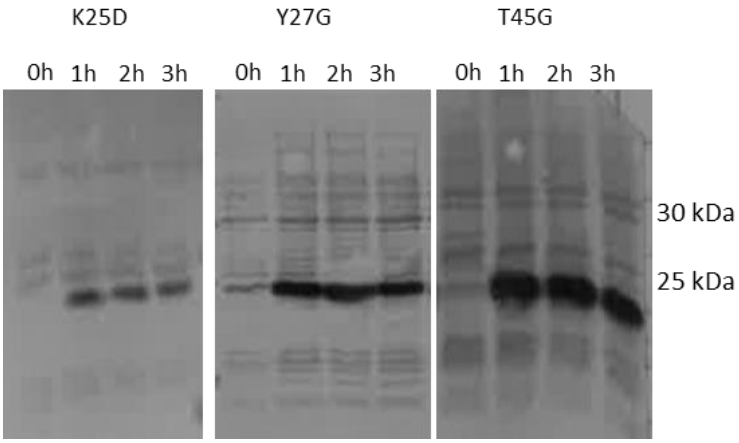


Figure 15: Western blot of hG0S2_Lys20-Ala52 mutants after 3 hours of expression at 30°C. The cultures were induced with 0.5M IPTG. The samples were separated then on an 18% gel. Unstained protein marker with sizes in kDa. Primary antibody: mouse-anti-HIS-N-terminal; secondary antibody: goat-anti-mouse-HRP

b.) TGH assay

When investigating the mutants on their inhibitory effect on ATGL, the T45G mutant shows an activity similar to the full-length protein and the Lys20-Ala52 WT, respectively (Figure 16). It seems that the mutation does not affect the inhibition of ATGL, in contrast to the Y27G mutation. The Y27G mutant is only half active, which suggests an important role of Tyr27 in the biochemical reaction. The K25D mutant also does not show complete inactivation of ATGL, but this might also be due to its low expression level observed on the western blot (Figure 15).

It was not clear whether the low expression level of K25D and not the mutation itself is the reason for its inhibiting incapability. Thus, to prove if the efficiency of Lys20-Ala52_K25D in inhibiting ATGL depends on the applied dose, a two- and a threefold amount of the normal concentration (which is 25µg) was used.

Figure 17 clearly shows that the inhibition occurs in a dose dependent manner; whereas a complete inactivation of ATGL could also not be established with a threefold amount of the K25D mutant.

These results give already a clue about G0S2-ATGL interaction. The residues Lys25 and Thr45 seem not to be required for ATGL binding or inhibition, whereas Tyr27 plays an important role in the inhibitory mechanism of G0S2 towards ATGL. However, if Tyr27 is needed for binding ATGL or if it is part of inhibiting the catalytic reaction cannot be stated. To go deeper into this characterization and to determine distinct residues performing specific functions, further mutations in this hydrophobic stretch of G0S2 need to be generated, as well as mutations on ATGL's binding region for G0S2.

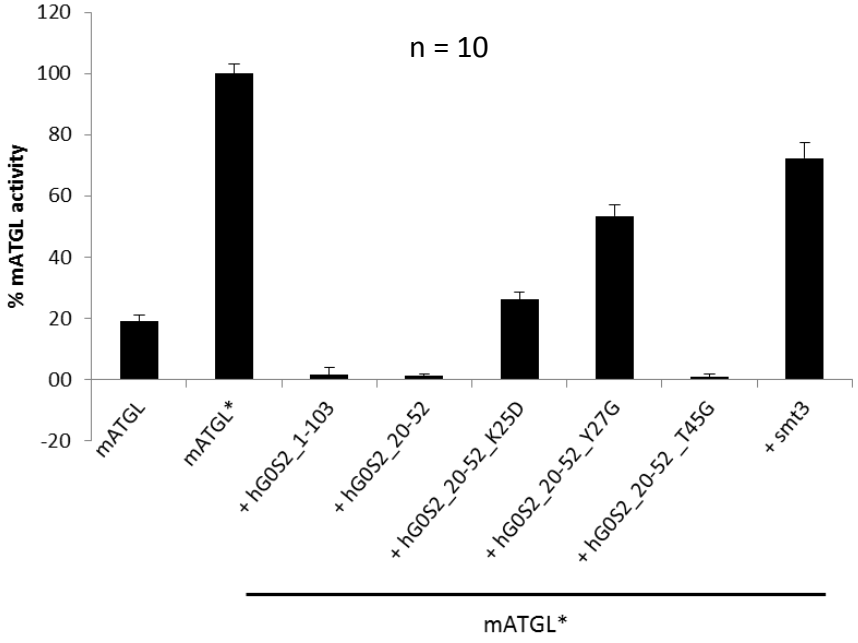


Figure 16: TGH assay: the Lys20-Ala52 mutants were tested for their inhibitory activity on ATGL in vitro. Equal amounts of mATGL and hG0S2 were used, 25µg each. The measured FA in the samples were calculated into nmol FA/h*mg protein to obtain the percentage of mATGL hydrolysing activity.

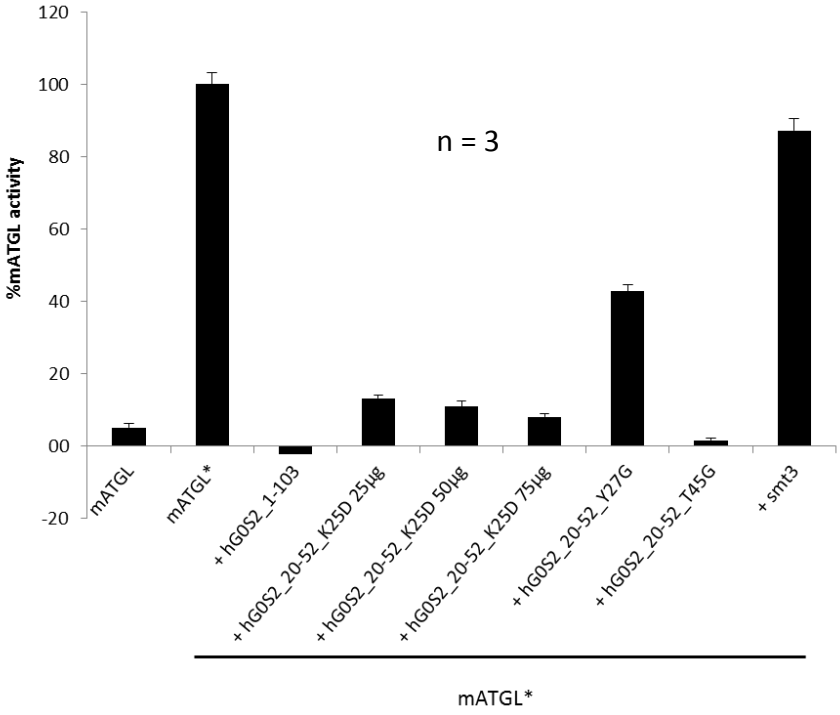


Figure 17: TGH assay: Different concentrations of the hG0S2_20-52_K25D mutant were tested to to investigate a possible dose-dependence of its inhibiting ability towards mATGL. Equal amounts of mATGL and hG0S2 were used, 25µg each, additionally twofold (50µg) and threefold (75µg) amounts of hG0S2_20-52_K25Dwere used. The measured FA in the samples were calculated into nmol FA/h*mg protein to obtain the percentage of mATGL hydrolyzing activity

4.3 G0S2 from evolutionary distant species

In our *in vitro* setup, we use human G0S2 to inhibit mouse ATGL in presence of mouse CGI-58. Additionally, we wanted to test whether evolutionary distant genes are also able to inhibit mATGL. Synthetic genes coding for G0S2 orthologues from chicken, Japanese rice fish and duckbill were purchased, cloned in the pet28a_SMT_tev expression vector and tested for their inhibiting activity *in vitro*. The constructs including their sizes are shown in Table 7.

Table 7 Generated constructs with G0S2 from different species

Construct	Bp	kDa	Numbers of residues from hG0S2 protein
Smt3-tev_chickenG0S2	711	26.7	99
Smt3-tev_japanesericefishG0S2	783	29.5	123
Smt3-tev_platypusG0S2	750	27.9	112

a.) SDS-PAGE and Western blot

The SDS-PAGE (Figure 18) shows clearly that the G0S2 proteins are overexpressed properly, which is also was confirmed on the western blot shown in Figure 19. G0S2 from chicken, Japanese rice fish and platypus differ a bit in size compared to the human G0S2. The chicken G0S2 consists out of 99 aa, the Japanese rice fish G0S2 is 123 aa in length and G0S2 from platypus 112 aa, that is why they run on different heights on the gel.

The chicken and Japanese rice fish G0S2 also show expression in the non-induced status. It should be mentioned here that the pet28a vector sometimes show leaky expression, which is also observed in this case.

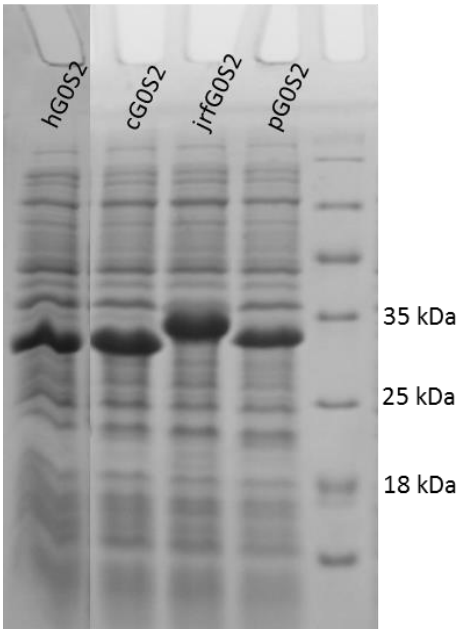


Figure 18: SDS-Page of G0S2 from chicken, Japanese rice fish and platypus f.l.t.r. compared to humanG0S2, after 3 hours of expression at 30°C. The cultures were induced with 0.5M IPTG. The samples were separated then on an 18% gel. Unstained protein marker with sizes in kDa

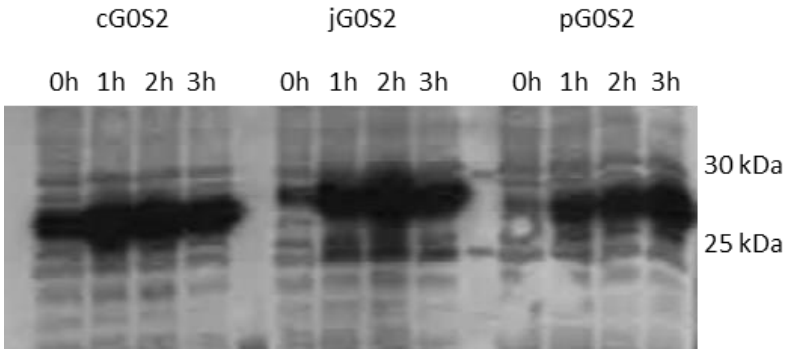


Figure 19: Western blot of G0S2 from chicken, Japanese rice fish and platypus f.l.t.r., after 3 hours of expression at 30°C. The cultures were induced with 0.5M IPTG. The samples were separated then on an 18% gel. Prestained protein marker with sizes in kDa. Primary antibody: mouse-anti-HIS-N-terminal; secondary antibody: goat-anti-mouse-HRP

b.) TGH assay

The TGH assay (Figure 20) clearly shows that the G0S2 proteins from chicken, Japanese rice fish and platypus are also able to fully inhibit mouse ATGL's hydrolysing activity comparable to human G0S2. Although the sequences are less identical compared to mouse and human G0S2, it seems that the N-terminal conservation is sufficient for complete ATGL inactivation (

Figure 21).

As all efforts to crystallize the hG0S2 were not successful yet, these findings could provide a novel approach in finding out the 3D structure of the G0S2 protein.

It would also be interesting to truncate these G0S2 proteins similar to the hG0S2_20_44 variant to confirm the conserved regulatory module.

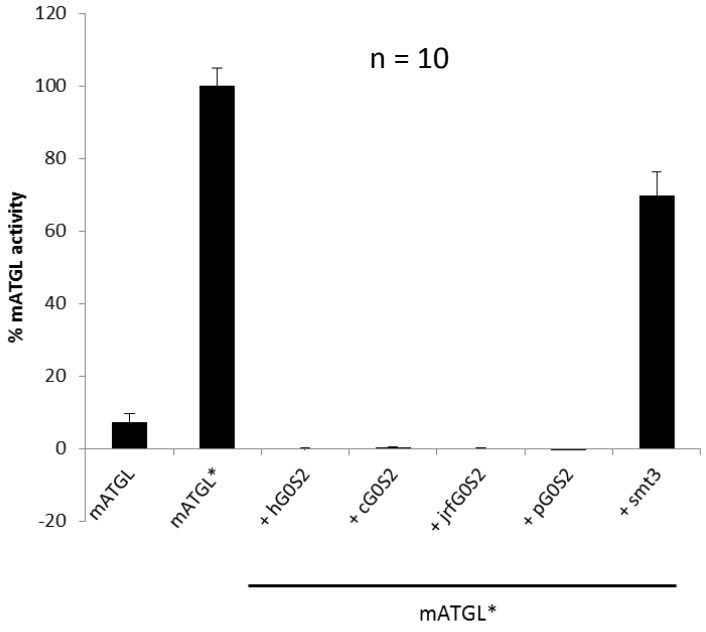


Figure 20: TGH assay: genes of G0S2 from chicken, Japanese rice fish and platypus were tested on their inhibitory effect on mATGL. Equal amounts of mATGL and hG0S2 were used, 25µg each. The measured FA in the samples were calculated into nmol FA/h*mg protein to obtain the percentage of mATGL hydrolyzing activity

A

Human	100 %				
Mouse	78 %	100 %			
Platypus	42 %	42 %	100 %		
Gallus	44 %	40 %	48 %	100 %	
Oryhzias	23 %	26 %	29 %	31 %	100 %
	Human	**Mouse**	**Platypus**	**Gallus**	**Oryhzias**

B

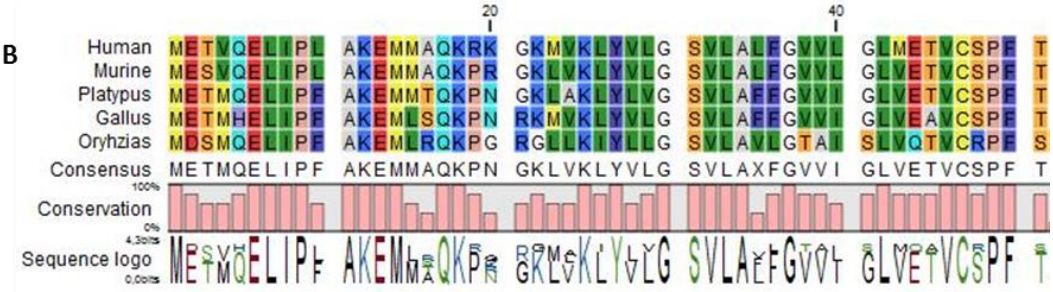


Figure 21: Sequence identities and conservation of human, mouse, chicken, Japanese rice fish and platypus G0S2. A: alignment done by SIAS. B: alignment generated using CLC-workbench (trial-version)

5 Conclusion and outlook

The shortest, active G0S2 peptide identified in this work corresponds to the Lys20-Glu44 variant, consisting out of 25 aa, which represents a fourth of the full-length hG0S2 protein. However, significant inhibitory capacity could also be observed for the 21-44 variant (in our experiments, the Gly21-Ala52's ability to inhibit ATGL *in vitro* varies between 94 – 98 %, compared to the Lys20-Ala52 and full-length protein, which are between 98-100%).

Determination of the shortest C-terminal truncations was more straight-forward. Interestingly, the shortest variants also show a better protein expression, as detected with SDS-PAGE. Coomassie staining often was not sufficient to clearly identify samples from N-terminal truncations.

For a next step it would be crucial to order synthetic peptides with sequence stretches corresponding to the G0S2 sequence of Lys20-Cys44 (and potentially 21-44) to investigate their efficiency in ATGL inhibition.

Furthermore the generation of a fusion construct of the shortest G0S2 fragment with adipose tissue specific targeting sequence for experiments in cell culture would be also of

fundamental interest. If the experiments show success, animal experiments need to be established prospectively, as the initial aim is the development of a tissue specific drug.

To find out more about the underlying interaction mechanism of ATGL and G0S2, further mutation studies need to be done. To get a deeper look upon conserved residues of G0S2, a HMM logo using PFAM (http://pfam.xfam.org/family/PF15103.1#tabview=tab4) was generated. The profile shows the level of conservation of each amino acid within the G0S2 proteins from 34 different organisms. Lys25 seems to be highly conserved, Tyr27 even more, which suggests that it might play an important role in G0S2 activity, which was also shown in my work. Ala34 also appears to be much conserved also Gly39, Gly37 and Gly40.

Thus, additional mutation studies are necessary to investigate how these changes of single residues influence the inhibitory activity from G0S2 on ATGL.

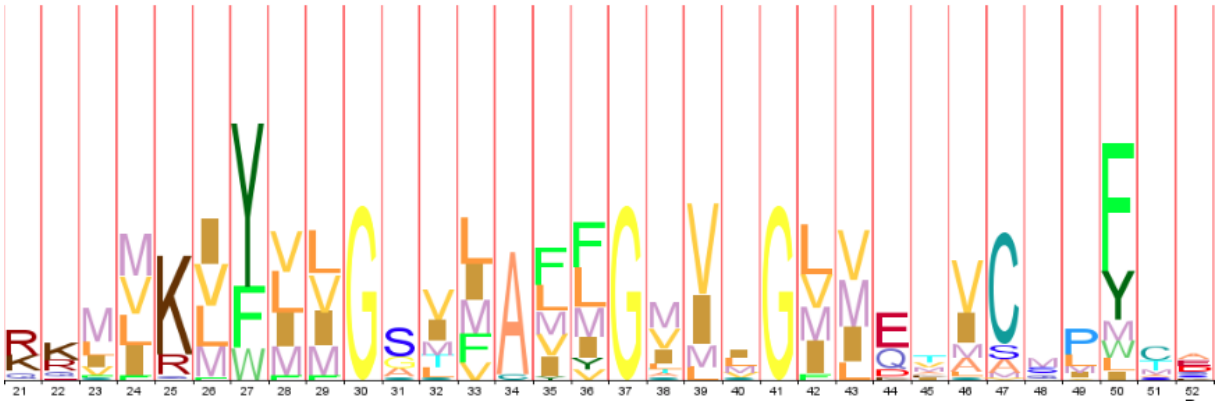


Figure 22: HMM-Logo from G0S2 from 34 different species using pfam.xfam.org. The size of the letters gets along with the level of conservation. It is only a zoomed out view from the whole logo, showing amino acids corresponding to the hydrophobic core of G0S2.

As G0S2 from distant species are highly expressed and also show complete activity towards ATGL, it could also be tried to truncate them in a similar manner comparable to the shortest hG0S2 construct, as the conservation is the highest in this region.

It could also be attempted to do crystallization studies with these G0S2 proteins, as it was not possible yet to crystallize human and mouse ATGL.

6 Bibliography

1. Cordain L, Eaton SB, Sebastian A, Mann N, Lindeberg S, et al. (2005) Origins and evolution of the Western diet: health implications for the 21st century. The American Journal of Clinical Nutrition 81: 341-354.
2. Bartz R, Li WH, Venables B, Zehmer JK, Roth MR, et al. (2007) Lipidomics reveals that adiposomes store ether lipids and mediate phospholipid traffic. J Lipid Res 48: 837-847.
3. Kuerschner L, Moessinger C, Thiele C (2008) Imaging of lipid biosynthesis: how a neutral lipid enters lipid droplets. Traffic 9: 338-352.
4. Stone SJ, Levin MC, Zhou P, Han J, Walther TC, et al. (2009) The endoplasmic reticulum enzyme DGAT2 is found in mitochondria-associated membranes and has a mitochondrial targeting signal that promotes its association with mitochondria. J Biol Chem 284: 5352-5361.
5. Brasaemle DL (2007) Thematic review series: adipocyte biology. The perilipin family of structural lipid droplet proteins: stabilization of lipid droplets and control of lipolysis. J Lipid Res 48: 2547-2559.
6. Holm C (2003) Molecular mechanisms regulating hormone-sensitive lipase and lipolysis. Biochem Soc Trans 31: 1120-1124.
7. Lass A, Zimmermann R, Haemmerle G, Riederer M, Schoiswohl G, et al. (2006) Adipose triglyceride lipase-mediated lipolysis of cellular fat stores is activated by CGI-58 and defective in Chanarin-Dorfman Syndrome. Cell Metab 3: 309-319.
8. Zimmermann R, Strauss JG, Haemmerle G, Schoiswohl G, Birner-Gruenberger R, et al. (2004) Fat mobilization in adipose tissue is promoted by adipose triglyceride lipase. Science 306: 1383-1386.
9. Fredrikson G, Belfrage P (1983) Positional specificity of hormone-sensitive lipase from rat adipose tissue. J Biol Chem 258: 14253-14256.
10. Haemmerle G, Zimmermann R, Hayn M, Theussl C, Waeg G, et al. (2002) Hormone-sensitive lipase deficiency in mice causes diglyceride accumulation in adipose tissue, muscle, and testis. J Biol Chem 277: 4806-4815.
11. Karlsson M, Contreras JA, Hellman U, Tornqvist H, Holm C (1997) cDNA cloning, tissue distribution, and identification of the catalytic triad of monoglyceride lipase. Evolutionary relationship to esterases, lysophospholipases, and haloperoxidases. J Biol Chem 272: 27218-27223.
12. Jenkins CM, Mancuso DJ, Yan W, Sims HF, Gibson B, et al. (2004) Identification, cloning, expression, and purification of three novel human calcium-independent phospholipase A2 family members possessing triacylglycerol lipase and acylglycerol transacylase activities. J Biol Chem 279: 48968-48975.
13. Villena JA, Roy S, Sarkadi-Nagy E, Kim KH, Sul HS (2004) Desnutrin, an adipocyte gene encoding a novel patatin domain-containing protein, is induced by fasting and glucocorticoids: ectopic expression of desnutrin increases triglyceride hydrolysis. J Biol Chem 279: 47066-47075.
14. Wilson PA, Gardner SD, Lambie NM, Commans SA, Crowther DJ (2006) Characterization of the human patatin-like phospholipase family. J Lipid Res 47: 1940-1949.
15. Shewry PR (2003) Tuber storage proteins. Ann Bot 91: 755-769.
16. Dessen A, Tang J, Schmidt H, Stahl M, Clark JD, et al. (1999) Crystal structure of human cytosolic phospholipase A2 reveals a novel topology and catalytic mechanism. Cell 97: 349-360.

17. Rydel TJ, Williams JM, Krieger E, Moshiri F, Stallings WC, et al. (2003) The crystal structure, mutagenesis, and activity studies reveal that patatin is a lipid acyl hydrolase with a Ser-Asp catalytic dyad. Biochemistry 42: 6696-6708.
18. Lass A, Zimmermann R, Oberer M, Zechner R (2011) Lipolysis - a highly regulated multi-enzyme complex mediates the catabolism of cellular fat stores. Prog Lipid Res 50: 14-27.
19. Lu X, Yang X, Liu J (2010) Differential control of ATGL-mediated lipid droplet degradation by CGI-58 and G0S2. Cell Cycle 9: 2719-2725.
20. Li LO, Klett EL, Coleman RA (2010) Acyl-CoA synthesis, lipid metabolism and lipotoxicity. Biochim Biophys Acta 1801: 246-251.
21. Unger RH, Clark GO, Scherer PE, Orci L (2010) Lipid homeostasis, lipotoxicity and the metabolic syndrome. Biochim Biophys Acta 1801: 209-214.
22. Das SK, Eder S, Schauer S, Diwoky C, Temmel H, et al. (2011) Adipose triglyceride lipase contributes to cancer-associated cachexia. Science 333: 233-238.
23. Haemmerle G, Lass A, Zimmermann R, Gorkiewicz G, Meyer C, et al. (2006) Defective lipolysis and altered energy metabolism in mice lacking adipose triglyceride lipase. Science 312: 734-737.
24. Schweiger M, Schreiber R, Haemmerle G, Lass A, Fledelius C, et al. (2006) Adipose triglyceride lipase and hormone-sensitive lipase are the major enzymes in adipose tissue triacylglycerol catabolism. J Biol Chem 281: 40236-40241.
25. Fruhbeck G, Mendez-Gimenez L, Fernandez-Formoso JA, Fernandez S, Rodriguez A (2014) Regulation of adipocyte lipolysis. Nutr Res Rev 27: 63-93.
26. Egan JJ, Greenberg AS, Chang MK, Londos C (1990) Control of endogenous phosphorylation of the major cAMP-dependent protein kinase substrate in adipocytes by insulin and beta-adrenergic stimulation. J Biol Chem 265: 18769-18775.
27. Greenberg AS, Egan JJ, Wek SA, Garty NB, Blanchette-Mackie EJ, et al. (1991) Perilipin, a major hormonally regulated adipocyte-specific phosphoprotein associated with the periphery of lipid storage droplets. J Biol Chem 266: 11341-11346.
28. Patel RT, Soulages JL, Hariharasundaram B, Arrese EL (2005) Activation of the lipid droplet controls the rate of lipolysis of triglycerides in the insect fat body. J Biol Chem 280: 22624-22631.
29. Collins S, Surwit RS (2001) The beta-adrenergic receptors and the control of adipose tissue metabolism and thermogenesis. Recent Prog Horm Res 56: 309-328.
30. Krintel C, Osmark P, Larsen MR, Resjo S, Logan DT, et al. (2008) Ser649 and Ser650 are the major determinants of protein kinase A-mediated activation of human hormone-sensitive lipase against lipid substrates. PLoS One 3: e3756.
31. Zechner R, Zimmermann R, Eichmann TO, Kohlwein SD, Haemmerle G, et al. (2012) FAT SIGNALS--lipases and lipolysis in lipid metabolism and signaling. Cell Metab 15: 279-291.
32. Yamaguchi T, Omatsu N, Morimoto E, Nakashima H, Ueno K, et al. (2007) CGI-58 facilitates lipolysis on lipid droplets but is not involved in the vesiculation of lipid droplets caused by hormonal stimulation. J Lipid Res 48: 1078-1089.
33. Cornaciu I, Boeszoermenyi A, Lindermuth H, Nagy HM, Cerk IK, et al. (2011) The minimal domain of adipose triglyceride lipase (ATGL) ranges until leucine 254 and can be activated and inhibited by CGI-58 and G0S2, respectively. PLoS One 6: e26349.
34. Heckmann BL, Zhang X, Xie X, Saarinen A, Lu X, et al. (2014) Defective adipose lipolysis and altered global energy metabolism in mice with adipose overexpression of the lipolytic inhibitor G0/G1 switch gene 2 (G0S2). J Biol Chem 289: 1905-1916.

35. Schweiger M, Paar M, Eder C, Brandis J, Moser E, et al. (2012) G0/G1 switch gene-2 regulates human adipocyte lipolysis by affecting activity and localization of adipose triglyceride lipase. J Lipid Res 53: 2307-2317.
36. Wang Y, Zhang Y, Qian H, Lu J, Zhang Z, et al. (2013) The G0/G1 switch gene 2 is an important regulator of hepatic triglyceride metabolism. PLoS One 8: e72315.
37. Yang X, Zhang X, Heckmann BL, Lu X, Liu J (2011) Relative contribution of adipose triglyceride lipase and hormone-sensitive lipase to tumor necrosis factor-alpha (TNF-alpha)-induced lipolysis in adipocytes. J Biol Chem 286: 40477-40485.
38. Festuccia WT, Laplante M, Berthiaume M, Gelinas Y, Deshaies Y (2006) PPARgamma agonism increases rat adipose tissue lipolysis, expression of glyceride lipases, and the response of lipolysis to hormonal control. Diabetologia 49: 2427-2436.
39. Kershaw EE, Schupp M, Guan HP, Gardner NP, Lazar MA, et al. (2007) PPARgamma regulates adipose triglyceride lipase in adipocytes in vitro and in vivo. Am J Physiol Endocrinol Metab 293: E1736-1745.
40. Kim JY, Tillison K, Lee JH, Rearick DA, Smas CM (2006) The adipose tissue triglyceride lipase ATGL/PNPLA2 is downregulated by insulin and TNF-alpha in 3T3-L1 adipocytes and is a target for transactivation by PPARgamma. Am J Physiol Endocrinol Metab 291: E115-127.
41. Jocken JW, Langin D, Smit E, Saris WH, Valle C, et al. (2007) Adipose triglyceride lipase and hormone-sensitive lipase protein expression is decreased in the obese insulin-resistant state. J Clin Endocrinol Metab 92: 2292-2299.
42. Kralisch S, Klein J, Lossner U, Bluher M, Paschke R, et al. (2005) Isoproterenol, TNFalpha, and insulin downregulate adipose triglyceride lipase in 3T3-L1 adipocytes. Mol Cell Endocrinol 240: 43-49.
43. Xie X, Langlais P, Zhang X, Heckmann BL, Saarinen AM, et al. (2014) Identification of a novel phosphorylation site in adipose triglyceride lipase as a regulator of lipid droplet localization. Am J Physiol Endocrinol Metab 306: E1449-1459.
44. Pagnon J, Matzaris M, Stark R, Meex RC, Macaulay SL, et al. (2012) Identification and functional characterization of protein kinase A phosphorylation sites in the major lipolytic protein, adipose triglyceride lipase. Endocrinology 153: 4278-4289.
45. Mason RR, Meex RC, Lee-Young R, Canny BJ, Watt MJ (2012) Phosphorylation of adipose triglyceride lipase Ser(404) is not related to 5'-AMPK activation during moderate-intensity exercise in humans. Am J Physiol Endocrinol Metab 303: E534-541.
46. Zhang L, Godzik A, Skolnick J, Fetrow JS (1998) Functional analysis of the Escherichia coli genome for members of the alpha/beta hydrolase family. Fold Des 3: 535-548.
47. Lefevre C, Jobard F, Caux F, Bouadjar B, Karaduman A, et al. (2001) Mutations in CGI-58, the gene encoding a new protein of the esterase/lipase/thioesterase subfamily, in Chanarin-Dorfman syndrome. Am J Hum Genet 69: 1002-1012.
48. Cygler M, Schrag JD, Sussman JL, Harel M, Silman I, et al. (1993) Relationship between sequence conservation and three-dimensional structure in a large family of esterases, lipases, and related proteins. Protein Sci 2: 366-382.
49. Schrag JD, Cygler M (1997) Lipases and alpha/beta hydrolase fold. Methods Enzymol 284: 85-107.
50. Granneman JG, Moore HP, Granneman RL, Greenberg AS, Obin MS, et al. (2007) Analysis of lipolytic protein trafficking and interactions in adipocytes. J Biol Chem 282: 5726-5735.

51. Gruber A, Cornaciu I, Lass A, Schweiger M, Poeschl M, et al. (2010) The N-terminal region of comparative gene identification-58 (CGI-58) is important for lipid droplet binding and activation of adipose triglyceride lipase. J Biol Chem 285: 12289-12298.
52. Russell L, Forsdyke DR (1991) A human putative lymphocyte G0/G1 switch gene containing a CpG-rich island encodes a small basic protein with the potential to be phosphorylated. DNA Cell Biol 10: 581-591.
53. Yang X, Lu X, Lombes M, Rha GB, Chi YI, et al. (2010) The G(0)/G(1) switch gene 2 regulates adipose lipolysis through association with adipose triglyceride lipase. Cell Metab 11: 194-205.
54. Schweiger M, Schoiswohl G, Lass A, Radner FP, Haemmerle G, et al. (2008) The C-terminal region of human adipose triglyceride lipase affects enzyme activity and lipid droplet binding. J Biol Chem 283: 17211-17220.
55. Zandbergen F, Mandard S, Escher P, Tan NS, Patsouris D, et al. (2005) The G0/G1 switch gene 2 is a novel PPAR target gene. Biochem J 392: 313-324.
56. Nielsen TS, Vendelbo MH, Jessen N, Pedersen SB, Jorgensen JO, et al. (2011) Fasting, but not exercise, increases adipose triglyceride lipase (ATGL) protein and reduces G(0)/G(1) switch gene 2 (G0S2) protein and mRNA content in human adipose tissue. J Clin Endocrinol Metab 96: E1293-1297.
57. Kusakabe M, Watanabe K, Emoto N, Aki N, Kage H, et al. (2009) Impact of DNA demethylation of the G0S2 gene on the transcription of G0S2 in squamous lung cancer cell lines with or without nuclear receptor agonists. Biochem Biophys Res Commun 390: 1283-1287.
58. Welch C, Santra MK, El-Assaad W, Zhu X, Huber WE, et al. (2009) Identification of a protein, G0S2, that lacks Bcl-2 homology domains and interacts with and antagonizes Bcl-2. Cancer Res 69: 6782-6789.
59. Siderovski DP, Blum S, Forsdyke RE, Forsdyke DR (1990) A set of human putative lymphocyte G0/G1 switch genes includes genes homologous to rodent cytokine and zinc finger protein-encoding genes. DNA Cell Biol 9: 579-587.
60. Cerk IK, Salzburger B, Boeszoermenyi A, Heier C, Pillip C, et al. (2014) A peptide derived from G0/G1 switch gene 2 acts as noncompetitive inhibitor of adipose triglyceride lipase. J Biol Chem 289: 32559-32570.
61. Schweiger M, Eichmann TO, Taschler U, Zimmermann R, Zechner R, et al. (2014) Measurement of lipolysis. Methods Enzymol 538: 171-193.

7 Appendix

This section comprises general protocols for the preparation of competent cells used in our laboratory, sequences and sequencing files, respectively.

7.1 Protocols

Competent cells are capable to take-up plasmid DNA and represent a useful tool for cloning and expression of genes of interest. The uptake of DNA can be initiated via electroporation or heat shock in electrically or chemically competent cells, respectively.

For this thesis two different *E. coli* strains were made competent: One Shot TOP 10 cells (Invitrogen life technologies, Paisley, UK) for cloning, and BL21(DE3) RILP CodonPlus cells (Stratagene Agilent Technologies, Cedar Creek, USA) for expression.

7.1.1 Preperation of chemically competent cells

- Inoculate 3 ml LB medium with the appropriate E. coli strain and incubate the culture overnight at 37°C
- Add the overnight culture to 500 ml LB medium and incubate the culture at 30°C until the absorbance at 600nm was approximately 0.5
- Chill the culture for at least 10 min on ice (in the following steps the suspension should be kept on ice as much as possible
- Centrifuge the cell suspension for 10 min at 3000 rpm (JA-10 rotor) at 4°C
- Gently resuspend the pellet in 100 ml ice-cold TB buffer
- Incubate the cell suspension on ice for 10 min.
- Centrifuge for 10 min at 3000 rpm at 4°C
- Gently resuspend the pellet in 18.6 ml ice-cold TB buffer and add 1.4 ml DMSO
- Incubate the cell suspension on wet ice for at least 10 min.
- Aliquot the cell suspension at 200 µl per tube
- Shock-freeze the cell suspension in liquid nitrogen and store the tubes at -80°C

7.1.2 Preperation of electro competent cells

- Inoculate 30 ml LB with single colony from a fresh plate and grow them overnight at 37°C.
- Inoculate 1l LB with 10 ml of the ONC and grow it at 37°C on a shaker with 200 rpm until it reaches an OD600 of 0.4-0.5 (approximately after 4h). Continue with all working steps on ice or at 4°C
- Centrifuge cells for 20 min at 3000 rpm (JA-10 rotor) at 4°C
- Wash with 1 volume 10% glycerol
- Centrifuge cells for 20 min at 3000 rpm at 4°C
- Wash with 0.5 volume 10% glycerol
- Centrifuge cells for 20 min at 3000 rpm at 4°C
- Wash with 0.1 volume 10% glycerol
- Centrifuge cells for 20 min at 3000 rpm at 4°C

- Add 0.5 – 1 ml 10% glycerol (per litre liquid culture started with), resuspend pellet and prepare aliquots of 50 µl in 1.5 ml reaction tubes and freeze them immediately in liquid nitrogen
- Store the electro competent cells at -80°C.

7.2 Sequences

This unit consists of sequences relevant for this work devided into wild-type sequences, template sequences used for PCR, sequences of proteins used in the assays and obtained sequencing results. The sequence alignements were done with ClustalW or ClustalO.

7.2.1 Wild-type sequences

This chapter contains the wild-type sequences of G0S2 from different species and sequences of mATGL and mCGI-58. All sequences are taken from NCBI.

Human (Homo sapiens) G0S2

DNA

```
>gi|609453|gb|M69199.1|HUMG0S2A Human G0S2 protein gene, complete cds
GGAGTCTCCAACTGGGAGAGCTGCAGCTGCCGAGAGGAGGAGAACGCTGAGGTCGGTCGGACCAACGGACGCGCTGACCGCTGCCAACTGCAGC
TCGCGCTGCCTCCTGCTCGCGCCGTGCCACTAAGGTCACTCCCGCCTCCGAGAGCCCAGAGCCGAGATGGAAACGGTCCAGGAGCTGATCCCCC
TGGCCAAGGAGATGATGGCCCAGAAGCGCAAGGGGAAGATGGTGAAGCTGTACGTGCTGGGCAGCGTGCTGGCCCTCTTCGGCGTGGTGCTCGG
CCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCGCCAGACGTCTGCGGGACCAGGAGGCAGCCGTGGCGGAGCTGCAGGCCGCCCTGGAGCGA
CAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACACGGTCCTCGGCGGCCGGGCCCTGTCCAACCGGCAGCACGCCTCCT
AGGAACTGTGGGAGACCAGCGGAGTGGGAGGGAGACGCAGTAGACAGAGACAGACCGAGAAGGAAGGGAGAGACAGAGGGGGCGCGCGCACAGG
AGCCTGACTCCGCTGGGAGAGTGCAGGAGCACGTGCTGTTTTTTATTTGGACTTAACTTCAGAGAAACCGCTGACATCTAGAACTGACCTACCA
CAAGCATCCACCAAAGGAGTTTGGGATTGAGTTTTGCTGCTGTGCAGCACTGCATTGTCATGACATTTCCAACACTGTGTGAATTATCTAAATG
CGTCTACCATTTTGCACTAGGGAGGAAGGATAAATGCTTTTTATGTTATTATTATTAATTATTACAATGACCACCATTTTGCATTTTGAAATAA
AAAACTTTTTATACCAT

>gb|M69199.1|HUMG0S2A:161-472 Human G0S2 protein gene, complete cds
ATGGAAACGGTCCAGGAGCTGATCCCCCTGGCCAAGGAGATGATGGCCCAGAAGCGCAAGGGGAAGATGGTGAAGCTGTACGTGCTGGGCAGCG
TGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCGCCAGACGTCTGCGGGACCAGGAGGCAGCCGT
GGCGGAGCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACACGGTCCTCGGCGGCCGG
GCCCTGTCCAACCGGCAGCACGCCTCCTAG
```

Protein

```
>gi|182851|gb|AAB04044.1| G0S2 protein [Homo sapiens]
METVQELIPLAKEMMAQKRKGKMVKLYVLGSVLALFGVVLGLMETVCSPFTAARRLRDQEAAVAELQAAL
ERQALQKQALQEKGKQQDTVLGGRALSNRQHAS
```

Chicken (Gallus gallus) G0S2

DNA

```
>gi|312283675|ref|NM_001190924.3| Gallus gallus G0/G1switch 2 (G0S2), mRNA
CGGGCGGAGCCGCGGGCAGCGCTGTGCGGGGCGAAAGAGCTGAGCGCCGCAACGGCACCGCAACCGCACCGCGGCCTCAGAACGGAGCTCTTCC
TCCGCTGCTGTGCAGAGATGGAAACCATGCACGAGCTGATCCCCTTCGCCAAAGAGATGCTCAGCCAGAAGCCCAACAGGAAGATGGTGAAGCT
GTACGTGCTGGGCAGCGTGCTGGCGTTCTTCGGCGTGGTCATCGGTCTGGTGGAGGCAGTGTGCAGCCCTTTCACCTCCGAAGGGAATATAGAG
GAGGAGAAGAGACCGAGCCCATCTCGAGAGCCAGCGCTTCCTCGGAAGCGGGAGGATTTGGTGTTGGAGCAGAGCAAGAAGTCGTCAGCGGTGC
AGCGGGGGGTGGTGACCAGGCAGCATGCATCCTAAGGGACCCCGTGCTCAGCCCAGCACACGGCTCTGCTGCCCCGATGTCCCCGGAGGTGGTG
GTCCCAGCGATGGTGGTTCCAGTGAGGAGGATGAAGGTTTGGCTGAATGTGGAGGCTGCGAAGTGCAGGACTTGAAAGAAAGGAGCGGTTTTAT
TTGCTGCTGTGCAGCAGTAGGAAAAAATACCTTTTGTTGGTATAAAAATGTGCAAGATGCACGGCGTGGTTTCTTATTTTTACTACAGATGCGT
TTTACAACGTTTTGCAAAGTCAATAAATATTTTCTATGGTA
```

>gi|312283675:112-411 Gallus gallus G0/G1switch 2 (G0S2), mRNA
ATGGAAACCATGCACGAGCTGATCCCCTTCGCCAAAGAGATGCTCAGCCAGAAGCCCAACAGGAAGATGGTGAAGCTGTACGTGCTGGGCAGCG
TGCTGGCGTTCTTCGGCGTGGTCATCGGTCTGGTGGAGGCAGTGTGCAGCCCTTTCACCTCCGAAGGGAATATAGAGGAGGAGAAGAGACCGAG
CCCATCTCGAGAGCCAGCGCTTCCTCGGAAGCGGGAGGATTTGGTGTTGGAGCAGAGCAAGAAGTCGTCAGCGGTGCAGCGGGGGGTGGTGACC
AGGCAGCATGCATCCTAA

Protein
>gi|300116722|ref|NP_001177853.1| G0/G1switch 2 [Gallus gallus]
METMHELIPFAKEMLSQKPNRKMVKLYVLGSVLAFFGVVIGLVEAVCSPFTSEGNIEEEKRPSPSREPALPRKREDLVLEQSKKSSAVQRGVVT
RQHAS

Japenese rice fish (Oryzias latipes) G0S2

DNA
>gi|765120492|ref|XM_004068910.2| PREDICTED: Oryzias latipes G0/G1 switch 2 (g0s2), mRNA
TTTAAAGAGCTTTCTGGAAACAGACGTGTGAGAGCTGCTGGATCTTGATGGAGGATTAGTCCACACAGCGCCTGCCAAGAAGTGGGTCAAAGGT
TCTGGCTCAGAGGCCGCAGCGTCTGATAAATATGCAGTGTGGGGGTGTGGGAGGGCAGAGCAGGGCTATAAAGGAGCTCAGCCTCTGTAACAGG
ACTGTTGTGTCCTCAGATTGGATCAGAGGAGCGGACTGAAAACAGGAACCAGCACCAGATTCTCTCACAAGCAGCTGAAGACATGGACAGCATG
CAGGAGCTGATCCCATTTGCCAAAGAGATGCTGAGACAGAAACCGGGCCGGGGCCTGCTGAAGATCTACCTGCTGGGCTCTGTGCTGGCGGTTC
TGGGAACAGCCATCAGCCTGGTCCAGACTGTGTGCCGGCCCTTCTCCTCTGGAGATCCGGTGGACCCAGAAATGCTGCTCATGCTGGCCCGGGT
GCGGAATGAGGCTGAATCTGGGACAAAAAACAGTCTGGAGTGGTTCACAGAGGAAGAGGAGGAGGAGGTGGTGCTGGATGAAAACCGCTTTGCC
AAGACCCAGATTCTGCACAGCTCCAAAAGCCACACGTTCAGCCCACGAAACCAAATCAACCGCCTCCATGCGTCCTAGTAACAGCTGGATCCTA
GCAGGGGACTTTTTATTGGAGGCTGCCCTTCACAGATAGTCTGACTGCAGCGCCCCCTTGTGCTGCAGGGCCGTTCCTGCACCTGTGTCAAGGT
TTAGGCTTACGTGATGCTCAGGCAGAGACGGGGCGGAGCCTCAGTTGTGCACTTTGATGTCCAGCAGGAAAAGATGTATGAGAGCAAACCAGAG
TCCAGGAAAAGGTTCTATTAGGACTTTTCCAGTTCTTTCTGCACTGAAGGTTGAGGTCATCGGTTCTGACTGGAGAACCCTGTTTGCTTTGAAG
GAGCCAAAACTGATTTAGAGTCAGGAGTTCAAGAAAAGCGCTTAAAGGCAGTTTACGTCTGTATCAGACAAATAGTACTTTGTGAGAATGCAAA
CACTGTGATTTTTTTATACCAATAAAGTTTTTATATATATTTAA

>gi|765120492:271-642 PREDICTED: Oryzias latipes G0/G1 switch 2 (g0s2), mRNA
ATGGACAGCATGCAGGAGCTGATCCCATTTGCCAAAGAGATGCTGAGACAGAAACCGGGCCGGGGCCTGCTGAAGATCTACCTGCTGGGCTCTG
TGCTGGCGGTTCTGGGAACAGCCATCAGCCTGGTCCAGACTGTGTGCCGGCCCTTCTCCTCTGGAGATCCGGTGGACCCAGAAATGCTGCTCAT
GCTGGCCCGGGTGCGGAATGAGGCTGAATCTGGGACAAAAAACAGTCTGGAGTGGTTCACAGAGGAAGAGGAGGAGGAGGTGGTGCTGGATGAA
AACCGCTTTGCCAAGACCCAGATTCTGCACAGCTCCAAAAGCCACACGTTCAGCCCACGAAACCAAATCAACCGCCTCCATGCGTCCTAG

Protein
>gi|432858828|ref|XP_004068958.1| PREDICTED: G0/G1 switch protein 2 [Oryzias latipes]
MDSMQELIPFAKEMLRQKPGRGLLKIYLLGSVLAVLGTAISLVQTVCRPFSSGDPVDPEMLLMLARVRNEAESGTKNSLEWFTEEEEEEVVLDE
NRFAKTQILHSSKSHTFSPRNQINRLHAS

Platypus (Ornithorhynchus anatinus) G0S2

DNA
>gi|620971648|ref|XM_001520993.2| PREDICTED: Ornithorhynchus anatinus G0/G1 switch 2 (G0S2),
mRNA
CCCTATCTGCCCACAGCCACCACTCCTGCTGCCCGAAATGGAAACCATGCAAGAGCTGATCCCCTTCGCTAAGGAGATGATGACCCAGAAGCCC
AACGGAAAGCTAGCGAAGCTGTACTTGGTGGGCAGCGTGCTGGCCTTCTTTGGGGTCGTGATCGGCCTGGTGGAGACCGTGTGCAGCCCCTTCA
CCGCCAGAGTCTGGCCGCGGAAGGAAGAGGAGGAGGAGGAGGAGGCGGCAGCGGCAGCCCCGGCCGTGGCTCAGCGAGCCGCCAAGATCCAAGC
ACAGAAGCAGAGGGAGCTCATCTGGGAGAAGGCCAAACTGCAGCCGCAGGCTGTGGGGGGCAGGAGCCTGACCAACAGGCTGCATGCGTCCTAG
GAGCGAGTGGACAGTCCTGCCAGGGAGAGTGGGTGCGAGAGAGAGAGTGTGAGCAAGAGGAAACGGAGGGAGAGAGTGAGGTCAGATAGAGGCG
GGTGGATACCTCTTGGGACGGAGCTGATTTCTGAAGAGGGAAA

>gi|620971648:38-376 PREDICTED: Ornithorhynchus anatinus G0/G1 switch 2 (G0S2), mRNA
ATGGAAACCATGCAAGAGCTGATCCCCTTCGCTAAGGAGATGATGACCCAGAAGCCCAACGGAAAGCTAGCGAAGCTGTACTTGGTGGGCAGCG
TGCTGGCCTTCTTTGGGGTCGTGATCGGCCTGGTGGAGACCGTGTGCAGCCCCTTCACCGCCAGAGTCTGGCCGCGGAAGGAAGAGGAGGAGGA
GGAGGAGGCGGCAGCGGCAGCCCCGGCCGTGGCTCAGCGAGCCGCCAAGATCCAAGCACAGAAGCAGAGGGAGCTCATCTGGGAGAAGGCCAAA
CTGCAGCCGCAGGCTGTGGGGGGCAGGAGCCTGACCAACAGGCTGCATGCGTCCTAG

Protein
>gi|149592500|ref|XP_001521043.1| PREDICTED: G0/G1 switch protein 2 [Ornithorhynchus anatinus]
METMQELIPFAKEMMTQKPNGKLAKLYLVGSVLAFFGVVIGLVETVCSPFTARVWPRKEEEEEEEAAAAAPAVAQRAAKIQAQKQRELIWEKAK
LQPQAVGGRSLTNRLHAS

CLUSTAL 2.1 multiple sequence alignment of G0S2 from different species on protein level

```
Homo                METVQELIPLAKEMMAQKRKGKMVKLYVLGSVLALFGVVLGLMETVCSPFTA-------- 52
Ornithorhynchus     METMQELIPFAKEMMTQKPNGKLAKLYLVGSVLAFFGVVIGLVETVCSPFTAR------- 53
Gallus              METMHELIPFAKEMLSQKPNRKMVKLYVLGSVLAFFGVVIGLVEAVCSPFTS-------- 52
Oryzias             MDSMQELIPFAKEMLRQKPGRGLLKIYLLGSVLAVLGTAISLVQTVCRPFSSGDPVDPEM 60
                    *::::****:****: **    : *:*::*****.:*..:.*:::** **::

Homo                ----ARRLRDQEAAVAELQAALERQALQKQALQ--------EKGKQQDTVLGGRALSNRQ 100
Ornithorhynchus     --VWPRKEEEEEEEAAAAAPAVAQRAAKIQAQKQR--ELIWEKAKLQPQAVGGRSLTNRL 109
Gallus              ----EGNIEEEKRPSPSREPALPRKREDLVLEQ----------SKKSSAVQRG--VVTRQ 96
Oryzias             LLMLARVRNEAESGTKNSLEWFTEEEEEEVVLDENRFAKTQILHSSKSHTFSPRNQINRL 120
                           .: :         . ..  .    .           . .  .       .*

Homo                HAS 103
Ornithorhynchus     HAS 112
Gallus              HAS 99
Oryzias             HAS 123
                    ***
```

Mouse (Mus musculus) ATGL

DNA

```
>gi|58759052|gb|AY894805.1| Mus musculus adipose triglyceride lipase mRNA, complete cds
ATGTTCCCGAGGGAGACCAAGTGGAACATCTCATTCGCTGGCTGCGGCTTCCTCGGGGTCTACCACATTGGCGTGGCCTCCTGCCTCCGTGAGC
ACGCGCCCTTCCTGGTGGCCAACGCCACTCACATCTACGGAGCCTCGGCAGGGGCGCTCACCGCCACAGCGCTGGTCACTGGGGCCTGCCTGGG
TGAAGCAGGTGCCAACATTATTGAGGTGTCCAAGGAGGCCCGGAAGCGGTTCCTGGGTCCTCTGCATCCCTCCTTCAACCTGGTGAAGACCATC
CGTGGCTGTCTACTAAAGACCCTGCCTGCTGATTGCCATGAGCGCGCCAATGGACGCCTGGGCATCTCCCTGACTCGTGTTTCAGACGGAGAGA
ACGTCATCATATCCCACTTTAGCTCCAAGGATGAGCTCATCCAGGCCAATGTCTGCAGCACATTTATCCCGGTGTACTGTGGCCTCATTCCTCC
TACCCTCCAAGGGGTGCGCTATGTGGATGGCGGCATTTCAGACAACTTGCCACTTTATGAGCTGAAGAATACCATCACAGTGTCCCCATTCTCA
GGCGAGAGTGACATCTGCCCTCAGGACAGCTCCACCAACATCCACGAGCTTCGCGTCACCAACACCAGCATCCAGTTCAACCTTCGCAATCTCT
ACCGCCTCTCGAAGGCTCTCTTCCCGCCAGAGCCCATGGTCCTCCGAGAGATGTGCAAACAGGGCTACAGAGATGGACTTCGATTCCTTAGGAG
GAATGGCCTACTGAACCAACCCAACCCTTTGCTGGCACTGCCCCCAGTTGTCCCCCAGGAAGAGGATGCAGAGGAAGCTGCTGTGGTGGAGGAG
AGGGCTGGAGAGGAGGATCAATTGCAGCCTTATAGAAAAGATCGAATTCTAGAGCACCTGCCTGCCAGACTCAATGAGGCCCTGCTGGAGGCCT
GTGTGGAACCAAAGGACCTGATGACCACCCTTTCCAACATGCTACCAGTGCGCCTGGCAACGGCCATGATGGTGCCCTATACTCTGCCGCTGGA
GAGTGCAGTGTCCTTCACCATCCGCTTGTTGGAGTGGCTGCCTGATGTCCCTGAAGATATCCGGTGGATGAAAGAGCAGACGGGTAGCATCTGC
CAGTATCTGGTGATGAGGGCCAAGAGGAAATTGGGTGACCATCTGCCTTCCAGACTGTCTGAGCAGGTGGAACTGCGACGTGCCCAGTCTCTGC
CCTCTGTGCCACTGTCTTGCGCCACCTACAGTGAGGCCCTACCCAACTGGGTACGAAACAACCTCTCACTGGGGGACGCGCTGGCCAAGTGGGA
AGAATGCCAGCGTCAGCTACTGCTGGGTCTCTTCTGCACCAATGTGGCCTTCCCGCCGGATGCCTTGCGCATGCGCGCACCTGCCAGCCCCACT
GCCGCAGATCCTGCCACCCCACAGGATCCACCTGGCCTCCCGCCTTGCTGA
```

Protein

```
>gi|58759053|gb|AAW81963.1| adipose triglyceride lipase [Mus musculus]
MFPRETKWNISFAGCGFLGVYHIGVASCLREHAPFLVANATHIYGASAGALTATALVTGACLGEAGANIIEVSKEARKRFLGPLHPSFNLVKTI
RGCLLKTLPADCHERANGRLGISLTRVSDGENVIISHFSSKDELIQANVCSTFIPVYCGLIPPTLQGVRYVDGGISDNLPLYELKNTITVSPFS
GESDICPQDSSTNIHELRVTNTSIQFNLRNLYRLSKALFPPEPMVLREMCKQGYRDGLRFLRRNGLLNQPNPLLALPPVVPQEEDAEEAAVVEE
RAGEEDQLQPYRKDRILEHLPARLNEALLEACVEPKDLMTTLSNMLPVRLATAMMVPYTLPLESAVSFTIRLLEWLPDVPEDIRWMKEQTGSIC
QYLVMRAKRKLGDHLPSRLSEQVELRRAQSLPSVPLSCATYSEALPNWVRNNLSLGDALAKWEECQRQLLLGLFCTNVAFPPDALRMRAPASPT
AADPATPQDPPGLPPC
```

Mouse (Mus musculus) CGI-58

DNA

```
>gb|BC037063.1|:54-1109 Mus musculus abhydrolase domain containing 5, mRNA (cDNA clone
MGC:46776 IMAGE:4985992), complete cds
ATGAAAGCGATGGCGGCGGAGGAGGAGGTGGACTCGGCAGACGCCGGTGGAGGGTCAGGATGGCTGACAGGGTGGCTTCCTACCTGGTGTCCCA
CATCTACATCACACCTTAAAGAAGCTGAAGAGAAAATGTTAAAATGTGTCCCCTGCACTTACAAGAAAGAGCCTGTGCGCATATCCAATGGAAA
CAGAATATGGACGCTGATGTTCTCTCACAACATTTCTAGTAAGACGCCACTTGTCCTCCTTCATGGTTTTGGAGGAGGTCTTGGACTTTGGGCC
CTGAATTTTGAAGATCTAAGCACCGATAGGCCTGTCTATGCCTTTGACCTATTGGGCTTCGGAAGAAGTAGTAGACCTAGGTTTGACAGTGATG
CGGAAGAAGTGGAGAATCAGTTTGTGGAATCCATTGAAGAGTGGAGATGTGCCCTCAGGTTGGACAAAATGATCTTGCTTGGACACAACCTGGG
AGGGTTCTTGGCTGCCGCTTACTCACTGAAGTACCCATCAAGGGTTAGTCACCTCATTTTAGTAGAGCCATGGGGTTTTCCTGAGCGACCAGAT
CTTGCTGATCAAGAGAGACCAATTCCAGTTTGGATCAGGGCCCTAGGGGCAGCATTGACTCCCTTTAACCCCTTGGCTGGCCTCAGGATTGCAG
GACCTTTTGGGTTAAGTCTAGTGCAGCGTTTGAGGCCTGATTTCAAGCGGAAGTACTCCTCTATGTTTGAAGATGACACGGTGACAGAGTACAT
CTACCACTGTAATGTACAAACCCCAAGTGGTGAGACAGCTTTCAAAAACATGACGATTCCTTATGGGTGGGCCAAACGGCCAATGCTTCAGCGG
ATAGGTGGCTTGCATCCTGACATTCCAGTTTCAGTGATCTTTGGAGCCCGATCCTGCATAGATGGCAACTCTGGAACCAGCATCCAGTCACTGC
GACCGAAGTCCTACGTGAAGACAATTGCCATCCTCGGGGCGGGGCATTATGTGTATGCAGATCAGCCAGAAGAATTCAACCAGAAAGTCAAGGA
GATCTGCCACACAGTAGACTGA
```

Protein

```
>gi|22477988|gb|AAH37063.1| Abhydrolase domain containing 5 [Mus musculus]
MKAMAAEEEVDSADAGGGSGWLTGWLPTWCPTSTSHLKEAEEKMLKCVPCTYKKEPVRISNGNRIWTLMFSHNISSKTPLVLLHGFGGGLGLWA
LNFEDLSTDRPVYAFDLLGFGRSSRPRFDSDAEEVENQFVESIEEWRCALRLDKMILLGHNLGGFLAAAYSLKYPSRVSHLILVEPWGFPERPD
```

LADQERPIPVWIRALGAALTPFNPLAGLRIAGPFGLSLVQRLRPDFKRKYSSMFEDDTVTEYIYHCNVQTPSGETAFKNMTIPYGWAKRPMLQR
IGGLHPDIPVSVIFGARSCIDGNSGTSIQSLRPKSYVKTIAILGAGHYVYADQPEEFNQKVKEICHTVD

7.2.2 Template sequences

These DNA sequences were derived by a sequencing of the used constructs. This chapter shows the sequences used in the cloning experiments. For a better understanding the following colour code was used 6xhis smt3 TEV GOI STOP mutation

pSMT-tev_hG0S2

DNA

ATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATG
GCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAGGTGTCCG
ATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGA
CTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCAC
AGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCATGGAAACGGTCCAGGAGCTGATCCCCC
TGGCCAAGGAGATGATGGCCCAGAAGCGCAAGGGGAAGATGGTGAAGCTGTACGTGCTGGGCAGCGTGCTGGCCCTCTTCGGCGTGGTGCTCGG
CCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCGCCAGACGTCTGCGGGACCAGGAGGCAGCCGTGGCGGAGCTGCAGGCCGCCCTGGAGCGA
CAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACACGGTCCTCGGCGGCCGGGCCCTGTCCAACCGGCAGCACGCCTCTT
AGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACT
AGCATAACCCCTTGGGG

Protein

FCLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMD
SLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSMETVQELIPLAKEMMAQKRKGKMVKLYVLGSVLALFGVVLG
LMETVCSPFTAARRLRDQEAAVAELQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*LEHHHHHH*DPAANKARKEAELAAATAEQ*L
A*PLG

pSMT-tev_hG0S2_20-52

DNA

ATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATG
GCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAGGTGTCCG
ATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGA
CTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCAC
AGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCAAGGGGAAGATGGTGAAGCTGTACGTGC
TGGGCAGCGTGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGCAGCCCCTTCACGGCC

Protein

CLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDS
LRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVKLYVLGSVLALFGVVLGLMETVCSPFTA*RRLRDQEQ
PWRSCRPPWSDRLSRSKPCRRKASSRTRSSAAGPCPTGSTPLSSSTTTTTTEIRLLTKPERKLSWLLPPLSNN*HNPLGPLN

7.2.3 Sequencing results

This chapter shows the sequencing files for all obtained constructs from the cloning experiments. The sequences were aligned using ClustalO. The sequences were also translated into the aa sequence. For a better understanding the following colour code was used 6xhis smt3 TEV GOI STOP mutation

a) N- and C-terminal truncations

hG0S2_21-52

DNA

TTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGC
TAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAGGTGTCCGAT
GGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACT
CCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCACAG
AGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCGGGAAGATGGTGAAGCTGTACGTGCTGGGC
AGCGTGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCTAAAGACGTCTGCGGGACCAGGAGGCAG
CCGTGGCGGAGCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACACGGTCCTCGGCGG
CCGGGCCCTGTCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCT
GAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTA
TATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGC
GCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAG
GGTTCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAACTTGATTAGGGTGATGGTTCACGTAGTGG

Protein

CLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDS
LRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSGKMVKLYVLGSVLALFGVVLGLMETVCSPFTA*RRLRDQEAA
VAELQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*

hG0S2_22-52

DNA

TTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGC
TAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAGGTGTCCGAT
GGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACT
CCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCACAG
AGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCAAGATGGTGAAGCTGTACGTGCTGGGCAGC
GTGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCTAAAGACGTCTGCGGGACCAGGAGGCAGCCG
TGGCGGAGCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACACGGTCCTCGGCGGCCG
GGCCCTGTCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAG
TTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATAT
CCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCC
CTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGT
TCCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGAT

Protein

CLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDS
LRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSKMVKLYVLGSVLALFGVVLGLMETVCSPFTA*RRLRDQEAAV
AELQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*

hG0S2_23-52

DNA

ATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATG
GCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAGGTGTCCG
ATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGA
CTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCAC
AGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCATGGTGAAGCTGTACGTGCTGGGCAGCG
TGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCTAAAGACGTCTGCGGGACCAGGAGGCAGCCGT
GGCGGAGCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACACGGTCCTCGGCGGCCGG
GCCCTGTCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGT
TGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATC
CGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCC
TAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTT
CCGATTTAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAG

Protein

CLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDS
LRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSMVKLYVLGSVLALFGVVLGLMETVCSPFTA*RRLRDQEAAVA
ELQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*

hG0S2_24-52

DNA

TTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCT
AGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAGGTGTCCGATG
GATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTC
CTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCACAGA
GAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCGTGAAGCTGTACGTGCTGGGCAGCGTGCTGG
CCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCTAAAGACGTCTGCGGGACCAGGAGGCAGCCGTGGCGGA
GCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACACGGTCCTCGGCGGCCGGGCCCTG
TCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTG
CTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATT
GGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGC
CCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATT
TAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGG

Protein

CLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDS
LRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSVKLYVLGSVLALFGVVLGLMETVCSPFTA*RRLRDQEAAVAE
LQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*

hG0S2_25-52

DNA

ATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATG
GCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAGGTGTCCG
ATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGA
CTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCAC
AGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCAAGCTGTACGTGCTGGGCAGCGTGCTGG
CCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCTAAAGACGTCTGCGGGACCAGGAGGCAGCCGTGGCGGA
GCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACACGGTCCTCGGCGGCCGGGCCCTG
TCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTG
CTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATT
GGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGC
CCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATT
TAGTGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGG

Protein

CLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDS
LRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSKLYVLGSVLALFGVVLGLMETVCSPFTA*RRLRDQEAAVAEL
QAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*

hG0S2_26-52

DNA

ATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATG
GCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAGGTGTCCG
ATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGA
CTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCAC
AGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCCTGTACGTGCTGGGCAGCGTGCTGGCCC
TCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCTAAAGACGTCTGCGGGACCAGGAGGCAGCCGTGGCGGAGCT
GCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACACGGTCCTCGGCGGCCGGGCCCTGTCC
AACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTG
CCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGC
GAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGCTACACTTGCCAGCGCCCTAGCGCCCG
CTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAATCGGGGGCTCCCTTTAGGGTTCCGATTTAG
TGCTTTACGGCACCTCGACCCCAAAAAACTTGATTAGGGTGATGGTTC

Protein

CLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDS
LRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSLYVLGSVLALFGVVLGLMETVCSPFTA*RRLRDQEAAVAELQ
AALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*

hG0S2_20-46

DNA

TAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCA
GCCATATGGCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAA
GGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAG
GAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTG
AGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCAAGGGGAAGATGGTGAAGCT
GTACGTGCTGGGCAGCGTGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGAAGCCCCTTCACGGCCGCCAGACGTCTGCGG
GACCAGGAGGCAGCCGTGGCGGAGCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACA
CGGTCCTCGGCGGCCGGGCCCTGTCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGC
CCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTG
AAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCGC
TACACTTGCCAGCGCCCTAGCGG

Protein

*K*FCLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGK
EMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVKLYVLGSVLALFGVVLGLMETV*SPFTAARRLR

hG0S2_20-45

DNA

TCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGC
AGCCATATGGCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAA
AGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAA
GGAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATT
GAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCAAGGGGAAGATGGTGAAGC
TGTACGTGCTGGGCAGCGTGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTTAGTGCAGCCCCTTCACGGCCTAAAGACGTCTGCG
GGACCAGGAGGCAGCCGTGGCGGAGCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGAC
ACGGTCCTCGGCGGCCGGGCCCTGTCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAG
CCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCT
GAAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTG

Protein

LE*FCLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGK
EMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVKLYVLGSVLALFGVVLGLMET*CSPFTA*RRLR

hG0S2_20-44

DNA

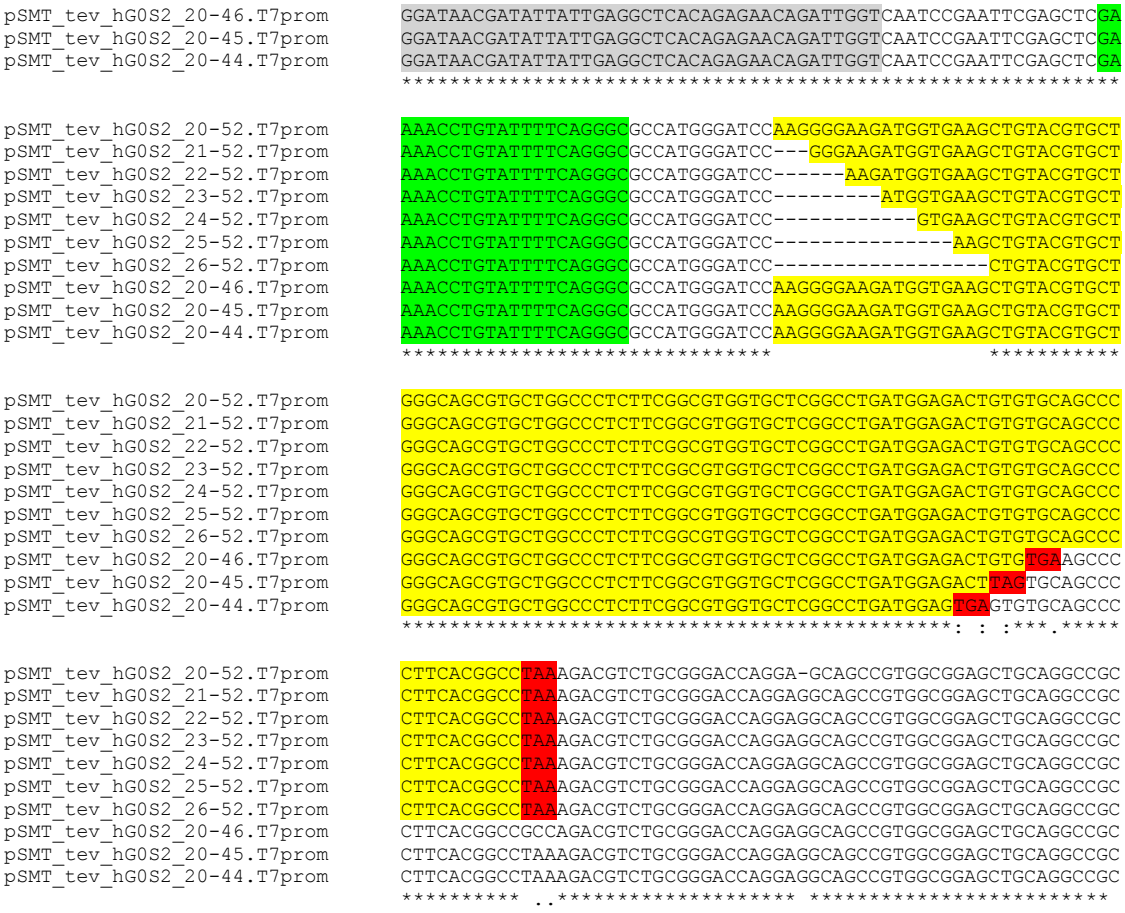
TCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGC
AGCCATATGGCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAA
AGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAA
GGAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATT
GAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCAAGGGGAAGATGGTGAAGC
TGTACGTGCTGGGCAGCGTGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGTGAGTGTGCAGCCCCTTCACGGCCTAAAGACGTCTGCG
GGACCAGGAGGCAGCCGTGGCGGAGCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGAC
ACGGTCCTCGGCGGCCGGGCCCTGTCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAG
CCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCT
GAAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAACGGCGCATTAAGCGC

Protein

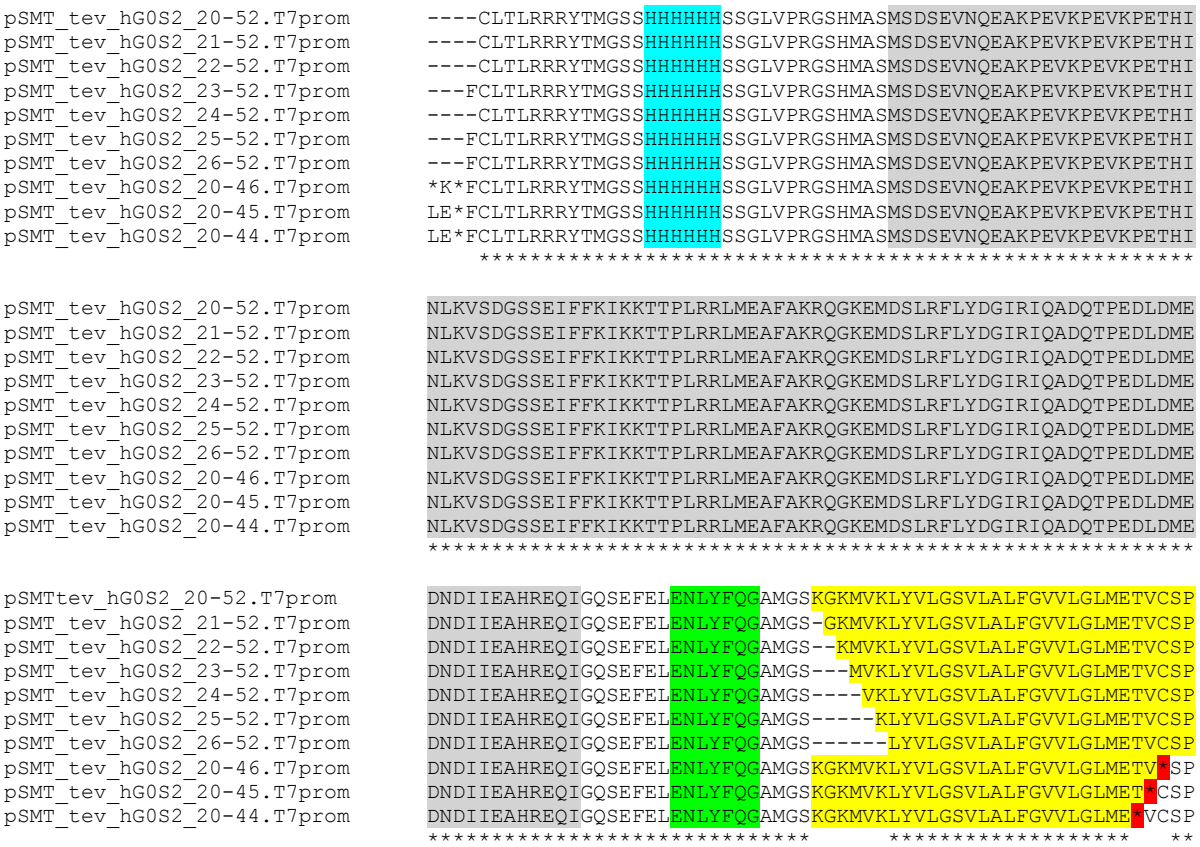
LE*FCLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGK
EMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVKLYVLGSVLALFGVVLGLME*VCSPFTA*RRLR

CLUSTAL O (1.2.1) multiple sequence alignment on nucleotide level:

```
pSMT_tev_hG0S2_20-52.T7prom      -----------TTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCA
pSMT_tev_hG0S2_21-52.T7prom      -----------TTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCA
pSMT_tev_hG0S2_22-52.T7prom      -----------TTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCA
pSMT_tev_hG0S2_23-52.T7prom      ---------ATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCA
pSMT_tev_hG0S2_24-52.T7prom      ------------TTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCA
pSMT_tev_hG0S2_25-52.T7prom      ---------ATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCA
pSMT_tev_hG0S2_26-52.T7prom      ---------ATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCA
pSMT_tev_hG0S2_20-46.T7prom      -TAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCA
pSMT_tev_hG0S2_20-45.T7prom      TCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCA
pSMT_tev_hG0S2_20-44.T7prom      TCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCA
                                                 **********************************************

pSMT_tev_hG0S2_20-52.T7prom      TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTC
pSMT_tev_hG0S2_21-52.T7prom      TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTC
pSMT_tev_hG0S2_22-52.T7prom      TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTC
pSMT_tev_hG0S2_23-52.T7prom      TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTC
pSMT_tev_hG0S2_24-52.T7prom      TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTC
pSMT_tev_hG0S2_25-52.T7prom      TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTC
pSMT_tev_hG0S2_26-52.T7prom      TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTC
pSMT_tev_hG0S2_20-46.T7prom      TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTC
pSMT_tev_hG0S2_20-45.T7prom      TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTC
pSMT_tev_hG0S2_20-44.T7prom      TCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTC
                                 ************************************************************

pSMT_tev_hG0S2_20-52.T7prom      AGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACAT
pSMT_tev_hG0S2_21-52.T7prom      AGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACAT
pSMT_tev_hG0S2_22-52.T7prom      AGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACAT
pSMT_tev_hG0S2_23-52.T7prom      AGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACAT
pSMT_tev_hG0S2_24-52.T7prom      AGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACAT
pSMT_tev_hG0S2_25-52.T7prom      AGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACAT
pSMT_tev_hG0S2_26-52.T7prom      AGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACAT
pSMT_tev_hG0S2_20-46.T7prom      AGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACAT
pSMT_tev_hG0S2_20-45.T7prom      AGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACAT
pSMT_tev_hG0S2_20-44.T7prom      AGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACAT
                                 ************************************************************

pSMT_tev_hG0S2_20-52.T7prom      CAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCC
pSMT_tev_hG0S2_21-52.T7prom      CAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCC
pSMT_tev_hG0S2_22-52.T7prom      CAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCC
pSMT_tev_hG0S2_23-52.T7prom      CAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCC
pSMT_tev_hG0S2_24-52.T7prom      CAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCC
pSMT_tev_hG0S2_25-52.T7prom      CAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCC
pSMT_tev_hG0S2_26-52.T7prom      CAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCC
pSMT_tev_hG0S2_20-46.T7prom      CAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCC
pSMT_tev_hG0S2_20-45.T7prom      CAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCC
pSMT_tev_hG0S2_20-44.T7prom      CAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCC
                                 ************************************************************

pSMT_tev_hG0S2_20-52.T7prom      TTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
pSMT_tev_hG0S2_21-52.T7prom      TTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
pSMT_tev_hG0S2_22-52.T7prom      TTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
pSMT_tev_hG0S2_23-52.T7prom      TTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
pSMT_tev_hG0S2_24-52.T7prom      TTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
pSMT_tev_hG0S2_25-52.T7prom      TTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
pSMT_tev_hG0S2_26-52.T7prom      TTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
pSMT_tev_hG0S2_20-46.T7prom      TTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
pSMT_tev_hG0S2_20-45.T7prom      TTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
pSMT_tev_hG0S2_20-44.T7prom      TTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
                                 ************************************************************

pSMT_tev_hG0S2_20-52.T7prom      ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGA
pSMT_tev_hG0S2_21-52.T7prom      ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGA
pSMT_tev_hG0S2_22-52.T7prom      ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGA
pSMT_tev_hG0S2_23-52.T7prom      ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGA
pSMT_tev_hG0S2_24-52.T7prom      ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGA
pSMT_tev_hG0S2_25-52.T7prom      ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGA
pSMT_tev_hG0S2_26-52.T7prom      ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGA
pSMT_tev_hG0S2_20-46.T7prom      ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGA
pSMT_tev_hG0S2_20-45.T7prom      ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGA
pSMT_tev_hG0S2_20-44.T7prom      ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGA
                                 ************************************************************

pSMT_tev_hG0S2_20-52.T7prom      GGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGA
pSMT_tev_hG0S2_21-52.T7prom      GGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGA
pSMT_tev_hG0S2_22-52.T7prom      GGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGA
pSMT_tev_hG0S2_23-52.T7prom      GGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGA
pSMT_tev_hG0S2_24-52.T7prom      GGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGA
pSMT_tev_hG0S2_25-52.T7prom      GGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGA
pSMT_tev_hG0S2_26-52.T7prom      GGATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGA
```



CLUSTAL O (1.2.1) multiple sequence alignment on protein level:



```
pSMTtev_hG0S2_20-52.T7prom      FTA*RRLRDQEQPWRSCRPPWSDRLSRSKPCRRKASS------------RTRSS------
pSMT_tev_hG0S2_21-52.T7prom     FTA*RRLRDQEAAVAELQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*LEHHH
pSMT_tev_hG0S2_22-52.T7prom     FTA*RRLRDQEAAVAELQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*LEHHH
pSMT_tev_hG0S2_23-52.T7prom     FTA*RRLRDQEAAVAELQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*LEHHH
pSMT_tev_hG0S2_24-52.T7prom     FTA*RRLRDQEAAVAELQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*LEHHH
pSMT_tev_hG0S2_25-52.T7prom     FTA*RRLRDQEAAVAELQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*LELRG
pSMT_tev_hG0S2_26-52.T7prom     FTA*RRLRDQEAAVAELQAALERQALQKQALQEKGKQQDTVLGGRALSNRQHAS*LEHHH
pSMT_tev_hG0S2_20-46.T7prom     FTAARRLR----------------------------------------------------
pSMT_tev_hG0S2_20-45.T7prom     FTA*RRLR----------------------------------------------------
pSMT_tev_hG0S2_20-44.T7prom     FTA*RRLR----------------------------------------------------
                                *** ****
```

b) Lys20-Ala52 mutants

pSMT-tev_hG0S2_20-52_K25D

DNA

```
CTCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCG
GCAGCCATATGGCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTT
AAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGT
AAGGAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTA
TTGAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCAAGGGGAAGATGGTGGA
CCTGTACGTGCTGGGCAGCGTGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCTAAAGACGTCTG
CGGGACCAGGAGGCAGCCGTGGCGGAGCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGG
ACACGGTCCTCGGCGGCCGGGCCCTGTCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAA
AGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTG
CTGAAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGAC
CGCTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTTTCTCGCCACGTTCGCCGGCTTTCCCCGTCAAGCTCTAAAT
CGGGGGCTCCTTTAGGTT
```

Protein

```
L*K*FCLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQG
KEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVDLYVLGSVLALFGVVLGLMETVCSPFTA*RRL
```

pSMT-tev_hG0S2_20-52_Y27G

DNA

```
TCTAGAATAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGC
AGCCATATGGCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAA
AGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAA
GGAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATT
GAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCAAGGGGAAGATGGTGAAGC
TGGGCGTGCTGGGCAGCGTGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGACTGTGTGCAGCCCCTTCACGGCCTAAAGACGTCTGCG
GGACCAGGAGGCAGCCGTGGCGGAGCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGAC
ACGGTCCTCGGCGGCCGGGCCCTGTCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAG
CCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCT
GAAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGCGCAGCGTGACCG
CTACACTTGCCAGCGCCCTAGCGCCCGCTCCTTTCGCTTTCTTCCCTTCCTT
```

Protein

```
K*FCLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKE
MDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVKLGVLGSVLALFGVVLGLMETVCSPFTA*
```

pSMT-tev_hG0S2_20-52_T45G

DNA

```
AACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAGCCATATGGCTAGCATG
TCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAGGTGTCCGATGGATCTT
CAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGGAAATGGACTCCTTAAG
ATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGAGGCTCACAGAGAACAG
ATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCAAGGGGAAGATGGTGAAGCTGTACGTGCTGGGCAGCG
TGCTGGCCCTCTTCGGCGTGGTGCTCGGCCTGATGGAGGGTGTGTGCAGCCCCTTCACGGCCTAAAGACGTCTGCGGGACCAGGAGGCAGCCGT
```

GGCGGAGCTGCAGGCCGCCCTGGAGCGACAGGCTCTCCAGAAGCAAGCCCTGCAGGAGAAAGGCAAGCAGCAGGACACGGTCCTCGGCGGCCGG
GCCCTGTCCAACCGGCAGCACGCCTCTTAGCTCGAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGT
TGGCTGCTGCCACCGCTGAGCAATAACTAGCATAACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATC
CGGATTGGCGAATGGGACGCGCCCTGTAGCGGCGCATTAAGCGCGGCGGGTGTGGTGGTTACGGCGCAGCGTGACCGCTACACTTGCCAGCGCC
CTAGCGCCCGCTCCTTTCGCTTTCTTCCTTCCTTTCTCGCCACGTTCGCGGCTTTCCCGTCAAGCTCTAAATCGGGGGCT

Protein

TLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLR
FLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVKLYVLGSVLALFGVVLGLMEGVCSPFTA*

CLUSTAL O(1.2.1) multiple sequence alignment on protein level:

```
pSMT-tev_hG0S2_20-52          -----CLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETH
pSMT-tev_hG0S2_20-52_K25D     L*K*FCLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETH
pSMT-tev_hG0S2_20-52_Y27G     --K*FCLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETH
pSMT-tev_hG0S2_20-52_T45G     -------TLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETH
                                     **************************************************

pSMT-tev_hG0S2_20-52          INLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDM
pSMT-tev_hG0S2_20-52_K25D     INLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDM
pSMT-tev_hG0S2_20-52_Y27G     INLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDM
pSMT-tev_hG0S2_20-52_T45G     INLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPEDLDM
                              ************************************************************

pSMT-tev_hG0S2_20-52          EDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVKLYVLGSVLALFGVVLGLMETVCS
pSMT-tev_hG0S2_20-52_K25D     EDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVDLYVLGSVLALFGVVLGLMETVCS
pSMT-tev_hG0S2_20-52_Y27G     EDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVKLGVLGSVLALFGVVLGLMETVCS
pSMT-tev_hG0S2_20-52_T45G     EDNDIIEAHREQIGQSEFELENLYFQGAMGSKGKMVKLYVLGSVLALFGVVLGLMEGVCS
                              ************************************ * ***************** ***

pSMT-tev_hG0S2_20-52          PFTA*RRLRD
pSMT-tev_hG0S2_20-52_K25D     PFTA*RRL--
pSMT-tev_hG0S2_20-52_Y27G     PFTA*-----
pSMT-tev_hG0S2_20-52_T45G     PFTA*-----
                              *****
```

c) G0S2 from other species

pSMT-tev_cG0S2

DNA

TCTAGAATATTTTGTTTACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCGCGCGGCAG
CCATATGGCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCAATTTAAAG
GTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACAGGGTAAGG
AAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGATATTATTGA
GGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCATGGAAACGATGCACGAACTC
ATTCCGTTTGCGAAAGAGATGCTGTCCCAGAAACCGAATCGCAAGATGGTGAAACTGTATGTGTTGGGCTCTGTACTGGCCTTCTTCGGTGTCG
TGATTGGCCTGGTTGAAGCGGTTTGCTCGCCCTTTACCAGCGAAGGGAACATCGAAGAAGAGAAACGCCCTTCACCGAGTCGTGAACCGGCTTT
ACCACGCAAACGGGAGGATCTGGTCCTTGAACAGAGCAAGAAAAGCAGTGCCGTACAACGTGGTGTTGTGACTCGTCAGCATGCATCGTAACTC
GAGCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAGCAATAACTAGCAT
AACCCCTTGGGGCCTCTAAACGGGTCTTGAGGGGTTTTTTGCTGAAAGGAGGAACTATATCCGGATTGGCGAATGGGACGCGCCCTGTAGCGGC
GCATTAAGCGCGGCGGGT

Protein

*NILFTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKE
MDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSMETMHELIPFAKEMLSQKPNRKMVKLYVLGSVLAFFGVV
IGLVEAVCSPFTSEGNIEEEKRPSPSREPALPRKREDLVLEQSKKSSAVQRGVVTRQHAS*LEHHHHHH*DPAANKARKEAELAAATAEQ*LA*
PLGASKRVLRGFLLKGGTISGLANGTRPVAAH*ARR

pSMT-tev_jrfG0S2

DNA

AGGAATCCCTCTAGATATTTTGTTTACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCCTGGTGCCG
CGCGGCAGCCATATGGCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGACTCACATCA
ATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCTAAAAGACA
GGGTAAGGAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGGATAACGAT
ATTATTGAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCATGGATAGCATGC
AAGAGCTGATTCCGTTTGCCAAAGAGATGCTGCGTCAGAAACCAGGACGTGGCTTGCTGAAAATCTATCTTCTTGGGAGTGTACTGGCAGTGTT

AGGCACAGCGATTTCCTTGGTCCAAACTGTCTGCCGTCCTTTCTCCAGTGGTGATCCGGTTGATCCCGAAATGCTGCTCATGTTAGCTCGCGTT
CGGAATGAAGCGGAAAGCGGTACGAAGAACTCTCTGGAATGGTTCACCGAAGAGGAAGAAGAGGAAGTGGTGCTGGACGAAAACCGCTTTGCCA
AAACCCAGATTCTGCATTCGAGCAAATCACACACCTTTTCTCCGCGCAACCAGATCAATCGCCTCCATGCGTCGTAACTCGAGACCTGCTTTTG
CTCGCTTGGATCCATGGATAGCATGCAAGAGCTGATTCCGTTTGCCAAAGAGATGCTGCGTCAGAAACCAGGACGTGGCTTGCTGAAAATCTAT
CTTCTTGGGAGTGTACTGGCAGTGTTAGGCACAGCGATTTCCTTGGTCCAAACTGTCTGCCGTCCTTTCTCCAGT

Protein

GIPLDILFTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQ
GKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSMDSMQELIPFAKEMLRQKPGRGLLKIYLLGSVLAVL
GTAISLVQTVCRPFSSGDPVDPEMLLMLARVRNEAESGTKNSLEWFTEEEEEEVVLDENRFAKTQILHSSKSHTFSPRNQINRLHAS*LETCFC
SLGSMDSMQELIPFAKEMLRQKPGRGLLKIYLLGSVLAVLGTAISLVQTVCRPFSS

pSMT-tev_pG0S2

DNA

CCCGACCAACTTCCCTCTAGAATAATTTTGTTTACTTTAAGAAGGAGATATACCATGGGCAGCAGCCATCATCATCATCATCACAGCAGCGGCC
TGGTGCCGCGCGGCAGCCATATGGCTAGCATGTCGGACTCAGAAGTCAATCAAGAAGCTAAGCCAGAGGTCAAGCCAGAAGTCAAGCCTGAGAC
TCACATCAATTTAAAGGTGTCCGATGGATCTTCAGAGATCTTCTTCAAGATCAAAAAGACCACTCCTTTAAGAAGGCTGATGGAAGCGTTCGCT
AAAAGACAGGGTAAGGAAATGGACTCCTTAAGATTCTTGTACGACGGTATTAGAATTCAAGCTGATCAGACCCCTGAAGATTTGGACATGGAGG
ATAACGATATTATTGAGGCTCACAGAGAACAGATTGGTCAATCCGAATTCGAGCTCGAAAACCTGTATTTTCAGGGCGCCATGGGATCCATGGA
AACCATGCAGGAACTGATTCCGTTTGCGAAAGAGATGATGACGCAGAAACCGAACGGCAAGTTGGCGAAGCTGTATCTGGTAGGGAGCGTGTTA
GCCTTCTTTGGTGTCGTCATTGGCCTTGTGGAAACTGTGTGCAGTCCCTTCACAGCTCGCGTATGGCCTCGGAAAGAGGAAGAAGAGGAAGAAG
AAGCAGCGGCTGCAGCACCAGCCGTTGCGCAACGTGCAGCCAAAATCCAGGCGCAGAAACAGCGCGAACTGATCTGGGAGAAAGCCAAACTGCA
ACCGCAAGCGGTTGGAGGTCGTTCGCTGACCAATCGCCTCCATGCTTCCTAACTCGAGCACCACCACCACCACCAACTGAGATCCGGCCTGCTA
ACAAACCCGAAAGGAAGCTGAGTTGGCTGCTGCCACCGCTGAAGCAATAACTAGCATAACCCCCATGGGGCCTCTAAACGGGTCTTGGAGGGGG
TTTTTTTGCCTGAAATAGGAGGAAACCTATA

Protein

PDQLPSRIILFTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVKPETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFA
KRQGKEMDSLRFLYDGIRIQADQTPEDLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSMETMQELIPFAKEMMTQKPNGKLAKLYLVGSVL
AFFGVVIGLVETVCSPFTARVWPRKEEEEEEEAAAAAPAVAQRAAKIQAQKQRELIWEKAKLQPQAVGGRSLTNRLHAS*LEHHHHHQLRSGLL
TNPKGS*VGCCHR*SNN*HNPHGASKRVLEGVFLPEIGGN

CLUSTAL O (1.2.1) multiple sequence alignment on protein level:

```
pSMT-tev_hG0S2        --------FCLTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVK
pSMT-tev_cG0S2        ------*NILFTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVK
pSMT-tev_jrfG0S2      ---GIPLDILFTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVK
pSMT-tev_pG0S2        PDQLPSRIILFTLRRRYTMGSSHHHHHHSSGLVPRGSHMASMSDSEVNQEAKPEVKPEVK
                                ********************************************

pSMT-tev_hG0S2        PETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPE
pSMT-tev_cG0S2        PETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPE
pSMT-tev_jrfG0S2      PETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPE
pSMT-tev_pG0S2        PETHINLKVSDGSSEIFFKIKKTTPLRRLMEAFAKRQGKEMDSLRFLYDGIRIQADQTPE
                      ************************************************************

pSMT-tev_hG0S2        DLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSMETVQELIPLAKEMMAQKRKGKMVK
pSMT-tev_cG0S2        DLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSMETMHELIPFAKEMLSQKPNRKMVK
pSMT-tev_jrfG0S2      DLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSMDSMQELIPFAKEMLRQKPGRGLLK
pSMT-tev_pG0S2        DLDMEDNDIIEAHREQIGQSEFELENLYFQGAMGSMETMQELIPFAKEMMTQKPNGKLAK
                      ***********************************    **** ****  **      *

pSMT-tev_hG0S2        LYVLGSVLALFGVVLGLMETVCSPFTAARRLRDQEAAVAELQAALE--------RQALQK
pSMT-tev_cG0S2        LYVLGSVLAFFGVVIGLVEAVCSPFTSEGNIEEEKRPSPSREPALPRKREDLVLEQS---
pSMT-tev_jrfG0S2      IYLLGSVLAVLGTAISLVQTVCRPFSSGDPVDPEMLL-MLARVRNEAESGTKNSLEWFTE
pSMT-tev_pG0S2        LYLVGSVLAFFGVVIGLVETVCSPFTARVWPRKEEEEEEEAAAAAPAVAQRAAKIQAQKQ
                        *  *****  *    *   ** **

pSMT-tev_hG0S2        ---------QALQEKGKQQDTVLGGRALSNRQHAS*LEHHHHHH*DPAANKARKEAELAA
pSMT-tev_cG0S2        ------------------KKSSAVQRGVVTRQHAS*LEHHHHHH*DPAANKARKEAELAA
pSMT-tev_jrfG0S2      EEEEEVVLDENRFAKTQILHSSKSHTFSPRNQINRLHAS*LETCFCSLGSMDSMQELIPF
pSMT-tev_pG0S2        RELIWEKAKLQPQAVGGRSLTNRLHAS*LEHHHHHQLRSGLLTNPKGS*VGC-------C
```