



**TECHNISCHE UNIVERSITÄT MÜNCHEN**

Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt

Lehrstuhl für Molekulare Ernährungsmedizin

**A novel SP1/SP3 dependent intronic enhancer governing transcription of  
the UCP3 gene**

Christoph Hoffmann

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzende: Univ.-Prof. Dr. A. Schnieke

Prüfer der Dissertation:

1. Univ.-Prof. Dr. M. Klingenspor
2. Univ.-Prof. Dr. B. Küster

Die Dissertation wurde am 06.05.2014 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 17.10.2014 angenommen.

## A. Table of contents

A.	Table of contents.....	I
B.	List of figures, tables and appendices .....	VI
C.	Abbreviations .....	VII
1.	Introduction.....	1
1.1.	Macronutrient metabolism, energy balance and obesity.....	1
1.2.	Brown adipose tissue, skeletal muscle and UCP3 .....	1
1.3.	Brown adipose tissue – function and recruitment.....	2
1.4.	The uncoupling protein family .....	2
1.4.1.	UCP3 .....	4
1.4.1.1.	UCP3 function.....	4
1.4.1.2.	Expression pattern and physiological regulation of UCP3 expression .....	7
1.4.1.3.	Molecular regulation of UCP3 expression.....	7
1.4.2.	UCP1 .....	9
1.4.3.	UCP2 .....	9
1.4.4.	Transcriptional regulation of brown adipose tissue and skeletal muscle determination and differentiation .....	9
1.4.5.	White and Brite adipose tissue .....	10
1.5.	A naturally occurring mutation leads to BAT specific UCP3 deficiency in <i>Phodopus sungorus</i> 11	
1.5.1.	The IVS1+1505 G/A polymorphism .....	12
1.6.	Goals of the thesis .....	13
2.	Material and Methods.....	14
2.1.	Culturing and handling of bacteria.....	14
2.1.1.	Bacterial transformation .....	14
2.1.2.	Plasmid amplification, extraction and quantification .....	14
2.1.3.	Long term storage of plasmid carrying bacteria .....	15
2.2.	Cloning.....	15
2.2.1.	Standard techniques of cloning.....	15
2.2.1.1.	Agarose gel electrophoresis and DNA recovery .....	15
2.2.1.2.	PCR amplification for cloning .....	16
2.2.1.3.	Restriction enzyme digest .....	16
2.2.1.4.	Phosphorylation of DNA.....	17
2.2.1.5.	Dephosphorylation of DNA .....	17
2.2.1.6.	Blunting of DNA .....	17
2.2.1.7.	Ligation reaction.....	18
2.2.1.8.	Validation of ligation products .....	18
2.2.2.	Viral vectors.....	19
2.2.2.1.	miRNA expression.....	19

2.2.2.2.	Tagging of proteins .....	19
2.2.2.3.	Overexpression constructs .....	19
2.2.3.	Reporter gene constructs .....	19
2.2.3.1.	Generation of pGL3-basic-GLuc and pGL3-CMV-PLuc.....	20
2.2.3.2.	Generation of deletions in the Phodopus sungorus reporter gene vector .....	20
2.2.3.3.	Site directed mutagenesis in the Phodopus sungorus reporter gene vector.....	20
2.2.3.4.	Generation of reporter gene constructs containing different promoter sizes .....	20
2.3.	Protein extraction and quantification .....	20
2.3.1.	RIPA protein extract .....	21
2.3.2.	BCA protein quantification .....	21
2.3.3.	Boiling SDS extraction.....	22
2.3.4.	Protein precipitation .....	22
2.4.	Western Blot.....	22
2.4.1.	SDS-PAGE.....	22
2.4.2.	Protein transfer .....	23
2.4.2.1.	Antibody incubation and detection.....	24
2.4.3.	Coomassie stain.....	25
2.4.4.	Silver stain .....	25
2.5.	Cell culture.....	25
2.5.1.	General workflow .....	26
2.5.2.	Passaging cultured cells.....	26
2.5.3.	Counting cultured cells.....	27
2.5.4.	Thawing, proliferation and freezing of cultured cells .....	27
2.5.5.	HIB1b cells .....	27
2.5.6.	C2C12 cells.....	28
2.5.7.	iBPAs.....	28
2.5.8.	Oil Red O Stain.....	29
2.5.9.	Collagen coating of culture plates and dishes.....	29
2.6.	Transfection.....	30
2.6.1.	Calcium phosphate transfection .....	30
2.6.2.	Liposome mediated transfection .....	30
2.6.3.	Nucleofection .....	31
2.6.4.	Viral transduction .....	31
2.6.4.1.	PlatE cells.....	32
2.6.4.2.	Production of retroviral particles .....	33
2.6.4.3.	Infection.....	33
2.6.4.4.	Selection .....	33
2.7.	EMSA .....	33
2.7.1.	Nuclear protein extraction .....	33
2.7.2.	Probe preparation .....	34
2.7.3.	Protein-DNA incubation .....	35

2.8.	Reporter gene assay .....	36
2.9.	Semi-quantitative and quantitative real time PCR.....	37
2.9.1.	RNA extraction und quantification.....	37
2.9.1.1.	RNA quality control .....	37
2.9.1.2.	cDNA Synthesis.....	38
2.9.2.	Primer design for qPCR.....	38
2.9.3.	qPCR .....	38
2.10.	DNA affinity chromatography .....	39
2.10.1.	Resin Preparation .....	39
2.10.2.	EMSA optimisation of binding conditions .....	40
2.10.3.	Oligonucleotide trapping.....	40
2.10.4.	Heparin affinity chromatography.....	40
2.10.5.	Preparative EMSA.....	41
2.10.6.	Magnetic bead purification using biotin/avidin interaction .....	41
2.10.7.	Promoter-trapping .....	42
2.10.7.1.	Probe preparation .....	42
2.10.7.2.	Purification .....	42
2.11.	Candidate identification via mass spectrometry.....	43
2.12.	Chromatin immunoprecipitation.....	43
2.12.1.	Primer selection.....	43
2.12.2.	Sample preparation and crosslinking .....	44
2.12.3.	Chromatin preparation and shearing.....	44
2.12.3.1.	Sonification.....	44
2.12.3.2.	Enzymatic digestion.....	44
2.12.3.3.	Optimisation of chromatin fragmentation .....	45
2.12.4.	Readout, quantification, controls.....	45
2.13.	Statistical analysis.....	46
2.14.	Bioinformatics: Genomatix.....	46
3.	Results .....	47
3.1.	From <i>Phodopus sungorus</i> to cell culture.....	47
3.1.1.	Morphological comparison of cell lines .....	47
3.1.2.	Allele specific differences in protein-DNA interaction.....	49
3.1.3.	EMSA complexes are formed with nuclear extract from all cell lines.....	50
3.1.4.	Transcription factor expression in different cell lines.....	51
3.1.5.	Reporter gene activity – allele specific difference .....	53
3.1.6.	Reporter gene activity – agonist screen .....	54
3.2.	Candidate finding: DNA affinity chromatography.....	56
3.2.1.	EMSA optimization .....	56
3.2.2.	Oligonucleotide trapping.....	57
3.2.3.	Preparative EMSA.....	58
3.2.4.	SDS-PAGE and Silver-Stain.....	59

3.2.5.	Mass spectrometry – candidate list .....	60
3.2.6.	Heparin affinity chromatography .....	65
3.3.	Candidate finding: Bioinformatics .....	66
3.3.1.	Bioinformatics and consensus motifs.....	66
3.4.	Candidate Validation .....	67
3.4.1.	Validation of candidates from bioinformatics using EMSA cold competition.....	67
3.4.2.	Transcript of SP1, SP2 and SP3, but not of CDX proteins, is present in all cell lines .....	70
3.4.3.	SP1 and SP3 bind the IVS1+1505G probe in EMSA .....	70
3.4.4.	SP1 and SP3 are enriched in heparin affinity chromatography eluates.....	71
3.4.5.	RNAi mediated knockdown and overexpression .....	72
3.4.5.1.	Reduction of SP1/3 abundance by virus delivered miRNAs .....	73
3.4.5.2.	Overexpression of human SP1 and SP3.....	74
3.5.	An intronic DR1 element is dependent on the IVS1+1505G element.....	75
3.5.1.	Targeted mutagenesis reveals importance of both intronic elements.....	75
3.5.2.	Chemical inhibition of SP1/3 binding suppresses the effects of PPAR-agonists.....	76
3.5.2.1.	Reporter gene assays .....	76
3.5.2.2.	Endogenous UCP3 expression .....	77
3.6.	DR1/GC-Box modules are present in several mammalian species .....	78
3.6.1.	Sequence analysis.....	78
3.6.2.	EMSA competition experiments.....	78
3.6.3.	Reporter gene constructs: <i>Mus musculus</i> and <i>Homo sapiens</i> .....	79
3.7.	Deletion screens and data mining uncover additional regulatory sites.....	80
3.7.1.	Data mining reveals binding of MyoD, Myogenin and p300 to the intronic enhancer..	82
3.7.2.	Bioinformatic search for complex modules .....	83
3.8.	DNase I hypersensitive sites supports relevance of the intronic enhancer.....	85
3.9.	Discovery of an additional upstream DR1 element binding PPAR $\gamma$ .....	86
3.10.	Tissue specific regulation of UCP3 expression .....	87
3.10.1.	The putative upstream enhancer elements does not influence reporter activity.....	87
3.10.2.	Synergism of rosiglitazone, all-trans retinoic acid and T3 in C2C12 cells.....	89
3.10.3.	PPAR agonist induced UCP3 expression is sensitive to mithramycin in C2C12 cells.....	90
3.11.	Chromatin Immunoprecipitation .....	90
3.11.1.	Methodical problems: Neither SP1 and SP3 antibodies, nor published positive control antibodies precipitate the UCP3 intron 1 enhancer module .....	92
3.11.2.	ChIP: Overexpression of tagged proteins .....	94
3.12.	Outlook: Enhancer complex purification.....	96
3.13.	Summary of results.....	99
4.	Discussion .....	100
4.1.	Baseline characterisation of the cell lines used .....	100
4.1.1.	All four cell lines exhibit basic features required for research on UCP3 expression ...	100
4.1.2.	Initial experiments question tissue specificity of the protein binding the IVS1+1505 element	101

4.2.	Candidate identification by affinity chromatography .....	102
4.2.1.	Purification of DNA-binding proteins: An optimised strategy.....	103
4.3.	Bioinformatics and sequence analysis – identification of the SP transcription factors .....	105
4.3.1.	The SP/KLF family of transcription factors .....	106
4.3.2.	Validation of SP1/SP3 binding .....	107
4.4.	From a single binding site to a complex enhancer: Additional intronic binding elements. ....	108
4.4.1.	Intronic DR1/GC-Box interdependence and the promoter DR1 element.....	108
4.4.2.	Identification of additional binding elements: A putative NF1/STAT element, and binding of MyoD, myogenin and p300.....	109
4.5.	Relevance of the intronic enhancer in skeletal muscle and heart .....	110
4.5.1.	Tissue specific binding to the four putative elements of the enhancer.....	111
4.5.2.	Different effect of the IVS1+1505A allele in muscle and C2C12 cells .....	112
4.5.3.	Summary: The enhancer in brown adipose tissue, skeletal muscle and heart.....	114
4.6.	Presence of the enhancer module in other species .....	115
4.7.	Regulation of UCP3 expression – interaction of promoter and enhancer .....	116
4.8.	Identification of a putative upstream enhancer .....	117
4.9.	Regulation of UCP3 transcription: A refined model.....	118
5.	Outlook.....	120
6.	Summary – PhD thesis Christoph Hoffmann .....	122
D.	Appendix.....	123
6.1.1.	Appendix 1: Chemicals .....	123
6.1.2.	Appendix 2: Plastic ware .....	125
6.1.3.	Appendix 3: Enzymes.....	126
6.1.4.	Appendix 4: Cell lines and organisms.....	126
6.1.5.	Appendix 5: Antibodies .....	127
6.1.6.	Appendix 6: Kit Systems .....	127
6.1.7.	Appendix 7: Oligonucleotide sequences .....	128
6.1.8.	Appendix 8: Machines and devices .....	132
6.1.9.	Appendix 9: Bioinformatic resources .....	133
6.1.10.	Appendix 10: Vector maps .....	137
6.1.11.	Appendix 11: Software and internet resources .....	141
6.1.12.	Appendix 12: Accession numbers ChIP-seq data .....	141
6.1.13.	Appendix 13: Complete list of proteins identified by mass spectrometry.....	142
E.	Literature.....	145
F.	Acknowledgements .....	154
	Eidesstattliche Erklärung .....	155
	Lebenslauf .....	156

## B. List of figures, tables and appendices

#	page	List of figures
1	3	Mitochondrial substrate and proton flux
2	6	UCP functions and their regulation
3	8	General model of regulation of UCP3 transcription
4	11	Organisation of the <i>Phodopus sungorus</i> UCP3 Gene
5	12	The IVS1505 polymorphism is responsible for the absence of UCP3 mRNA in BAT
6	40	Column affinity chromatography setup
7	48	Morphological comparison of C2C12, HIB1b and the two iBPA cell lines
8	49	Allele-specific complex formation with the IVS1+1505 element in EMSA
9	50	The IVS1+1505G complexes are formed with nuclear extracts from all four cell lines
10	52	mRNA expression of selected regulators of transcription in different cell lines
11	53	Only the IVS1+1505G reporter is responsive to PPAR agonists in all assayed cell lines
12	55	Agonist screen in HIB1b and C2C12
13	55	Only PPAR $\gamma$ agonists induce activity of the UCP3 reporter gene construct in HIB1b cells
14	56	EMSA optimization of dsDNA competitor concentration
15	57	IVS1+1505G binding activity in fractions collected during affinity chromatography
16	58	Preparative EMSA
17	59	Silver stained SDS-PAGE of EMSA eluates
18	61	Filtering candidate lists of proteins identified by mass spectrometry
19	65	Heparin affinity chromatography
20	67	Competitor screen using consensus binding sequences
21	68	Comparison of sequences used in competition experiments
22	69	Expression of candidate proteins
23	71	SP1 and SP3 bind the IVS1+1505G Probe in EMSA
24	71	Heparin affinity chromatography eluates contain SP1 and SP3
25	72	Viral packaging and infection
26	74	Knockdown of SP1 and SP3 reduces reporter gene activity of the IVS1+1505G reporter
27	75	The intronic DR1/GC-Box tandem element is required for PPAR agonist induction
28	76	Mithramycin [...] blocks the effect of PPAR agonists on the IVS1+1505G reporter
29	77	Endogenous UCP3 expression in iBPA-L2 cells is sensitive to mithramycin treatment
30	79	DR1/GC-Box modules are found within the first intron of UCP3 in several species
31	81	Stepwise deletion [...] uncovers an additional region required for UCP3 expression
32	82	Sequence analysis and publicly available ChIP-seq data reveal a MyoD/myogenin binding site directly upstream of the DR1/GC-Box module
33	84	Putative enhancer regions in different species
34	85	DNAseI hypersensitive (DHS) sites in the vicinity of the UCP3 Gene
35	86	The first intron of <i>Phodopus sungorus</i> UCP3 resembles the first 4000 bp of its <i>Mus musculus</i> counterpart
36	88	Neither inclusion of the CoupTF2 nor the upstream PPAR $\gamma$ element can rescue the effect of the IVS1+1505A allele
37	89	In C2C12 cells, the P2000 reporter requires a cocktail [...] of agonists for full activity
38	30	Endogenous UCP3 transcript is responsive to PPAR $\gamma$ - and $\delta$ -agonists in C2C12 cells
39	91	Chromatin shearing and preparation
40	93	[...] positive control antibodies fail to precipitate the IVS1+1505G enhancer
41	95	The Ty1-tagged overexpressed transcription factors bind the IVS1+1505G probe and can be bound by a Ty1-antibody
42	97	Optimization of enhancer complex purification
43	98	Lack of differential binding in enhancer trapping
44	104	Proposed optimized strategy for affinity purification of transcription factors
45	106	Complex formation pattern for SP1 and SP3
46	119	Refined model of regulation of UCP3 transcription

#	page	List of tables
1	64	Selected proteins identified in mass spectrometry
2	66	Binding motifs identified for the IVS1+1505G element using bioinformatics
3	133	Criteria and models used for bioinformatics
4	134	Description of the binding sites identified by MatInspector
5	135	Putative binding sites identified by MatInspector
6	136	Putative enhancer regions identified by ModellInspector

#	page	List of appendices
1	123	Chemicals
2	125	Plastic ware
3	126	Enzymes
4	126	Cell lines and organisms
5	127	Antibodies
6	127	Kit Systems
7	128	Oligonucleotide sequences
8	132	Machines and devices
9	133	Bioinformatic resources
10	137	Vector maps
11	141	Internet resources
12	141	Accession numbers ChIP-seq data
13	142	Complete list of proteins identified by mass spectrometry

## C. Abbreviations

4-HNE	4-hydroxy-nonanal
9cRA	9-cis retinoic acid
AA	amino acids
ANOVA	analysis of variance
ANP	Atrial natriuretic peptide
ANT	adenine nucleotide translocator
ARE	antioxidant response element
ATRA	all-trans retinoic acid
$\beta$ -AR	beta adrenoreceptor
BAT	brown adipose tissue
BMP	bone morphogenetic protein
bp	base pairs
CACT	carnitine-acylcarnitine translocase
cDNA	complementary DNA generated by reverse transcription
CDS	coding sequence
C/EBP	CCAAT-enhancer-binding protein
CHAPS	3-[(3-cholamidopropyl)dimethylammonio]-1-propanesulfonate
ChIP	chromatin immunoprecipitation
CMV	cytomegalovirus
CoA	Coenzyme A
COUP	chicken ovalbumin upstream promoter
CPT	carnitine palmitoyl transferase
DHFR	dihydrofolate reductase
DHS	DNase I hypersensitivity
DIO2	type II iodothyronine deiodinase
DMSO	dimethylsulphoxid
dNTP	desoxyribonucleotid triphosphate
DR	direct repeat
EMSA	electrophoretic mobility shift assay
EF1 $\alpha$	Elongation factor 1 $\alpha$
FABP	fatty acid binding protein



FFA	free fatty acid
FGF	fibroblast growth factor
gDNA	genomix DNA
HIB1b	Hibernoma 1b
iBPA	immortalized brown preadipocytes
IMM	inner mitochondrial membrane
IVS	intervening sequence
kDa	kiloDalton
KLF	Krüppel-like factor
MACP	mitochondrial anion carrier proteins
MAPK	mitogen-activated protein kinases
MC4R	melanocortin 4 receptor
MIM	mitochondrial inner membrane
miRNA	micro RNA
MMLV	moloney murine leukemia virus
MS	mass spectrometry
NA	noradrenaline
NF1	nuclear factor 1
Nrf2	nuclear response factor 2
nt	nucleotides
oligo	oligonucleotide
ORF	open reading frame
PAGE	polyacrylamide gelelectrophoresis
PBS	phosphate buffered saline
PC	pyruvate carrier
PGC	PPAR $\gamma$ coactivator
PGK	phosphoglycerat kinase
PLB	passive lysis buffer
PPAR	peroxisome proliferator activated receptor
PPRE	PPAR responsive element
qPCR	quantitative PCR
RAR	retinoid acid receptor
RNAi	RNA interference
ROS	reactive oxygen species
RXR	retinoid X receptor
SDS	sodium dodecyl sulphate
shRNA	short hairpin RNA
SKTM	skeletal muscle
SP	specificity protein
STAT	signal transducer and activator of transcription
TAE	tris acetic acid EDTA
TBE	tris boric acid EDTA
TBS	tris buffered saline
TBST	tris buffered saline supplemented with Tween-20
TEMED	tetramethylethylenediamine
TF	transcription factor
TFBS	transcription factor binding site
TR	thyroid receptor
Tris	tris(hydroxymethyl)aminomethane
TSS	transcriptional start site
UCP	Uncoupling Protein
VLDL	very low density lipoprotein
WAT	white adipose tissue

# 1. Introduction

## 1.1. Macronutrient metabolism, energy balance and obesity

Food is composed of several different components that can be grouped into the following categories: macronutrients, micronutrients, dietary fibre and water. Macronutrients are needed as a source of energy, as building material and play a role in signalling.

Inadequate food intake can lead to severe disease conditions. While in modern western societies an undersupply of nutrients rarely is a problem, overnutrition has become a serious health issue. In cases of long term positive energy balance, fat tissue is built up. The positive energy balance originates from excessive macronutrient ingestion and a high content of lipids and carbohydrates in our diet, combined with a low ingestion of fibre and a sedentary lifestyle lacking activity. Ultimately this leads to overweight and obesity, which then increases the risk for diabetes and cardiovascular disease. Obesity, diabetes and cardiovascular disease are often combined by the term metabolic syndrome. Besides inducing obesity, a high load of carbohydrates and lipids also directly increases the risk of diabetes and cardiovascular disease. Furthermore other components of our nutrition, gastrointestinal microbiota, social environment and genetic and epigenetic factors contribute to these phenomena.

The metabolic syndrome has advanced to one of the main health problems in industrialized countries [1,2]. Understanding macronutrient metabolism is a key step to understand, prevent and treat those diseases. Three main issues seem to be responsible for the disease character of the metabolic syndrome: Ectopic and excessive lipid accumulation, abnormal blood chemistry and tissue inflammation. All these three issues act synergistically and promote each other, and in all three fields lipids play a central role. Lipids supply the majority of excess energy. Increased plasma triglyceride, low density lipoproteins and cholesterol are causative for cardiovascular complications. Lipid storage in white adipose tissue (WAT) as well as ectopic storage in other tissues has been identified to play a central role in impairment of glucose clearance, insulin secretion and insulin sensitivity. Thus, one approach to understanding and tackling the metabolic syndrome is to understand and enhance lipid metabolism. Achieving this, one can hope to reduce the deleterious effects of excess lipid intake.

Two tissues are in the focus of this approach: Brown adipose tissue (BAT) and skeletal muscle (SKTM), both tissues with immense lipid consumption capacity. BAT, initially believed to be limited to small mammals, hibernators and new-borns (including human infants[3]), has recently come into focus due to identification of active and functional brown adipose tissue in adult humans [4-6]. In humans, BAT mass and activity negatively correlates with BMI and age [4]. Interestingly, both tissues express uncoupling protein (UCP) 3 [7,8].

## 1.2. Brown adipose tissue, skeletal muscle and UCP3

Initially WAT and BAT were believed to share a common developmental origin, but newer studies demonstrate that BAT in fact is related to SKTM. Both tissues share a common precursor cell that is positive for the muscle transcription factor Myf5 [9]. These precursor cells arise from the dermomyotome, while the precursors of WAT arise from the sclerotome (reviewed in [10]). At first glance, BAT and SKTM seem remarkably different in morphology and function, but in their energy metabolism they share several characteristics. Both tissues are able take up and catabolise high amounts of free fatty acids (FFAs) and glucose, both tissues have a very high mitochondrial content

and both tissues express UCP3. The main reason why we and others are interested in UCP3 is that we believe that these commonalities are functionally connected. Fittingly, in both tissues UCP3 abundance increases in situations of high  $\beta$ -oxidation [11-14], where UCP3 is believed to aid free fatty acid breakdown [15] and protect mitochondrial function [16]. Apart from SKTM and BAT, lower amounts of UCP3 protein can also be found in heart muscle, a third side of lipid consumption.

### **1.3. Brown adipose tissue – function and recruitment**

BAT is morphologically characterized by displaying multilocular fat storage and an increased amount of mitochondria. The tissue shows extensive vascularization and innervation by the sympathetic nervous system. BAT mitochondria are very rich in respiratory chain complexes and UCP1 [17]. UCP1 is required to fulfil the main function of BAT: releasing the energy stored in macronutrients as heat [17] to defend the body temperature in cold environments [18,19]. Both recruitment and activity of BAT are induced by cold exposure, and this is controlled via the sympathetic nervous system by activation of  $\beta_3$ -adrenergic receptors through noradrenaline (NA). Subsequently to binding of NA a  $G_s$ -protein is activated which activates an adenylate cyclase. The following increase in cAMP activates protein kinase A (PKA) and the p38-MAPK pathway [20]. PKA and the p38-mitogen-activated protein kinases (p38-MAPK) pathway lead to increased lipolysis, supplying peroxisome-proliferator activated receptor (PPAR) $\alpha$  ligands. Together with other signals, most notably T3 generated from circulating T4 by intracellular type II iodothyronine deiodinase (DIO2), PPAR $\alpha$  and thyroid receptor (TR) activate mitochondrial biogenesis and UCP1 expression [21-23]. Recently, atrial and bone natriuretic peptide (ANP and BNP) [24,25] have been identified to induce brown fat expansion in response to high fat feeding and exercise. Both act via protein kinase G and converge into the p38-MAPK pathway.

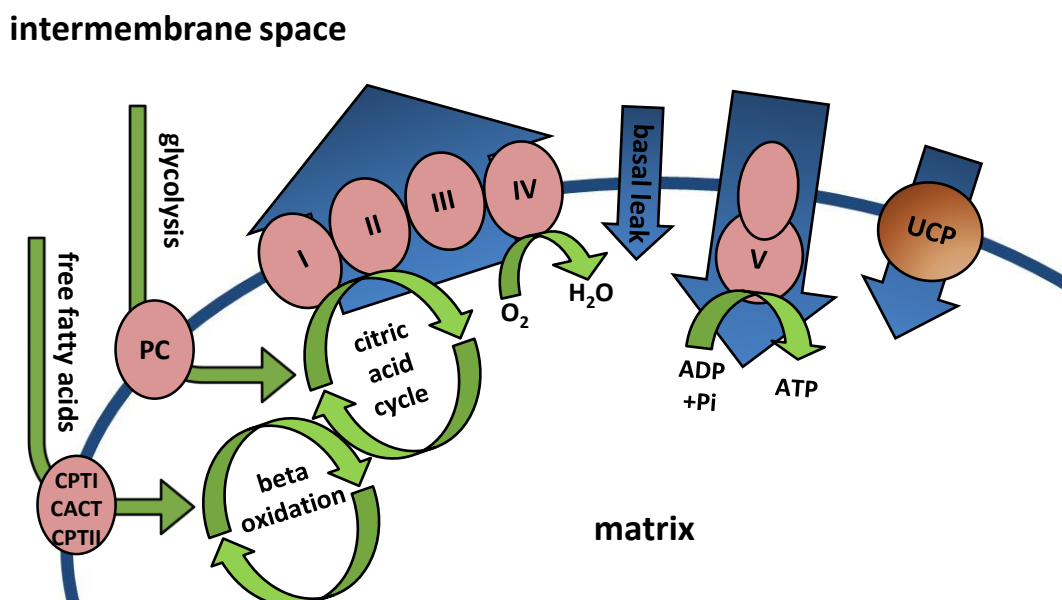
### **1.4. The uncoupling protein family**

The uncoupling proteins belong to the superfamily of mitochondrial anion carrier proteins (MACPs). The family has three members, UCP1, 2 and 3, in mammals. UCP2 and UCP3 are most closely related, sharing 73% protein identity, while UCP1 has 59 and 57% identity to the other UCPs [7,26]. The UCP2 and UCP3 genes are located directly adjacent in the genome (Chromosome 7 in mouse, Chromosome 11 in human), arising from a recent gene duplication. Two other putative UCPs, named UCP4 [27] and brain mitochondrial carrier protein (BMCP) [28], are not considered family members of the mammalian UCPs.

All MACPs are located in the mitochondrial inner membrane (MIM). Structurally they contain six transmembrane helices, possibly originating from a triplication of a roughly 100 aminoacid (AA) two helix domain [29]. All MACPs, including the UCPs, thus have a molecular mass of around 30 kDa. UCP-specific AA-motifs that are absent in non-UCP MACPs have been identified and are suggested to participate in fatty acid binding [30]. FA binding is a common feature of all UCPs that is known to be required for activation of uncoupling activity [31]. Fatty acids *in vivo* are released after  $\beta$ -adrenergic stimulation of lipolysis. A second motif of interest is a nucleotide binding pocket [32], preferably binding GDP and ADP, which is important for inhibition of activity [33]. While of crucial relevance for regulation of UCP activity, GDP and ADP binding activity is not unique for UCPs.

Based on sequence comparison, UCP genes outside of the mammalian class could be identified. UCP homologues can be found in vertebrates, plants, insects, nematodes and fungi, demonstrating ancient evolutionary origin. Sequence comparison of putative UCPs from *Arabidopsis*, *Drosophila*, *Caenorhabditis* and *Dictyostelium* revealed that UCP4 most likely is closest to the ancestral MCAP from which all UCPs originate [34]. Comparison of UCP-homologues of the mammalian UCPs 1-3 in non-mammalian vertebrates suggests that their three progenitors already were present before the divergence of lobe and ray finned fish more than 400 million years before [35].

The name “UCP” stems from a proposed mechanism where UCPs allow protons to cross the MIM [36] without passing the  $F_1F_0$  ATP-synthase. Such a mechanism would “uncouple” respiratory chain activity from ATP generation. Three main mechanisms are discussed. The cofactor model proposes UCPs to have an incomplete proton channel which is not able to confer proton flow over the inner membrane by itself. Upon binding, the carboxyl group of a FFA supplies a binding site for protons, thereby completing the channel [37]. The FFA cycling model proposes UCPs to not directly transfer protons, but rather to flip FFAs from the inner leaflet of the MIM to the outer leaflet [38]. There the carboxyl group gets protonated and can then flip back to release the proton. Lastly, the conformational change method predicts that binding of a purine diphosphate pushes the UCPs into an inactive conformation while fatty acid binding induces an active conformation. Which of the three proposed mechanisms actually is correct is still a matter of debate, but the most recent publication on this discussion favours the cofactor model [39]. One should note that most of the research on the molecular mechanism of uncoupling has been carried out on UCP1. The question whether UCP2 and UCP3 have a physiologically relevant thermogenic uncoupling activity is still a matter of debate and discussed extensively below.



**Figure 1: Mitochondrial substrate and proton flux:** Free fatty acids from lipolysis and pyruvate from glycolysis enter the mitochondrial matrix via the pyruvate carrier (PC) and the carnitine palmitoyl transferase(CPTI) – carnitin acylcarnitin translocase (CACT) – CPTII shuttle, respectively. Both substrates are broken down to acetyl-CoA and fed into the citric acid cycle which fuels the proton pumping activity of the respiratory chain (Complex I to IV). The protons pumped into the intermembrane space can return to the matrix via either the  $F_0F_1$  ATPase (Complex V), the basal proton leak or via the induced proton leak (UCP1-3). Only the path via Complex V yields ATP. The ratio between ATP producing proton flow and proton leak can range from >5:1 (only basal leak) to <1:10 (fully activated UCP1 in BAT).

### 1.4.1. UCP3

UCP3 was identified shortly after UCP2 and 19 years after UCP1 [7]. In human, two different mRNA variants are described which originate from alternative splicing. When translated, those forms would yield a long UCP3 protein of 312 AA, and a short one of only 275 AA which is devoid of the C-terminus. The short form lacks the purine dinucleotide binding motif and one trans-membrane domain. While expression of both isoforms is well established on mRNA level, presence of translated protein has been doubted for the short isoform [40]. Rodent UCP3 is equivalent to the human UCP3 long form.

#### 1.4.1.1. UCP3 function

The function of UCP3 has not been finally resolved yet. Contrary to UCP1, UCP3 is unlikely to confer non shivering thermogenesis [19]. One complication of the research on UCP3 function is that overexpression in non-mammalian systems has been demonstrated to cause artefacts, thereby preventing the use of yeast or comparable simple assay systems [41,42].

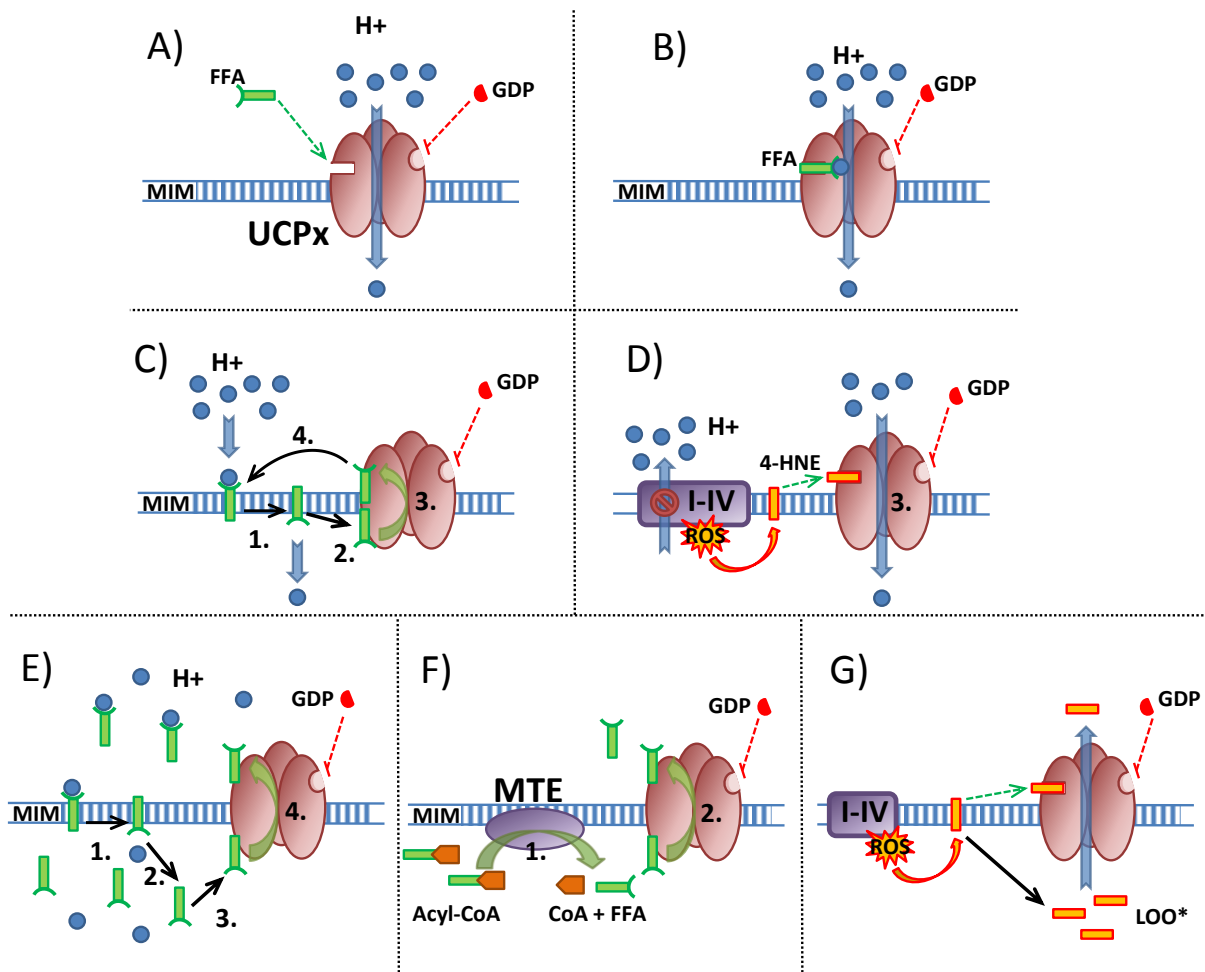
Uncoupling by UCP3 is considered mild uncoupling, and thus may provide a valve to reduce ROS generation [43]. Knockout mice showing a reduced proton leak and increased ATP/ADP ratio support this [44], although some authors fail to reproduce those effects in their animal models [45,46]. The one hypothesis for this divergence is that uncoupling activity is tightly regulated, so that in absence of specific activators no uncoupling can be detected [46]. Furthermore forced expression of UCP3 in mammalian systems seems likely to cause artefacts even in mammalian systems [46]. A role of UCP3 in ROS damage mitigation is supported by a wide variety of evidence. Firstly, UCP3 expression is induced by H<sub>2</sub>O<sub>2</sub> treatment [47]. Secondly, UCP3 activity is activated by 4-hydroxy-nonenal (4-HNE) [48], a molecule generated when lipid membranes get damaged by ROS. 4-HNE is considered a messenger molecule for oxidative stress [49]. When activated, UCP3 reduces membrane potential and thus ROS production and the effect can be inhibited by purine di- and triphosphates [43]. All ROS defence hypotheses assign the same molecular function to UCP3 as for UCP1, uncoupling of respiratory chain action from ATP synthesis, except in a different physiological context. Based on basic physics, uncoupling is always accompanied by heat generation, but due to the low protein amount compared to UCP1, thermogenesis by UCP3 is not relevant.

The second complex of functional hypotheses deals with cellular lipid catabolism and -tolerance. These hypotheses are well in line with physiological regulation of UCP3 by increased FFA levels and the observation that tissues with low  $\beta$ -oxidation capacity show a higher UCP3 expression (see 1.4.1.2, page 7). The physiological effect of UCP3 in situations of high FFA concentrations is still a matter of debate, but several hypotheses are based on UCP3 being a fatty acid exporter. This would fit to the FFA cycling model proposed for UCP1 (see 1.4). As FFAs cannot be metabolised inside mitochondria they would need to be exported, and since no mitochondrial FFA transporter has been described as of yet, UCP3 might fill this vacant role. Himms-Hagen and Harper suggest that UCP3 helps to prevent a Coenzyme A (CoA) shortage in situations of high beta oxidation [15]. Such a shortage is believed to occur when medium chain acyl-CoAs accumulate, which thereby would trap CoA and limit the amount of free CoA. CoA inside the mitochondrial matrix is required for steps in beta-oxidation and citric acid cycle. In a situation of limited free CoA, a mitochondrial thioesterase would release trapped CoA and UCP3 would subsequently export the released free fatty acid into the intermembrane space. Notably, a futile cycle consisting of activation, import, release and re-export of free fatty acids would translocate protons. Furthermore, the ATP consumption by this futile cycle would be indistinguishable from uncoupling when only measuring oxygen consumption.

Later, Schrauwen and Hesselink propose a protective function in the more general situation of lipotoxicity [16]. As in the hypothesis of Himms-Hagen and Harper, in their hypothesis the function of UCP3 would be to export FFAs from the mitochondrial matrix. The main difference is the reason why FFAs get inside the mitochondria: In case of an excessive free fatty acid load within a cell, protonated fatty acids could cross the mitochondrial inner matrix and deprotonate, thereby being trapped. Again, UCP3 could re-export the FFAs to prevent accumulation in the matrix and allow loading them onto CoA. As above, this mechanism is the same as the flip-flop mechanism and would lead to a proton transfer into the matrix.

Beside protons and free fatty acids, other molecules are considered to be transport substrates of UCP3. Export of lipid radicals would protect the mitochondrial matrix proteins and genome from ROS damage. In absence of both inhibition and induction, an outward transport of  $\text{Cl}^-$  has been suggested for UCP3 [50] and demonstrated for UCP1 [51]. Other reports consider pyruvate as a transport substrate, possibly as a mechanism to balance membrane potential,  $\beta$ -oxidation and glycolysis [50]. Recently UCP2, the closest relative to UCP3, has been demonstrated to exchange for  $\text{C}_4$  metabolites for protons and phosphate ions across the MIM [52], thereby also supporting a putative transport function for UCP3. Lastly, a direct participation in  $\text{Ca}^{2+}$  transport has been suggested [53], but a recent study suggests that this effect is indirect via modulation of ATP/ADP ratio [54].

An influence on whole body metabolism and body mass has extensively been evaluated. Evidence from knockout models does not show an influence on weight or fat mass [44], but two naturally occurring mutations, one in human, one in the djungarian hamster (*Phodopus sungorus*), support relevance in body mass regulation. In human, the promoter -55C/T polymorphism influences body fat distribution and obesity in several populations [55,56], although a recent meta-analysis revokes part of the effects [57]. Other polymorphisms exist in the UCP3 gene, so some effects might originate from linkage to a functional single nucleotide polymorphism (SNP). In hamster, the intronic +1505G/A polymorphism [58] leads to absence of UCP3 in BAT and to increased body mass and reduced expression of genes from several core metabolic pathways [59]. Taken together, an influence on body mass and whole body energy metabolism is likely but the effect most likely is subtle, partly compensated for, or dependent on other factors.



**Figure 2: UCP functions and their regulation: A-D) Four models for activation of uncoupling activity:** Depending on UCP, tissue and condition, uncoupling can act in thermogenesis, reduce ROS production or modify ATP/ADP ratio. **E-G) Proposed functions not related to uncoupling.** **A)** GDP and FFAs (free fatty acids) act as allosteric effectors and change conformation of UCP to enable or prevent proton transport. When neither is bound, transport of non-proton molecules is proposed. That this non-ligated situation exists *in vivo* is considered unlikely and is yet to be demonstrated. **B)** FFAs complete an otherwise incomplete tunnel in UCP by providing proton binding sites. **C)** FFAs get protonated in the intermembrane space and flip into the matrix where they deprotonate. UCP re-exports them, thereby completing a cycle. **D)** In cases of high proton gradient the respiratory chain (I-IV) stalls and electrons escape, generating lipid radicals. 4-HNE (4-hydroxynonenal), a product of lipid peroxidation, activates UCP thereby reducing the proton gradient. **E)** In situations of high FFA supply and limited usage or beta oxidation capacity the FFAs accumulate in the matrix where they exert toxic effects. UCP3 is able to transfer them back out. **F)** A high amount of  $\beta$ -oxidation intermediates reduces free CoA levels, thereby constraining  $\beta$ -oxidation and citric acid cycle. MTE (mitochondrial thioesterase) releases CoA and FFA, of which the latter is exported by UCP3. **G)** Under conditions as in D), electrons escape the respiratory chain into the MIM and generate lipid radicals (LOO\*) which can damage mitochondrial DNA and matrix enzymes. UCP2 and 3 export them, reducing the damage. I-IV: respiratory chain; MIM: mitochondrial inner membrane.

#### **1.4.1.2. Expression pattern and physiological regulation of UCP3 expression**

UCP3 protein is primarily present in BAT and SKTM and can be found, although at much lower abundance, in heart. For WAT and kidney [7] presence of low amounts of mRNA has been demonstrated.

Expression of UCP3 is induced in conditions of increased FFA usage. Cold exposure, a situation where brown adipose tissue combusts fatty acids, increases expression of UCP3 in BAT [60]. Upon birth, UCP3 expression is induced by intake of the fat-rich mothers milk [61,62]. Fasting, a condition where the muscle heavily relies on fatty acids while BAT activity is reduced, increases UCP3 expression in SKTM [11] and reduces it in BAT [63]. Acute exercise, which mobilises stored fat, increases UCP3 expression [64] in SKTM. A diet rich in fatty acids increases UCP3 expression in SKTM [13]. The theory that elevated FFAs are the common inducer of UCP3 expression in all those conditions is supported by the induction of muscle UCP3 expression after lipid infusion [14] and induction of UCP3 expression in L6 cell cultures fed oleic acid [65]. This effect can also be seen in adult rat cardiomyocyte where oleic acid showed the strongest induction of five different fatty acids [66]. Besides the sheer amount of FFA, the ability to break down FFAs influences UCP3 expression: For SKTM, UCP3 expression is highly dependent on fibre type with a high expression in glycolytic muscle fibres and a lower expression in oxidative fibre [7,40]. Endurance training, increasing oxidative capacity, leads to a decrease in UCP3 expression [67] and acute reduction of beta-oxidation via CPT1 inhibition (Etomoxir) increases UCP3 expression [68].

#### **1.4.1.3. Molecular regulation of UCP3 expression**

The molecular regulation of UCP3 expression supports a function in lipid metabolism. Key regulators of UCP3 expression are the peroxisome proliferator-activated receptors, PPARs. These proteins are known to bind free fatty acids breakdown products, unsaturated fatty acids, modified fatty acids and free fatty acid derived messenger molecules, like prostaglandins. Being nuclear hormone receptors, they translocate to the nucleus upon ligand binding and activate transcription via PPAR response elements (PPREs). Although artificial specific ligands for all three PPARs ( $\alpha$ ,  $\gamma$  and  $\delta$ ) are commercially available, it is still debated which PPARs activate UCP3 transcription in which tissue.

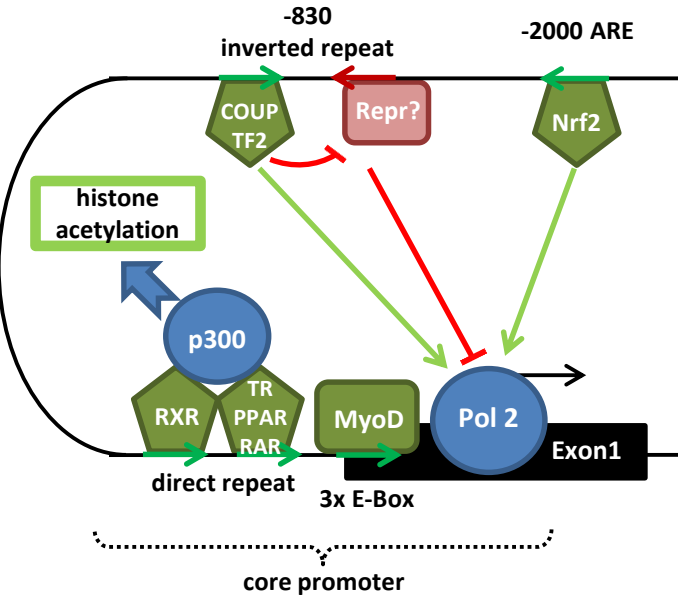
PPAR $\gamma$  action on UCP3 expression is supported by a plethora of reports for BAT [69-72] in both mouse, rat and cell culture. ChIP data [73,74] demonstrate direct binding to the UCP3 gene. PPAR $\gamma$  is known to exist in two splice variants ( $\gamma$ 1,  $\gamma$ 2) in mouse and man, of which  $\gamma$ 2 is considered a key regulator of BAT differentiation, but none of them is exclusively found in BAT. It is not known if a specific PPAR isoform is responsible for the effect on UCP3 transcription. For SKTM, the situation is less clear and involvement of PPAR $\gamma$  is still in discussion. In cell culture, different studies are reporting diametrically opposed results [75-79]. These inconsistencies might stem from different PPAR $\gamma$  agonists being capable of activating different subsets of the PPAR $\gamma$  effect repertoire [80], or from differences in readout strategy. Reporter gene assays using vectors only including the promoter never see a PPAR $\gamma$  effect [78,79]. Experiments assaying endogenous expression yield varying results [65,75,76,81], possibly due to usage of different cell lines and agonists. In vivo, PPAR $\gamma$  agonist effects in SKTM seem to be age dependent. For mice in the suckling state, no PPAR $\gamma$  effects could be found [62], while in adults effects have been demonstrated [82]. Taken together, PPAR $\gamma$  is a key regulator of UCP3 transcription in BAT. For SKTM, the situation seems to be dependent on additional parameters. Little is known about the effects of PPAR $\gamma$  agonists on heart UCP3 expression. Despite several artificial PPAR $\gamma$  ligand being well described, few endogenous PPAR $\gamma$  ligands are known yet. Prostaglandins, in particular 15-deoxy- $\Delta^{12,14}$ -prostaglandin-J2, are prime candidates [83].



PPAR $\alpha$  binds lipolytic products and thereby ties expression of UCP3 to lipolysis [84]. Involvement of PPAR $\alpha$  in BAT is well supported by studies in animal [72] and cell culture experiments [70,71], although one study fails to see effects [69]. In muscle, PPAR $\alpha$  activity on UCP3 expression seems to be limited to the suckling age [62,78] where it relays the stimulating effect of milk intake of UCP3 expression. While a bezafibrate effect on UCP3 expression has been claimed to act via PPAR $\alpha$  [85], bezafibrate has known PPAR $\gamma$  and  $\delta$  activity. Experiments using a highly specific agonist (Wy-14643) or PPAR $\alpha$  knockout neither show an effect in adult animals [69,75], nor in cell culture [65,75,77]. In contrast, Wy-14643 has profound effects on heart muscle UCP3 expression [78,86,87]. In summary, PPAR $\alpha$  is an important regulator of UCP3 transcription in BAT and heart. In SKTM, apart from directly after birth, PPAR $\alpha$  does not regulate UCP3 expression.

PPAR $\delta$  (synonymous with PPAR $\beta$ ) has been shown to mediate the effects of endurance training, and its overexpression induces a shift to oxidative muscle fibres [88,89]. It is a central regulator of UCP3 expression in SKTM both in animal and cell culture [65,75,78,81,89,90] and ChIP data demonstrate its binding to the core promoter [91]. Effects on transcription of UCP3 in BAT are poorly investigated and thus no conclusion can be drawn yet. In heart, PPAR $\delta$  regulation of UCP3 expression has been described [92]. In conclusion, PPAR $\delta$  mediates regulation of UCP3 transcription in SKTM.

PPARs act as heterodimers in combination with retinoid X receptor (RXR)  $\alpha$ , another nuclear receptor known to stimulate UCP3 transcription [81]. RXR $\alpha$  binds 9-cis-retinoic acid (9cRA) [93] which stimulate UCP3 expression in BAT [70] and SKTM [65]. This dimer then recruits the histone acetylase p300 which opens chromatin and acetylates other transcriptional activators thereby facilitating UCP3 transcription [79]. Apart from PPARs, RXR can dimerise with other nuclear receptors including retinoid acid receptor (RAR), binding all-trans retinoic acid (ATRA) [93] and 9cRA, thyroid receptor (TR), binding T3, and vitamin D receptor (VDR), binding 1 $\alpha$ ,25-dihy-droxyvitamin D3. At least in heart, TR activity is required for PPAR $\alpha$  agonist activity [94]. All those proteins also induce UCP3 expression [12,65,95-97]. The RXR/TR dimer is responsible for the induction of UCP3 transcription by T3 [65], a hormone also stimulating BAT expansion, differentiation and energy metabolism. Most of the heterodimers mentioned above are believed to alternatively bind a single DR1 element in the promoter. The binding element for VDR is unknown.



**Figure 3: General model of regulation of UCP3 transcription:** The core promoter of UCP3 is bound by MyoD via 3 non-canonical E-Boxes close to the transcriptional start site. A direct repeat element nearby can be bound by different nuclear hormone receptors, mainly PPARs and RXR as a heterodimer. These proteins then recruit histone acetylase p300 which opens the chromatin. Two upstream enhancer elements are described: One binding COUP-TF2 and a yet unnamed repressor via an inverted repeat, and the other one binding Nrf2 via an antioxidant response element. Numbers denoted for the COUP-TF2 and ARE elements note the distance to the TSS in mouse.

- ▮ Nuclear hormone receptor
- ▮ Repressing TF
- ▮ Activating TF
- Enzyme

Another nuclear hormone receptor, chicken ovalbumin upstream promoter transcription factor 2 (COUP-TF2), binds an enhancer element upstream of the core promoter [98]. In SKTM, MyoD and myogenin stimulate transcription of the UCP3 gene [95,99]. In BAT, their role is questionable, as both are considered muscle specific and absent in fully differentiated BAT cells. A link to ROS defence is found in the transcription factor nuclear response factor 2 (Nrf2) [47], a factor known to induce a wide array of stress response genes. Nrf2 binds upstream of the UCP3 promoter upon H<sub>2</sub>O<sub>2</sub> exposure in both SKTM and heart cell lines and stimulates UCP3 transcription [47]. While some studies also suggest activation of UCP3 transcription by the beta-adrenoreceptor ( $\beta$ -AR) pathway via CREB, this effect is most likely secondary due to the increased lipolysis supplying PPAR ligands [84].

#### **1.4.2. UCP1**

UCP1, also known as Thermogenin, was first described to be responsible for thermogenesis in 1978 [17] and purified in 1980 [100]. Apart from the related brown adipocytes [101] and thymocytes [102] brown adipose tissue is the only major tissue expressing UCP1 in mammals, and the amount of UCP1 protein in BAT dwarfs the amount found in non-BAT tissues [100], at least on a per cell basis. UCP1 is crucial for the heater function of BAT [18,19]. Heat generation is possible due to a futile cycle (see 1.4, page 2) which enables  $\beta$ -oxidation, glycolysis, citric acid cycle and respiratory chain to break down nutrients at high rate without the generation of ATP. Other than UCP2 and 3, UCP1 allows enough protons to cross the MIM [36] to drastically impair ATP synthesis, owing to its high abundance.

#### **1.4.3. UCP2**

UCP2 was identified 1997 in mouse and human [26], measures 309 AA and is expressed in most tissue and cell types. Current hypotheses on UCP2 function assume that UCP2 confers mild uncoupling which only slightly reduces ATP production and membrane potential. This would decrease ROS production. Furthermore it would influence glucose sensing by fine tuning ADP/ATP-ratio, modifying insulin secretion from the pancreas [103] and calcium homeostasis [54,104]. Apart from these uncoupling related functions, UCP2 has been suggested to transport non-proton substrates. One report proposes pyruvate and chloride transport [50] while another publication proposes antiport of different C4-metabolites against phosphate and protons [52]. This antiport en passant would also explain the uncoupling activity observed.

#### **1.4.4. Transcriptional regulation of brown adipose tissue and skeletal muscle determination and differentiation**

BAT and WAT differentiation share some transcriptional regulators, but there also are several described to be more abundant in brown adipocytes when compared to white adipocytes. In addition to general adipose tissue and BAT specific factors, muscle specific TFs play a role in BAT determination and differentiation. Only few factors are truly unique to brown adipose tissue. Thus the most likely hypothesis is that it is the combination of factors that is unique and specific for an UCP1 expressing brown adipocyte. The key transcriptional regulators orchestrating BAT differentiation are currently regarded to be C/EBPs, PPARs, RXR $\alpha$ , RAR, TR, PGC1 $\alpha$  and PRDM16. Upstream of this network, Mef2, TGF $\beta$  signalling, the sympathetic nervous system (SNS) and a number of endocrine and paracrine peptides are thought to be determining factors. The most notable of the latter category are Noggin and bone morphogenetic protein (BMP) 4 (see below), BMP 7 [105] and fibroblast growth factors (FGF) 19 and 21 [106,107].

In the very early stages of determination, SKTM and BAT originate from a common precursor that is positive for the factor myf5 [9], which was initially believed to be muscle specific. These precursor cells

can then either become myogenic or adipogenic precursors. The earliest steps that determine BAT and SKTM precursors are poorly investigated, but some evidence points out that BMP4 and Noggin might be very early decisionmakers [108-111]. Furthermore, classical determination pathways like Hedgehog and Wnt signalling are believed to play a role (reviewed in [112]). Two main anti-adipogenic/pro-myogenic determination pathways are known. Firstly, Mef2c, a central muscle determination factor, blocks PRDM16 expression via myomir-133 [113]. Secondly, TGF $\beta$  signalling blocks BMP7 signalling and expression of C/EBP $\beta$  via miR-155 [114]. Oppositely, there are three pro-adipogenic pathways: Firstly the SNS stimulates  $\alpha_2$ - and  $\beta_3$ -adrenoreceptors.  $\beta_3$ -adrenoreceptor activation leads to suppression of Mef2c and thus myomir-133 [113] via a PKA dependant pathway. Activation of  $\alpha_2$ -adrenoreceptors leads to an increased expression of DIO2 via protein kinase C (PKC) and CREB. Secondly, BMP7 stimulates p38-MAPK via TAK1 [105], thereby inducing expression of PGC1 $\alpha$ , PPAR $\gamma$  and C/EBP $\alpha$ . Thirdly, fibroblast growth factors induce activation of PKC signalling which reduces Mef2c activity and activates PGC1 $\alpha$  by deacetylation via AMPK and SIRT1 [115]. While the initial, endocrine FGF21 production originates from the liver triggered by milk intake after birth [107], in the adult state FGF21 is secreted upon activation of BAT-thermogenesis as a para-/autocrine feedback loop [116].

When the decision is made that a cell is to become a brown adipocyte, the five regulators are believed to form the central differentiation network. PRDM16, PGC1 $\alpha$ , PPAR $\alpha/\gamma$  and C/EBP $\beta$  start the transcriptional differentiation cascade. Of these factors, PRDM16 is considered the main SKTM versus BAT switch [9,117]. This is achieved by facilitating expression of pro-adipogenic transcription factors, and partially by directly repression of pro-myogenic transcription factors, most notably runx1t1, via miR 193b-365 [118]. Either directly via miRNA193b-365 activity or indirectly via loss of runx1t1, expression of myf5, MyoD and myogenin diminishes. Furthermore PRDM16 acts as a co-factor in regulation of BAT specific gene expression. PGC1 $\alpha$  leads to increased expression of PPAR $\gamma$ , TR [119] and DIO2, which are all considered important inducers of BAT differentiation and NST capacity. Both PRDM16 and PGC1 $\alpha$  act in complexes with the PPARs [120,121]. Further downstream effects, including mitochondrial biogenesis, are then induced by RAR and ER which integrate hormonal signals and act as transcriptional active nuclear hormone receptor dimers with RXR $\alpha$ .

In case of a pro-myogenic situation, PRDM16, BMP7 downstream signalling and the C/EBPs remain repressed. Without miR193b-365, the muscle-transcription factor runx1t1 [118] protein level remains high and represses expression of PPAR $\gamma$  and PRDM16. Myf5, MyoD and Myogenin then induce myogenic gene expression and repress expression of adipogenic genes. These three transcription factors and the factor MRF4 are commonly termed as the four muscle regulatory factors (MRFs) and orchestrate muscle differentiation. For detailed insight on the earliest SKTM determination processes a review should be consulted [122].

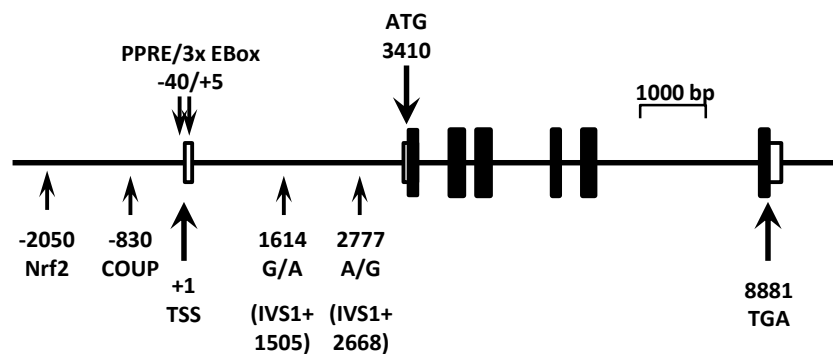
#### **1.4.5. White and Brite adipose tissue**

Despite being different in origin [9,123] and function, BAT and WAT share characteristics in their metabolism. As both tissues store lipids in one (WAT) or many (BAT) intracellular lipid droplets, they also share anabolic and catabolic capacities as well as protein expression related to lipid storage. Fittingly, they also share key transcription factors, most notably PPAR $\gamma$  and C/EBPs. Differential transcriptional regulation between BAT and WAT thus believed to be achieved via different co-regulators. Of these PGC1 $\alpha$  is considered the main pro-BAT switch [124], while its pro-WAT counter players are RIP140 [125], p107 [126] and RB (retinoblastoma protein) [127]. Function, origin and regulation of a third class of adipocytes, the recently discovered brown adipocytes in interspersed WAT (brite) adipocytes [101], are still debated.

### 1.5. A naturally occurring mutation leads to BAT specific UCP3 deficiency in *Phodopus sungorus*

The core promoter model presented in figure 3 was commonly accepted for transcriptional regulation of UCP3 expression until a naturally occurring mutation was discovered in *Phodopus sungorus* [59,128]. This mutation shifted our focus to the first intron and marked the starting point for this project. *Phodopus sungorus*, the djungarian hamster, is native to central asia. Due to the continental climate, animals have to tolerate a wide range of temperatures ranging from below -30°C to above 30°C. They achieve this by building well isolated burrows and exhibiting pronounced seasonal adaptations, including high non-shivering thermogenesis capacity, body mass reduction, and torpor. These adaptations are the reason why the animal is a model organism for research on metabolism and brown adipose tissue.

The mutation of interest was identified in the breeding colony of the Philipps Universität Marburg where a subpopulation of animals completely lacked UCP3 expression in BAT, but had close to normal expression levels in SKTM [59,128]. Equivalent subpopulations could later be identified in several other independent breeding colonies, demonstrating that the mutation makes up a relevant proportion of the native population [59]. The trait was heritable and led to increased body mass [59] and decreased cold tolerance [58] in UCP3 deficient animals when compared to UCP3 expressing littermates. Sequencing of the UCP3 gene revealed two base exchanges located in the first intron of UCP3, intervening sequence (IVS) 1+1505G/A and IVS1+2668A/G [59]. Both were in linkage disequilibrium, and the 1505A/2668G haplotype segregated with the lack of UCP3 in BAT. Figure 4 shows a scheme of the *Phodopus sungorus* UCP3 gene including features relevant for regulation of its expression.

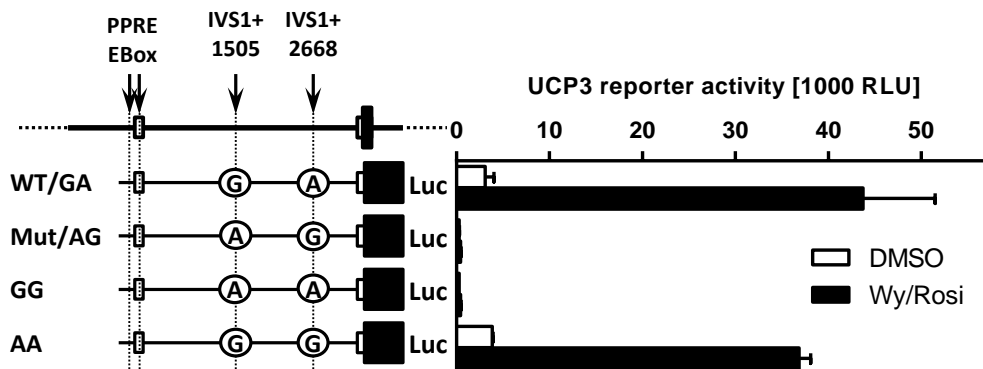


**Figure 4: Organisation of the *Phodopus sungorus* UCP3 Gene:** Shown are the seven exons (boxes), transcriptional start site (TSS), start codon (ATG) and stop codon (TGA). Open boxes mark untranslated exons, filled boxes mark translated exons. Numbers are base position in relation to the TSS. At position 1614 (IVS1+1505) and 2777 (IVS1+2668) the two base exchanges segregating with the BAT specific UCP3 deficiency are located. Note that the IVS1 nomenclature differs from base numbering in relation to the TSS by the size of exon 1, 109 bp. NRF2 and Coup-TF2 sites are located upstream of the TSS, PPAR/RXR sites (PPRE) and MyoD sites (3x EBox) are directly adjacent to TSS.

### 1.5.1. The IVS1+1505 G/A polymorphism

Using reporter gene constructs carrying all four possible combinations possible between the two base exchanges (figure 5), it was demonstrated that the first polymorph site, IVS1+1505G/A is responsible for the effect on UCP3 expression. Interestingly, addition of PPAR $\alpha$  and PPAR $\gamma$  agonists amplifies the difference between the two alleles because only the G-allele was agonist responsive. Transfection into a SKTM cell line, C2C12, did not yield differences between the alleles in non-stimulated cells, and yielded only a minor allele difference after stimulation of reporter gene activity (not shown,[59]). This is consistent with the initial observation in *Phodopus sungorus*.

Electrophoretic mobility shift assays (EMSA) subsequently demonstrated formation of a protein-DNA complex with the G-allele. Complex formation was clearly diminished with the A-Allele (figure 8, page 49). Together, this suggest an activator binding transcription factor binding site (TFBS), containing the IVS1+1505 position, which is impaired by the A-Allele [59]. This TFBS seems to be indispensable for UCP3 transcription in BAT and is required for PPAR agonist action.



**Figure 5: The IVS1505 polymorphism is responsible for the absence of UCP3 mRNA in BAT:** 4 reporter gene constructs containing 250 bp of the *Phodopus sungorus* UCP3 promoter and the complete first intron fused to a luciferase gene were transfected into HIB1b cells. The 4 constructs covered all combinations between the two base exchanges in the first intron. Both constructs carrying the G-Allele at the IVS1+1505 site yield higher luciferase activity under non-stimulated conditions compared to their IVS1+1505A counterparts and are responsive to PPAR agonists. The IVS1+2668 polymorphism has no influence on reporter gene activity. Modified after [59].

## 1.6. Goals of the thesis

Experiments in cell culture and hamster suggest that a TFBS is located within the first intron of the UCP3 gene. Part of this element is the IVS1+1505 base exchange, and the allele at this position influences protein binding. We expected the binding proteins to be transcriptional regulators, and to participate in BAT-specific regulation of UCP3 expression. The first goal of this study thus was to identify proteins binding to the IVS1+1505 element. We hoped to uncover a yet unknown regulator of UCP3 expression that acts as a switch allowing differential expression regulation between BAT and SKTM. In a best case scenario, the factor might also uncover a pathway not yet known to play a role in regulation of UCP3 expression. Lastly we wondered whether the TFBS was a lone regulatory element, or part of a more complex enhancer region. If indeed additionally elements were uncovered, we hoped to be able to elucidate the interplay between the elements and proteins.

We aimed to identify the proteins forming the IVS1+1505G specific complex by two different experimental strategies: Firstly, we wanted to purify the binding proteins with the help of oligonucleotide trapping DNA affinity chromatography [129,130]. Subsequently, protein identification by mass spectrometry (MS) was planned to be carried out by the Proteomics and Bioanalytics department (Technical University of Munich, headed by Professor Dr. Bernhard Küster). Secondly, we aimed to use bioinformatics for binding site prediction. Subsequently we wanted to validate the candidates from both strategies in EMSA, via addition of specific antibodies and cold competitors, and in cell culture, using overexpression and RNAi mediated knockdown experiments. To search for additional TFBS in the vicinity of the IVS1+1505 element we planned to perform deletion and mutation screens. Aided by the tools mentioned above, we hoped to be able to pinpoint distinct binding sites and uncover the binding factors.

## 2. Material and Methods

For a complete list of all chemicals including manufacturer and order number, consult appendix 1. Consumerable plastic ware is listed in appendix 2. All enzymes, cell lines and antibodies used can be found in appendix 3, 4 and 5, respectively. Kit systems are listed in appendix 6. Oligonucleotide sequences are available in appendix 7. All devices and machines are listed in appendix 8. Appendix 9 contains detailed information on bioinformatics. Appendix 10 contains relevant vector maps. Lastly, appendix 11 and 12 list internet resources, tools, software and online data resources. For all buffers final concentrations are given unless stated otherwise. HCl or NaOH were used for pH-adjusting unless specified otherwise.

### 2.1. Culturing and handling of bacteria

#### 2.1.1. Bacterial transformation

For amplification of plasmids the bacterium *Escherichia coli* was exploited. A suitable amount of DNA (usually 1 ng of supercoiled or 10 ng of circular, non-supercoiled DNA) in 1-4  $\mu$ l volume was added to 50  $\mu$ l chemically competent bacteria (NEB5 $\alpha$  competent *E. coli*, NEB, Ipswich, USA) and incubated on ice for 30 min. The mixture was then heat-shocked by immediate transfer to a 41°C water bath for 45 seconds and cooled down again on ice for 3 min. For regeneration 250  $\mu$ l room temperature SOC medium was added and the bacteria incubated for 1h on a shaker at 37°C. After streaking the bacteria onto LB-agar plates, cultures were incubated top down at 37°C overnight. To prevent growth of non-transformed bacteria, the plates usually contained an antibiotic substance. Plates can be stored top down at 4°C for several weeks when sealed with parafilm.

Lysogene broth*		Antibiotics	
5 g/l	Yeast extract	Ampicillin	50 $\mu$ g/ml
10 g/l	Tryptone	Spectinomycin	50 $\mu$ g/ml
5 g/l	NaCl	Kanamycin	50 $\mu$ g/ml

adjust pH to 7.0, autoclave  
\* for agar plates add 15 g/l agar-agar

SOC medium	
5 g/l	Yeast Extract
20 g/l	Tryptone
10 mM	NaCl
2.5 mM	KCl
10 mM	MgCl <sub>2</sub>
10 mM	MgSO <sub>4</sub>
20 mM	Glucose*

\* Autoclave before adding glucose. Allow to cool down, add glucose and filter sterile (0.2  $\mu$ M)

#### 2.1.2. Plasmid amplification, extraction and quantification

To generate preparative amounts of a plasmid, a colony of bacteria containing the respective plasmid (usually from a previous bacterial transformation) was picked with a sterile pipette tip and dropped into a test tube containing 2 ml LB medium. As each colony is expected to clonally rise from a single bacterium, all vectors in each colony should be alike. Loss of the plasmid was prevented by addition of the respective selective antibiotic. For preparation of 10-30  $\mu$ g of DNA, the test tubes were incubated shaking (220 rpm) at 37°C overnight. For preparation of 200-1000  $\mu$ g, the tubes were incubated for 8

to 10 h and then diluted 200 to 1000 fold, depending on optical density, into 200 ml LB medium (plus antibiotics) in an Erlenmeyer flask. After incubation (shaking, 37°C, overnight) plasmids were extracted using the PureYield Plasmid Miniprep System or the PureYield Plasmid Midiprep System (both Promega, Fitchburg, USA) for 2 ml or 200 ml cultures, respectively. Purification was carried out according to the protocol supplied by the manufacturer. Afterwards, the concentration of DNA was assayed by measuring extinction at 260 nm using the Nano-Quant plate and a Tecan Infinite M200 reader (Tecan, Männedorf, Switzerland). Extinction measurement at 230 nm and 280 nm allowed estimation of organic solvent and protein contamination, respectively. Purified DNA was stored at -20°C.

### **2.1.3. Long term storage of plasmid carrying bacteria**

To store plasmids for extended periods glycerol stocks were generated. 400 µl LB medium containing bacteria carrying the plasmid of interest were mixed with 600 µl 80% glycerol and then stored at -80°C. The bacterial solution was usually prepared by transformation and amplification as described above. When desired, a small fraction of the frozen stock was streaked onto a LB-agar plate using an inoculation loop. After overnight cultivation (37°C, top down), colonies were picked and amplified further as described above.

## **2.2. Cloning**

Cloning encompasses several techniques to obtain, generate, modify, separate, amplify and combine fragments of DNA to generate a construct, often a plasmid, that either influences or assays processes in a model organism. The main tasks in this project were forced protein expression, mRNA depletion by RNA interference (RNAi) and measuring transcriptional activity of promoter and enhancer elements. Cloning techniques commonly used are polymerase chain reaction (PCR), restriction enzyme digestion, phosphorylation and dephosphorylation of DNA, blunting of single stranded DNA overhangs and ligation of DNA fragments. Intermediates and products of the cloning procedure were usually analysed and purified by agarose gel electrophoresis. Correctness of generated constructs was tested by sequencing or restriction digestion. For all cloning procedures, double distilled water or water of comparable quality was used. All enzyme amounts are typically given in activity units.

### **2.2.1. Standard techniques of cloning**

#### **2.2.1.1. Agarose gel electrophoresis and DNA recovery**

For separation, purification and visualisation of double stranded DNA agarose gel electrophoresis was used. This method is based on the principle that DNA moves in the electric field towards the cathode due to its negatively charged backbone. When passing through a matrix, small DNA molecules can move faster than large ones.

Depending on the size of DNA fragments to be resolved, the gel matrix consisted of 0.7 to 2.0 % w/v agarose resolved in 1x TAE buffer. This was achieved by boiling the agarose-buffer mix in a microwave until no visible agarose particles were left. Roti-Safe (Carl Rot, Karlsruhe, Germany), a minor groove binding fluorophore, was added in a dilution of 1:20,000 to the agarose mixture to visualize DNA. The gel was cast in a horizontal gel apparatus and left to cool down for polymerisation after a comb with the desired amount of teeth was inserted. For separation, polymerised gels were overlaid with buffer and loading buffer was added to the DNA sample which subsequently was loaded onto the gel. To estimate fragment size, a marker containing a mix of defined size DNA fragments was also loaded. Separation was carried out at 100V and room temperature until desired separation was achieved. DNA was visualized by excitation of the Roti-Safe-DNA complex using a UV-transilluminator (Intas, Göttingen, Germany). When desired, blocks containing DNA were excised using a scalpel and purified with the



Wizard SV PCR and Gel Clean-Up system (Promega, Fitchburg, USA) according to the respective manual.

**6x DNA Loading buffer**

- 10 mM Tris pH 7.6
- 60 mM EDTA
- 60% V/V Glycerol
- 0.2% W/V Orange G\*

\* for a blue loading buffer, 0.03% bromophenol blue and 0.03% xylene cyanol can be added instead

**50x TAE Puffer**

- 2 M TrisBase
- 1 M Essigsäure
- 50 mM EDTA pH 8
- adjust pH 8.3 using acetic acid

**2.2.1.2. PCR amplification for cloning**

To amplify a sequence of interest for cloning, Q5 Polymerase (NEB, Ipswich, USA) was used. Primer sequences flanking the region of interest were generated with the primer3 tool of the SDSC workbench. Restriction sites were added manually to the 5' end of both primer sequences. Melting temperatures between 60 and 69°C were preferred for binding of the complementary region, yielding a melting temperature above 70°C with added restriction sites. Temperatures were checked by the respective tool on the NEB homepage. After the run, loading buffer was added and presence and size of PCR products was analysed using agarose gel electrophoresis. The PCR product of interest then was ligated into pJET1.2 using the CloneJET PCR cloning kit (Fermentas, Vilnius, Lithuania) according to the standard protocol.

**typical Q5 PCR reaction**

- 4 µl Q5 Puffer (5x)
- 200 µM each dNTP
- 1 µl template\*<sup>1</sup>
- 500 nM each Primer
- 0.5 U Q5 Enzyme
- ad 20 µl water

**typical Q5 PCR protocol**

98°C	30 sec	
98°C	10 sec	<b>3 x</b>
X°C* <sup>2</sup>	15 sec	
72°C	X sec* <sup>3</sup>	
98°C	10 sec	<b>32 x</b>
72°C	X sec* <sup>3</sup>	
72°C	60 sec	
4°C	hold	

\*<sup>1</sup> Template amount per reaction:

\*<sup>2</sup> Annealing temperature Cycle 1-3:

\*<sup>3</sup> Elongation time:

0.1 ng Vector, 0.5 ng BAC or cDNA from 50 ng total RNA  
 calculated TM +3°C for primers >20bp, TM for primers ≤20bp  
 20 sec/kb for vector or BAC template, 40 sec/kb for cDNA

**2.2.1.3. Restriction enzyme digest**

Restriction enzyme digests were used to excise fragments from plasmids or to open vectors for ligation of DNA fragments into a vector. Usually, 0.5 µl of enzyme and 2 µl buffer were added to 2 µg of DNA in a total volume of 20 µl. The mixture then was incubated for 3h at 37°C followed by heat inactivation by incubation of 20 min at 80°C. When multiple Enzymes were mixed for digestion, buffer was selected according a tool on the respective manufacturer's homepage. If necessary, enzyme amounts used were adjusted depending on activity in the buffer. Total enzyme volume per 20 µl reaction never exceeded 2 µl.

**typical restriction digest**

5-10 U enzyme 1  
5-10 U enzyme 2  
2 µl suitable Buffer  
2000 ng DNA  
ad 20 µl water

**2.2.1.4. Phosphorylation of DNA**

Phosphorylation of DNA ends is the process of adding a phosphate group to the 5' end of double or single stranded DNA or RNA. DNA fragments generated by PCR amplification or annealing of oligonucleotides usually are non-phosphorylated, while DNA fragments from restriction digest retain the phosphate from the sugar phosphate backbone. Phosphorylation reactions were carried out using the enzyme T4 polynucleotide kinase (PNK) (Fermentas, Vilnius, Lithuania) in presence of the respective buffer and 1 mM ATP. Reactions were carried out for 30 min at 37°C and stopped by heating to 75°C for 10 min.

**typical PNK reaction**

≤20 pmol DNA fragments\*  
2 µl 10x Buffer A  
10 U T4 PNK  
ad 20 µl water

\* The amount of DNA fragments is usually limited by either the available concentration for long DNA fragments or adjusted to the up-/downstream procedures for phosphorylation of oligonucleotides.

**2.2.1.5. Dephosphorylation of DNA**

DNA was dephosphorylated to suppress unwanted ligation reactions. Usually linearized vector was dephosphorylated to prevent circularisation without an insert. For restriction digestions, 1U FastAP (Fermentas, Vilnius, Lithuania) was added directly to the digestion. If dephosphorylation was carried out after digestion, the reaction was incubated at 37°C for 30 min. Reactions were heat inactivated at 75°C for 10 min.

**typical FastAP reaction**

1 µg linear vector DNA  
2 µl 10x FastAP Buffer  
1 U FastAP  
ad 20 µl water

**2.2.1.6. Blunting of DNA**

When a blunt end DNA fragment, usually generated by PCR or by excision using a blunt end generating restriction enzyme, was to be ligated into a vector that was linearized with sticky end generating enzymes, or when fragments of incompatible ends were to be ligated, the single stranded overhangs were removed (3' overhang) or filled in (5' overhang) using T4 DNA polymerase (Fermentas, Vilnius, Lithuania). Reactions were incubated at room temperature for 5 min and heat inactivated at 75°C for 10 min.

#### **typical T4 DNA Pol reaction**

1 µg	vector DNA
4 µl	5x T4 Pol buffer
100 µM	each dNTP
1 U	T4 DNA Pol
ad 20 µl	water

#### **2.2.1.7. Ligation reaction**

Ligations are reactions where compatible DNA ends are joined by an enzyme, usually T4 Ligase (Fermentas, Vilnius, Lithuania), generating a phosphodiester bond. ATP is consumed during that reaction. One of the two ends to be joined has to be phosphorylated for a ligation. Ligation reactions were usually incubated in a PCR cyclor for better temperature and time control. For blunt end ligations, polyethylene glycol (PEG) 4000 was added. For a typical ligation, 25 ng vector and insert in molar excess (usually 1:3 vector:insert) were mixed in 10 µl reaction volume. In cases of very small (<200bp) or very large (<5000bp) insert, different vector to insert ratios (1:1 to 1:5) were tested. Ligation was carried out for 30 min at 21°C, 1 hour at 14 °C followed by a 4°C hold. Ligations were never heat inactivated. Ligation reactions into the CloneJet 1.2 blunt vector (Fermentas, Vilnius, Lithuania) deviated from this protocol and were carried out according to the manufacturer's protocol. Of the fragments ligated the one carrying the bacterial origin of replication is considered as vector. The vector was usually dephosphorylated before ligation. Ligation reactions were either directly transformed into bacteria or stored at -20°C for later transformation.

#### **typical T4 DNA Ligase reaction**

25 ng	linearized vector
1 µl	10x T4 Pol buffer
X ng* <sup>1</sup>	insert
0.5/2.5U* <sup>2</sup>	T4 DNA Pol
0/5%* <sup>3</sup>	PEG 4000
ad 10 µl	water

\*<sup>1</sup> Insert is used in a molar excess over vector of 1:3

\*<sup>2</sup> 0.5U for sticky end ligation, 2.5U for blunt end ligation

\*<sup>3</sup> 5% PEG 4000 final for blunt end ligations, no PEG for sticky end ligations

#### **2.2.1.8. Validation of ligation products**

After transformation and amplification of ligated vector DNA, the correctness of the products was tested by control digestion and Sanger Sequencing. Usually three to six colonies were picked per ligation of which all were tested by digestion. If promising clones were identified, the sequence of one clone was validated using the GAP4 software of the Staden Package. For control digestions, a standard restriction digest was performed with a set of enzymes cutting once in vector, once in inserts. If two fragments were generated, the vector contains the insert. If the restriction site was not located centrally, information on the orientation of the insert could be gained. Sequencing was performed by MWG Eurofins Operon according to their sequencing guide. Sequencing primers were generated using the tool provided by MWG biotech.

### **2.2.2. Viral vectors**

Viral vectors used in this thesis were based on the Moloney Murine Leukemia Virus (MMLV) genome. Two vector backbones containing the LTRs and the packaging signal of said virus were the basis of all constructs generated: pMXs-IRES-Puro (Cell Biolabs, San Diego, USA) uses the retroviral LTR promoter for transcription of both viral genome mRNA in the packaging cell line and expression of the bicistronic mRNA containing puromycin resistance and expression cassette of interest in target cells. The self-inactivating pMXs-EF1-PGK-BSD (Cell Biolabs, San Diego, USA) uses the LTR to generate the genome mRNA in the packaging cell, but relies on the eukaryotic promoters EF1 for transgene expression and PGK for BSD resistance. For the ease of assaying transfection and infection, pMXs-EF1-PGK-GFP::BSD was generated by PCR amplifying emGFP pcDNA6.2-GW/EmGFP-miR (Invitrogen, Carlsbad, USA), ligating it into pJET1.2 blunt (Fermentas, Vilnius, Lithuania), excising it with NcoI (Fermentas, Vilnius, Lithuania), and ligating the fragment into the NcoI site in front of the BSD gene. The PCR step added restriction sites and a short spacer to prevent steric hindrance between the two subunits of the generated fusion protein.

#### **2.2.2.1. miRNA expression**

For each target gene two different miRNA sequences were generated using the BlockIT RNAi design tool (Invitrogen, Carlsbad, USA). Oligonucleotides of the respective sequences were obtained from MWG Biotech, annealed and cloned into pcDNA6.2-GW/EmGFP-miR according to the BLOCK-iT Pol II miR RNAi Expression Vector Kit protocol (Invitrogen, Carlsbad, USA). If desired, miRNAs were concatemered. Afterwards, the miRNA cassette and the emGFP gene were PCR amplified, gel purified and ligated into pJET1.2 blunt. Two BsaI sites, introduced during PCR, were cut to release the cassette and generate EcoRI and BamHI overhangs. In parallel, pMXs-IRES-Puro was opened with EcoRI and BamHI for integration of the gel purified emGFP-miRNA cassette.

#### **2.2.2.2. Tagging of proteins**

If tagged proteins were to be generated, oligonucleotides containing a Kozak consensus sequence and the sequence of the Tag were ordered and annealed. The destination vector, pMXs-EF1-PGK-GFP::BSD, was opened using BamHI and NotI and the Kozak-Tag fragment was inserted by ligation. The oligonucleotides were designed to contain compatible overhangs for ligation into the MCS and restriction sites for subsequent insertion of the cDNA of interest.

#### **2.2.2.3. Overexpression constructs**

The first step to generation of overexpression constructs is RNA extraction from the tissue of interest followed by cDNA synthesis. Primers were designed, containing suitable restriction enzyme sites (Esp3I in case of CREB, SP1 and SP3, BamHI and NotI in case of SP2), and used to amplify the gene of interest from cDNA. The PCR product was gel purified and ligated into pJet1.2 blunt, followed by digestion using the respective enzymes. This generated BamHI and NotI overhangs. The excised fragment was gel purified and ligated into the destination vector, which was linearized with BamHI and NotI, followed by control digestion and sequencing. If tagged proteins were to be generated, primers were designed to avoid amplification of 5'UTR and fit to the restriction sites and frame of the Kozak-Tag sequence.

### **2.2.3. Reporter gene constructs**

Reporter gene constructs contain the coding sequence (CDS) of an easy-to-measure protein, in most cases a luciferase, fused to a regulatory region of interest. Regions of interest can range from a small enhancer element of less than 100bp to a full promoter region of several thousand base pairs. Reporter gene constructs usually are constructed in specific vectors that are sequence optimised to contain a

reduced amount of putative transcription factor binding sites and can also contain a basic promoter or enhancer, depending on the nature of the element of interest. All reporter vectors used were assembled in a modified pGL3-basic (Promega, Fitchburg, USA) whose *Photinus* luciferase was replaced by a *Gaussia* luciferase to increase sensitivity. pGL3-basic is devoid of promoter and enhancer.

#### **2.2.3.1. Generation of pGL3-basic-GLuc and pGL3-CMV-PLuc**

To replace the *Photinus* luciferase of pGL3-basic by a *Gaussia* luciferase, the GLuc gene from pCMV-GLuc (NEB, Ipswich, USA) was PCR amplified and cloned into cloneJet 1.2 blunt (Fermentas, Vilnius, Lithuania). Bsal sites were added by PCR and, upon restriction, generated NcoI and XbaI overhangs. From pGL3-basic, the *Photinus* luciferase was removed using NcoI and XbaI (Fermentas, Vilnius, Lithuania) and the empty backbone was ligated with the *Gaussia* cassette.

For generation of pGL3-CMV, the cytomegalovirus (CMV) promoter was excised from pCMV-GLuc using BglII and HindIII. pGL3-basic was linearized using the same enzymes and the CMV promoter was inserted.

#### **2.2.3.2. Generation of deletions in the *Phodopus sungorus* reporter gene vector**

Four *Phodopus* UCP3 reporter constructs, containing the intron in all four possible combinations of both polymorph sites and 250bp of the core promoter, were already available from a previous study [59]. Their *Photinus* luciferase gene was exchanged for a *Gaussia* luciferase gene as described above. To delete regions of interest, primers flanking the region to be deleted were designed in a way that they amplified outwards, generating a product containing the whole vector with exception of the enclosed region. The amplicon was gel purified, phosphorylated and re-circularized. Luciferase, promoter and intronic enhancer were sequenced to avoid undesired PCR-generated mutations.

#### **2.2.3.3. Site directed mutagenesis in the *Phodopus sungorus* reporter gene vector**

For introduction of point mutations and small deletions/insertions, PCR-based mutagenesis was employed. The QuikChange II mutagenesis kit (Agilent Technologies, Santa Clara, USA) was used according to manual. Primers were generated by a tool provided by the manufacturer. If whole TFBS were to be deleted, they were replaced by a restriction site for easy validation of the introduced mutation.

#### **2.2.3.4. Generation of reporter gene constructs containing different promoter sizes**

To assay the influence of promoter elements outside of the core promoter, elements upstream of the 250bp promoter fragment already present in the reporter vectors [59] were added. To do so, a reporter vector from another study [98] was used to obtain the sequences of interest. This vector, -2244UCP3luc, contained bases -2244 to +38 of the *Phodopus sungorus* UCP3 gene in pGL3 basic. A 1000bp and a 2000bp promoter vector were generated. The 1000bp version was generated by excising the respective promoter fragment using BglII and ligation into the BglII site of the initial reporter vectors. The 2000bp construct was generated by PCR amplification of the respective region, adding Esp3I restriction sites. Restriction with Esp3I generated MluI and BglII overhangs which fit into the MluI and BglII sites of the 250bp promoter/intron vectors.

### **2.3. Protein extraction and quantification**

Protein can be extracted in native and active conformation or in denatured and inactive conformation. Native extraction usually relies on mechanical disruption of a cell, sometimes aided by non-ionic, non-denaturing detergents. Depending on the detergents, method of cell disruption and centrifugation procedures different proteins from different compartments can be obtained. Native extraction of

nuclear protein is described in the EMSA part of material and methods. Denaturing extraction usually relies on strong denaturing detergents, disrupts all membranes and thereby yielding a total protein extract. Two more or less denaturing methods of protein extraction are described below: RIPA (Radio-immunoprecipitation assay) protein extracts retain some of the native protein conformation while still efficiently solubilising most proteins while boiling SDS extraction perfectly extracts all protein but completely unfolds and denaturizes them.

### 2.3.1. RIPA protein extract

Extracts from a larger number of cells were usually generated using RIPA buffer. In RIPA extracts protein concentration can be determined by the bichinonic acid assay (BCA), thus providing an ideal starting point for western blot experiments.

For RIPA extracts, cells were grown on multiwell plates. Medium was removed by suction and the plate was placed on ice. After washing with ice cold phosphate buffered saline (PBS), RIPA buffer was added. Cells were scraped using a sterile pipette tip or a cell scraper, depending on well size, and transferred to a reaction tube. After 20 min of vigorous shaking at 4°C, the debris was pelleted for 10 min at 25,000g in a precooled benchtop centrifuge. The supernatant was transferred to a new tube and stored at -80°C. An aliquot of RIPA buffer was saved for protein quantification. If extracts were to be carried out at a later time point, the cell culture plates were frozen at -80°C after removal of PBS. In that case, RIPA buffer was added after removing the plate from the freezer and the protocol was carried out as described above.

RIPA buffer*		RIPA used per well	
50 mM	Tris	6-well	200 µl
1% V/V	NP-40 substituent	12-well	100 µl
0.25% W/V	Na-desoxycholate	24-well	50 µl
150 mM	NaCl	48-well	25 µl
1 mM	EDTA		
adjust pH to 7.4			

\* before use, protease and phosphatase inhibitor cocktails were added in a dilution of 1:200.

### 2.3.2. BCA protein quantification

The bichinonic acid (BCA) protein assay (Pierce Biotechnology, Rockford, USA), is an assay used to quantify protein concentration in an aqueous solution. The principle is based on chelation of Cu<sup>+</sup> by BCA. This complex exhibits a strong absorbance at 562 nm. The Cu<sup>+</sup> ions are generated by a reaction named biuret reaction where peptides of three or more aminoacids (AA) react with Cu<sup>2+</sup> in a suitable buffer. Besides the sheer amount of peptide bonds, the AA composition of the peptides and proteins influences the reaction. The extinction measured is converted into protein concentration using a set of known protein standards. While complex comparable protein mixtures can be compared with high accuracy, for non-complex or highly different proteins comparisons may be slightly off due to differences in AA composition.

For the standard curve 10 µl protein standards solved in water were prepared in duplicates in a 96-well plate. 13 µl of water were added to each standard well while two wells received 23 µl water as background measurement. To both background and standard wells, 2 µl of the sample buffer were added. For each sample, two wells containing 23 µl of water were prepared and 2 µl of sample were added. 200 µl of BCA working reagent (196 µl Reagent A plus 4 µl Reagent B) were added to the prepared samples and the plate was incubated for 30 min at 37°C. Extinction was measured and a

standard curve was calculated. Reagent A and B were commercially available from Thermo Fischer Scientific. Standard concentrations were 1/0.7/0.5/0.3/0.15/0.075 µg BSA per µl.

### **2.3.3. Boiling SDS extraction**

Boiling SDS extraction usually was carried out when only small samples of cells were to be processed. As quantification of protein concentration is not possible due to interfering substances in the buffer, defined numbers of cells were used as starting material.

Cells were collected by trypsinisation and counted. An aliquot of cells containing the desired number was transferred to a fresh cup, pelleted at 500g for 1 min, resolved in PBS, pelleted again and then resolved in 1x SDS sample buffer. After brief vortexing the samples were boiled at 95°C for 5 min and briefly spun down (16,000g) to pellet the released genomic DNA.

#### **4x SDS sample buffer**

125 mM	Tris pH6.8
3.2%	SDS
16%	Glycerol
200 mM	DTT
0.05% W/V	bromophenol blue

### **2.3.4. Protein precipitation**

Proteins in aqueous solution can be precipitated by addition of an organic solvent. Addition of solvent reduces the dielectric constant of the solution and thereby favours protein aggregation. Those aggregates can then be pelleted by centrifugation.

To one volume of protein solution, 4 volumes of cold acetone (-20°C) were added. After mixing by vigorous vortexing for 10 seconds, the mixture was incubated 30 min to overnight at -20°C. Subsequently, proteins were pelleted by centrifugation in a precooled (4°C) centrifuge at 25,000g for 20 min. The supernatant then was discarded and the pellet was washed with 500 µl 75% acetone (-20°C) by gentle inversion (five times) and a 20 min incubation at -20°C. After centrifugation (5 min, 25,000g, 4°C) the supernatant was carefully removed and the pellet air dried. Proteins were resolved by boiling in 1x SDS sample buffer for 5 min at 95°C while vigorously shaking (1200 rpm).

## **2.4. Western Blot**

The western blot methodology [131,132] combines methods to separate proteins by their size, transfer them to a membrane and detect specific epitopes using antibodies. Those antibodies, and thereby their target epitopes, can then be visualised using a second antibody coupled to a detection chemistry. The current gold-standard of detection is the use of fluorophores, as different wavelengths allow for detection of several different proteins on the very same membrane without individual signals interfering.

### **2.4.1. SDS-PAGE**

If proteins were to be resolved by their molecular size, SDS polyacrylamide gel electrophoresis was used. Beforehand proteins were boiled in SDS sample buffer, achieving three things: Firstly disulphide bridges are broken by a reducing agent (dithionthreitol, DTT). Secondly proteins are unfolded by SDS. Thirdly the charge of the unfolded protein chain is covered by the negative charge of the SDS. The proteins then can be separated through a gel matrix made of polyacrylamide strands that are cross-linked by bisacrylamide. The range of resolution depends on the concentration of acrylamide used. By combining a low percentage, acidic and thus fast-migration stacking gel with a high percentage, basic,

slow running separation gel and a tris-glycine running buffer samples can be focussed into distinct bands.

SDS gels were cast in a horizontal electrophoresis chamber. After setting up the casting stand, glass plates were cleaned by rubbing the plates with gloved hands while rinsing them with water. Plates and spacers were assembled and leak tightness was tested. 5 or 20 ml of separation gel mixture was prepared for each mini- or midi-gel, respectively. The separation gel was cast, leaving about 1 cm for the stacking gel, and carefully layered with isopropanol. After 45 min, the stacking gel mixture was prepared (2.5 ml per mini-gel, 5 ml per midi-gel), isopropanol was poured off and the stacking gel was cast. A comb was inserted and the gel was left to polymerize for 45 min. When preparing gels the day before the actual separation, gels were wrapped in wet tissue paper and kitchen foil and stored at 4°C overnight. For protein separation, gels were inserted in the chamber and tris-glycine running buffer was added to match the fill line.

The desired amount of sample was prepared and boiled for 5 min at 95°C. In case of RIPA buffer, an aliquot containing a set amount of total protein was diluted to a desirable volume using water and mixed with SDS sample buffer. Depending on epitope, 10-20 µl sample containing 10-50 µg protein in case of RIPA extracts or boiling SDS extracts corresponding to 10,000 to 100,000 cells were loaded per lane. 1.5 µl of prestained molecular weight marker was loaded in one well of each gel to monitor the separation process and estimate molecular weight in the downstream readouts.

<b>10 % separation gel*</b>		<b>5 % stacking gel*</b>	
10%	mixed acrylamide*	5%	mixed acrylamide*
0.10%	SDS	0.10%	SDS
375 mM	Tris pH 8.8	125 mM	Tris pH 6.8
0.05%	TEMED	0.10%	TEMED
0.05%	AMPS* <sup>2</sup>	0.06%	AMPS* <sup>2</sup>

\* A 30% solution of acrylamide and bisacrylamide mixed in a ratio of 37.5:1 was used. Final percentage of total acrylamide varied between 7.5 to 12.5% for the separation gel depending on epitope.

\*<sup>2</sup> AMPS was always added last, afterwards gels were cast immediately.

<b>1x SDS running buffer</b>	
25 mM	Tris
250 mM	Glycin
0.10%	SDS

adjust pH to 8.3 using glycine

#### **2.4.2. Protein transfer**

After separation by electrophoresis, proteins were transferred to nitrocellulose membrane by applying an electrical field. Two methods were used: the semidry transfer protocol and the tank transfer protocol. For both protocols the gel was placed on a membrane and sandwiched between filter papers drained in transfer buffer. All layers of the filter/gel/membrane stack were carefully assembled to prevent air bubbles and cut to the same size to ensure a homogenous electric field. As proteins are migrating to the cathode, in the final setup the membrane was placed between the cathode and the gel.



For the semidry blot, two layers of filter paper were placed on each side of the gel/membrane assembly. The stack was placed on the flat electrode (anode) of the transfer cell. A second flat electrode (cathode) was gently pressed on the stack. Transfer was carried out at 1 mA per cm<sup>2</sup> membrane surface at room temperature for 60 min per mm gel thickness.

For the tank blot protocol, three layers of filter were used. The stack was packed between two fibre pads and locked in a gel cassette. This gel cassette was then assembled into an electrode assembly and submerged into transfer buffer. A cooling pack was inserted into the transfer tank to keep temperature close to 4°C. Transfer was carried out at 100 V for 60 min per mm gel thickness.

**1x transfer buffer**

48 mM Tris  
1.3 mM SDS  
20% Methanol  
adjust pH to 9.2 using glycine

**2.4.2.1. Antibody incubation and detection**

To detect a specific protein of interest in a sample, an antibody solution was added to the membrane. The antibody binds the protein of interest and can be then targeted by a secondary antibody raised against the constant region of the antigen-specific, so called primary, antibody. The secondary antibody itself is covalently coupled to a detectable molecule, usually a fluorophore.

After transfer, the membrane briefly was rinsed in tris buffered saline (TBS) to remove residual SDS and methanol. TBS was replaced by blocking solution, 3 % bovine serum albumin (BSA) in TBS (TBS-BSA), completely covering the membrane. The membrane was gently rocked at room temperature for 2h to saturate the membrane with protein. A suitable buffer for the primary antibody incubation was prepared and used for dilution of the antibody. Different antibodies required different dilutions and different buffers, but all buffers used were based on TBS supplemented either with or without 0.1% Tween-20 and/or 3% BSA. The membrane was briefly washed with the residual antibody incubation buffer and afterwards sealed in a small plastic bag along with the diluted antibody. This bag was then incubated gently rocking at 4°C overnight. The day after the bag was opened, the primary antibody discarded and the membrane washed five times by 5-minute-washes with TBS containing 0.1% Tween20 (TBST). The respective fluorescence-labelled secondary antibody was diluted 1:20,000 in TBST and incubated with the membrane for 2h at room temperature rocking gently. After removal of the secondary antibody the membrane was washed thrice for 5 min per wash with TBST and twice with TBS. At that point, the membrane was either scanned using a suitable device or stored in TBS for later scanning. All steps including and subsequent to the secondary antibody dilution were carried out in darkened vessels to reduce photobleaching.

**tris buffered saline**

20 mM Tris  
140 mM NaCl  
adjust pH to 7.6

### 2.4.3. Coomassie stain

If proteins were to be stained in gel, one common method is the coomassie stain. The dye binds to the protein via van der Waals forces, preferably via basic amino acids. Upon binding, the dyes absorption at 595 nm increases. The same dye and colour shift is also used in the Bradford protein assay.

To stain an SDS-PAGE using coomassie, it was submerged in at least 5 gel volumes of coomassie staining solution. After incubation at room temperature and gentle rocking, the stain was removed and replaced by destaining solution. Depending on desired sensitivity, the staining incubation ranged from 30 minutes to overnight. For destaining, the gel was washed four times in staining solution, firstly for 30 seconds, then for 5 minutes, for 20 minutes and lastly for one hour. Afterwards, the gel could be scanned using a standard flatbed scanner or the Odyssey scanners IRD700 channel. The lower limit of detection was 5 ng per band.

coomassie staining solution		coomassie destaining solution	
0.035%	Coomassie G250	40%	Methanol
40%	Methanol	10%	Acetic acid
10%	Acetic acid		

### 2.4.4. Silver stain

A second method for staining protein is the Silver Stain. Silver ions get complexed by proteins and subsequently reduced in a developing buffer, leading to formation of a visible precipitate of metallic silver. This precipitate then auto-catalytically induces formation of more precipitate, thereby leading to a development over time.

SDS-PAGE gels were agitated in fixation solution for 30 min and washed by two quick two min washed and 1 hour wash with water. The gel was sensitised using 0.02% sodium thiosulfate for 2 minutes and then washed twice for 10 sec each with water. Next the gel was submerged in 0.1% Silver nitrate for 30 min at 4°C. After two more 30 sec washes with water, the gel was developed until the desired level of staining was achieved. Development was stopped by removing developer and washing and incubating the gel with 1% acetic acid.

silver fixation solution		silver developing solution	
50%	Methanol	0.04%	Formaldehyde
5%	Acetic acid	2%	Sodium carbonate

## 2.5. Cell culture

Cultured cells provide a readily accessible and easily to handle model system. Cells can be derived from several tissues and thus correspond to different cell types. They often exist in two different states, a proliferating state that can be suppressed by chemical inducers or upon confluence and a differentiated state that is induced by suppressing proliferation and addition of suitable hormones and small molecules. The proliferating cell state usually is easier to manipulate and handle. Cell cultures fall in three categories: Primary cells, immortalised primary cells and clonal cell lines. Primary cells are directly derived from animal tissue. They are obtained by tissue section followed by mechanic fragmentation and collagenase digestion. Different cell types can be separated by centrifugation and plating procedures, albeit contamination with non-desired cell types is common. Primary cells usually have a low proliferation potential and constantly need to be replaced but are considered to be very close to their parent tissue. Primary cells can be immortalised by viral insertion of genes that enforce proliferation, in most cases the SV40 large T-antigen (SV40-LTA). This procedure increases their

proliferation capacity but the viral infection- and integration procedure, especially when combined with selective pressure, can influence cell features that might be of interest. Random integration also generates a pool of cells that are genetically dissimilar due to different integration loci and for some applications expression of virus-harboured genes might not be desirable. Immortalised cells tend to genetically drift because some subpopulations are more proliferative than others. This drift often can change the cells characteristics over time. Clonal cell lines are clones arising from a single immortalised cell by clonal expansion. This immortality can be naturally occurring (for example, a cancerous cell), selected for by repeated passaging, or induced by the experimenters influence. As they usually have been in culture for years or decades, they often have lost few or several characteristics of their parent tissue. On the other hand, they are well characterised, easy to handle, stable and easily accessible.

### 2.5.1. General workflow

All cell culture work was performed in a laminar flow bench or in closed, sterile vessels outside of the bench. The bench was kept sterile by UV irradiation and use of 70% alcohol. All work was performed with gloved, sterilized hands. All reagents and plastic ware used were either gamma irradiated, autoclaved, filtered sterile or prepared from sterile components inside the laminar flow bench. Buffers and media coming into contact with cells were pre-warmed to 37°C unless stated otherwise. All cells were incubated in a dedicated cell culture incubator at 37°C, 5% CO<sub>2</sub> and 100% relative humidity. General handling procedures, centrifugation, microscopy or other manipulations were performed at room temperature.

### 2.5.2. Passaging cultured cells

Due to their permanent division, cells need to be subcultured periodically. As all cells used were growing as an adherent monolayer, cells had to be detached and re-plated. When cells were to be subcultured, medium was removed by suction and cells were briefly washed with phosphate buffered saline (PBS). PBS was removed and a trypsin-EDTA solution was added to the cells. Cells were put back in the incubator and checked for detachment in regular intervals. Usually cells were detached after 1-3 min at 37°C. Upon detachment, 2-3 times the trypsin volume of fresh proliferation medium were added to suppress trypsin activity and detachment was aided by pipetting up and down three times using a sterile pipette. Cells were transferred into a sterile falcon tube and pelleted by centrifugation (200g, 2 min). After careful resuspension using a 1000 µl pipette tip, cells were either counted and seeded at defined numbers or split by a fixed ratio according to growth area. A split ration of 1:10 indicates that either 10% of the cells detached were replated onto a vessel of same growth area or that all cells were seeded on an area ten times bigger than before detachment.

Vessel	growth area	Medium	PBS wash	Trypsin
15 cm dish	152 cm <sup>2</sup>	20 ml	10 ml	5 ml
10 cm dish	58 cm <sup>2</sup>	7 ml	5 ml	2 ml
6 cm dish	21 cm <sup>2</sup>	4 ml	2 ml	1 ml
3.5 cm dish	8 cm <sup>2</sup>	2 ml	1 ml	0.5 ml
single 6-well	9.6 cm <sup>2</sup>	2 ml	1 ml	0.5 ml
single 12-well	3.5 cm <sup>2</sup>	1 ml	1 ml	0.25 ml
single 24-well	2.0 cm <sup>2</sup>	0.75 ml	0.5 ml	125 µl
single 48-well	0.75 cm <sup>2</sup>	0.4 ml	250 µl	75 µl
single 96-well	0.35 cm <sup>2</sup>	200 µl	150 µl	30 µl

### 2.5.3. Counting cultured cells

To count cells they were detached, pelleted and resuspended as described above. Depending on growth area, cells had to be diluted to a convenient density. Two different methods of cell counting were used.

The first method of counting was using a Neubauer counting chamber. The chamber provides a chamber of 0.1 mm height if a cover slip is correctly placed. Two different grids yield areas of defined size. 15  $\mu$ l of diluted cell suspension was added next to the cover slip from where they entered the chamber by capillary force. Using a microscope, five or more squares in a suitable grid were counted to yield about 50 to 100 counted cells. Knowing cell count, chamber height, grid area and dilution factor, the number of cells per volume could be calculated.

The second method of counting was based on the coulter principle. If a cell solution is pulled through an orifice, the electrical resistance between two electrodes at opposing sides of the orifice changes whenever a cell passes. The magnitude of change is mainly dependent on the size of the cell. For counting cells using that principle, the handheld Scepter Cell Counter (Millipore, Billerica, USA) was used according to manual.

### 2.5.4. Thawing, proliferation and freezing of cultured cells

For long time storage, cells were frozen in the gaseous phase of liquid nitrogen at around  $-150^{\circ}\text{C}$ . To prevent freezing damage, cells were frozen in DMSO (dimethylsulphoxid) containing medium.

For thawing cells, a cryo-stock of cells was removed from the nitrogen storage tank and rapidly thawed in a  $37^{\circ}\text{C}$  water bath. As soon as the medium was thawed, cells were carefully removed from the vial using a pipette containing 6 ml proliferation medium. To remove traces of cryomedium, cells were pelleted in a falcon tube (200g, 2 min). From that point, cells were seeded as done after trypsin detachment.

For preparing cryo-stocks, cells were detached according to the subculturing protocol. After pelleting they were resuspended in cold freezing medium and aliquoted into cryo-vials at 1 ml per vial. Usually  $5 \cdot 10^5$  to  $10^6$  cells were frozen per vial. The stocks were then pre-frozen at  $-80^{\circ}\text{C}$  overnight and transferred to the nitrogen storage tank the day after.

	<b>freezing medium</b>	<b>Vessel</b>	<b># of cryo-stocks</b>
70%	basal medium*	15 cm dish	8 - 15 vials
20%	FCS	10 cm dish	3 - 5 vials
10%	DMSO	6 cm dish	2 - 3 vials
*	DMEM of HIB1b basal medium	3.5 cm/6-well	1 - 2 vials

### 2.5.5. HIB1b cells

HIB1b cells [133,134] are a clonal cell line derived from a mouse hibernoma, a tumor in brown fat tissue. The tumors were induced by generating transgene mice expressing SV40 early genes (including LTA) under the aP2 promoter. Initially cells were shown to express UCP1 protein after differentiation and stimulation, but expression was only minute and cannot be reproduced any more, possibly due to extensive passaging since the time of cell line generation. HIB1b cells were kindly provided by the Spiegelman lab (Dana-Farber Cancer Institute, Harvard Medical School). HIB1b cells were proliferated in a specific basal medium supplemented with 10% fetal calf serum (FCS). The medium is based on Dulbecco's Modified Eagle medium (DMEM) and the F12 nutrient mixture. Cells were passaged every 2-3 days, shortly before reaching confluence. Cells are usually split 1:5 to 1:10. For differentiation cells

were grown to confluence and exposed to induction medium for 48 hours. Afterwards, cells were exposed to differentiation medium for six more and stimulated with agonists of choice for another 4 to 24 hours. During differentiation fresh medium was supplied every other day. For reporter gene assays usually a shorter differentiation protocol was used. For HIB1b, vessels were collagen coated.

<b>Hib1b basal medium</b>		<b>proliferation medium</b>	
12 g/l	DMEM:F12*	90%	HIB1b basal medium
14 mM	NaHCO <sub>3</sub>	10%	FCS
16.4 µM	Biotin	<b>differentiation medium</b>	
4.2 µM	Calciumpanthotenate	93%	HIB1b basal medium
2.25 mM	Glutamine	7%	fetal calf serum
15 mM	Glucose	20 µg/ml	Gentamycin
13.5 mM	HEPES pH7.4	1 nM	T3
	adjust pH to 7.4, filter sterile	20 nM	Insulin
	*contains 17.5 mM Glucose	<b>additional, for induction medium</b>	
		125 µM	Indomethacin
		500 µl	IBMX
		5 µM	Dexamethasone

### 2.5.6. C2C12 cells

C2C12 cells are a fibroblast-like, mononuclear myoblast clonal cell line. The cell line was generated by serial and selective passaging of mouse muscle cells from C3H mouse 3 days after a crush injury. The cell line is clonal and differentiates into multinuclear myotubes upon confluency. Differentiation is increased by replacing fetal serum by adult serum and/or reducing serum concentration. C2C12 cells were proliferated in DMEM (high glucose, +L-glutamin, +NaHCO<sub>3</sub>, +pyridoxine-HCl) supplemented with 10% FCS. Cells were passaged every 2-3 days at a dilution of 1:10 to 1:20. Cells always were split before reaching confluence to prevent loss of differentiation capacity. If cells were to be differentiated, they were grown to 70-100% confluence and treated with differentiation medium. After 8 days of differentiation with medium replacement every other day, cells were fully differentiated and could be stimulated.

<b>C2C12 proliferation medium</b>		<b>differentiation medium</b>	
90%	DMEM	98%	DMEM
10%	fetal calf serum	2%	adult horse serum
		20 µg/ml	Gentamycin

### 2.5.7. iBPAs

Immortalised brown preadipocytes (iBPAs) are a pool population of stromavascular fraction cell from BAT of new-born mice. After section and collagenase digestion, the cells were infected with an integrating virus expressing the large T-antigen. Two independent batches of iBPAs were used: iBPA-L1 was kindly provided by the Patrick Seale lab (Institute for Diabetes, Obesity and Metabolism, Philadelphia) and iBPA-L2 by the Alexander Pfeifer lab (Institute of Pharmacology and Toxicology, University of Bonn). The main difference between those two batches of iBPAs is that iBPA-L1 were generated using a retrovirus and puromycin selection while iBPA-L2 was generated using a lentivirus without any selection step.

iBPA cells were essentially cultured like HIB1b cells apart from three distinct differences: The medium composition was different, the cell numbers used per area were lower due to larger cell size and the differentiation after induction only lasted 4 days. The protocol for iBPA-L2 cells slightly differentiated from the iBPA-L1 protocol. At day -4 cells were seeded at 3000 cells/cm<sup>2</sup> in proliferation medium. If the cells were just used for proliferation, they were split again at day -1. If cells were to be differentiated, medium was changed to differentiation medium at day -2, to induction medium at day 0 and supplied with fresh differentiation medium at day 2 and day 4. At day 6, cells were fully differentiated.

<b>iBPA proliferation medium</b>		<b>additionally, for induction</b>	
90%	DMEM	125 µM	Indomethacin
10%	fetal calf serum	500 µl	IBMX
<b>additionally, for differentiation</b>		5/2 µM*	Dexamethasone
20 µg/ml	Gentamycin	* 5 µM for L1, 2 µM for L2	
1 nM	T3		
20 nM	Insulin		

### **2.5.8. Oil Red O Stain**

Oil Red O is a fat soluble dye used for staining lipids in cell culture and tissue sections. Fat accumulation is considered a differentiation marker for BAT and WAT cell lines.

A stock solution of 5 g Oil Red O per litre of isopropanol was prepared. At the day of staining 3 volumes of Oil Red O stock solution were diluted with 2 volumes of water and incubated at room temperature until needed. Right before usage, the diluted Oil Red O solution was filtered through Whatman paper. Medium was removed, cells were washed with PBS and fixed in 3.7% formaldehyde in water for 1 hour. Formaldehyde solution was removed and the cells were briefly washed with PBS. Then cells were layered with the diluted and filtered Oil Red O solution and incubated for 1 hour at room temperature. Afterwards, cells were rinsed extensively with warm tap water and covered with 80% glycerol. At that point pictures were taken.

### **2.5.9. Collagen coating of culture plates and dishes**

Collagen is a matrix molecule that facilitates attachment of cells to the coated surface. To coat cell culture vessels with collagen a 50 µg/ml solution of collagen was prepared from a commercial available collagen solution. 1 ml of the diluted solution was added for each 10 cm<sup>2</sup> growth area and incubated overnight at room temperature. The day after the solution was removed, the plates were left to dry and afterwards briefly washed with water. Prepared that way they could be used for about 2 weeks.

## 2.6. Transfection

Transfection is the process of introducing DNA into a eukaryotic cell of interest. Several different methods are known and each cell line has a different susceptibility to different transfection methods. Three main categories of transfection methods are known: Precipitate-forming transfection, electroporation and viral infection, each having different benefits and drawbacks. During all transfection procedures, antibiotics were omitted.

### 2.6.1. Calcium phosphate transfection

Calcium phosphate transfection [135] is a precipitate forming transfection. DNA is mixed with a  $\text{CaCl}_2$  solution and then added to a phosphate buffer. A calcium-phosphate precipitate is formed that settles on the ground of the culture vessel where it can be taken up by the cells. The method was primarily used for transfection of HEK293 and PlatE cell lines where it yields very high efficiency and allows transfection of huge cell numbers for at almost negligible cost.

One day before transfection, cells were split to yield 50-70% confluency at the day of transfection. The next morning cells were in their logarithmic growth phase and were split again onto the desired transfection vessel at a density of 100,000 cells per  $\text{cm}^2$ . 3-6 h later two reaction tubes were prepared for each transfection. For generation of 1 ml precipitate solution, in one tube 25  $\mu\text{g}$  of DNA was diluted to 450  $\mu\text{l}$  in 0.1x TE buffer. Afterwards, 50  $\mu\text{l}$  2.5 M  $\text{CaCl}_2$  solution was added mixed by vortexing. In the second tube 500  $\mu\text{l}$  2x HEBS buffer was prepared. The  $\text{CaCl}_2$ -DNA mixture then was added dropwise to the HEBS buffer while gently vortexing to ensure permanent and even mixing. The mixture was incubated for 5 min at room temperature. Afterwards, the mixture was briefly vortexed again and added dropwise to the cells. 16 hours later cells received fresh medium. Maximal transgene expression was usually reached 48 hours after addition of precipitate.

1x TE for transfection		2x HEBS		solution used per vessel	
10 mM	Tris pH 7.4	11 mM	Glucose	1000 $\mu\text{l}$	each 10 cm dish
1 mM	EDTA pH 8	42 mM	HEPES	400 $\mu\text{l}$	each 6 cm dish
	yields a pH of 7.6	10 mM	KCl	200 $\mu\text{l}$	each 6-well
CaCl <sub>2</sub> for transfection		276 mM	NaCl	100 $\mu\text{l}$	each 12-well
10 mM	Tris pH 7.2	1.9 mM	$\text{Na}_2\text{HPO}_4$	30 $\mu\text{l}$	each 48-well
2.5 M	$\text{CaCl}_2$		adjust pH to 7.05 – 7.10		
	yields a pH of 7.2				

### 2.6.2. Liposome mediated transfection

Another precipitate forming method of transfection is lipofection [136]. In this method a lipid formulation, usually partly based on cationic lipids, is complexed with DNA to form lipid vesicles loaded with DNA. Those vesicles are added to the vessel and taken up by the cells. A wide variety of lipid formulations is available all containing different proprietary and secret components. For experiments carried out during this thesis, Lipofectamin LTX and PlusReagent (Invitrogen, Carlsbad, USA) were used. The protocol presented is optimised for high throughput transfection in 96 well plates but can be scaled up or down if desired. Liposome mediated transfection was used for HIB1b, C2C12 and iBPA cells. As for calcium phosphate transfections, two vessels were required but for complex transfection schemes mixtures were often prepared in multiwell plates with the help of multichannel pipette.

Cells were prepared as for calcium phosphate transfection, although different cell numbers were seeded. In a multiwell plate for each transfection DNA was diluted in 10  $\mu\text{l}$  OptiMEM reduced serum medium. After all DNAs of interest were added, 1  $\mu\text{l}$  PlusReagent per  $\mu\text{g}$  DNA was added. Lipofectamin

LTX was diluted at 10  $\mu$ l using OptiMEM and briefly mixed. Within 5 min the Lipofectamin mixture was added to the DNA mixture and incubated 30 min at room temperature. 20  $\mu$ l of that mixture were added to each 96 well. The day after cells received fresh medium.

Cell line	Cells/well	ng DNA	$\mu$ l Lf-LTX
HIB1b	30,000	100	0.25
iBPA	20,000	100	0.5
C2C12	20,000	150	0.5

### 2.6.3. Nucleofection

Nucleofection is an electroporation method able to directly deliver nucleic acids into the nucleus. Nucleofection treats cells in defined salt solution with different electrical pulses. Under suitable conditions the pulses briefly open pores in both cellular membrane and nuclear membrane, allowing nucleic acids to enter. Exact composition of buffer and the electrical parameters of the pulses are of essence and differ for different cell lines. Nucleofection was used for non-proliferating cells or in situation where very high efficiency was required.

Nucleofection was carried out according to the standard protocol. Briefly, DNA was mixed with the respective nucleofection buffer while cells were detached by trypsin digestion. Defined numbers of cells were pelleted (90g, 10 min) and medium was removed. Cells were resuspended in DNA containing buffer, transferred to the electroporation cuvette and electroporated in the nucleofector device. Two different devices (Nucleofector II and Nucleofector 96) for different cell numbers were available.

Cell line	device	# of Cells	ng DNA	buffer	program
HIB1b differentiated	NF96	500,000	500	SE	CM137
HIB1b proliferating	NF96	500,000	500	SE	CM137
iBPA 24h after induction	NF96	250,000	500	SE	DS138
iBPA proliferating	NF96	250,000	500	SE	DS138
NIH-3T3 proliferating	NF96	250,000	500	SG	EN158
HEK293 proliferating	NF96	500,000	500	SE	CM130
HIB1b differentiated	NFII	5,000,000	2500	T	T020
iBPA 24h after induction	NFII	2,500,000	2500	T	T020

### 2.6.4. Viral transduction

Transfection by viral infection, sometimes also referred to as transduction, is a method to deliver nucleic acids into a cell using artificial retrovirus particles. These particles are able to enter a cell and release their nucleic acid genome. Viral particles are generated using a packaging cell line where all components of the virus are produced and assembled. Four virus types are most commonly used: basal retroviridae, lentiviridae (a complex subspecies of retroviridae), adenoviridae and adeno-associated viridae. In this project the viral vectors used were based on moloney murine leukemia virus (MMLV), a basal retrovirus.

Retrovirae are ssRNA viridae that contain two copies of mRNA genome. A classic retroviral genome contains two long terminal repeats (LTRs), which are promoter sequences needed for genome replication and integration, a packaging signal ( $\Psi$ ) and two open reading frames (ORFs), gag-pol and env, carrying the information for production of retroviral proteins. The gag-pol genes share a common reading frame while the env gene is translated from its own reading frame with the help of alternative splicing. The gag gene encodes the inner structure of the virus, including capsid and matrix



proteins, while pol encodes the functional proteins, including reverse transcriptase, integrase and protease. Env encodes the outer hull proteins including a receptor-binding glycoprotein and thereby influences the host range, the tropism.

Viral entry into a cell begins with binding of the glycoprotein to a surface protein of a suitable host cell. Upon binding the virus fuses with the cell and releases its contents into the cytoplasm where the reverse transcriptase generates cDNA from the genomic mRNA. The cDNA now can be integrated by the integrase into the host genome. Once inserted into the host genome, the host cell's transcription and translation machinery produces the protein products of gag-pol and env as well as the full length mRNA genome of the virus. The genomic mRNA, identified by the presence of  $\psi$ , is then packaged in the viral capsid and hull. Furthermore integrase, reverse transcriptase and protease proteins are included as well. As for some species of retroviridae (including MMLV) the integration machinery cannot pass the nuclear membrane they are reliant on cell division for their integration.

#### **2.6.4.1. PlatE cells**

Platinum E cells are a cell line that is used for generating ecotropic MMLV particles. The cell line is based on the HEK293 cell line but carries two additional gene clusters in its genome. These clusters contain parts of the retroviral genome fused to a mammalian strong constitutive promoter and a resistance gene and thereby produce the viral proteins in selectable, bicistronic ORFs. The two clusters are pEF1 $\alpha$ -gag/pol-IRES-bs<sup>R</sup> and pEF1 $\alpha$ -env-IRES-puro<sup>R</sup> with pEF1 $\alpha$  being the promoter of elongation factor 1 $\alpha$ , gag/pol and env being viral structure proteins, IRES being an internal ribosome entry site and bs<sup>R</sup>/puro<sup>R</sup> being resistance genes against blasticidin and puromycin, respectively. This transfer of viral genes to the nucleus of a cell allows production viral particles transporting a mRNA of interest as long as the mRNA contains a  $\Psi$  and the LTR sequences required for integration. Furthermore, these viral particles cannot leave an infected target cell due to the absence of viral protein sequences. The genomic viral mRNA usually is transcribed from a plasmid that carries  $\Psi$  and the LTRs as well as other sequences (promoters, resistance genes) of interest and a multiple cloning site (MCS).

Platinum E cells were cultured according to standard procedures, very much like HIB1b cells and iBPA cells. Transfection was performed by calcium phosphate transfection. There is no differentiation or induction of PlatE cells. Once a month cells were subjected to selection pressure for two passages to eliminate cells that have lost expression of gag-pol and env.

#### **PlatE proliferation medium**

90%	DMEM
10%	fetal calf serum
1 $\mu$ g/ml	Puromycin*
10 $\mu$ g/ml	Blasticidin*

\* only during selective passages

#### **2.6.4.2. Production of retroviral particles**

For production of viral particles, PlatE cells were seeded onto collagen coated 6-cm dishes and transfected with 10 µg retroviral genome plasmid according to the calcium phosphate transfection protocol. 16 hours after transfection, cells received 2 ml fresh medium. 48 hours after transfection, the medium, now containing the viral particles, was carefully removed and filtered through a 0.45 µm cellulose acetate filter. The filtered supernatant medium then could be used for infection while the cells were discarded.

#### **2.6.4.3. Infection**

24 hours before infection, cells were seeded onto 6 well plates at a density of 1500-3000 cells/cm<sup>2</sup>. For each supernatant three wells were prepared. The harvested supernatant was diluted with fresh proliferation medium in three dilutions: 800 µl + 1100 µl, 200 µl + 1700 µl and 50 µl + 1850 µl. To these mixtures 100 µl medium containing 50 µg/ml polybrene were added, yielding a final concentration of 2.5 µg/ml. After mixing, the cell's medium was replaced by the polybrene-supplemented supernatants. 16 hours later supernatants were removed and the cells received fresh medium.

#### **2.6.4.4. Selection**

24 hours after removal of viral supernatants, most viral particles have completed infection and integration. Cells received fresh medium including a selective antibiotic to get rid of non-infected cells. When reaching confluence, cells were subcultured using medium free of selective agents. 3 hours after subculturing the selective agent was added. 72 hours after infection, wells showing comparable infection efficiency were selected by their GFP fluorescence and then used for experiments. When cells reached confluence again cultures were free of non-infected cells and the selective antibiotic was omitted.

### **2.7. EMSA**

The electrophoretic mobility shift assay (EMSA) is an assay to test DNA-protein interaction in vitro. A non-denatured protein sample, usually a nuclear extract, is incubated with a synthetic short DNA probe that resembles a genome region of putative regulatory function. After incubation, the reaction mixture is run on a non-denaturing PAGE to separate unbound, fast migrating probe from slower migrating protein complexes. Detection of probe and probe containing complexes is possible via a label bound to the probe, usually a radioactive phosphate or a fluorophore.

#### **2.7.1. Nuclear protein extraction**

To obtain protein extracts for EMSA, cell nuclei were purified and protein from nuclei was released. Two features of nuclei were utilised for purification: Their very high density compared to other organelles and their robustness against shearing and detergents. Two protocols were used for generation of nuclear extract: One high throughput protocol for quick preparation of extracts from several small samples [137] and one high yield protocol for obtaining high amounts of nuclear extract from high cell numbers. All steps of nuclear extraction were performed on ice using precooled solutions and devices.

For small scale nuclear protein extractions, cells obtained from a single 6-well or 6 cm dish were used. Cells were washed with PBS and scraped in 1 ml scraping solution. The scraped cells were pelleted (1500g, 5 min), washed in 500 µl scraping solution, pelleted again (12,000g, 15 sec), resuspended in 400 µl swelling solution and allowed to swell for 15 min. Afterwards, 25 µl of 10% Igepal-CA630 were added and the samples were vortexed vigorously for 10 seconds. This treatment disrupted the cytoplasmic membrane but left the nuclei intact, which then were pelleted (15 sec, 12,000g). The

supernatant, containing RNA and cytoplasmic proteins, was discarded, the pellet was resuspended in 50 to 150 µl nuclei elution buffer (about 2 pellet volumes) and vigorously rocked at 4°C for 15 min. Due to the high salt concentration in the buffer, transcription factors detach from chromatin and can leave the nucleus whose pores were widened during the swelling step. Nuclear debris containing the membrane fraction and chromatin was pelleted by centrifugation (5 min 25,000g) and the supernatant was aliquoted and stored at -80°C.

For large scale extraction from HIB1b cells, usually from one to several 15 cm dishes, cells were washed with PBS and scraped in homogenisation solution (500 µl per 15 cm dish). Cells were transferred to a glass potter and homogenised by ten strokes with a glass pistil. The homogenate was aliquoted into reaction tubes (1 ml per tube), incubated for 10 min and nuclei were pelleted at 3300g for 15 min. The pellets were resuspended in 250 µl low salt buffer per 15 cm dish and allowed to swell for 1 min. An equal volume of high salt buffer was added and the samples were vigorously rocked at 4°C for 30 min. Afterwards, debris was pelleted by centrifugation (30 min 25,000g) and supernatants were pooled, aliquoted and stored at -80°C.

<b>scraping solution</b>		<b>swelling solution</b>		<b>nuclei elution solution</b>	
1x	PBS	10 mM	HEPES pH7.9	20 mM	HEPES pH7.9
1:1000	inhibitors*	10 mM	KCl	400 mM	NaCl
		0.1 mM	EDTA	1 mM	EDTA
		0.1 mM	EGTA	1 mM	EGTA
		1 mM	DTT	1 mM	DTT
		1:200	inhibitors*	25%	Glycerol
				1:200	inhibitors*
<b>homogenisation buffer</b>		<b>low salt buffer</b>		<b>high salt buffer</b>	
10 mM	HEPES pH7,9	20 mM	HEPES pH7,9	20 mM	HEPES pH7.9
1.5 mM	MgCl <sub>2</sub>	1.5 mM	MgCl <sub>2</sub>	1.5 mM	MgCl <sub>2</sub>
10 mM	KCl	20 mM	KCl	1200 mM	KCl
0.5 mM	DTT	0.5 mM	DTT	0.5 mM	DTT
20 mM	NaF	20 mM	NaF	20 mM	NaF
1:1000	inhibitors*	0.2 mM	EDTA	0.2 mM	EDTA
		25%	Glycerol	25%	Glycerol
		1:200	inhibitors*	1:200	inhibitors*

\* protease inhibitor cocktail and phosphatase inhibitor cocktail

### 2.7.2. Probe preparation

Probes for EMSA usually are double stranded DNA molecules of 20 to 60bp. The single, complementary oligonucleotides were synthesized separately and annealed to double strands.

Complementary DNA oligonucleotides resembling the region of interest were ordered from MWG biotech with one strand being end-labelled with a Cy5-dye. After delivery, oligonucleotides (oligos) were resolved at a concentration of 100 µM in water. 9 µl of each oligonucleotide and 2 µl oligo annealing buffer were combined in a reaction tube, placed in a falcon tube and covered by aluminium foil to prevent photobleaching. 2 l water was heated to 85°C in an erlmeyer flask and the assembly containing the probes to be annealed was submerged. The temperature was kept at 85°C for 5 min and then allowed to slowly cool down to 4 °C overnight. Next day, 4 µl DNA loading buffer were added and the probes were resolved on a 20x20 cm 12 % TBE-PAGE at 4°C, protected from light. After resolution, DNA was visualised by placing the gel on a fluorophore (F254) imbued silica plate and

irradiation with a 260 nm UV lamp (DNA shading). Shading gel pieces containing the double stranded probe were excised, transferred to a fresh reaction tube and grinded using a tube pestle in 150 µl of TE buffer. After shaking the mixture overnight at 4°C, the probe had left the gel pieces by diffusion and was separated from gel by centrifugation (6000g, 1 min) through the filter of a filter tip. DNA concentration was measured by absorptiometry, probes were stored at -20°C.

<b>Oligo annealing buffer</b>		<b>12% TBE-PAGE</b>		<b>5x TBE buffer</b>	
100 mM	Tris pH 8.0	12%	mixed acrylamide*	445 mM	Tris
10 mM	EDTA pH 8.0	0.5x	TBE	445 mM	Boric acid
1 M	NaCl	2.50%	Glycerol	10 mM	EDTA pH 8.0
		0.05%	TEMED		yields a pH of 8.3
		0.08%	AMPS* <sup>2</sup>		

\* A 40% solution of acrylamide and bisacrylamide mixed in a ratio of 37.5:1 was used.

\*<sup>2</sup> AMPS was always added last as it started polymerisation.

### 2.7.3. Protein-DNA incubation

Incubation of probe and nuclear extract will usually lead to formation of several complexes. How many complexes are formed not only depends on the sequence of the probe and the nuclear extract composition, but also on concentration of both, buffer conditions and presence of competitors and cofactors. Thus, it is crucial to test different conditions and throughoutly test which complexes are specific and which are not. Complexes formed are resolved on a 5.3% TBE-PAGE and detected via the fluorescence of the probe of interest while other components like unbound protein or non-labelled competitor DNA and its complexes remain invisible. All steps are performed on ice or in a 4°C environment, all components are prepared and kept on ice.

Binding buffer and poly dl:dC are diluted to yield a final volume, after protein and probe, of 10 µl. Usually 5-10 µg of protein per lane is added and incubated for 10 min. Labelled probe and, if needed, unlabelled competitor were mixed and added after the 10 min. After further 20 min of incubation, 1 µl of 10x EMSA loading buffer was added and the mixture was loaded onto a vertical TBE-PAGE where it was resolved. If addition of antibodies was desired, they were added 10 min after addition of probe. In these cases, the mixture was added for an additional 10 min. Depending on the desired resolution, page formats from 6x10 to 20x20 were used. After resolution, the gel was scanned on a Typhoon TRIO scanner for Cy5-fluorescence, usually at high sensitivity, 100 µm resolution and photomultiplier setting 650.

<b>5x EMSA binding buffer</b>		<b>5.3% TBE-PAGE</b>		<b>10x EMSA Loading buffer</b>	
50 mM	Tris pH 7.5	10%	mixed acrylamide*	250 mM	Tris pH 7.5
5 mM	MgCl <sub>2</sub>	0.5x	TBE	0.20%	Orange G* <sup>3</sup>
250 mM	NaCl	2.50%	Glycerol	40%	Glycerol
2.5 mM	EDTA	0.05%	TEMED		
2.5 mM	DTT	0.08%	AMPS* <sup>2</sup>		
20%	Glycerol				

\* A 40% solution of acrylamide and bisacrylamide mixed in a ratio of 37.5:1 was used.

\*<sup>2</sup> AMPS was always added last as it started polymerisation.

\*<sup>3</sup> Orange G was used because it does not interfere with the Cy5-signal. For other fluorophores, using another loading dye might be mandatory.

## 2.8. Reporter gene assay

For Reporter gene assays started with transfection of plasmids by lipofection or nucleofection. A mix of vectors was transfected and the mix always contained one assay vector carrying a *Gaussia* luciferase driven by the transcriptional activity of the region of interest and one vector expressing a *Photinus* luciferase under a constitutive promoter active under all conditions in all cell lines used, the CMV promoter. The latter of these vectors is needed as a normalisation vector to control for transfection efficiency and cell number. Depending on the experiment, expression vectors were included, either expressing a gene of interest or a miRNA targeting the transcript of interest. As for lipofection always a minimum amount of DNA was required, these transfections had to be filled up with unrelated DNA, in our case a modified pcDNA3 vector that had its CMV promoter removed.

Depending on the type of experiment, the readout was carried out after different amounts of time. For overexpression, agonist treatment experiments or comparison of reporter vectors, reporter activity was measured after 48 hours. For experiments using RNAi mediated knockdown of target genes, readout happened 96-120 hours after transfection, thereby allowing for depletion of the target protein by its turnover. To prepare cells for measurement of luciferase activity, medium was removed by suction and cells were washed with PBS. To every well 30 µl 1x passive lysis buffer (PLB) was added. At that time, cells could be stored at -80°C for later measurement.

For the actual measurement, frozen or fresh cells covered with 1x PLB were rocked gently at room temperature for 20 minutes. Using an 8-channel pipette, 20 µl of lysate was transferred to a white, non-transparent 96-well plate. The actual measurement took place in a Tecan Infinite M200 device that had two injectors, of which one was used to inject LAR2 buffer while the other injected SNG buffer. Measurements were carried out well by well. After each injection, luminescence was measured. Both reporters could be measured in the same well as buffer conditions in the buffers were chose to only allow activity of one of both enzymes. PLB, LAR2 and SNG buffers were components of the Dual Luciferase Assay System (Promega, Fitchburg, USA).

typical reporter gene transfection*		luciferase measurement protocol	
15 ng	UCP3-promoter driven <i>Gaussia</i> vector	50 µl	injection LAR2
5 ng	CMV driven <i>Photinus</i> vector	2 sec	orbital shaking
50 ng	expression plasmid	1 sec	resting time
30 ng	empty/modified pcDNA3	5 sec	measurement: <i>Photinus</i>
* for transfection of one 96-well of HIB1b cells using lipofection		50 µl	injection SNG
		2 sec	orbital shaking
		1 sec	resting time
		5 sec	measurement: <i>Gaussia</i>

## 2.9. Semi-quantitative and quantitative real time PCR

In many cases abundance of mRNA for a transcript correlates with the transcription of the respective gene. Furthermore, high mRNA levels usually lead to high protein levels and are an indicator for high protein amounts present in the cell. Despite several factors are known that can dissociate the link between mRNA abundance, transcription and protein abundance, the fact that the correlation holds true in many cases and the ease of measuring mRNA abundance make it a common choice for estimation of gene expression.

mRNA abundance is measured by PCR. The method requires preparation of mRNA from the sample and reverse transcription of mRNA into complementary DNA (cDNA) using a virus-derived reverse transcriptase. This, usually single stranded, cDNA can then be amplified in a standard PCR using gene specific primers. As the starting amount of template, at least for a saturating number of cycles, does not correlate with the final yield of product, the amount of product is either measured at set cycles during the logarithmic phase (semiquantitative) or after every cycle (quantitative, real time). While the readout for semiquantitative PCR is agarose gel electrophoresis, in real time quantitative PCR (qPCR) an intercalating dye, SybrGreen, is added to the reaction mixture. This fluorophore binds to double stranded DNA and then shows greatly increased fluorescence. Fluorescence is measured after every cycle. A serial dilution of template allows for an estimation of amplification efficiency and thus allows comparing different pairs of primers. Unless using a dilution series with defined, known amounts of template, all values obtained are relative. Besides mRNA abundance, abundance of DNA (genomic, mitochondrial, vector) can be measured. In that case, the cDNA synthesis step is omitted and the DNA sample is directly used for PCR.

### 2.9.1. RNA extraction und quantification

RNA extraction was performed by a combined phenol extraction/column purification protocol. Cells grown on multiwell plates had their medium removed and were washed with PBS. After removal of PBS, 1 ml TriSure (BioLine reagents, London, UK) was added per 10 cm growth area. TriSure, a mixture of phenol and an aqueous buffer containing guanidinium thiocyanate, lyses cells, unfolds protein und thereby extracts nucleic acids. The cells were briefly incubated, scraped and collected and 200µl chloroform was added. After vigorous vortexing (15 sec) and centrifugation (15 min, 12,000g, 4°C), the aqueous top phase (~450 µl) was transferred to a fresh tube. The interphase and bottom phase contain the majority of protein and DNA and were discarded. 500 µl 75% Ethanol were added to the top phase, vortexed and then purified further using the SV total RNA isolation system (Promega, Fitchburg, USA) according to manual. RNA concentration was determined by its extinction at 260 nm as was done for plasmid DNA. RNA then was stored at -80°C.

#### 2.9.1.1. RNA quality control

RNA quality was accessed by measuring integrity of ribosomal RNA. As total RNA was extracted, the protocol yields a mixture of rRNA and mRNA. For rRNA, yielding three distinct bands (28S: 4718 nt, 18S: 1874 nt and 5S/5.8S: 120/160 nt), degradation of RNA can easily be estimated by checking for band integrity and presence of breakdown products. Breakdown products appear as a lower molecular weight smear directly below the rRNA bands. Due to the fact that rRNA and mRNA are equally prone to degradation, rRNA intactness allows a good estimate of mRNA intactness.

Two methods were used to visualize the bands: either a conventional agarose gel electrophoresis or a chip-based capillary electrophoresis. Agarose gel electrophoresis was carried out as described above, except for a heating step to melt self-complementary secondary structures (65°C 10 min) followed by

quenching on ice before loading the gel. Capillary electrophoresis was carried out using the RNA 6000 nano chip and the bioanalyzer device (both Agilent Technologies, Santa Clara, USA) as described in the manual. While evaluation using a conventional gel was done by eye, the bioanalyzer uses basic calculations to estimate the breakdown of mRNA, yielding a RNA integrity number where 10 means no rRNA degradation.

### 2.9.1.2. cDNA Synthesis

As RNA is not a suitable template for PCR amplification, it has to be transcribed into DNA. To do so, a mix of two reverse transcriptases and a mix of poly-A tail binding primers and random hexamers was employed. Before cDNA synthesis, a genomic DNA (gDNA) digestion step was carried out to remove gDNA contaminations. All components were part of the QuantiTect Reverse Transcription Kit (Qiagen, Venlo, Netherlands) and were used according to manual. 500 ng total RNA were used in a 10 µl reaction. The 42°C synthesis step was extended to 30 min to increase cDNA yield. cDNA was stored at -20°C.

### 2.9.2. Primer design for qPCR

Primers were designed employing the Primer3 tool of the SDSC workbench. Target sequences were obtained from ENSEMBL and primers were designed to span one or several large introns to prevent amplification of gDNA. If possible, one of the last introns was picked as the 3' end is reverse transcribed more efficiently due to the use of oligo-dT primers. Primers were designed to have a melting temperature close to 60°C, a length of 20 nt, a GC content near 50% and yield an amplicon size of 70 to 250bp. Primers were ordered from MWG biotech as unmodified DNA oligonucleotides and tested and optimised in PCR. For optimisation PCRs, the very same reaction mixture and temperature protocol as used in qPCR was used. Typically, for each primer pair, four annealing temperatures (53, 55, 57 and 59 °C) and 2 primer concentrations (250 nM and 400 nM) were tested. A condition amplifying only a product of the expected size and good yield was selected and used for all subsequent experiments.

### 2.9.3. qPCR

qPCR was performed on 386-well plates in a LightCycler 480 (Roche Applied Science, Penzberg, Germany) device. Samples were measured in technical triplicates, standards in duplicates. All PCR components were supplied as a premixed solution, SensiMix SYBR no Rox (BioLine), containing SYBR green, buffer, nucleotides and a hot start polymerase. Master mixes were prepared from SensiMix, Primers and Water, distributed across the plate, and 2.5 µl template, diluted between 1:10 and 1:50 in water, was added using a matrix pipette. The plate was sealed by foil and centrifuged 10 sec at 1000g to collect all liquid material at the bottom of the plate. All dilutions and reaction setups were carried out on ice. After 45 cycles of amplification, a melting curve was generated to validate presence of a single product. When desired, some wells from the plate were opened and the product was separated and visualized on an agarose gel.

qPCR reaction		typical qPCR protocol	
6.25 µl	SensiMix	95°C	7 min
250 nM	each Primer	97°C	10 sec
ad 10 µl	water	X°C*1	15 sec
2.5 µl	template	72°C	20 sec*2

**45x**

melting curve 60 to 95°C

\*1 melting temperature according to optimisation

\*2 fluorescence quantification was carried out at the end of every elongation step

## 2.10. DNA affinity chromatography

DNA affinity chromatography is a method where DNA-binding proteins can be purified via their DNA-binding activity. To do so, a DNA molecule is immobilised on a resin. Nuclear extracts are incubated with the DNA molecule before immobilisation or with the DNA-loaded resin and excess and unbound protein is removed by washing steps. Afterwards, bound proteins can be eluted from the resin and either further purified or identified by mass spectrometry.

### 2.10.1. Resin Preparation

A frequently used method to prepare affinity resins is the activation of crosslinked agarose beads using cyanogen bromide. These activated resins then can be used to covalently bind molecules containing primary amines. Several molecules can be immobilised that way with proteins, DNA and heparin being the most common choices. In the process, the cyanogen bromide (CNBr) activates the hydroxyl sidechains of the agarose forming highly reactive imidocarbonates and cyanate esters. When keeping the pH in a desirable range between 11 and 11.5, the more stable imidocarbonate groups are preferably generated.

First, ligands were solved in coupling buffer. Agarose beads were fined three times as following: The desired amount of beads, usually 15-20 g (50 ml settled bed volume), were added to 200 ml water and inverted a few times. After 5 min, the intact beads had settled while small fragments were still floating. The water containing the floating fragments was decanted. After the last wash, the fined beads, a stirring bar and about 50 ml water were transferred to a beaker where they were slowly stirred at room temperature. Temperature and pH were monitored using an electrode. 5 g CNBr were crushed using a mortar and slowly added to the mixture while gently stirring the beads. pH was kept between 11 and 11.5 by addition of 1M NaOH, temperature was kept between 20 and 25 °C by addition of small pieces of ice. When the drop in pH slowed, the reaction was close to finishing and the beads were quickly transferred to a suction filter, washed with 3-4 vol of water and coupling buffer, and, as soon as the beads reached the state of a moist compact cake, transferred to a falcon tube containing the ligand in coupling buffer. The falcon tube was filled to 50 ml using coupling buffer and then incubated overnight at 4°C mixing head-over-end. Next day, the liquid was removed using a suction funnel and the resin was washed with 3 volumes of blocking buffer. The concentration of ligand in coupling buffer and washing buffer was measured to calculate coupling efficiency. Unreacted groups were blocked by incubation in blocking buffer at 4°C overnight. At the third day, the resin was washed with 5 volumes coupling buffer, 5 volumes acetate buffer and 5 volumes TRIS-EDTA (TE) buffer. The resin then was stored at 4°C in TE containing 10 mM sodium azide.

<b>coupling buffer</b>		<b>blocking buffer</b>		<b>acetate buffer</b>	
0.1 M	NaHCO <sub>3</sub>	0.2 M	Glycin	0.1 M	Sodium acetate
0.5 M	NaCl		adjust pH to 8		adjust pH to 4: acetic acid
	adjust ph to 8.4				
<b>typical amounts of ligand per ml resin</b>					
50 nmol	Oligonucleotide				
2000 U	Heparin				



### 2.10.2. EMSA optimisation of binding conditions

To optimise binding conditions for affinity chromatography, different parameters were sequentially optimised in electrophoretic mobility shift assay (EMSA). Essentially, parameters were evaluated by running a dilution series of the component of interest. Of all dilutions assayed, the one with the best ratio of specific complex formation over non-specific binding was chosen. The procedure was carried out as described by Moxley et. al. [129]. Conditions optimised were DNA and protein concentration, DNA- and non-DNA-competitors and detergents.

### 2.10.3. Oligonucleotide trapping

A purification strategy proposed by Gadgil et. al. [130] is oligonucleotide trapping. Trapping means that complex formation between DNA and protein is done in solution and that the formed complex then is captured by an anchor molecule on the beads. Oligonucleotide trapping refers to that anchor being a 10 nt ssDNA molecule that can capture the complex via a complementary overhang of the DNA probe. The anchor is coupled to CNBr activated resins via an amino-modification added during synthesis. Trapping was performed in column-scale (see Figure 6), allowing high sample amounts and yields but only a low amount of samples in parallel with limited reproducibility. 0.5 ml settled resin volume was trapped between two glass-fibre frits in a small plastic column. A connector and tubing allowed buffer flow from a reservoir to the column assembly. Buffer flow was either driven by gravity or by a syringe-based perfusion pump. For gravity based setups, a cock allowed to stop the buffer flow. Before preparation of the binding reaction the column assembly was washed with 20 ml binding buffer, ensuring that the assembly is free of air bubbles. 50 ml reaction mixture was prepared according to EMSA optimisation. After probe incubation, the sample was loaded into a buffer reservoir, usually a 50 ml syringe, and slowly allowed to pass the column at rates around 1-2 ml per minute. The assembly was then washed with 20 ml of binding buffer and eluted using 10 ml binding buffer containing 350 mM NaCl. During the process, every 5 min a 0.5 ml fraction is collected, with exception of the elution where all fractions are constantly taken. Fractions were assayed for activity using EMSA. Of each fraction, 2.5  $\mu$ l were incubated with 12.5  $\mu$ l binding mixture containing binding buffer and probe, but no NaCl.

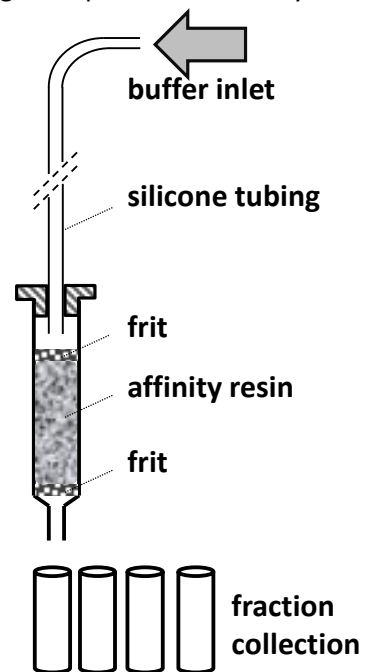


Figure 6: Column affinity chromatography setup

### 2.10.4. Heparin affinity chromatography

Heparin is a polymer of repeating disaccharide units ranging from 6 kDa to 30 kDa. In affinity chromatography, heparin is used for two properties: Its polyanionic structure is bound by many DNA- and RNA-binding proteins and it has a limited cation exchange capacity. Heparin, immobilised on beads, is a common tool to pre-purify DNA binding proteins by fractionation and background reduction. Washing steps remove non-DNA-binding proteins, while stepwise elution helps to separate different DNA-binding proteins into different fractions. Depending on the for subsequent purification steps this separation of transcription factors into different fractions can be beneficial or detrimental.

To pre-purify DNA binding proteins, 1-2 ml nuclear extract was diluted 7-fold with heparin affinity buffer without KCl (HA-0) and ran over the equilibrated column containing 1 ml of heparin agarose beads as done for oligonucleotide trapping. After washing with 10 ml HA buffer with 100 mM KCl (HA-

100), the proteins were eluted by nine 1 ml steps of increasing KCl in binding buffer ranging from HA-200 to HA-1000. Fraction collection and activity monitoring was done as in oligonucleotide trapping.

**Heparin affinity (HA) buffer**

20 mM	HEPES pH 7.9
0-1000 mM	KCl *
5 mM	MgCl <sub>2</sub>
8%	Glycerol
0.5 mM	DTT
1:1000	Inhibitors * <sup>2</sup>
100 µg/ml	Insulin

\* 0 mM for HA-0 (nuclear extract dilution), 100 mM for washing, 200-1000 mM for elution

\*<sup>2</sup> protease inhibitor cocktail and phosphatase inhibitor cocktail

**2.10.5. Preparative EMSA**

After oligonucleotide trapping, EMSA was used to further purify the eluate from column chromatography. The two highest activity fractions were pooled and diluted five times with binding buffer containing 50 nM probe. Usually, the pool was split in two halves of which one received a probe binding the proteins of interest, while one contained a mutated probe. The samples were run on a 20x20 cm page of 1.5 mm gel thickness with one large preparative pocket and two small analytic pockets which were loaded with standard nuclear extract binding reactions for orientation. After separation, the gels were scanned and the scans were printed out at 1:1 scaling and used as a positioning pattern to accurately cut out the complex formed with the binding probe and the respective region for the mutated probe gel. The gel fragments were grinded in 3 volumes of PAGE extraction buffer and incubated shaking at 4°C overnight. Next day, gel fragments were removed by filtering and the protein solution was concentrated by acetone precipitation.

**PAGE extraction buffer**

0.1 M	Sodium acetate
0.1%	SDS
50 mM	DTT

yields a pH of 8.4

**2.10.6. Magnetic bead purification using biotin/avidin interaction**

Magnetic beads provide an affinity support that can easily be handled using a strong magnet. Instead of pelleting them by centrifugation, they can be drawn to the side of a reaction tube, allowing quick, easy and reproducible washing. Their main drawback is their lower bead volume and thus bead surface, which is why they usually are used in small scale one step batch purification strategies. All steps were carried out at 4°C.

For Magnetic bead purification streptavidin-coated Dynabeads MyOne-T1 (Invitrogen, Carlsbad, USA) were used. Two different biotin-labelled probes resembling the 31bp around the IVS1+1505 element were annealed, one with a 14 bp spacer, one without. 1000 fmol probe were immobilised on 250 µg beads according to the manufacturers protocol. Afterwards, they were washed three times with 500 µl 1x EMSA binding buffer and incubated for 30 minutes with 1.5 ml reaction mixture. Beads were washed three more times and eluted by incubation for 5 min in 1x binding buffer containing 350 mM NaCl. Activity was assayed as in oligonucleotide trapping

### MagBead reaction mixture

50 µl/ml	Nuclear Extract
0.10%	CHAPS
2 ng/µl	Salmon sperm DNA
1x	EMSA binding buffer

### 2.10.7. Promoter-trapping

Promoter-trapping [138] is a form of affinity chromatography using larger DNA probes that cover several putative binding elements. As the binding was too complex to display well interpretable EMSA signals, optimisation was carried out in small batch purifications followed by western blot against putative binding proteins. The basic principle is similar to Oligonucleotide trapping, but a different immobilisations strategy was used to trap the probe on agarose beads (Pierce Biotechnology). The beads carried the NeutrAvidin protein which was able to bind the biotinylated DNA probe.

#### 2.10.7.1. Probe preparation

Probes were generated by PCR. A biotinylated and a Cy5-labelled primer were used to amplify the region of interest using the HerculaseII polymerase. The *Phodopus sungorus* UCP3 reporter gene constructs described above were used as template and thereby allowed generation of probes containing different mutations and deletions. Afterwards, the probes were gel-purified via an agarose gel.

#### HerculaseII probe amplification

100 µl	5x Hercul buffer
500 nM	each Primer
250 µM	each dNTP
125 ng	reporter vector
5 µl	HerculaseII

#### PCR protocol

95°C	2 min	
95°C	20 sec	
55°C	20 sec	<b>35x</b>
72°C	30 sec	
72°C	1 min	
4°C	hold	

#### 2.10.7.2. Purification

For each reaction 50 µl bead slurry, containing 25 µl settled beads, was washed twice by addition of 250 µl binding buffer, inverting four times, pelleting of beads (1 min, 1000g) and removal of supernatant. Afterwards, probe was diluted in 250 µl binding buffer, added to the beads and incubated rotating for 1h. Probe-bound beads were pelleted, supernatant was removed and beads were washed three times with 450 µl binding buffer containing 0.05% igepal CA-630. 150 µl reaction mixture was added, containing binding buffer, protein, 0.05% igepal CA630, Wy14643, rosiglitazone, GW0742, ATRA, T3 and other agents as described in the respective experiments. Beads were incubated at 4°C for 30 min with the reaction mixture rotating head over end, spun down, washed three times with 350 µl binding buffer containing 0.01% Igepal and finally eluted twice. For elution, a high salt buffer was added, incubated for 2 min, beads were spun down and supernatant was collected. Activity was checked in EMSA using the 31 bp IVS1+1505G probe. Binding of DNA to beads was assayed by measuring Cy5-fluorescence in the probe solution before and after bead incubation. All steps were carried out at 4°C.

## 2.11. Candidate identification via mass spectrometry

Candidate identification was carried out in collaboration with the Proteomics and Bioanalytics department of the Technical University of Munich, which was headed by Professor Dr. Bernhard Küster. All steps subsequent to purification and SDS-PAGE (sample preparation, mass spectrometry and peptide mapping) were performed by Hannes Hahne and Fiona Pachtl.

Samples were separated by SDS-PAGE prior to in-gel tryptic digestion. In-gel trypsin digestion was performed according to standard procedures. LC-MS/MS measurements were performed on an amaZon ETD mass spectrometer (Bruker Daltonik, Bremen, Germany) coupled to an easy-nLC (Proxeon, DK). Peptides were separated on a self-packed 0.075x20 cm reversed-phase column (Reprosil, Dr. Maisch, Ammerbuch, Germany) using a 110 minutes linear gradient (2-40% acetonitrile in 0.1 % formic acid, flow rate 300 nl/min). Intact masses of eluting peptides were determined in enhanced scan mode and the five (ten) most intense peaks were selected for further fragmentation by collision-induced dissociation (CID) and acquisition of fragment spectra in ultra-scan mode. Singly charged ions were rejected. Dynamic exclusion was enabled and dynamic exclusion duration was set to 10 seconds. Peaklist files were generated using DataAnalysis 4.0 (Bruker Daltonik, Bremen, Germany) and database searches were performed using the Mascot search engine version 2.3 (Matrix Science, London, UK) with a parent ion tolerance of 0.3 Da and a fragment ion tolerance of 0.5 Da against the IPI mouse database (v3.26, 52,735 sequences). Enzyme specificity was set to trypsin and up to two missed cleavage sites were allowed. Carbamidomethylation of cysteine residues (57.01 Da) was set as fixed modification, variable modifications included acetylation of protein N-terminus (42.01 Da) and oxidation of methionine (15.99 Da). Search result files were imported into Scaffold 3.0 (Proteome Software, Portland, Oregon). Threshold parameters were set as follows: protein probability, 95%; minimum number of peptides, 1; peptide probability, 0%, yielding a false-discovery rate of 1.6% on protein level and 1.95% on peptide-spectrum match level.

## 2.12. Chromatin immunoprecipitation

The most direct way to test interaction between a transcription factor and its putative binding element is chromatin immunoprecipitation (ChIP). The advantage is that interactions are assayed on level of nuclear genome and its natural chromatin environment. For ChIP, the protein-DNA interactions in a life cell are covalently fixed by exposing the cell to formaldehyde-containing medium. Then, after a crude extraction of nuclei the chromatin is sheared into small fragments. Using specific antibodies, fragments containing the transcription factor of interest are purified. Crosslinking is reversed by heating and the protein component is removed by proteinase K digestion from the purified chromatin fragments, leaving a mixture of DNA fragments. These fragments can then be assayed for elements of interest using PCR or next generation sequencing.

### 2.12.1. Primer selection

The basic primer design principles and tools were the same as for qPCR on mRNA, only the selected target regions were different. The template for qPCR on ChIP samples were fragmented gDNA, thus all processes of mRNA production and maturing were not relevant for target selection. To allow good sensitivity for PCR, primers with amplicons smaller than the desired fragment size were chosen, usually 50-200 bp. Primers were designed to always enclose the binding site of interest. Optimisation of PCR conditions was done as described above, but diluted and fragmented gDNA was used as template.

### **2.12.2. Sample preparation and crosslinking**

Cells for ChIP were grown on dishes or multiwell plates, depending on cell number required for the respective protocol and cell type. Fully differentiated iPAs were chosen as a suitable cell line and stimulated for 6 hours with 1  $\mu$ M Rosiglitazone and 5  $\mu$ M ATRA in addition to their usual differentiation medium. Cells were cross-linked either by addition of fresh medium containing 1% formaldehyde to the dish, or addition of PBS containing 1% formaldehyde to trypsin-detached cells. In both cases, crosslinking was carried out gently rocking for 10 min at room temperature. The crosslinking reaction was stopped by addition of Glycin to a final concentration of 125 mM and 5 min incubation. Afterwards, cells were washed with three washes of ice cold PBS according to the respective protocol. Three kits were tested for ChIP, the enzymatic SimpleChIP enzymatic system (Cell Signaling Technology, Cambridge, UK) and the sonification-based LowCell# and HighCell# ChIP kits (Diagenode, Liège, Belgium). Unless stated otherwise, all steps were done according to the standard protocol.

### **2.12.3. Chromatin preparation and shearing**

For optimal ChIP results the chromatin fragmentation needs to be in a certain range. If the fragments are low, the resolution is poor. If the fragments are too small, PCR sensitivity and fidelity might be impaired. For standard ChIP-PCR readouts, the large majority of fragments should be 200 to 800 bp long. Two different methods for chromatin shearing are commonly used.

#### **2.12.3.1. Sonification**

Sonification is a method commonly used to shear and disrupt biological structures. It can break down a various number of structures, beginning from small tissue clumps down to chromosomes. The efficiency of sonification is greatly increased by inclusion of detergents in the buffer, the most common ones being SDS, igepal CA-630 and sodium desoxycholate.

Nuclei preparation and sonification shearing were carried out according to the manual of the respective kit (LowCell# and HighCell#) using the BioRuptor (Diagenode, Liège, Belgium) device coupled to a cooling waterbath. All steps were carried out at 4°C. Briefly, defined numbers of cells were collected by trypsinisation, washed with PBS and lysed in a detergent containing buffer. After pelleting nuclei, they were resolved in SDS-containing buffer and sheared by sonification. Chromatin was sheared for 20 cycles (30 sec on/30 sec off) at high intensity. Debris was pelleted and the supernatant containing the chromatin was diluted to SDS concentrations of around 0.1%. Immediately, fragments were captured by incubation with antibody-loaded magnetic beads overnight. The day after, beads were washed and DNA was eluted by heating to 65°C, proteinase K digestion and boiling at 99°C. The exact protocol can be found in the respective manual.

#### **2.12.3.2. Enzymatic digestion**

A second common method to break down chromatin is using micrococcal nuclease. This enzyme can cut the DNA component of chromatin in between the nucleosomes, thus generating fragment sizes of a multiplicity of 150 bp, which is the DNA wound around one nucleosome. For ChIP, ideal fragment size is one to four nucleosomes, thus 150 to 600 bp.

Nuclei preparation and digestion shearing were done using the SimpleChIP enzymatic kit system protocol and components. All steps beside digestion were carried out at 4°C. Briefly, nuclei from  $4 \times 10^6$  cells were purified by detergent lysis and washed with digestion buffer. Nuclei containing chromatin were exposed to 3500 gel units micrococcal nuclease for 20 min at 37°C. Digestion was stopped by addition of EDTA and cooling to 4°C. Nuclei were pelleted, resuspended in ChIP buffer and ruptured by sonification using the BioRuptor device (10 cycles 30 sec on/30 sec off, high intensity). Debris was

pelleted and the supernatant was stored in aliquots at  $-80^{\circ}\text{C}$ . After measuring DNA content of chromatin, defined amounts of chromatin were diluted, incubated with antibody overnight and then immobilised by incubation with magnetic beads for 2 hours. After washing, chromatin was eluted, heated, proteinase K (Fermentas, Vilnius, Lithuania) digested and DNA was purified using spin columns.

### **2.12.3.3. Optimisation of chromatin fragmentation**

For optimisation of chromatin fragmentation, purification as described above was carried out. Two different optimisations were carried out. First, four different crosslinking times were tested: 8, 10, 12 and 14 minutes. For each crosslinking time shearing using a standard protocol and purification with SP1 antibody was carried out. After purification, enrichment was measured using a primer pair amplifying a known SP1 binding region. Chromatin crosslinked for 8 min did not yield any amplification in the subsequent PCR while crosslinking for 10 min gave satisfying PCR amplification. Thus all further experiments were carried out with 10 minutes of crosslinking. Second, after crosslinking and purification of nuclei, six aliquots were prepared. Each aliquot was sheared slightly different. For sonification, different amounts of shearing cycles were used (0, 12, 16, 20, 25, 30 cycles) on nuclei corresponding  $10^6$  cells. For enzymatic digestions, different amounts of enzyme were added (0, 2000, 3500, 5000, 10000, 15000 gel units) to nuclei of  $8 \cdot 10^6$  cells. Afterwards, DNA was purified as following: Chromatin was diluted to 480  $\mu\text{l}$  using PBS. 20  $\mu\text{l}$  5M NaCl and 25  $\mu\text{g}$  RNaseA/T1 (Fermentas, Vilnius, Lithuania) mix were added. After incubation at  $37^{\circ}\text{C}$  for 30 minutes, 20  $\mu\text{g}$  proteinase K was added and the mixture was further incubated for 2h at  $65^{\circ}\text{C}$  to reverse crosslink and digest protein. After cooling to room temperature, 900  $\mu\text{l}$  Phenol-Chloroform-Isomylalcohol (25:24:1) was added and mixed by vortexing. Phase separation by centrifugation (10 min 16,000g) yielded a clear aqueous top phase of 450  $\mu\text{l}$  containing the DNA. 45  $\mu\text{l}$  3 M sodium acetate pH 5.3 were added, followed by 900  $\mu\text{l}$  ice cold 96% ethanol. After vortexing and incubation at  $-80^{\circ}\text{C}$  for 30 min, DNA was precipitated by centrifugation (25,000g 20 min). The supernatant was removed and the pellet was washed by addition of 500  $\mu\text{l}$  ice cold 70% ethanol followed by gentle inversion and incubation at  $-20^{\circ}\text{C}$  for 30 min. After another round of centrifugation (25,000g 10 min) the supernatant as removed, the pellets were allowed to air dry and the DNA was resolved in 100  $\mu\text{l}$  water. DNA concentration was measured by absorptiometry and fragmentation was checked by agarose gel electrophoresis.

### **2.12.4. Readout, quantification, controls**

Enrichment of fragments was quantified using qPCR. 2.5  $\mu\text{l}$  undiluted eluate served as template. As input control 1-2% of the chromatin amount of a single ChIP reaction was diluted in elution buffer and treated in parallel to the eluates. The dilution curve needed for qPCR quantification was generated by serial dilution of fragmented DNA obtained during the optimisation of shearing. Despite the different template, qPCR otherwise was carried out as described above.

### **2.13. Statistical analysis**

To statistically evaluate data, SigmaStat 3.5 (Systat Software, Chicago, Illinois, USA) was used. Where indicated, data were log transformed by calculating  $\log_{10}$  of  $X+1$  with  $X$  being the respective data. Unless denoted otherwise, one-way or two-way analysis of variance (ANOVA) was used in combination with a post-hoc test by the Holm-Sidak method. Asterisks denote statistical difference to the respective control after adjustment for multiple testing. p-values in the text are only given if the difference was significant after adjusting, but the p-value given is unadjusted. When data were inter-day or inter-assay normalised, all associated values were divided by the mean of all common values unless stated otherwise.

### **2.14. Bioinformatics: Genomatix**

Transcription factor binding sites and complex binding site modules were identified using the Genomatix software package [Genomatix Software GmbH, Munich, Germany]. User defined matrices were derived by publication mining and alignment of competitor sequences from EMSA competitor screens using MatDefine. User defined matrices as well as standard matrix families from the Genomatix library were then combined into models using FastM. ModelInspector was used to search sequences retrieved from ENSEMBL for presence of modules matching to the defined models. Single TFBS were searched using MatInspector. The exact criteria used are stated in appendix 9.

### 3. Results

#### 3.1. From *Phodopus sungorus* to cell culture

In the beginning of the thesis, our working hypothesis was based on three findings: Firstly, a subpopulation of animals exists in the natural population of *Phodopus sungorus*. This subpopulation lacks UCP3 mRNA and protein in brown adipose tissue (BAT) while having close to normal UCP3 expression in skeletal muscle (SKTM). Secondly, the absence of UCP3 transcript and protein is dependent on a G to A exchange at position intervening sequence (IVS)1+1505 within the first intron of the UCP3 gene. A second intronic polymorphism, at position IVS1+2668, does not have any known influence on UCP3 expression. Thirdly, a yet unknown transcription factor binding site (TFBS) is located at or near position IVS1+1505. Binding of an activating transcription factor (TF) is reduced or lost in presence of the A-allele. Binding of this activator is required for both baseline and peroxisome proliferator activated receptor (PPAR)-agonist activated transcription in BAT. Furthermore it is likely that the TF is BAT-specific [59]. Our first goal was to find out which TF binds to the putative IVS1+1505G element.

As the phenotype of the IVS1+1505G/A polymorphism was predominantly observed in BAT, and only to a lesser amount in SKTM, the first step was to identify whether the three findings described above could be reproduced in cell culture. Three brown fat cell models were used: The clonal HIB1b cell line and two different batches of immortalised brown preadipocytes (iBPAs). For SKTM, the clonal C2C12 cell line was employed. All four cell lines are proliferating mouse cell lines that can be treated to obtain differentiated, non-proliferating adipocytes or myotubes. The cell lines were used for reporter gene assays, served as a source for nuclear extracts and were assayed for effects of different agonists for receptors known to be involved in regulation of UCP3 expression.

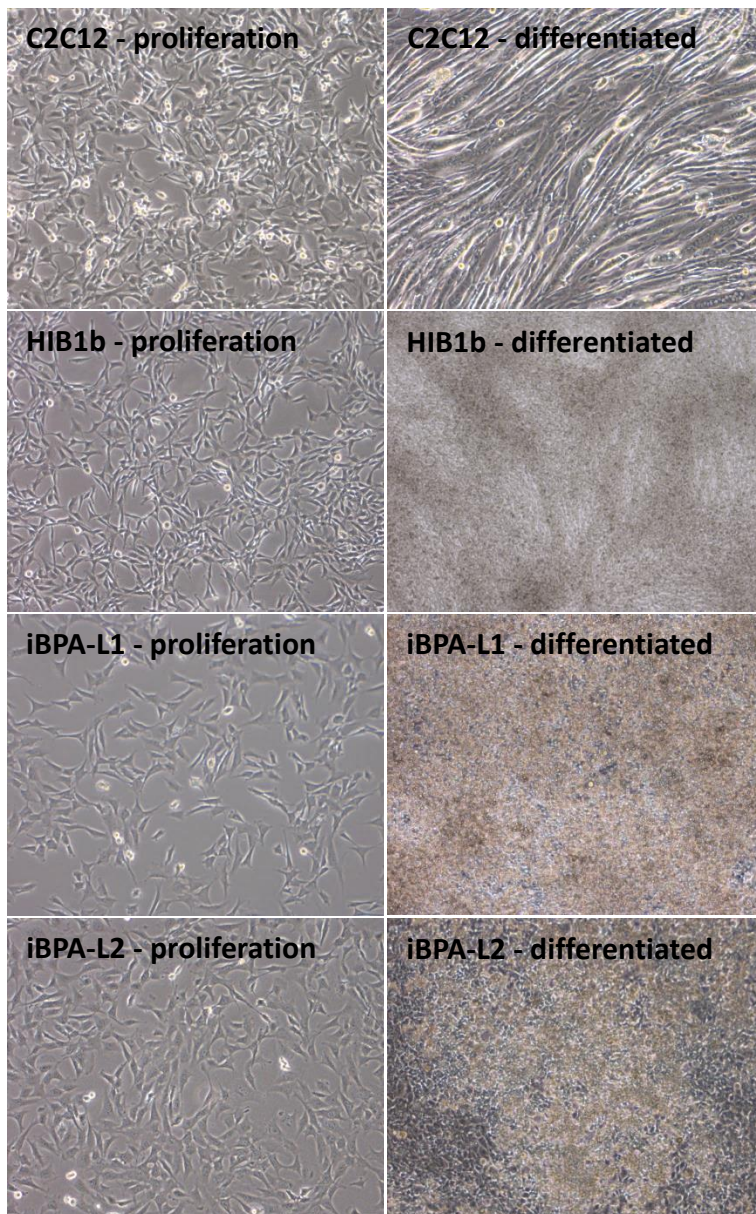
##### 3.1.1. Morphological comparison of cell lines

As initial characterisation all cells were cultured, differentiated and stained for fat accumulation. Cells were photographed using a microscope equipped with a camera. Pictures were taken before confluence and upon full differentiation (figure 7).

Before confluence, all cell lines had a fibroblast like appearance, but slightly differed in size and shape. From small to large, the size order was HIB1b, C2C12, iBPA-L1 and iBPA-L2. This difference in cell size was reflected in cell number when detaching and counting cells. HIB1b and C2C12 cells were more spindle-shaped, while both iBPA lines were more square-cut. These morphology differences intensified upon reaching confluence. After confluence, all cell lines, apart from HIB1b, showed contact inhibition, forming a monolayer. HIB1b cells are devoid of contact inhibition and can grow multilayered [133]. During differentiation, the spindle-shaped C2C12 cells fused to long and large myotube-like syncytia. The three fat cell lines started storing fat. The fat content was highest for iBPA-L1 and lowest for HIB1b cells. C2C12 cells did not store notable amounts of fat. All preadipocyte cell lines obtained a brownish tint during differentiation. Differentiation was uniform for C2C12 and iBPA-L1 and patchy for iBPA-L2 and HIB1b.

Taken together, all four cell lines proliferated and differentiated well in our hands. The BAT cell lines stored fat in multilocular droplets while the SKTM cell lines fused into myotubes. This meets the expectations for BAT and SKTM cell lines.

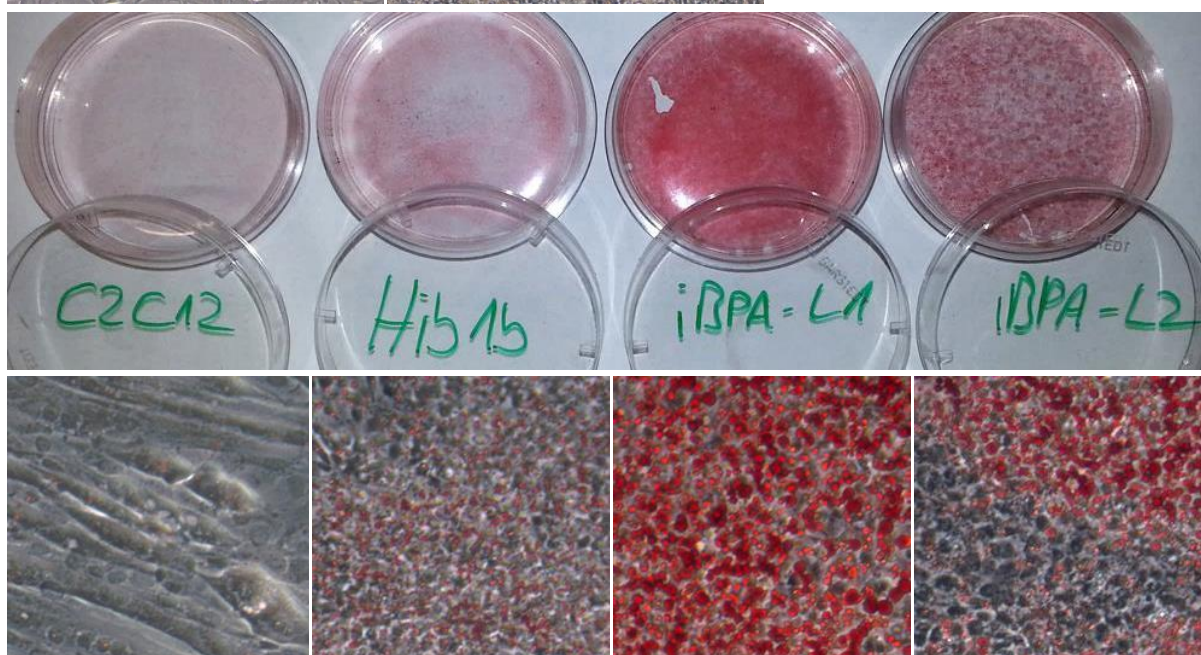




**Figure 7: Comparison of C2C12, H1B1b and the two iBPA cell lines.** Cells were proliferated and differentiated according to standard procedures. One day before confluence and at full differentiation images were taken using a camera-equipped microscope. Cells then were fixed and stained with Oil Red O, a lipophilic dye used to stain lipid droplets.

**Left:** Cells during proliferation (left column) and at full differentiation (right column).

**Bottom:** Oil Red O stained at full differentiation. For each cell line the whole dish was photographed (single picture, top row) and an image was taken with the microscope (four pictures, bottom row).

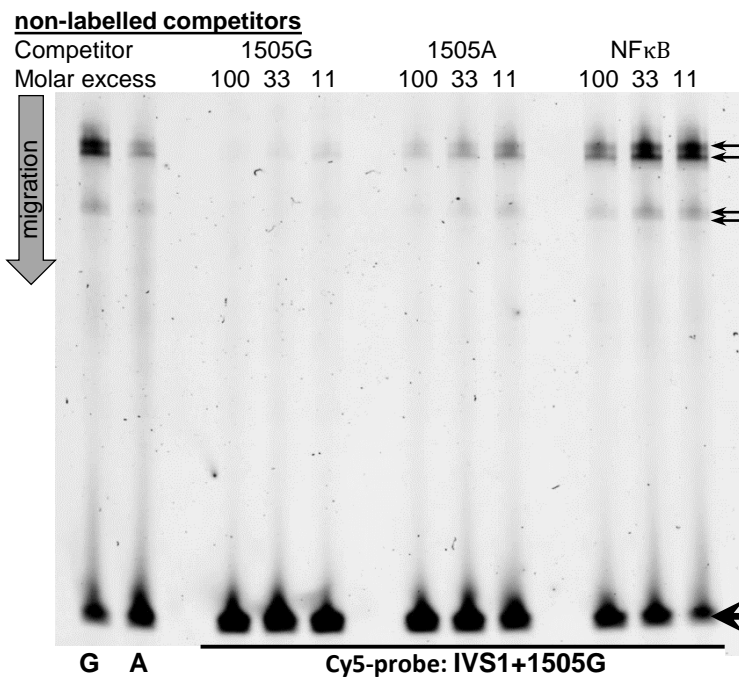


### 3.1.2. Allele specific differences in protein-DNA interaction

The starting point triggering this thesis was absence of UCP3 mRNA and protein in BAT of *Phodopus sungorus*, both on animal tissue and primary cell level [128]. The fact that the effect was heritable and depended on the intronic IVS1+1505 base exchange hinted towards a regulatory mechanism. One method to test a putative transcription factor binding site (TFBS) are electrophoretic mobility shift assays (EMSAs). In EMSA, a protein sample, usually a nuclear extract, is incubated with a labelled probe resembling the region of interest and subsequently separated on a gel. After separation, the migration distance distinguishes unbound from protein-bound probe. This way, presence and abundance of a transcription factor binding the element of interest can be assayed on the level of its DNA-binding activity. The advantage of this method is that it does not depend on knowledge about the identity of the factor.

To assay for allele-specific binding to the putative IVS1+1505 element, EMSA was carried out by incubating 10 µg HIB1b nuclear extract with 40 fmol Cy5-labelled probe in 10 µl binding buffer. Complexes were separated on a non-denaturing 5.3% tris-borate-EDTA polyacrylamid gel electrophoresis (TBE-PAGE). To distinguish unspecific DNA binding from specifically binding proteins, non-labelled probes either resembling the G-allele, the A-allele or an unrelated (here: NFκB) consensus motif were added in large molar excess. If the binding factor binds unspecifically to DNA, the visible protein-DNA-complex will be depleted by all three non-labelled probes, as they compete against the labelled probe for the binding factor. If complex formation is sequence specific, only the G-allelic probe will have this effect. Figure 8 shows the migration pattern of the probe. Four complexes with slow migration are formed with the IVS1+1505G probe, while complex formation with the IVS1+1505A probe is greatly diminished. Addition of unlabelled IVS1+1505G probe strongly reduces visible complex formation while IVS1+1505A probe has a reduced effect. The NFκB probe does only slightly impair complex formation at highest molar excess.

Together, this supports the hypothesis of an activator binding element present on the probe resembling the IVS1+1505 region. Affinity of the binding factor is high for the G-allele and low for the A-allele. The existence of four slow migrating signals indicates that of more than one protein-DNA complex is possible.



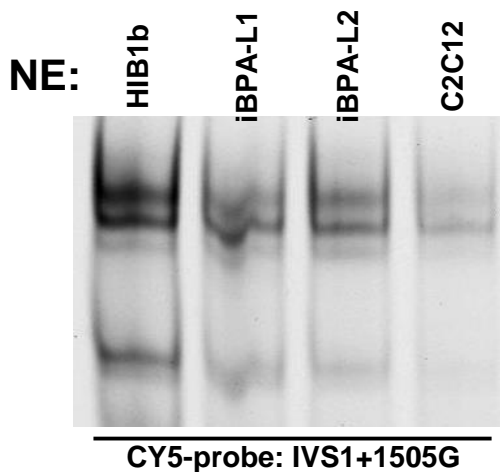
**Figure 8: Allele-specific complex formation with the IVS1+1505 element in EMSA:** HIB1b nuclear extract was incubated with a 31bp Cy5-labelled probe resembling the region around the IVS1+1505G/A polymorphism. Reactions were resolved on a 5.3% TBE-PAGE. The eleven lanes correspond to eleven binding reactions which contain either the IVS1+1505G or -A probe (denoted below). Lanes 1 and 2 compare complex formation on the G- and A-allelic probe. In lanes 3 to 11, non-labelled DNA oligonucleotides were added in 11 to 100-fold molar excess, as denoted above the lanes. The direction of migration was from top to bottom. The fast migrating signal (bold arrow, bottom) is unbound probe while the four slower migrating signals (sleek arrows, top) stem from protein-DNA complexes.

### 3.1.3. EMSA complexes are formed with nuclear extract from all cell lines

From our first experiments we knew that four allele-specific complexes are formed when incubating the IVS1+1505G probe with HIB1b nuclear extracts. We had no idea about the composition of those four complexes, but the pattern was distinct and well defined. We thus wanted to know whether the same pattern would appear when incubating the same probe with nuclear extracts from the other three cell lines. Differences in the observed pattern would hint to differences in complex composition or different modification states of the bound proteins.

Nuclear extracts were prepared from C2C12 and both iBPA cell lines in their differentiated state. Equal nuclear protein was incubated with 40 fmol 31bp IVS1+1505G probe and complexes were resolved on TBE PAGE. Figure 9 shows that all four cell lines contained the complex forming proteins. Abundance of DNA-binding activity was highest in HIB1b cells, intermediate in the two iBPA cell lines and lowest in C2C12 cells.

This demonstrates that most, if not all proteins involved in formation of the four complexes are present in all four cell lines. While differences in protein abundance were observed, no truly unique features can be found in the complex formation patterns of the different nuclear extracts.



**Figure 9: The IVS1+1505G complexes are formed with nuclear extracts from all four cell lines.** Nuclear extracts (NE) were prepared from fully differentiated HIB1b, iBPA-L1, iBPA-L2 and C2C12 cells. 10 µg of nuclear protein were incubated with 40 fmol of IVS1+1505G probe. The four reaction mixtures were resolved in four adjacent lanes. Complex formation in the area of interest of those lanes is shown left. The nuclear extract used is denoted above each lane.

### 3.1.4. Transcription factor expression in different cell lines

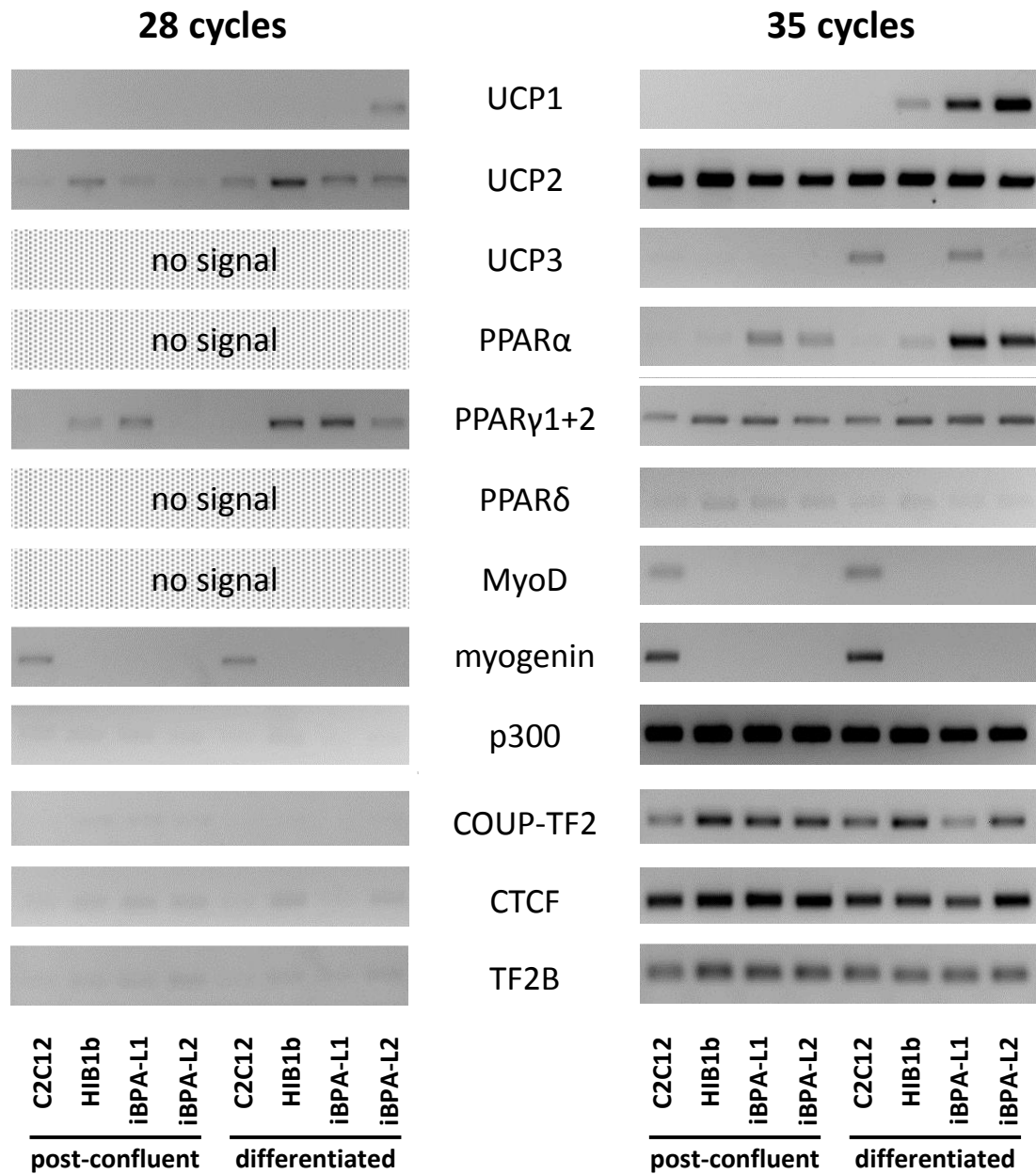
Several transcription factors have been proposed to be important for expression of UCP3. To validate expression of those factors in cell lines, the abundance of their mRNA was measured by semi-quantitative PCR. mRNA was prepared from C2C12, HIB1b, iBPA-L1 and iBPA-L2 cells at two time points: First, at confluence cells were exposed to differentiation medium for 24 hours and then harvested. Second, cells were subjected to the full induction and differentiation procedure before harvest. cDNA was generated of equal amounts of RNA. Two cycle numbers were chosen to estimate mRNA abundance: At 28 cycles highly abundant transcripts should already yield a detectable amount of product, while target mRNAs not yielding any signal after 35 cycles usually are expressed at only minute amounts if at all. All three UCPs and several transcriptional regulators including the three peroxisome proliferator-activated receptors (PPARs) were examined.

After resolution of PCR products on agarose gels (figure 10), all three UCPs could be detected. For UCP3, the highest amount of mRNA was present in differentiated C2C12 and iBPA-L1 cells, followed by a low amount in iBPA-L2 cells. UCP3 mRNA was neither detectable in HIB1b cells, nor in any non-differentiated cell. This is inconsistent with reporter gene data, where HIB1b cells show well detectable activity, even in non-stimulated conditions (figure 5, 11), and with EMSA data, where HIB1b nuclear extracts yield pronounced complex formation (figure 8). UCP1 mRNA was only detectable in differentiated brown adipocyte cell lines. The highest abundance was present in iBPA-L2 cells, the lowest in HIB1b. Lastly, all cDNA samples contained UCP2 mRNA. The amount was high, yielding detectable amounts of PCR product for all cell lines, 24 h post-confluent and differentiated, already after 28 cycles. UCP2 mRNA abundance increased with differentiation and was highest in HIB1b cells.

Of the PPARs, PPAR $\alpha$  was specific for BAT cell lines where it increased with differentiation. PPAR $\gamma$  was present in all cell lines, but abundance was higher in BAT cell lines. PPAR $\gamma$  abundance increased with differentiation. mRNA abundance for PPAR $\delta$  was barely detectable, yielding only a weak signal after 35 cycles. Abundance was comparable at both differentiation points and in all cell lines. In contrast to the PPARs, both MyoD and myogenin were strictly C2C12 specific. For both factors, no difference between differentiated and post-confluent cells was observed. Lastly, COUP-TF2 and p300 mRNA abundance was measured. Both factors were expressed at both differentiation states in all cell lines, although there was some variance for COUP-TF2. Part of this variance might be explained by variance in cDNA concentration or quality, as a comparable pattern was observed in abundance of two housekeeping mRNAs: CTCF and TF2B. The pattern of variance seen with the housekeepers was not seen for other amplicons, most likely because the variation was small compared to the physiological effect sizes or because product amounts were in saturation. For all amplicons negative controls were run in parallel. None of the negative controls yielded detectable signal (not shown).

In summary, transcription factor expression in the four cell lines was largely as expected. The pattern of MyoD, myogenin and PPARs fit to the tissue expression pattern described in literature, apart from the low abundance of PPAR $\delta$  mRNA in C2C12 cells. PPAR $\gamma$ , a key regulator for UCP3 expression, was present in all four cell lines. As all examined regulators of UCP3 expression were already expressed in post-confluent cells, we decided to do all reporter gene assays in that state. The only perturbing finding is the absence of UCP3 mRNA in HIB1b cells. In this thesis, HIB1b cells thus were only used for reporter gene assays and as a source for nuclear extracts.



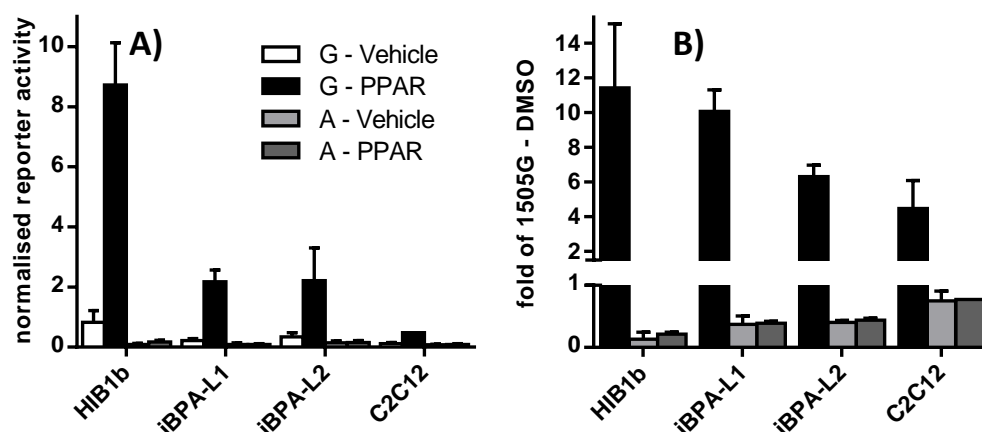


**Figure 10: Candidate factor mRNA abundance in different cell lines:** Cell lines were proliferated according to standard procedures. At confluence, one dish was treated with differentiation medium for 24 hours (post-confluent). A second dish was induced and differentiated until full differentiation (differentiated). RNA was prepared and cDNA synthesised. Doing so, eight different template cDNAs were obtained. Each image cropping shows the lineup of those eight templates. The order of templates in each cropping can be found at the bottom of each column. Every row shows two sets of PCR products for target genes denoted between the columns. The left column of pictures depicts the amount of product after 28 cycles. In the right column the amount after 35 cycles is shown. Transcripts of interest were the three UCPs, the nuclear hormone receptors PPAR $\alpha$ ,  $\gamma$ ,  $\delta$  and COUP TF2, the muscle transcription factors MyoD and myogenin as well as the histone acetylase p300. Variations seen might in part stem from difference in cDNA synthesis, as the two housekeeping genes CTCF and TF2B primers indicate, which yield slightly less signal in both C2C12 cDNAs and the differentiated iBPA-L1 cDNA. For every primer pair the negative control reaction was devoid of signal (not shown).

### 3.1.5. Reporter gene activity - allele specific difference

Next we wanted to assay whether the effect of the IVS1+1505 polymorphism on UCP3 expression can be reproduced in the different cell lines. Four *Phodopus sungorus* reporter gene constructs were available, only differing at the positions IVS1+1505 and IVS1+2668 [59]. All carried 250 bp of the core promoter, exon 1, and the whole first intron up to the start codon. As reporter gene, the constructs contained a *Gussia* luciferase (see 1.5.1.). We knew that only the IVS1+1505 polymorphism, but not the IVS1+2668 polymorphism was of relevance for activity in HIB1b and C2C12 cells (figure 5, [59]). Thus, henceforth we always compared the constructs only differing at the first position (IVS1+1505G/A) and carrying the same allele at the second position (IVS1+2668A). These two reporter gene vectors were transfected into all four cell lines. 16 hours after transfection, cells received their respective differentiation medium supplemented with T3 and Insulin. Additionally, half of the cells received PPAR agonists: BAT cell lines received Wy14643 and Rosiglitazone, C2C12 cells received Rosiglitazone and GW0742. The other half received DMSO (vehicle). After 24 additional hours, reporter gene activity was measured and activity of the UCP3 reporter was normalised to a co-transfected cytomegalovirus-promoter (CMV)-driven reporter vector.

Of all cell lines, HIB1b cells showed highest normalised UCP3 IVS1+1505G reporter activity while the two iBPA cell lines yielded intermediate activity. C2C12 cells exhibited the lowest activity (figure 11A). Interestingly, the activity of the IVS1+1505A vector was comparable in all four cell lines, and the non-stimulated activity of the IVS1+1505G construct in C2C12 cells is close to that level. In figure 11B, reporter gene activities were standardised to the activity of the respective, vehicle-treated IVS1+1505G reporter activity. This allows to easily compare allele-specific differences and PPAR agonist induction. Notably, this makes a decrease in non-stimulated IVS1+1505G activity appear as an increase of activity in all other conditions. Both agonist- and allele-effects were largest in HIB1b cells. Again, in iBPA cells we observed intermediate effects. In C2C12 cells PPAR-agonist induced induction was lowest and almost no allele-specific difference was seen in absence of PPAR agonist induction. The IVS1+1505A reporter gene vector was non-responsive to PPAR agonist treatment in all cell lines.



**Figure 11: Only the IVS1+1505G reporter is responsive to PPAR agonists in all assayed cell lines:** HIB1b, iBPA-L1 and -L2 and C2C12 cells were transfected with UCP3 reporter vectors and a control vector. 16 h later, at confluence, medium was changed to differentiation (including T3 and Insulin) medium. Half of the cells received medium containing PPAR agonists ( $\alpha+\gamma$  for BAT lines,  $\gamma+\delta$  for C2C12) the other half containing vehicle (DMSO). 24h later, reporter activity was measured. **A)** Normalized reporter activity in the four cell lines. For each cell line four bars exist, originating from two reporter gene constructs (IVS1+1505G and -A) and two treatments (vehicle and PPAR agonists) **B)** Activity standardised to the non-stimulated 1505G construct. The IVS1+1505G/vehicle bar is not shown as it always has, per definition, the value 1.

Taken together, the IVS1+1505G reporter vector was responsive to PPAR agonists in all four cell lines. This was not the case for its A-allelic counterpart. Under non-stimulated conditions, difference is absent in C2C12 cells and greatly reduced in all BAT cell lines. Importantly, the reason for this seems to be a reduced reporter activity of the IVS1+1505G construct in absence of agonist, and not an increased activity of the A-allelic construct.

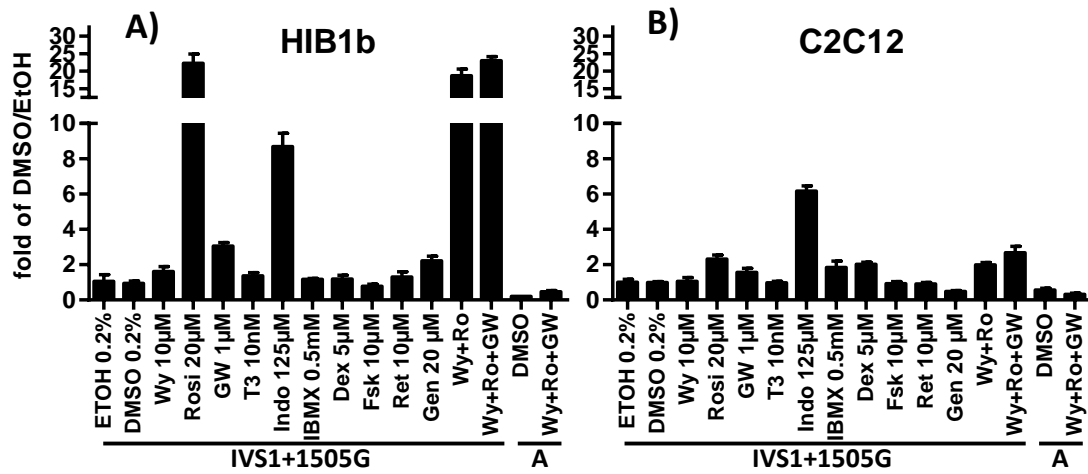
### 3.1.6. Reporter gene activity – agonist screen

As reports on the nuclear hormone receptors activating UCP3 transcription were conflicting, we assayed specific agonists for several putative regulators of UCP3 transcription for their ability to activate the UCP3 reporter gene constructs. HIB1b and C2C12 cells were transfected in their pre-confluent stage and, 16 hours later, stimulated for 24 hours.

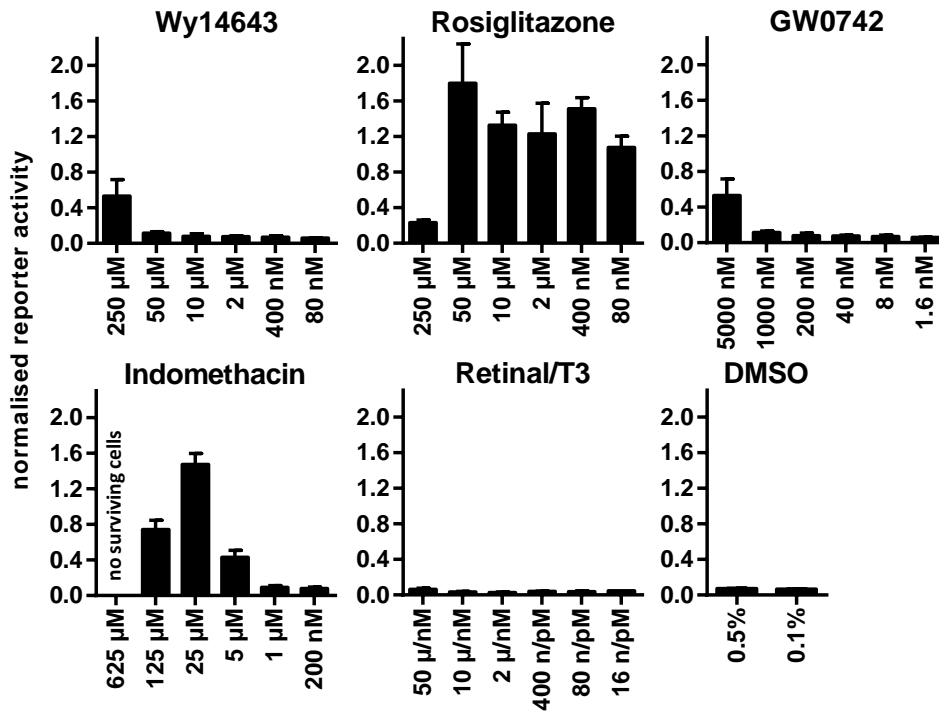
For HIB1b (Figure 12A), reporter activity strongly responded to rosiglitazone treatment. Contrary, forskolin, T3, retinal, Wy14643 and GW0742 did not influence or only mildly increased reporter activity. For forskolin, an activator of adenylyl cyclase, a transient effect was expected. Thus, shorter incubation times (4-8h) were tested, but did not lead to increased reporter activity either (not shown). Unexpected was the induction of reporter activity by indomethacin, a Cox-1 and -2 inhibitor. Cox-1 and Cox-2 are important for synthesis of prostaglandins which are believed to bind PPAR $\gamma$  in vivo [83]. For C2C12 (Figure 12B), indomethacin led to the highest increase of reporter gene activity. The rosiglitazone effect seen in HIB1b was lower in C2C12 cells. Interestingly, in C2C12 cells both agonist effects exhibited a high assay-to-assay variation with some experiments yielding a 10-fold induction of reporter gene activity, almost comparable to HIB1b cells, while others only demonstrated a 2- to 3-fold induction. The IVS1+1505A did not respond to any combination of agonists in any cell line tested.

To ensure all agonists were used at suitable concentrations, a wide range of concentrations was tested in HIB1b cells. Concentrations lower than the receptor binding constant will only yield poor activation, while at very high concentrations non-specific binding is possible. For rosiglitazone, GW0742 and T3 the binding constants are known, being 43 nM [139], 1.1 nM [140], and 0.1-1 nM [141] respectively. For Wy14643 activation of the mouse receptor is expected to happen at around 100 nM [142]. Additionally, Wy14643 and GW0742 activate PPAR $\gamma$  at micromolar concentrations [143]. For Indomethacin the inhibitory constant is species dependent and ranges from 0.1  $\mu$ M for ovine Cox-1 to 25  $\mu$ M for human recombinant Cox-2 [144,145], with the constant for Cox-1 generally being lower than for Cox-2. Retinal, or all-trans retinoic acid (ATRA), binds to RAR and activates transcription at 3-100 nM [93]. Figure 13 shows the normalised reporter gene activity for the selected agonists and vehicle (DMSO) at indicated concentrations. Wy14643 and GW0742 only stimulated reporter activity at 50  $\mu$ M and 5  $\mu$ M, respectively. In contrast, rosiglitazone stimulated reporter activity near its binding constant, at 80 nM. At 250  $\mu$ M Rosiglitazone only weakly induced reporter activity. For indomethacin, the highest induction of reporter activity happened at 25  $\mu$ M concentration. The combined treatment with T3 and retinal/ATRA did not have any effect.

Taken together, only rosiglitazone induced reporter activity at concentrations near its binding constant in HIB1b cells, thereby demonstrating that the effect indeed is mediated by PPAR $\gamma$ . Furthermore indomethacin induced activity. Unexpectedly, Wy14643 (PPAR $\alpha$ ), GW0742 (PPAR $\delta$ ), T3 (TR) and ATRA (RAR) had no significant effect alone. For high concentrations of Wy14643 and GW0742, the effects most likely stem from unspecific activation of PPAR $\gamma$ . In general, both reporter gene activity and induction were lower in C2C12 cells when compared to HIB1b cells.



**Figure 12: Agonist screen in HIB1b and C2C12:** HIB and C2C12 cells were transfected with the IVS1+1505G reporter and, 16 hours later, treated with differentiation medium containing the indicated agonists (X-axis) for 24 hours. Agonist concentrations are stated at the X-axis. For combination treatments, the same concentrations were used as for the single treatments. Data were normalised to a co-transfected, CMV-driven vector and standardised to the mean of the two vehicle controls, EtOH and DMSO. **A)** HIB1b cells. **B)** C2C12 cells. EtOH: ethanol, Wy: Wy14643, Rosi/Ro: rosiglitazone. GW: GW0742, Indo: indomethacin, Dex: dexamethasone, Fsk: forskolin, Ret: retinal, Gen: genestein.



**Figure 13: Only PPAR $\gamma$  agonists induce activity of the UCP3 reporter gene construct in HIB1b cells:** HIB1b cells were transfected with the IVS1+1505G reporter vector and a normalisation vector. 16h later, cells were stimulated with different concentrations of agonists for 24h. Besides the denoted agonists, medium contained 20 nM insulin but no T3. Reporter activity was normalised to a co-transfected, CMV-driven control vector. The six different graphs correspond to five different agonist concentration series and one vehicle control graph (DMSO). Agonists are indicated above each graph, concentrations are stated at the bottom of each column. Retinal and T3 were used in combination. The first, higher concentration given at the bottom of the retinal/T3 graph is for retinal, the latter for T3. The highest concentration of rosiglitazone and indomethacin reduced cell survival.



## 3.2. Candidate finding: DNA affinity chromatography

A classic way to purify and identify candidate proteins binding a transcription factor binding site (TFBS) is affinity chromatography coupled with mass spectrometry (MS). Affinity chromatography is used to purify the protein of interest via its immobilised binding element, while mass spectrometry can identify the proteins purified. Two different strategies were employed.

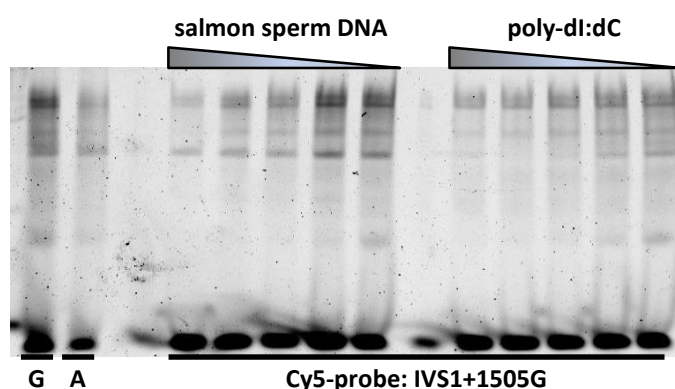
Oligonucleotide trapping used agarose beads carrying a 10 bp single stranded DNA overhang that was able to capture a DNA probe via its complementary overhang. Oligonucleotide trapping was carried out in a large scale column setup and with binding conditions carefully optimised for purity.

Biotin-avidin purification used streptavidin coated magnetic beads that were pre-loaded with biotinylated probe and then incubated with diluted nuclear extract. Magnetic bead purification was carried out in a small batch setup and mainly optimised for yield. While eluates from oligonucleotide trapping were subjected to an additional preparative EMSA and EMSA eluates were resolved on a 20x20cm SDS-PAGE, eluates from magnetic beads were directly precipitated and identified after running a short SDS-page for removal of low-mass contaminants. Exemplary, the complete workflow of an oligonucleotide trapping experiment is described in 3.2.1 to 3.2.5.

### 3.2.1. EMSA optimization

Due to the time- and reagent-consuming nature of column scale chromatography, optimisation of binding conditions was carried out in EMSA. Optimisation had the goal to reduce unspecific binding by addition of different competitors, while not interfering with specific binding. The optimisation was carried out according to Moxley et. al. [129]. Basically several concentrations for each substance of interest were tested and the highest non-interfering concentration was adopted for further experiments. One problem of such optimisation is that it has to be carried out for every batch of nuclear extracts due to different compositions of different extracts.

An exemplary optimisation gel is shown in figure 14. An exemplary set of optimised parameters are: binding constant: apparent transcription factor binding constant: 0.3 nM, final probe concentration: 3 nM, final heparin concentration: 1 ng/ $\mu$ l, poly-dI:dC: 6 ng/ $\mu$ l, single stranded T18: 1  $\mu$ M, 0.1% CHAPS, HIB1b nuclear extract: 20  $\mu$ l/ml reaction.

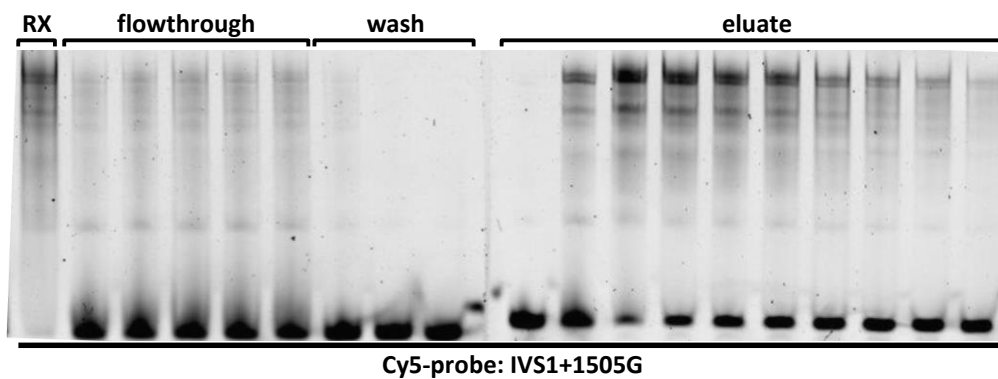


**Figure 14: EMSA optimization of dsDNA competitor concentration:** Binding reactions containing 0.2  $\mu$ l nuclear extract were mixed with binding buffer containing different concentrations of salmon sperm DNA or poly-dI:dC. Afterwards, 20 fmol probe was added. Final reaction volume was 10  $\mu$ l. Total amounts of competitors used were 50/25/12.5/6.25/3.125 ng. The first two lanes contain no competitors. Probe alleles are indicated below the gel, competitors above.

### 3.2.2. Oligonucleotide trapping

Using the parameters established in EMSA optimisation, 70 ml binding reaction containing buffer, nuclear extract, CHAPS and Heparin was prepared. After 10 minutes of incubation, 5 ml probe mix containing the probe, poly-dI:dC and T18 was added. After 20 minutes of incubation the mixture was run over the column (binding/flowthrough) which was subsequently washed with 20 ml binding buffer. Bound proteins were eluted with 10 ml elution buffer containing 350 mM NaCl. During column binding and washing, every 10 ml a few drops of flowthrough/wash fraction was collected. During elution, the whole 10 ml of eluate was collected in 0.5 ml fractions. All steps were carried out at 4°C.

All fractions were assayed for binding activity using EMSA. 2.5 µl reaction, flowthrough, wash or eluate were supplemented to 10 µl with fresh buffer containing Cy5-labelled IVS1+1505G probe, incubated and subsequently separated on TBE-PAGE. As seen in figure 15, only a fraction of the binding activity present in the reaction (RX) is present in the flowthrough that has passed the column. The wash fractions do contain almost no detectable binding activity. In contrast to the washing buffer, the eluates contain high amounts of specific binding activity. The activity steeply rises until eluate fraction three and then slowly decreases towards later fractions. The majority of the activity is eluted in fractions three to six, which were pooled and used for the preparative EMSA gels.

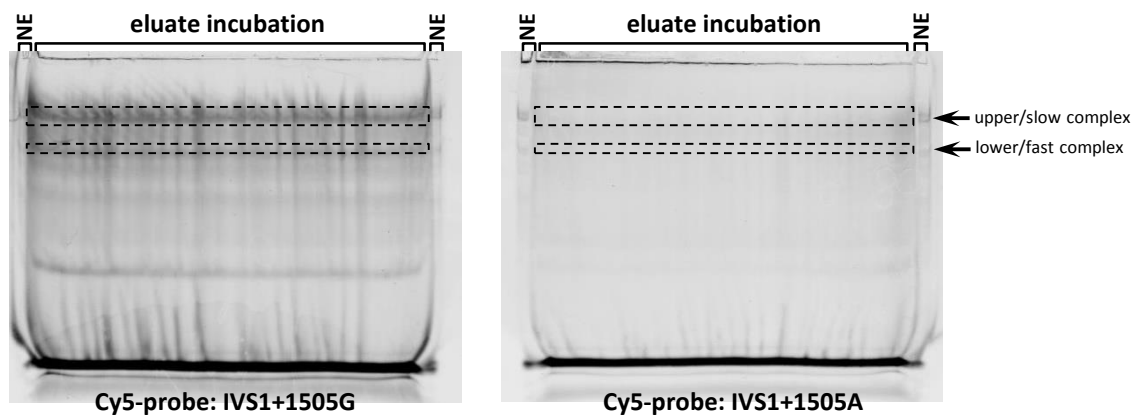


**Figure 15: IVS1+1505G binding activity in fractions collected during affinity chromatography:** Each lane contains 20 fmol probe diluted in 7.5 µl 1x binding buffer and 2.5 µl of the indicated fraction. Lane 1 contains a sample of the initial reaction mixture (RX). Lanes 2 to 6 contain samples of the reaction mix after passing the column (flowthrough). Lanes 7 to 9 contain samples from the washing process. Lanes 10 to 19 contain samples from the eluate fractions. Fractions are always ordered left to right. The leftmost eluate lane stems from the first eluate fraction, the rightmost eluate lane corresponds to the 10<sup>th</sup> eluate fraction. Binding reactions containing later eluate fractions are not shown, but contained successively less binding activity. The fraction added to the probe is indicated above the gels, the probe below.

### 3.2.3. Preparative EMSA

To increase purity for identification, an additional purification step was carried out. We decided to use preparative EMSA. The advantage of EMSA separation and complex elution is that it gets rid of contaminants binding to agarose beads and the single stranded DNA anchor used for oligonucleotide trapping. Furthermore it easily can be carried out with two different probes in parallel by running two gels, one with complex formation (by using the specific probe), one without (by using a mutated/unrelated probe or no probe at all). Proteins identified in eluates from both gels can be regarded as contaminations or background.

The active fractions from chromatography were pooled, diluted in binding buffer and split in two equal aliquots. To one, IVS1+1505G probe was added, to the other, IVS1+1505A probe. Figure 16 shows the two preparative EMSA gels used to separate the two binding reactions. The gel containing the G-probe shows higher complex formation compared to the A-probe reaction. Complexes from the IVS1+1505G gel were cut out along with the respective regions of the IVS1+1505A gel. Gel fragments were minced, proteins eluted with PAGE-extraction buffer and precipitated.



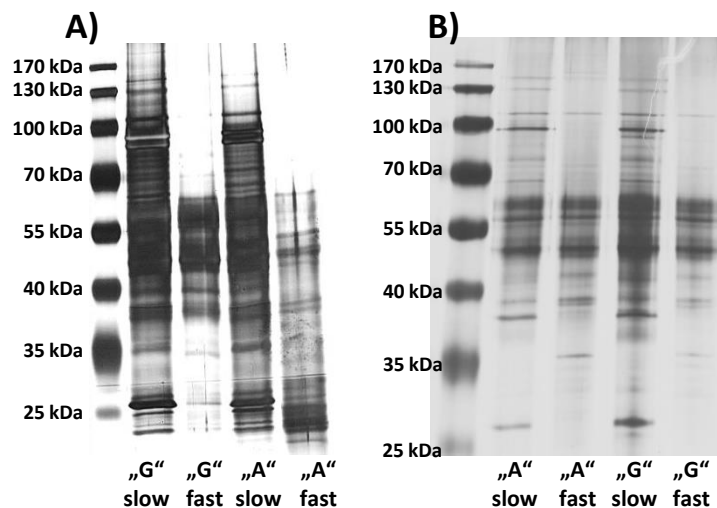
**Figure 16: Preparative EMSA:** Peak activity fractions from the previous oligonucleotide trapping step were pooled and divided. Half was incubated with IVS1+1505G, the other half with IVS1+1505A probe. Two preparative EMSA gels were used to resolve complexes formed. Each gel had one preparative well containing the eluate binding reaction flanked by two small ones containing nuclear extract binding (NE) reactions for orientation. Dashed lines and arrows indicate the complexes/regions where gels were cut for downstream processing.

### 3.2.4. SDS-PAGE and Silver-Stain

After acetone precipitation, the four protein extracts (two different probes, each two different complexes/regions cut) were resolved in 1x SDS sample buffer by vigorously shaking 10 min at 95°C and separated on SDS-PAGE.

Figure 17 shows silver stained SDS-PAGEs from two independent purifications. Both gels are dominated by two clusters of bands that are present in seven lanes of the eight lanes (two probes, two complexes, two experiments), one around 60 kDa and one around 50 kDa. Only in one lane, containing protein eluted from the fast-migrating, A-allelic complex in figure 17A, these clusters are absent. Other than that, several features are common between both gels, and thus both purifications, but seem to be complex/region specific. Despite differences in stain intensity, both gels show similar patterns for the slow migration complex with clear bands at around 150, 130, 105 and 95 kDa. Of these, the 150 kDa and the 105 kDa band were present in the fast migration complex, albeit at lower abundance. Furthermore, a cluster of weak bands is visible around 70 kDa, most being specific for the slow migration complex, some being present in the fast migration gel piece. Above 150 kDa, only few very weak protein bands can be seen. Below 50 kDa, protein patterns between the slow and the fast migration complex are different, apart from one band at 38 kDa.

None of the dominant bands shows a reproducible, clear difference between the two alleles, IVS1505G and IVS1505A. Part of that might stem from the limited dynamic range silver stain offers, thereby possibly masking differences. Another possible problem is the strong staining around 50-60 kDa as it may cover lower abundant signals. All clear distinct bands were cut and subjected to MS.



**Figure 17: Silver stained SDS-PAGE of EMSA eluates:** Eluates from preparative EMSA were precipitated, resolved on SDS page and stained by silver stain. **A)** and **B)** are two independent gels arising from two independent purifications. Each gel contains five lanes: one size marker lane and four eluate lanes corresponding to the four cut-out regions (see figure 16). The cut-out region is marked below the lanes: “A” and “G” denote the gel which the complex was cut from while “slow” and “fast” refer to the slower and the faster migrating complex from each EMSA gel, respectively. Marker band molecular weight is denoted left of the gels.

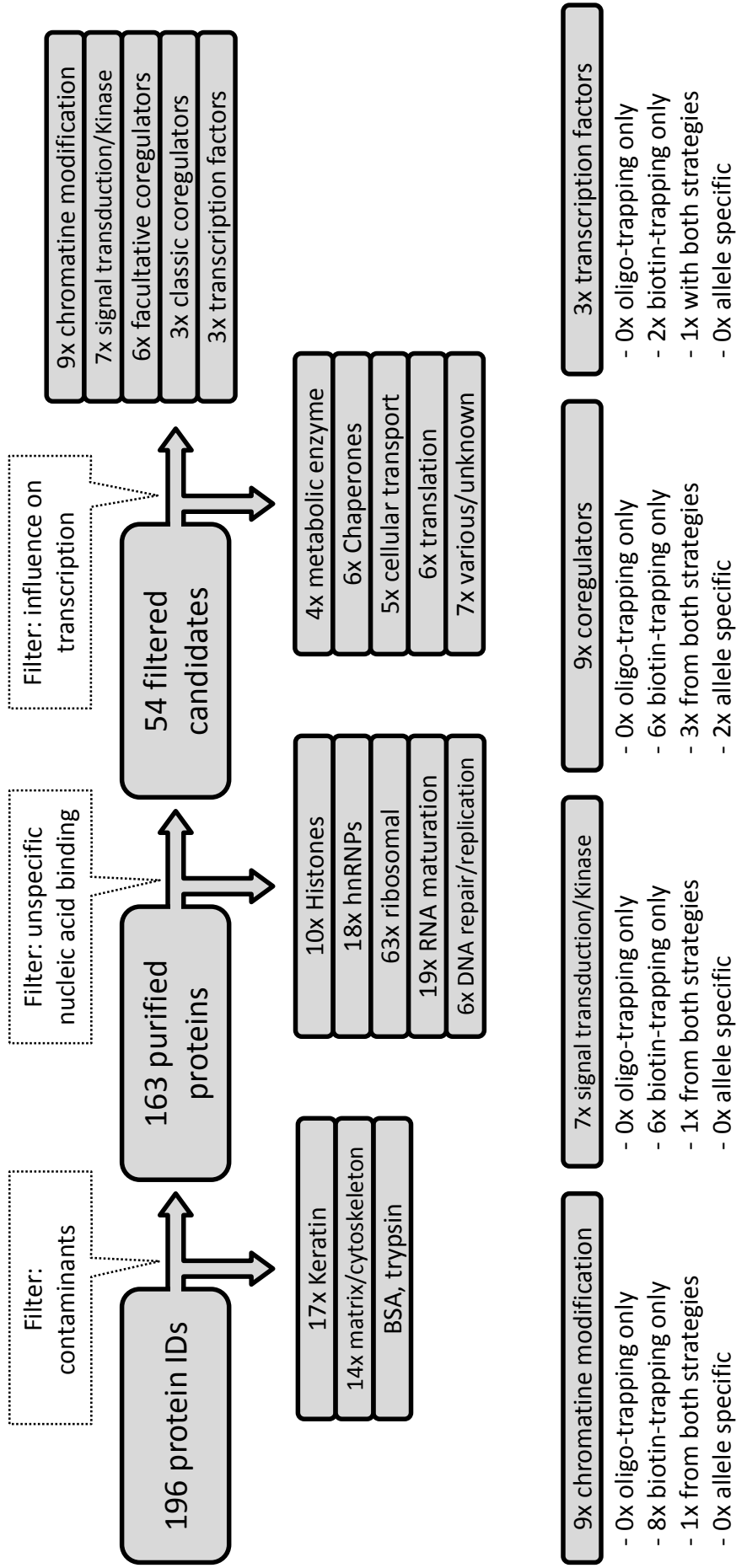
### 3.2.5. Mass spectrometry – candidate list

After separating the samples on SDS-PAGE, proteins were eluted from the PAGE gel. This processing was carried out by Fiona Pacht and Hannes Hahne of the Proteomics and Bioanalytics department (Technical University of Munich, headed by Professor Dr. Bernhard Küster). Briefly, protein was in-gel digested by trypsin and fragments were resolved on reverse-phase nano-Liquid Chromatography before analysis in an ETD-MS/MS mass spectrometer. Identified peptides were compared to IPI mouse database (v3.26, 52,735 sequences). A complete list can be found in the appendix 13.

Figure 18 depicts the filtering process the candidate list from MS underwent. Of 196 identified proteins, first obvious contaminants were excluded, including keratins, BSA, trypsin and matrix and cytoskeleton proteins. All these proteins are usually present in all samples as they are highly abundant in starting material and/or environment. Next, proteins known to non-specifically interact with nucleic acids were excluded. Histones, hnRNP proteins, ribosomal proteins, proteins involved in RNA maturation and DNA repair are commonly bound and eluted in DNA affinity chromatography, independent of the specific DNA sequence used. 54 candidates passed these two filters. Candidates were then manually sorted into two groups: Firstly, non-plausible candidates that have no direct or indirect influence on transcription regulation and no affinity to DNA or nuclear localisation. Secondly, plausible candidates, which can be transcription factors, transcriptional co-activators, chromatin/histone modifying enzymes and enzymes known to directly signal to plausible candidates. Co-activators were split into classic co-activators, that is proteins primarily known for their influence on transcription, and facultative co-activators which, besides their most established function in addition are believed to also possess some co-activator function.

As plausible candidates, three transcription factors, nine co-regulators (six facultative, three classic), nine chromatin modifying factors and seven signal transduction enzymes were found. For all plausible candidates allele specificity was estimated by comparing the number of identified fragments in the IVS1505G purifications to those from the IVS1505A purifications. Only two identifications were considered possibly allele specific (table 1). Furthermore table 1 shows whether a candidate was purified only by oligonucleotide trapping, by magnetic bead purification or by both methods.

The IPI mouse database was used to search for basic information on identified proteins. Table 1 shows all transcription factors and coregulators as well as selected candidates from other categories. Three transcription factors identified were: Interleukin-enhancer binding factor 2 (ILF2), nucleolar transcription factor 1 (UBF1) and nuclease-sensitive element binding factor 1 (Y-box transcription factor 1, YBX1). All three were present in the crude magnetic bead purification experiment eluates (with seven, five and four fragments respectively), but only UBF1 was identified after oligonucleotide trapping as well, and with only one fragment identification. Of the other categories, literature supports a possible involvement in transcriptional regulation for five other candidates: Staphylococcal nuclease domain-containing protein 1 (SND1), serine/threonine-protein kinase 36 (STK36), heat shock cognate 71 kDa protein (HSC71), probable ATP-dependent RNA helicase DDX5 (DDX5) and thyroid hormone receptor-associated protein 3 (THRAP3).



**Figure 18: Filtering candidate lists of proteins identified by mass spectrometry.** All protein identifications from mass spectrometry were filtered for contaminants and sorted by function. Proteins with a described function that makes an influence on transcription plausible are further sorted by the strategy used for their purification and identification and whether they are allele specific. Of all 196 candidates, only two are known to directly act as a classical transcription factor. Eight proteins had coregulator activity, of which three can be considered classic co-regulators while five are considered facultative co-regulators, meaning proteins mainly known for other functions, but also reported to be able to influence certain promoters or interact with certain transcription factors. The majority of proteins identified stems from the crude one-step magnetic bead purification employing biotin/avidin interaction. The sum of candidates in all categories exceeds 196 as some candidates fall in more than one category.

**ILF2**, usually as a dimer with ILF3, is a transcription factor and part of complexes involved in DNA replication and repair. Furthermore it plays a role in splicing as a component of a mRNP complex. For the  $\beta$ -globin gene, ILF2/ILF3 dimers are known to bind p300 and thereby facilitate histone acetylation [146]. The dimer has been found in the nucleolus of T-cells, possibly playing a role in ribosome biogenesis [147].

**YBX1** is a regulator of Pol2 transcription, is involved in splicing and DNA repair and is considered an oncogene. YBX1 acts as a dimer with EWS and as such plays a role in brown adipocyte development, most likely via binding to the BMP7 promoter [148]. A ChIP-seq experiment very recently revealed that Ybx1 commonly binds to introns, and interestingly, in BT474 breast cancer cells, the ChIP-seq peaks often also contained CREB1, SP1 and PPAR $\gamma$ /RXR $\alpha$  binding matrices [149]. YBX1 furthermore is part of a mRNP complex that also contains ILF2.

**UBF1** is a transcription factor and known to be relevant for PolI transcribed promoters where it exerts its function by counteracting repressors [150]. It is involved in expression in rRNA genes and thought to link ribosome biogenesis to glucose sensing [151]. Interaction with RUNX2 [152], ERK, IRS1 [153] and PI3K [154] has been shown. A single report also demonstrates binding to a Pol2 promoter [155].

**SND1** is a transcriptional coactivator also involved in miRNA maturation and splicing. It can stabilise mRNA and thus increase protein abundance by binding to the 3'UTR of an mRNA. SND1 is involved in interleukin-4 signalling where it binds and coactivates signal transducer and activator of transcription (STAT)6 via recruitment of p300 [156]. SND1 is found to be linked to several different cancer types. It was identified in several samples, both from both affinity strategies but also from both EMSA-cutout gels with a total of 58 fragments, ranking 18 of all 196 unfiltered candidates.

**STK36**, also known as Fused, is a kinase with a function in sonic hedgehog signalling. It phosphorylates the GLI transcription factors and is important for their activity, although this seems to be cell-type dependent [157]. Its mRNA is highly abundant in testis, has a low abundance in brain and is absent in most other tissues including SKTM. For BAT, no expression data are available. While little is known about Stk36 itself, active hedgehog signalling is an important pathway promoting osteo- and myogenesis [158,159] and acts anti-adipogenic [160]. These effects partly are dependent on GLI1-3 transcription factors.

**HSC71**, besides its function as a chaperone, is a co-repressor of transcription and an inducer of splicing. HSC71 is primarily found in SKTM [161]. Its co-repressor activity depends on its ability to inhibit CITED1 and p300 activity in SMAD signalling. Besides SMAD signalling, a link to WNT signalling has been described [162]. In mRNA metabolism, it is part of the spliceosome and component of the mRNP complex which also contains ILF2 and Ybx1. Furthermore it may play a role in a complex mediating histone H3 methylation. For HSC71, 26 fragments were identified using Oligonucleotide trapping, of which 13 IDs were in "G"-samples and only three in "A"-samples. Another 14 IDs came from magnetic bead purification.

**DDX5** has two independent functions. First, it has RNA helicase activity and participates in splicing. Second, it acts as a co-activator or -repressor. Its activating function is described in cooperation with AR, ESR [163] and RUNX2 [164] which most likely involves the histone acetylase p300. Furthermore DDX5 interacts with MyoD [165], a key regulator in muscle differentiation. Generally DDX5 is considered pro-myogenic. Its repressing action depends on interaction with HDAC1. Six fragments were found in magnetic bead purification, but only two in Oligonucleotide trapping. While those two

were only found in WT-EMSA-cutouts, the number of IDs is too low to be considered a reliable proof of allele specificity.

**THRAP3** was only identified in magnetic bead purification and only with two identified peptides. Besides being involved in splicing, it also acts as a co-activator of transcription. Thrap3 enhances PPAR $\gamma$  mediated activation [166] of transcription in presence of agonists, most likely via stabilising its DNA binding activity. A role in adipocyte differentiation has been proposed.

In summary, several candidates seemed interesting at first glance, being linked to BAT differentiation or interacting with proteins relevant for UCP3 expression in one way or another: YBX1 binds and activates the BMP7 promoter. DDX5 and THRAP co-activate nuclear hormone receptors, AR/ESR and PPAR $\gamma$  respectively. UBF1 interacts with IRS, PI3K and ERK. ILF2, SND1, HSC71 and DDX5 have been shown to interact with p300. SND1 is important for STAT signalling. STK36 and HSC71 have been linked to WNT and SMAD signalling, both involved in determination of SKTM and BAT. At second glance though, most of the links are very general or even contradictory. While p300 is important for regulation of UCP3 transcription, p300 is a very general histone acetylase. WNT signalling indeed is linked to determination of BAT, but the pathway, as well as DDX5 and HSC71, is considered pro-myogenic and anti-adipogenic. UBF1 is primarily known for regulating PolII mediated transcription and is known as an important co-activator for Runx2, a central pro-osteogenic transcription factor. The latter is also true for DDX5, which furthermore coactivates MyoD. Most proteins were identified with very few peptides, only in eluates from one affinity strategy or from purifications with both IVS1+1505G and –A probes. Last but not least, all candidates part of mRNP complexes and/or known to be involved in RNA maturation in one way or another. Only two proteins of interest remained after the extended literature search: SND1 and Thrap3. While they both lack direct DNA-binding, they might be indirectly part of the transcription factor complex binding the probe though, but the lack of allele specificity in case of SND1 and the identification of THRAP3 in only one single experiment and with only two fragments argues against them.

**Table 1: Selected proteins identified in mass spectrometry (next page):** All identifications obtained after oligonucleotide-trapping and magnetic bead purification were combined in one list and sorted. After the filtering (see figure 18), Uniprot and Pubmed were screened for information on identified proteins. For the 13 plausible candidates, including all transcription factors and co-regulators, full name of protein and gene (in brackets) are shown as well as a brief list of putative functions, interaction partners and relevant pathways are given. Furthermore the number of peptide IDs obtained in both strategies is listed. \*Only two proteins were preferably identified in samples from WT EMSA-cutouts, but due to the low number of IDs, the allele specificity is likely to be a of stochastic origin. All other proteins were identified both in eluates from purifications using the IVS1+1505G and the IVS1+1505A probe. A complete, unfiltered list is found in appendix 13.



**Table 1: Candidates from MS**

# of peptide IDs: Oligo-Trap Bio/Avi specific?

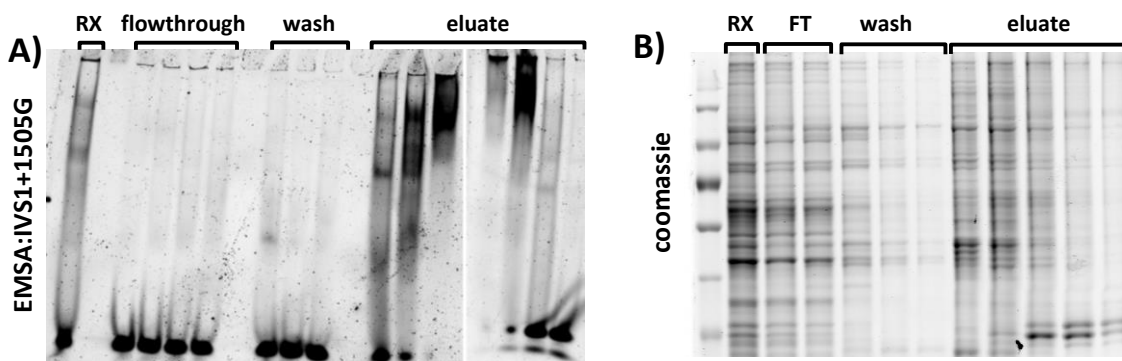
		# of peptide IDs: Oligo-Trap	Bio/Avi	specific?
<b>Staphylococcal nuclease domain-containing protein 1 (Snd1)</b>	Function: Coactivator and bridging factor for STAT5, STAT6 and general transcription factors. Interaction: GTF2E1, GTF2E2, STAT6, POLR2A, PIM1, p300 Pathway: Bioinformatic evidence for phosphorylation by PIM1 and responding to IL-4 signalling	39	19	no
<b>Serine/threonine-protein kinase 36 (Stk36)</b>	Function: regulates activity of GLI1-3, kinase activity, postnatal development Interaction: SPAG16, KIF27, GLI1-3 Pathway: sonic hedgehog (Shh) pathway, osteoblast and myoblast differentiation	31	4	no
<b>Heat shock cognate 71 kDa protein (Hspa8)</b>	Function: repressor of transcriptional activation: CITED1 and smad signalling, chaperone activity Interaction: PACRG, BAG1, DNAJC7, CITED1, IRAK1BP1, HSPH1/HSP105, TRIM5 Pathway: Part of mRNP granule complex (including Ybx1), WNT signalling,	16	14	yes*
<b>Proliferation-associated protein 2G4 (Pa2g4)</b>	Function: corepressor of AR, mRNA/ribosome associated, ERBB3 signal transduction Interaction: HDAC2, RB1, ERBB3, AR, nucleolin/NCL Pathway: Component of a ribonucleoprotein complex	0	13	no
<b>FACT complex subunit SSRP1 (Ssrp1)</b>	Function: mRNA elongation, DNA replication/repair, nucleosome organisation, coactivator for p63/TP63 Interaction: p63/TP63, FYTDD1/UIF, SRF, NEK9, Myog, CK2 Pathway: Phosphorylated by CK2 following UV irradiation, ubiquitinated and SUMOylated	0	11	no
<b>Probable ATP-dependent RNA helicase DDX5 (ddx5)</b>	Function: RNA helicase, coactivator for AR, ESR1, p53/TP53, RUNX2. Myo-/Osteoblast differentiation Interaction: AGO1/2, ESR1, AR, NCOA1-3, EP300, CREBBP, POLR2A, TP53, RUNX2, HDAC1, DDX17, BRDT Pathway: Sumoylation promoted by PIAS1, Muscle and osteoblast differentiation	2	6	yes*
<b>Interleukin enhancer-binding factor 2 (Ilf2)</b>	Function: Transcription regulation, DNA repair, chromatin modification Interaction: ILF3, YLPM1, KHDRBS1, RBMX, NCOA5, PPP1CA Pathway: Part of mRNP granule complex (including Ybx1). Part of DNA repair/replication complexes.	0	7	no
<b>Nucleolar transcription factor 1 (Ubtf)</b>	Function: Activator of transcription of rRNA-genes by PolI. Chromatin folding Interaction: RASL11A, PIK3CA, IRS1, TBP, TAF1A, TAF1D, ERK, Runx2 Pathway: Phosphorylated and activated by PIK3CA and ERK.	1	5	no
<b>Activated RNA polymerase II transcriptional coactivator p15 (Sub1)</b>	Function: Mediates interaction between g-TAFs and enhancers, unspecific ss- and dsDNA binding Interaction: CSTF2 Pathway: Phosphorylated by several kinases	0	6	no
<b>Nuclease-sensitive element-binding protein 1 (Ybx1)</b>	Function: alternative splicing, stabilizes mRNA, regulates transcription via Y-Boxes, DNA repair Interaction: EFI-B, PURA/B, APEX1, AGO1/2, EWS Pathway: Ubiquitinated by RBBP6, component of several large complexes, regulates BMP7 expression	0	4	no
<b>Scaffold attachment factor B1 (Safb)</b>	Function: ESR1 corepressor, can activate transcription of SREBF1 with RBMX Interaction: KHDRBS3, POLR2A, SAFB2, SFRS1, SFRS9, TRA2B/SFRS10, RBMX, SRPK1 Pathway: Sumoylated by PIAS1, desumoylated by SENP1	0	4	no
<b>Thyroid hormone receptor-associated protein 3 (Thrap3)</b>	Function: pre-mRNA splicing/decay, transcriptional activation, terminal adipocyte differentiation Interaction: HELZ2, PPARG, NXF1, SFPG1 Pathway: Component of SNARP, TRAP and exon junction complexes. BAT differentiation, PPAR $\gamma$ signalling.	0	2	no
<b>RNA polymerase II-associated factor 1 homolog (Paf1)</b>	Function: Paf1 complex: Hox and WNT transcription, histone modification, mRNA maturation Interaction: POLR2A, TCEA1, TTC37, KMT2A/MLL1, SUPT5H, RNF20, RNF40, UBE2E1 Pathway: Forms Par1 complex with: CDC73, PAF1, LEO1, CTR9, RTF1, WDR61	0	2	no

### 3.2.6. Heparin affinity chromatography

An alternative additional purification step is heparin affinity chromatography. Also carried out on column, it can be used to pre-purify DNA binding proteins in a sequence independent manner before sequence specific affinity chromatography. The advantages compared to preparative EMSA are that it allows for higher yields and avoids the contamination-prone EMSA cut-out and -elution.

Nuclear extract was diluted using salt free heparin affinity (HA-0) buffer, thereby adjusting to a KCl concentration of 100 mM. This concentration allows TFs to bind to heparin. After binding to the resin and subsequent washing, bound proteins were eluted using HA buffer containing increasing KCl concentrations (200-1000 mM). During flowthrough and wash, four and three fractions were collected, respectively. All eluate fractions were collected. Figure 19A shows the EMSA activity in the binding reaction (RX), the flowthrough, the wash and the eluate fractions. The last five eluate fractions are not shown as they did not contain any binding activity. In figure 19B a coomassie stained SDS-PAGE is shown. Each lane contains the same volume of the respective fraction. That way, total protein content could be estimated. Together, EMSA and coomassie stain demonstrate that most of the protein appears in the flowthrough and the first eluate fractions while most of the IVS1+1505G binding activity appears in the eluate fractions 3 to 5. While no exact measure was possible, most of the binding activity eluted at around 400 mM KCl.

In conclusion, heparin affinity chromatography led to a reduction of total protein content while drastically increasing the concentration of binding activity. With further slight optimisations in elution procedure, the method will be a powerful technique for pre-purification of nuclear extracts.



**Figure 19: Heparin affinity chromatography:** Nuclear extract was diluted six fold in heparin affinity buffer (RX) and run over a heparin agarose resin. After washing, bound proteins were eluted using an increasing KCl gradient. **A)** shows binding activity of the fractions using EMSA. While flowthrough (FT) and wash steps are devoid of binding activity, the eluate contains very high amounts of activity. **B)** shows an estimation of protein concentration in the respective fractions. Equal volumes of each fraction were loaded on an SDS page, resolved and stained using coomassie. Fractions are always ordered left to right, the leftmost sample being from the first eluate fraction containing the lowest KCl concentration.

### 3.3. Candidate finding: Bioinformatics

A quick approach to test the involvement of candidate transcription factors in a complex formed in EMSA is cold competition. In this strategy, an unlabelled, thus invisible probe is added in molar excess to the labelled probe of interest. If the unlabelled probe contains a binding motif for a protein participating in the complex formed with the labelled probe, the non-labelled probe will compete for the transcription factor. If molar excess and affinity to the competitor are high enough, the unlabelled probe will deplete the visible complex and form an invisible complex.

#### 3.3.1. Bioinformatics and consensus motifs

Three tools and four algorithms were used to find vertebrate transcription factor binding sites in the vicinity of the IVS1+1505 element: MATCH and PATCH of Biobase and MatInspector of Genomatix. Match was set for minimal false negatives, the other two were used with standard settings. The majority of Bioinformatics was performed by Tobias Fromme, who published parts of this dataset [59], but the validation experiments were part of this thesis.

Bioinformatics revealed elements for the following transcription factors: p53, PBX, Sry, TBP, Cdx, C/EBP, c-Myb, ETF, MSX as well as the rather large family of Forkhead transcription factors. Members of the Forkhead family showed up with all four algorithms (table 2).

For all these predicted binding TFs consensus matrices were obtained from Genomatix. These consensus motifs were embedded into the IVS1+1505 sequence context and then used for cold competition experiments (see 3.4.1 figure 21).

**Table 2: Bioinformatics**

MATCH	MATCH high quality	PATCH	MatInspector
C/EBP	XFD-2	ETF	FHKD family
Oct-1		TBP	
XFD-2		TF2D	
TATA		FoxM1a/b	
CdxA		HNF3a/b	
Pbx-1			
c-Myb			
SRY			
Msx-1			
p53			

**Table 2: Binding motifs identified for the IVS1+1505G element using bioinformatics:**

Four different algorithms were used to detect putative binding motifs in the vicinity of the IVS1+1505 base exchange. Besides other motifs, all algorithms report one or more forkhead factor family members: XFD-2, FoxM1, HNF3 and FHKD family. Bioinformatics were carried out by Tobias Fromme. Table 2 is modified from [58].

### 3.4. Candidate Validation

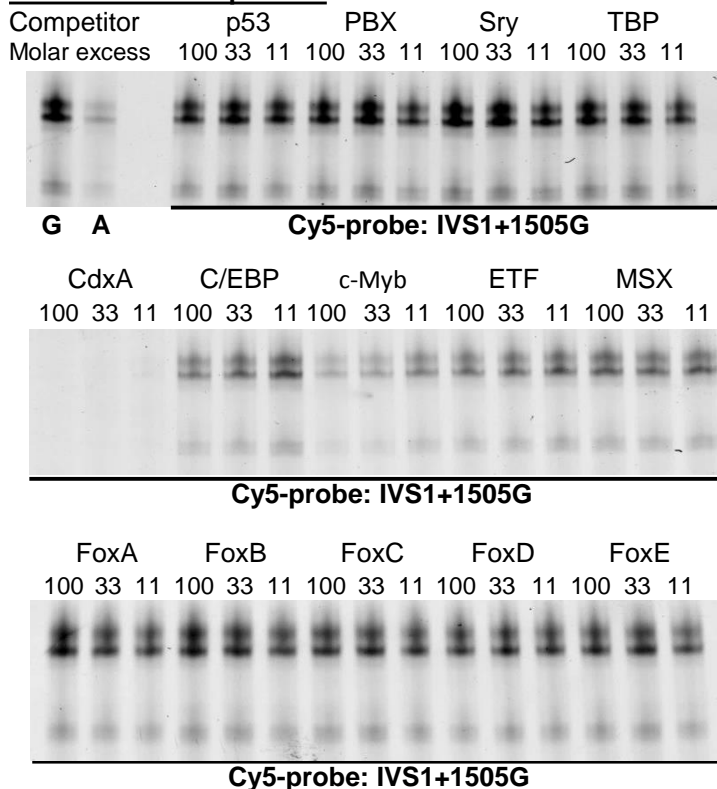
#### 3.4.1. Validation of candidates from bioinformatics using EMSA cold competition

Competitor oligonucleotides were purchased with the consensus sequence embedded in the 31bp IVS1+1505 sequence context and tested for their binding capacity. The Forkhead family was covered by five different pan-Forkhead Factor competitors [59].

The consensus competitors for P53, PBX, Sry, TBP, C/EBP, ETF, MSV, and the five pan-Forkhead matrices had no influence on complex formation (figure 20). The c-Myb competitor yielded a weak competition effect at 100- and 33-fold molar excess. Only the CdxA competitor strongly depleted complex formation at 100-, 33- and 11-fold excess, and a labelled version of the CdxA competitor produced a complex pattern comparable to the IVS1+1505G probe (not shown). Sequences of all used competitors, including competitors used in experiments not shown, are aligned in figure 21.

While this points to CdxA being the most likely candidate for binding the IVS1+1505G element, its expression pattern argued against its involvement. Three putative binding proteins are known for the CdxA element in mammals, CDX1, CDX2 and CDX4. None of these proteins is known to be expressed in muscle or BAT. We thus assayed our four cell lines for presence of CDX transcripts, but were unable to detect any (see 3.4.2, figure 22).

#### non-labelled competitors



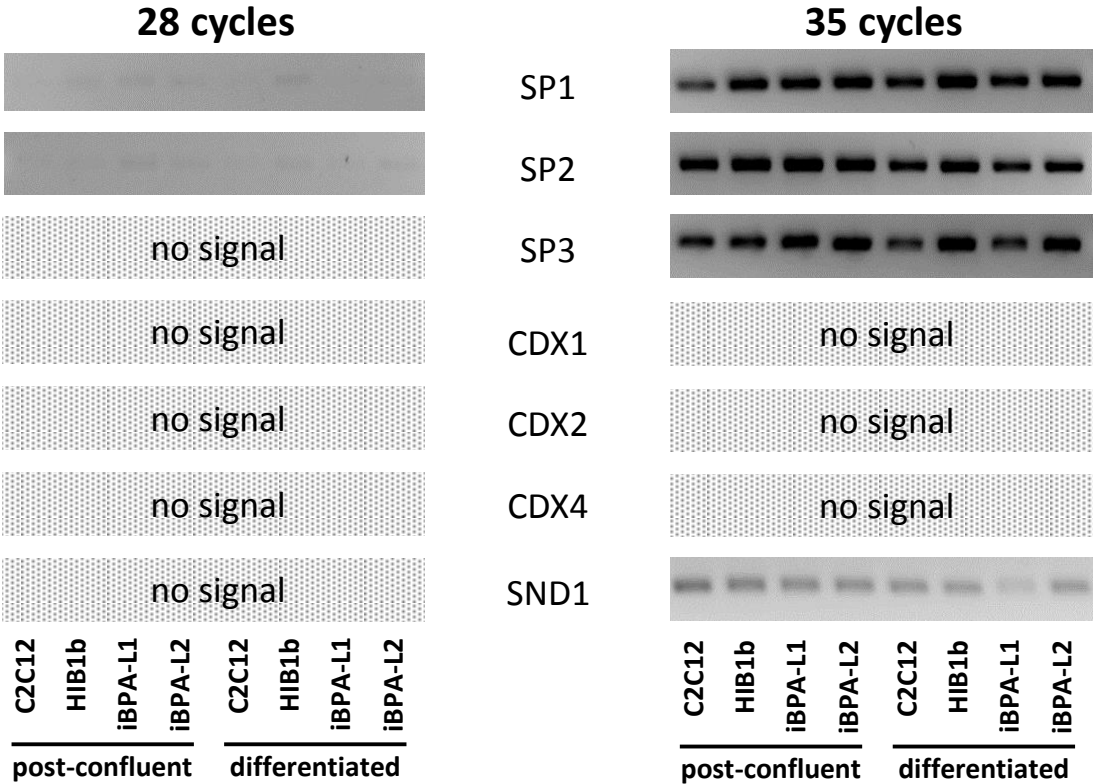
**Figure 20: Competitor screen using consensus binding sequences:** EMSA reactions were carried out using HIB1b nuclear extract and IVS1505G or IVS1505A probe. The labelled probe used is indicated below each lane. Aside from the first two lanes in the upper left corner, all lanes contained non-labelled competitor in molar excess. Competitors contained consensus binding motifs for specific transcription factors or five pan-Forkhead binding matrices integrating most Forkhead transcription factor binding sites. Competitors and molar excess are denoted above each image.

Competition	putative SP1/3 binding arm	putative Forkhead binding arm
Phodopus IVS1+1505G 31 bp	G T G	T T T
Phodopus IVS1+1505G 19 bp		T T C
Rattus putative element		T T C
Cdx consensus		T T A
Phodopus 19 bp IVS1+1497G		T T A
Mus putative element		T T A
Phodopus Mut1501-03 19 bp		T T A
Phodopus IVS1+1505G 17 bp		T T A
Phodopus IVS1+1505A 31 bp		T T A
c-Myb consensus		T T A
Phodopus Mut1506-08 19 bp		T T A
Putative Forkhead arm 17 bp		T T A
Phodopus IVS1+1504A 31 bp		T T A
Monodelphis putative element		T T A
TBP consensus		T T A
CEBP consensus		T T A
PBX1 consensus		T T A
Msx1 consensus		T T A
p53 consensus		T T A
Sry consensus		T T A
ETF consensus		T T A
pan-Forkhead matrix A		T T A
pan-Forkhead matrix B		T T A
pan-Forkhead matrix C		T T A
pan-Forkhead matrix D		T T A
pan-Forkhead matrix E		T T A

conserved base  
variations deleterious  
variations tolerated

**Figure 21: Comparison of sequences used in competition experiments:** All competitors used in EMSA were ordered by their competition activity, aligned and compared. Bases conserved in all strongly competing oligonucleotides are marked green. Base exchanges or deletion possibly having deleterious effect on competition capacity are marked red. Positions in which sequence variations are tolerated are not coloured. The putative SP1/3 binding arm and the putative Forkhead binding arm of the probe are encircled by bold lines.

We aligned the sequences of all competitors and ordered them by the strength of competition, including some from experiments not shown. By that, we were able to uncover which bases were crucial for complex formation, and which positions were less relevant. In figure 21, all positions marked green are common in all competitors fully active. Possibly deleterious base exchanges are marked red. This reveals that all strongly competing oligonucleotides contain a CACGCC motif in their centre, and that already single base deviations from this motif reduced competitive capacity. The IVS1+1505A allele changes this sequence to CACACC. Notably, a single base exchange at position IVS1+1504, which changed the motif sequence to CAAGCC, completely abolished competition. The IVS1+1504A allele thus was used in several later experiments as negative control. In the CdxA competitor, the CACGCC motif is completely unaltered. This putative consensus matrix hints towards a putative GC-Box binding motif (figure 21). Three GC-Box binding factors, SP1, SP2 and SP3, are known to be expressed ubiquitously. A fourth one, SP4, is mainly expressed in brain. SP1, SP3 and SP4 essentially bind the same sequences, a GC- or GT-Box.



**Figure 22: Expression of candidate proteins:** The cDNAs used for the transcription factor screen (figure 10) were also used to check expression of the GC-Box binding proteins SP1, SP2, SP3, and the CdxA-consensus binding proteins CDX1, CDX2 and CDX4. Furthermore we assayed tissue specificity of SND1. PCRs were carried out for 28 and 35 cycles (left and right column of pictures). Rows show the PCR products for the transcripts specified in the centre. For primers, except CDX4, the respective negative control PCR reactions yielded no product while the positive controls did. For CDX4 no positive signal could be generated under any condition.

### **3.4.2. Transcript of SP1, SP2 and SP3, but not of CDX proteins, is present in all cell lines**

A requirement for a transcription factor to participate in BAT specific regulation of UCP3 transcription is to be expressed in BAT. While identification of proteins by mass spectrometry is a reliable proof for their expression, expression needs to be validated for proteins found via bioinformatics means. As the both the GC-Box (figure 23) and the CdxA consensus competitors (figure 20) strongly influenced complex formation, we assayed the mRNA abundance for the respective binding transcription factors. Furthermore we wanted to know whether SND1, the candidate most reliably identified by mass spectrometry, was BAT-specific.

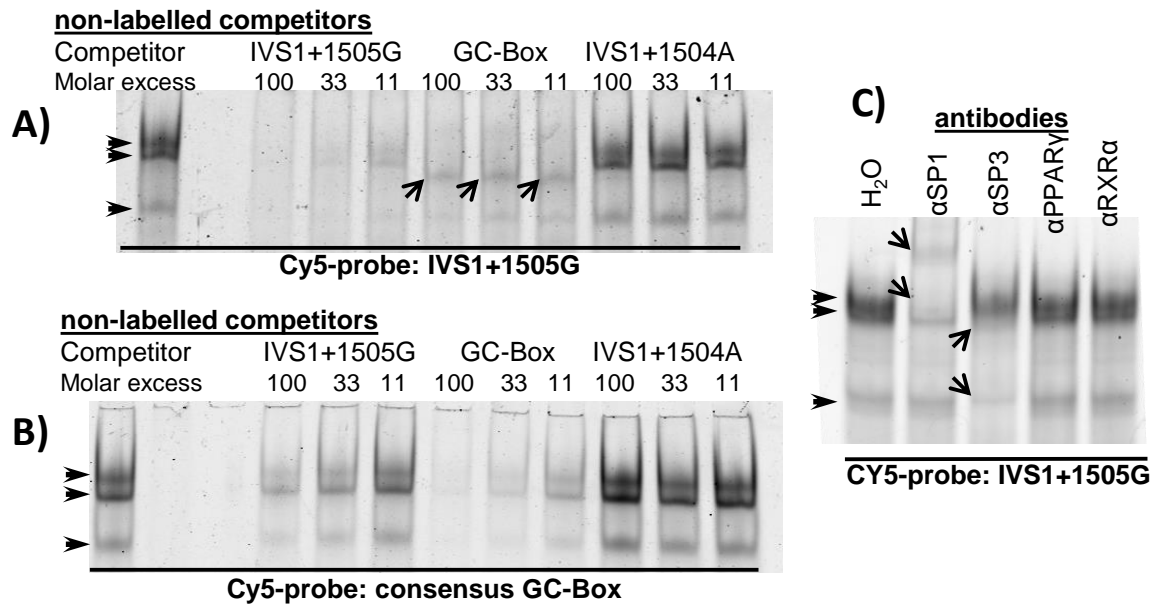
The general transcription factors SP1, SP2 and SP3 were expressed in all cell lines (figure 22, see page before), as was SND1. In contrast, neither CDX-transcript could be detected. While for CDX1 and CDX2 positive controls demonstrated function of the primer pairs, no cDNA tested yielded a positive signal for CDX4 (not shown). Thus, absence of CDX4 is not proven without doubt.

Of our candidate transcription factors only SP1, SP2 and SP3 were expressed in all four cell lines, as is the EMSA signal (figure 9). We therefore considered SP1 and SP3 as prime candidates and discarded the CDX proteins.

### **3.4.3. SP1 and SP3 bind the IVS1+1505G probe in EMSA**

A consensus GC-Box probe was ordered and used as a competitor. Figure 23A shows that this competitor was an even more efficient competitor than the IVS1+1505G probe itself. Figure 23B shows that an IVS1+1505G competitor was able to compete for the proteins bound to a labelled consensus GC-Box probe. Figure 23C shows the effects of antibodies added to the EMSA binding reaction. Antibodies binding RXR $\alpha$  and PPAR $\gamma$  did not influence complex formation with the IVS1-1505G probe, antibodies binding SP1 interfered with formation of the upper complex, and antibodies binding SP3 interfered with formation of the lower complexes. Antibodies binding SP2 and SP4 did not have an effect on complex formation (not shown). Furthermore, addition of SP1 antibody led to appearance of a supershifted complex. Results of the supershift experiments were successfully validated using several independent antibodies including crude immune sera (not shown, [167]). Interestingly, a novel complex formed with the IVS1+1505G probe was revealed in absence of SP1/SP3 binding (figure 23A). Later experiments demonstrated that this complex was absent in both iBPA cell lines and C2C12 cells (not shown). Some follow-up experiments were carried out, but none of those yielded reliable insight on the nature of this novel complex (not shown).

Together, this demonstrates that the *Phodopus sungorus* UCP3 IVS1+1505G probe contains a GC-Box element that is bound by SP1 and SP3 in EMSA. The competition capacity of the CdxA competitor most likely originates from fact that this GC-Box is unaltered in the competitor.



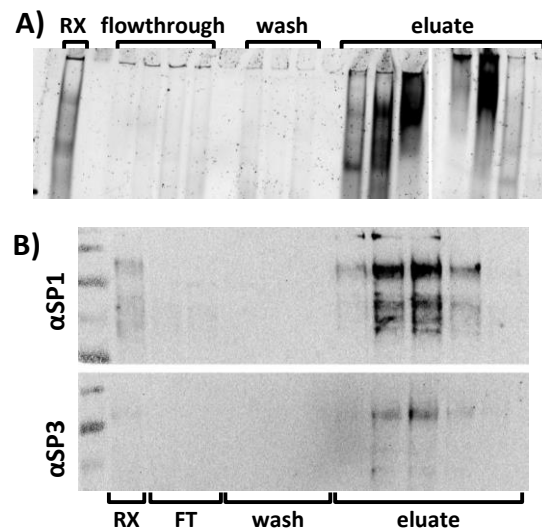
**Figure 23: SP1 and SP3 bind the IVS1+1505G Probe in EMSA.** **A)** EMSA using the IVS1+1505G probe and HIB1b nuclear extract. Labelled probe was added in parallel with the competitors indicated above the image in molar excess **B)** EMSA corresponding to A), but with a labelled consensus GC-Box probe. **C)** EMSA using the labelled IVS1+1505G probe and HIB1b nuclear extract. Antibodies (indicated above, 1  $\mu$ g each) were added to the binding mixture after addition of probe. Complexes of interest are marked by arrows.

#### 3.4.4. SP1 and SP3 are enriched in heparin affinity chromatography eluates

At the time we successfully carried out heparin affinity chromatography, first hints for the binding of SP1 and SP3 to the IVS1+1505G element (see figure 23) started accumulating. We thus assayed our affinity eluates for SP1 and SP3 using western blot.

Equal volumes of diluted nuclear extract, flowthrough, wash fractions and of the eluates containing the highest amount of EMSA binding activity (figure 24A) were resolved on two SDS-PAGES. The loading scheme is the same as seen in figure 19B. Two western blots (figure 24B) were carried out to detect SP1 and SP3. For both, immunoreactivity is low in flowthrough and wash, but high in the EMSA-active eluate fractions.

Together with the coomassie stained gel from figure 19B, this demonstrates that SP1 and SP3 immunoreactivity correlates with EMSA activity, but not with total protein content. This supports that indeed SP1 and SP3 are the EMSA active IVS1+1505G binding proteins.

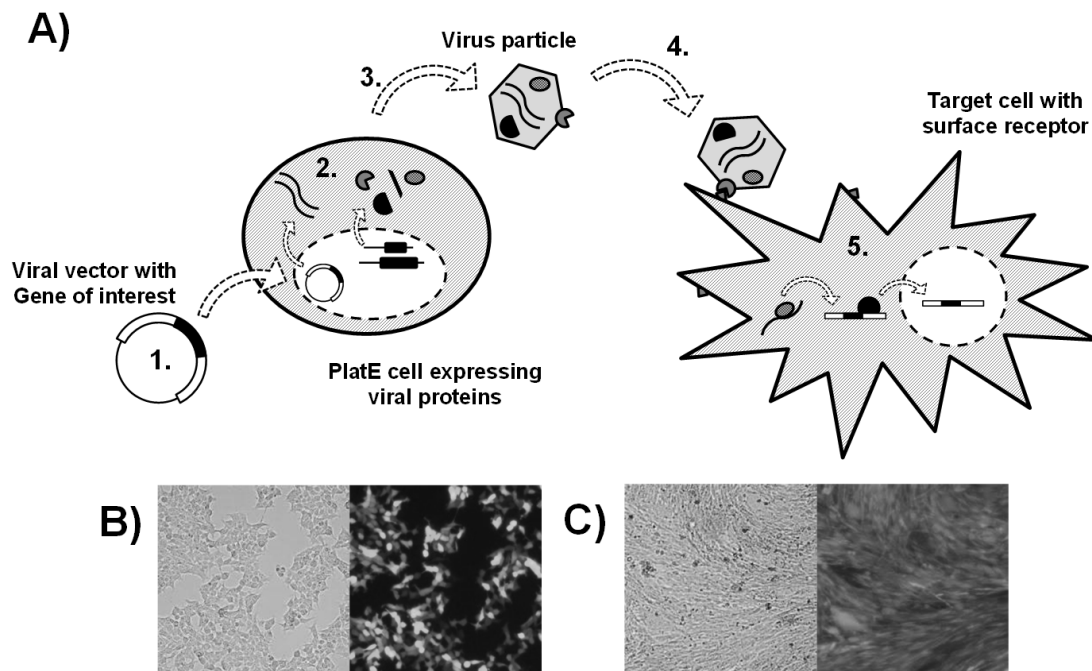


**Figure 24: Heparin affinity chromatography eluates contain SP1 and SP3: B)** The fractions collected in Heparin affinity chromatography were loaded on two SDS-PAGES, copying the one from figure 19B were run. Western blots followed by detection of SP1 and SP3 were carried out. **A)** shows a cropping from the EMSA already depicted in figure 19A.



### 3.4.5. RNAi mediated knockdown and overexpression

If SP1 and SP3 indeed would be positive regulators binding the IVS1+1505G element, reducing their expression level would lead to decreased reporter activity of the *Phodopus sungorus* reporter gene construct. Vice versa, increasing their expression might increase reporter activity, unless their basal expression level was already sufficient to achieve full occupancy of the respective binding element. Overexpression was achieved by cloning the coding sequence (CDS) of the respective gene of interest into a vector containing a strong constitutive promoter and then transfecting the vector into the target cells. For reduction of expression, RNA interference was used: For transient interference, pools of vectors expressing shRNAs were transfected. For stable knockdown, miRNAs were cloned into a viral vector which was then packed into retroviral particles suitable for stable integration. Selection of virus infected cells allowed generation of cell pools in which all cells expressed the inserted sequence, albeit at slightly different expression levels due to different insertion sites.



**Figure 25: Viral packaging and infection.** A) Schematic workflow of the viral transfection procedure. 1) A gene of interest cloned into a plasmid containing sequences for viral packaging and integration. This plasmid is transfected into a packaging cell by transient transfection. 2) The packaging cell produces viral enzymes, structural proteins and envelope proteins. The virus mRNA genome is copied from the vector by a RNA polymerase of the host cell. 3) Assembled viral particles containing the vector-derived genome are secreted into the medium, which is then collected. 4) After addition of the virus-containing medium to target cells, a viral envelope protein binds a protein on the target cell surface, allowing the virus to enter. 5) The virus is broken down and its mRNA genome is transcribed by viral reverse transcriptase. Upon cell division, reverse-transcribed genome and Integrase enter the nucleus and integrate into the host genome. B) PlatE cells expressing GFP coded by the viral vector. C) Target cells after infection and selection. All cells express GFP.

#### **3.4.5.1. Reduction of SP1/3 abundance by virus delivered miRNAs**

For stable knockdown, two sequences targeting SP1, two sequences targeting SP3, two sequences targeting UCP1 and two sequences targeting genes not expressed in mouse (shble and LacZ) were generated, inserted into a murine miR155 context and cloned behind an emGFP cassette. Cassettes carrying two miRNAs and one GFP were generated and inserted into retroviral vectors. Combinations generated are Ctr1 (both UCP1 miRNAs), Ctr2 (shble/LacZ), SP1 (both SP1), SP3 (both SP3) and SP1+3 a/b, carrying one miRNA against each SP1 and SP3. After infection, HIB1b cells were selected for integration and transiently transfected with reporter gene vectors. After 48h induction and 48h differentiation cells were treated with Wy14643 and Rosiglitazone for 24 hours.

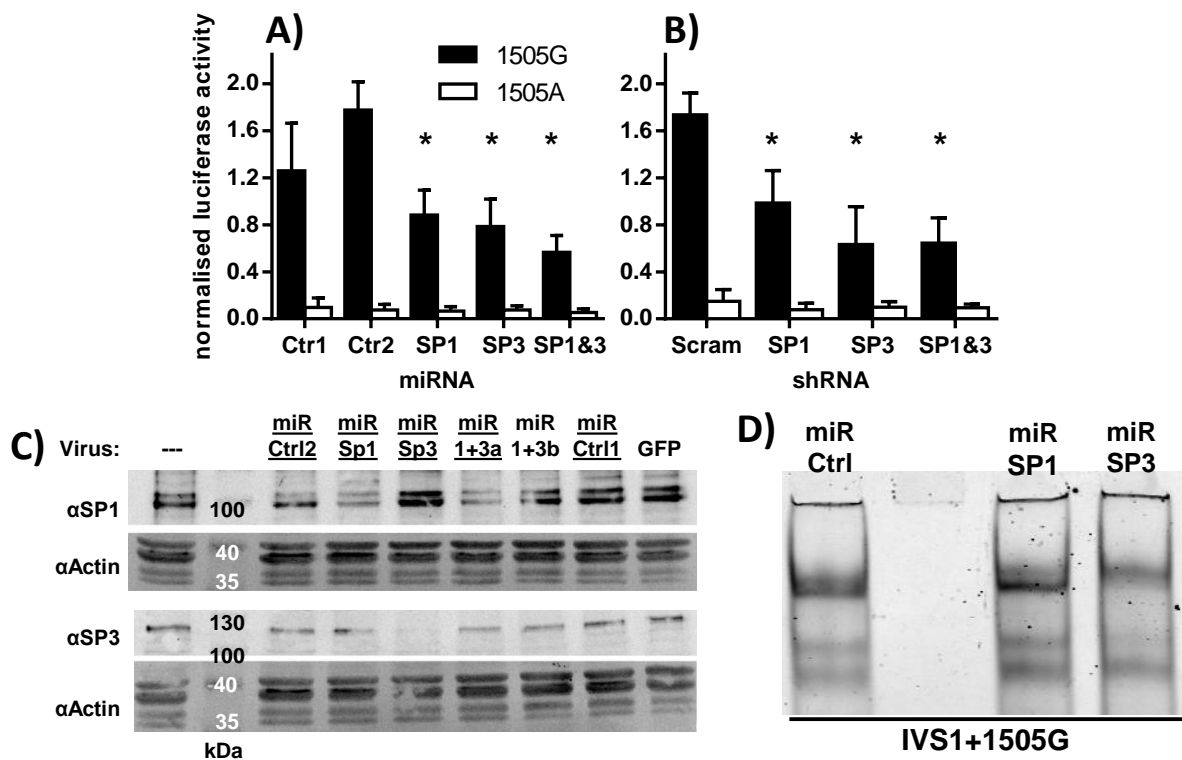
Knockdown was validated by western blot (figure 26C) and EMSA (figure 26D). Both single SP knockdown cell lines show reduced protein abundance of the respective SP-factor and formation of the respective complexes in EMSA. In some blots, compensatory upregulation of the non-depleted SP factor was observed. Of the double knockdowns, only the combination 1+3a led to depletion of SP1 and moderate depletion of SP3.

Three independent rounds of packaging, infection and selection were carried out and used for a total of eight rounds of reporter gene assays. SP1 and SP3 single knockdown reduced activity of the IVS1+1505G reporter by 30 (SP1 vs. Ctr2) to 56% (SP3 vs. Ctr2), depending on the control compared to (Figure 25A). Double knockdown of both SP-factors yielded an even stronger effect (55% vs Ctr1, 67% vs Ctr2). All effects are statistically significant ( $p < 0.05$ , Two-Way analysis of variance (ANOVA), Holm-Sidak post-hoc test). No significant effects of knockdown were observed for the IVS1+1505A construct. The reporter activity in Ctr1 infected cells showed strong variation between the different infection rounds.

To validate the results from the stable knockdown experiments and to rule out the off target effects of miRNA expression, infection and selection, an independent method of RNAi induced knockdown was chosen. Pools of four vectors expressing shRNAs either targeting SP1 or SP3 or carrying scrambled, non-target shRNAs were transiently transfected into HIB1b cells along with reporter gene constructs using the Nucleofector 96. The vectors were kindly supplied by the Guntram Suske group of the IMT Marburg. After electroporation, HIB1b cells were induced for 48h differentiated for another 24h. Then, cells were treated with Wy14643 and Rosiglitazone for 24 hours.

Knockdown of SP1 and SP3 reduced reporter activity of the IVS1+1505G by 43% and 64%, respectively (figure 26B). The effect of double knockdown was comparable to single knockdown of SP3. All effects are statistically significant ( $p < 0.05$ , Two-Way ANOVA, Holm-Sidak post-hoc test). No significant effects of knockdown were observed for the IVS1+1505A construct. Knockdown efficiency could not be validated due to moderate transfection efficiency but the results were reproduced with an independent transfection method (Lipofectamin LTX, not shown)

In conclusion, expression of SP1 and SP3 is required for activity of the IVS1+1505G construct. Reduced abundance of SP1 or SP3 leads to reduced reporter gene activity. This was demonstrated with two independent knockdown strategies employing independent RNAi-target sequences.



**Figure 26: Knockdown of SP1 and SP3 reduces reporter gene activity of the IVS1+1505G reporter.** **A)** HIB1b cells were infected with miRNA expressing retrovirae. After selection, reporter gene constructs were transfected and cells were differentiated for 96h before 24h Wy/Rosi stimulation. The different sets of columns correspond to the different miRNAs. **B)** Pools of shRNA expressing vectors were transfected into HIB1b cells along with reporter vectors. After 72h, cells were treated with Wy14643 and rosiglitazone. As in A), the column sets correspond to the different shRNA targets. **C)** Representative western blot for validation of miRNA-mediated knockdown of SP1 and SP3. Actin abundance served as loading control. Underlined cell lines were used for reporter gene assays. **D)** EMSA demonstrating the effect of SP1 and SP3 knockdown on complex formation. Ctrl1 and Ctrl2: Control/non-targeting miRNA; Scram: scrambled/non-targeting shRNAs. \* p<0.05, 2-way ANOVA for comparison to control miRNA/shRNA, Holm-Sidak post hoc test.

### 3.4.5.2. Overexpression of human SP1 and SP3

Overexpression for SP1 and SP3 was attempted by both stable and transient means. Vectors carrying the CDS of human SP1 and SP3 were kindly provided by Guntram Suske. Western blot revealed 5- to 30-fold overexpression depending on SP factor and method of transfection (not shown). In several cases, overexpression was accompanied by reduced expression of the non-overexpressed SP factor. Despite multiple attempts, no reliable effect of overexpression on reporter gene activity could be detected (not shown).

### 3.5. An intronic DR1 element is dependent on the IVS1+1505G element

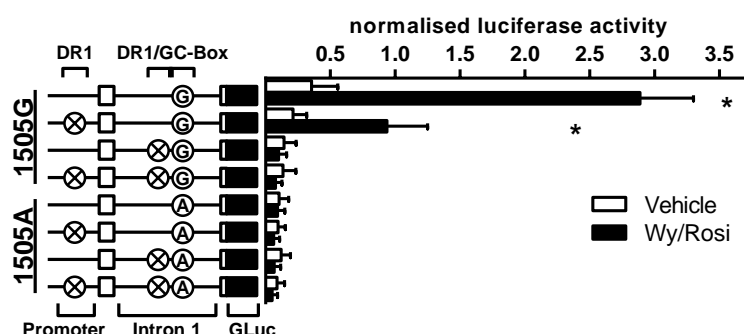
Initially we struggled to explain the effect of an intronic SP binding site on PPAR agonist induced reporter gene activity. This was partly due the fact that the DR1 element responsible for PPAR agonist action was being believed to reside in the core promoter. Furthermore this element was found to be functional in intron-less reporter gene constructs obviously not containing the intronic SP1/SP3 element [79]. This contradiction was resolved by the identification of a second DR1 element within the first intron of the *Mus musculus* UCP3 gene which binds PPAR $\gamma$  and RXR [73,74]. Two facts hint towards an interdependence of these two intronic TFBS. Firstly, the intronic DR1 element is in close proximity of the IVS1+1505 element described here, the first being only 40 bp upstream of the latter. Secondly, loss of binding to the IVS1+1505 element completely abolishes the effect of rosiglitazone on reporter gene activity of the G-allelic reporter. We thus set out to uncover the interrelationship between the two putative DR1 elements and the GC-Box.

#### 3.5.1. Targeted mutagenesis reveals importance of both intronic elements

The first step to understanding the relevance of the core promoter and intronic DR1 was understanding the contribution of both regulatory sequences. To do so, either or both elements were mutagenized in both IVS1+1505 reporter gene vectors. The eight reporter gene constructs generated were transfected into HIB1b cells and, after 16 hours, treated with either agonists or vehicle for 24 hours.

Of all reporter gene constructs, only two responded to PPAR agonist stimulation (figure 27). The largest fold increase was found with the all-WT construct, exhibiting 11-fold induction. The construct devoid of the promoter DR1 element yielded lower agonist stimulated activity and a trend towards a lower non-stimulated activity. Induction by PPAR agonist was reduced significantly to 5.6-fold. All constructs lacking either or both intronic elements were completely non-responsive to PPAR agonists and showed a trend towards reduced basal activity. Furthermore, no additional effect was seen when combining a second mutation with either intron mutation. Activity of the construct devoid of both DR1 elements and carrying the IVS1505G allele was the same as the activity of the constructs carrying only one intronic mutation.

In conclusion, both the intronic DR1 element and GC-Box are required to mediate effects of PPAR agonists on reporter gene activity. This was different from the promoter DR1 element. Constructs lacking the promoter DR1 still respond to PPAR agonists while constructs lacking either intronic element do not.



**Figure 27: The intronic DR1/GC-Box tandem element is required for PPAR agonist induction.** PCR mutagenesis was used to mutate the putative DR elements in promoter and intron. Together with the two allele of the GC-Box element this allowed generation of eight different reporter vectors containing different combinations of mutated and intact TFBS. The different constructs were transfected into HIB1b cells and treated with a combination of Wy14643 and rosiglitazone or with vehicle (DMSO) for 24 h. The bar chart shows normalised reporter activity for all reporter constructs in absence and presence of PPAR agonists. \*  $p < 0.05$ , 2-way ANOVA for Wy/Rosi compared with vehicle, Holm-Sidak post hoc test.

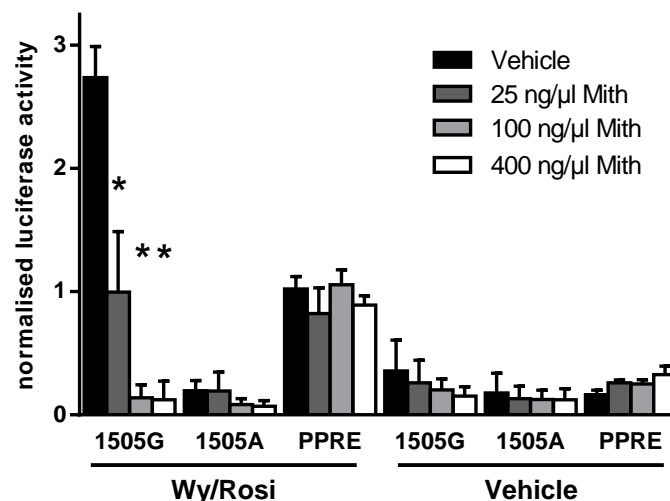
### 3.5.2. Chemical inhibition of SP1/3 binding suppresses the effects of PPAR-agonists

In addition to mutagenesis of the binding element and knockdown of the binding factor, chemical inhibition of transcription factor activity or binding is a common way to test for a factors involvement. Mithramycin is a GC-Box binding inhibitor that suppresses binding of SP transcription factors and of some of the related Krüppel-like factors (KLFs) to their respective GC-rich binding site. If the effect of PPAR agonist treatment can be abolished by addition of mithramycin, most likely a SP transcription factor is required for the response.

#### 3.5.2.1. Reporter gene assays

HIB1b cells were transfected with the IVS1+1505G, the IVS1+1505A or a consensus PPRE reporter vector. 16 hours after, the cells were treated with PPAR agonists in presence of vehicle or different concentrations of mithramycin. Figure 28 shows that both the consensus PPRE and the IVS1+1505G reporter construct were activated by PPAR agonists. For the IVS1+1505G construct, this activation was susceptible to parallel treatment with mithramycin. While 25 ng/μl mithramycin reduced the reporter activity by 64%, higher concentrations completely abolished PPAR agonist induced activation ( $p < 0.05$ , Two-Way ANOVA, Holm-Sidak post-hoc test) of the construct to the level of the A-allelic construct. The PPRE consensus vector was induced by PPAR agonists despite the presence of mithramycin and exhibited the same reporter activity at the highest mithramycin concentration as in absence of mithramycin. The IVS1+1505A construct did neither show a significant response to mithramycin nor to PPAR agonists.

These data are well in line with the result from our reportergene mutations (3.5.1), underlining that binding of SP transcription factors is indeed important for UCP3 expression and PPAR-agonist responsiveness. The resistance of the PPRE vector to mitramycin demonstrates that the mithramycin effect is not due to inhibition of PPAR binding.



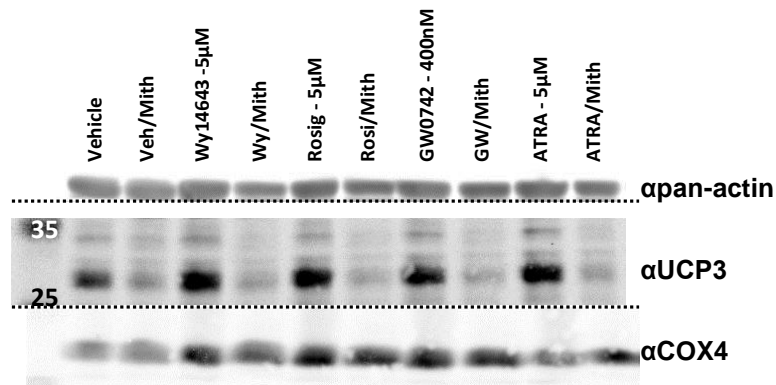
**Figure 28: Mithramycin, a GC-Box binding inhibitor, blocks the effect of PPAR agonists on the IVS1+1505G reporter.** The IVS1+1505G-, the IVS1+1505A- and a consensus PPRE reporter vector were transfected into HIB1b cells. 16 hours after transfection, the cells were treated with vehicle (DMSO) or Wy14643 and rosiglitazone in presence of different concentrations of mithramycin. After 24 hours of stimulation, reporter gene activities were measured and normalised to a CMV-driven vector. Different bar colours denote the different mithramycin concentrations. Construct and stimulation are designated at the x-axis. \*  $p < 0.05$ , 2-way ANOVA for difference to no mithramycin, Holm-Sidak post hoc test.

### 3.5.2.2. Endogenous UCP3 expression

In addition to reporter gene assays, the effect of mithramycin on UCP3 expression was assayed on the level of UCP3 protein expression. For these experiments, iBPA-L2 cells were used as, they were the only BAT cell line available that expressed a sufficient amount of UCP3 protein.

Cells were induced and differentiated for a total of six days and afterwards treated with either DMSO, Wy14643, rosiglitazone, GW0742 or ATRA in presence or absence of mithramycin for 30 hours. Afterwards, RIPA protein extracts were generated and 40 µg total protein was loaded and separated by SDS-PAGE. After transfer, UCP3, COX4 (cytochrome c oxidase subunit 4) and actin were detected (figure 29) by western blot. For all five conditions, the mithramycin-treated cells have less UCP3 protein than their non-treated counterparts. This pattern is neither seen for actin, nor for COX4. The vehicle treated cells furthermore show the lowest UCP3 protein abundance, although the COX4 signal is weakest in this sample as well, possibly pointing towards a lower mitochondrial abundance.

In conclusion, binding of SP transcription factors is important for transcription of the nuclear UCP3 gene. Thus, the effects described in 3.5.1 are neither limited reporter gene assays, nor to HIB1b cells.



**Figure 29: Endogenous UCP3 expression in iBPA-L2 cells is sensitive to mithramycin treatment.**

Cells were induced and differentiated. Upon full differentiation, cells were stimulated with PPAR or RXR/RAR agonists in presence or absence of 400 nM mithramycin. After 30 hours of treatment, RIPA extracts were generated and 40 µg protein per lane were resolved on SDS-PAGE. After transfer, the membrane was cut in three pieces which were used for three different antibody detections. Antibodies were used to detect pan-actin, UCP3 and COX4. Dotted lines mark cutting or cropping of membrane or picture.

### 3.6. DR1/GC-Box modules are present in several mammalian species

Next we wanted to know whether the intronic DR1/GC-Box module is *Phodopus* specific or a more general feature of the UCP3 gene. Initially we intended to uncover related modules in other species by sequence alignment using Dotplot (EMBOSS GUI) and ClustalW (SDSC Workbench). Using these two tools we were able to find conserved modules in *Rattus norvegicus* and *Mus musculus*. In contrast, this strategy was unsuccessful in non-rodent species. We thus used the Genomatix software package to search for GC-Boxes in close proximity to putative DR1 elements.

#### 3.6.1. Sequence analysis

The FastM tool was used to define a module consisting of a PPAR binding site and an SP1 binding site in close proximity. Next, the genomes of *Homo sapiens*, *Equus caballus* and *Sus scrofa* were screened for comparable modules downstream of the UCP3 TSS using ModellInspector. For human and pig only one DR1/GC-Box module was identified, although the for the pig module two putative SP binding components were predicted. For horse, four putative modules were predicted. A schematic sketch of the location of the SP-binding component as well as its sequence can be found in figure 30. Interestingly, despite large differences in intron 1 size, ranging from in 1840 in *Homo sapiens* to 7465 in *Equus caballus*, the DR1/GC-Box module is always found in a distance of 1600 to 2850 bp to the TSS. In human, where intron 1 is shortest of all inspected species, the module is located within the CDS in exon 2. The Genomatix part of this analysis was carried out by Tobias Fromme. All subsequent bioinformatics were carried out by Christoph Hoffmann.

#### 3.6.2. EMSA competition experiments

Cold competitors were generated for the putative SP-binding element of mouse, rat, human and pig (both elements) and tested for their ability to compete the binding activity of the hamster IVS1+1505G element in EMSA. Competitors were added in 16- and 50-fold molar excess to probe. Figure 30 shows an EMSA which includes competition by IVS1+1505G, +1505A and +1504A for comparison.

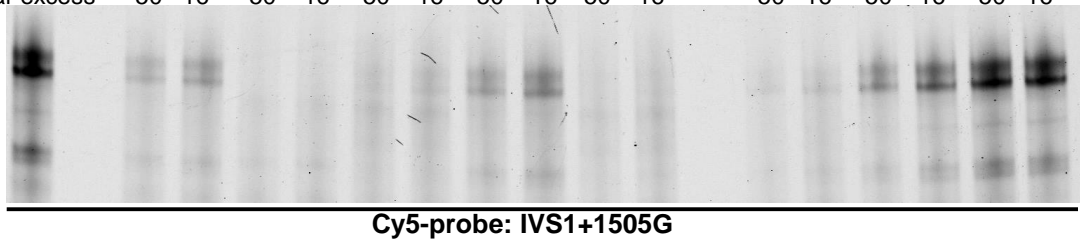
Of all elements, the human, rat and hamster IVS1+1505G element exhibit the strongest competition. The competition capacity of the mouse element is between that of the IVS1+1505A element and the strongest ones. The IVS1+1504A element (see 3.4.1) does not compete at all. For pig, element 1 strongly competes while element 2 only yields competition comparable to the IVS1+1505A element. Interestingly, ModellInspector assigns an orientation to both the DR1- and SP1/3-element, and in all cases (aside from the weaker binding, second pig element) both elements face each other.

These results suggest that DR1/GC-Box modules may be a common feature of mammalian UCP3 genes. Furthermore they hint to a functional relevance of the distance between core promoter and intronic enhancer module.

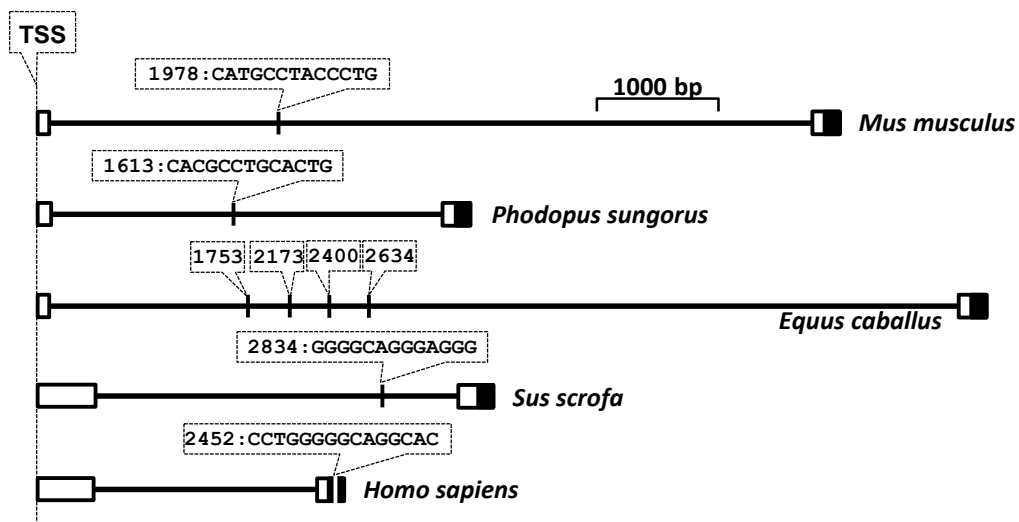
A)

**non-labelled competitor**

Competitor	Mus		Rattus		Sus E1		Sus E2		Homo		1505G		1505A		1504A	
Molar excess	50	16	50	16	50	16	50	16	50	16	50	16	50	16	50	16



B)



**Figure 30: DR1/GC-Box modules are found within the first intron of UCP3 in several species.** The FastM tool of the Genomatix package was used to define a matrix containing a DR1 site with a SP1/3 site in juxtaposition. Intronic sequences of *Mus musculus*, *Rattus norvegicus*, *Sus scrofa*, *Equus caballus* and *Homo sapiens* were screened for said module of a using Genomatix MatInspector. **A:** EMSA using the predicted SP1/3 binding components as cold competitor. Competitor and molar excess are stated above the image. Labelled, 31 bp IVS1+1505G probe was used. **B:** Location, distance to TSS and sequence of the putative SP1 binding modules found in vicinity of putative DR1 elements. The UCP3 genes of mouse, hamster, horse, pig and human are shown. Open boxes mark non translated exons, closed boxes specify coding sequence. The TSS is designated by a dotted vertical line. Intron and exon lengths are to scale.

**3.6.3. Reporter gene constructs: *Mus musculus* and *Homo sapiens***

As a second strategy to validate the DR1/GC-Box module function, reporter gene constructs for the mouse and human UCP3 gene were generated. Several different promoter sizes were generated but despite repeated attempts, none of the constructs expressed noteworthy amounts of luciferase. Different agonist stimulations were tested and constructs were cleared of possibly deleterious upstream ATG sequences, but no construct exhibited significant reporter activity in any condition assayed. Thus, function of the putative DR1/GC-Box modules could not be validated in reporter gene assays.



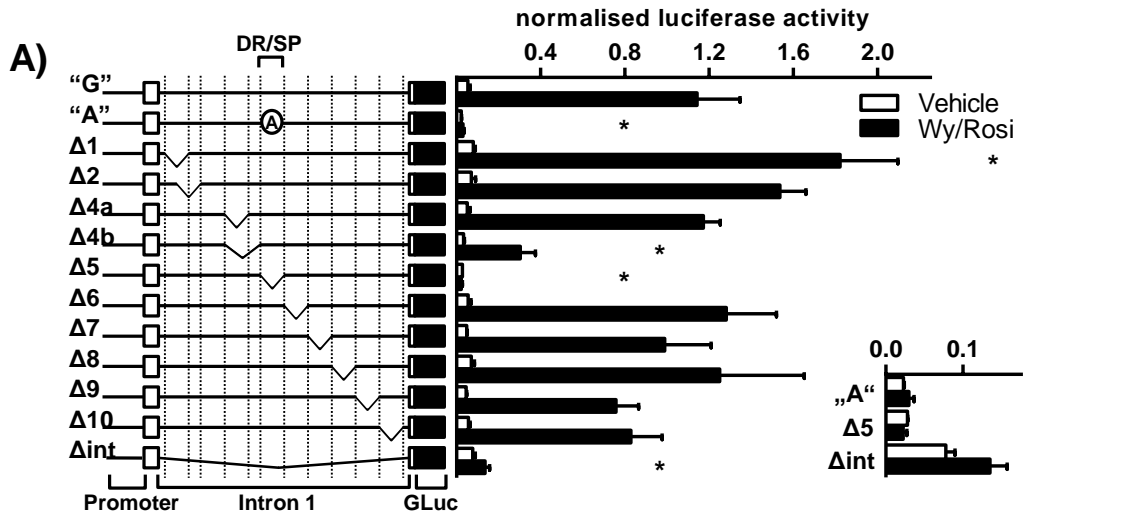
### 3.7. Deletion screens and data mining uncover additional regulatory sites

After identifying two elements located within the first intron, we set out to screen for additional regulatory sequences located downstream of the core-promoter. The whole intron was sub-divided into ten parts, and nine reporter gene constructs were generated, each lacking one part (figure 31A). For one of the ten parts, no deletion construct could be generated ( $\Delta 3$ ). Additionally, the whole first intron was deleted ( $\Delta \text{Int}$ ). Deletion  $\Delta 5$  only removed the DR1/GC-Box module, thereby being the smallest deletion. Deletions  $\Delta 4$  and  $\Delta 6$  directly flanked the two elements, ending 10 bp upstream the DR1-element and beginning 10 bp downstream of the GC-Box, respectively. Deletion  $\Delta 9$  covered the region around the IVS1+2668 base exchange (see 1.5 and figure 4). After initial experiments, a modified version of the  $\Delta 4$  construct was generated, retaining additional 36 bp of the intron directly upstream of the DR1 element.

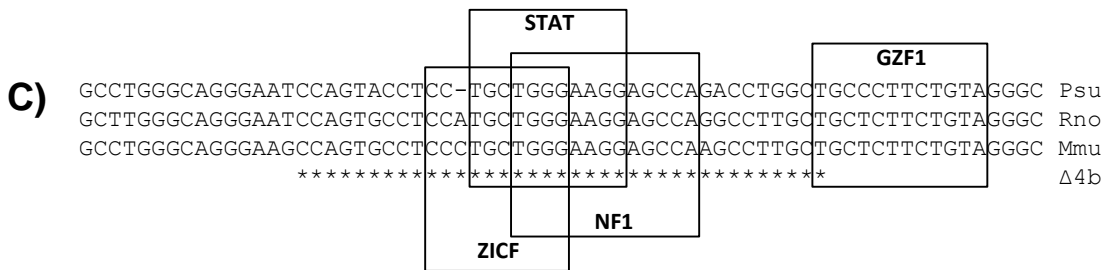
Reporter gene assays were carried out in HIB1b cells (figure 31A). Reporter activity of constructs  $\Delta 2$  and  $\Delta 6$ - $\Delta 10$  was not different from the IVS1+1505G full length construct while the construct  $\Delta 5$  was completely non-responsive to PPAR agonists. Interestingly, deletion of the whole intron generated a reporter vector seemingly more active than the IVS1+1505A and  $\Delta 5$  constructs, although the difference did not reach statistical significance (figure 31A, small insert). Deletion  $\Delta 1$  led to 60% higher reporter gene activity than the unmodified construct. Deletion  $\Delta 4$  ( $\Delta 4b$  in figure 31A) exhibited a 74% reduced reporter activity. The construct  $\Delta 4a$ , differing from  $\Delta 4b$  by only 36 bp, had full-length like reporter activity.

Sequence analysis (MatInspector) of these 36 bp revealed putative binding sites of nine different Genomatix matrix families: NRSF, RU49, HAND, STAT, ZICF, NF1, HEAT, SMAD and GZF1 (see figure 31B for full names). Of these only four were conserved in mouse and rat: STAT, ZICF, NF1 and GZF1 (figure 31C). As the GZF1 element is only mildly impaired by the deletion and the ZIC1 family transcription factors are brain specific [168], the STAT and NF1 elements remained as most likely candidates. Importantly, all binding predictions are based on Genomatix matrix families which may deviate from the transcription factor families described in other contexts. In general, the Genomatix families are rather wide. For example, the MyoD family includes several other basic Helix-Loop-Helix transcription factors known to partly share binding preferences with MyoD, Myogenin and the Myf transcription factors.

In summary, the deletion screen identifies two further intronic regions with putative regulatory elements. On one of the two regions, the region deleted in construct  $\Delta 1$ , to date no further research was done, but it may contain a repressor binding site. Instead we focussed on the element deleted in construct  $\Delta 4a$ , due to it being located adjacent to the DR1/GC-Box module. This element contains putative binding sites for activators of transcription. Candidate proteins belong to the Nuclear Factor 1 (NF1) or the Signal transducer and activator of transcription (STAT) family. The putative three-element enhancer is shown in figure 31D.



B)	Family	Full name	Strand	Genomatix name
	GZF1	GDNF-inducible zinc finger gene 1	plus	V\$GZF1
	HAND	Twist subfamily of class B bHLH transcription factors	minus	V\$HAND
	HEAT	Heat shock factors	plus	V\$HEAT
	NF1	Nuclear factor 1	minus	V\$NF1F
	NRSF	Neuron-restrictive silencer factor	plus	V\$NRSF
	RU49	Zinc finger transcription factor RU49 zinc finger proliferation 1 - Zipro1	plus	V\$RU49
	SMAD	Vertebrate SMAD family of transcription factors	minus	V\$SMAD
	STAT	Signal transducer/activator of transcription	minus	V\$STAT
	ZICF	Members of ZIC-family, zinc finger protein of the cerebellum	minus	V\$ZICF



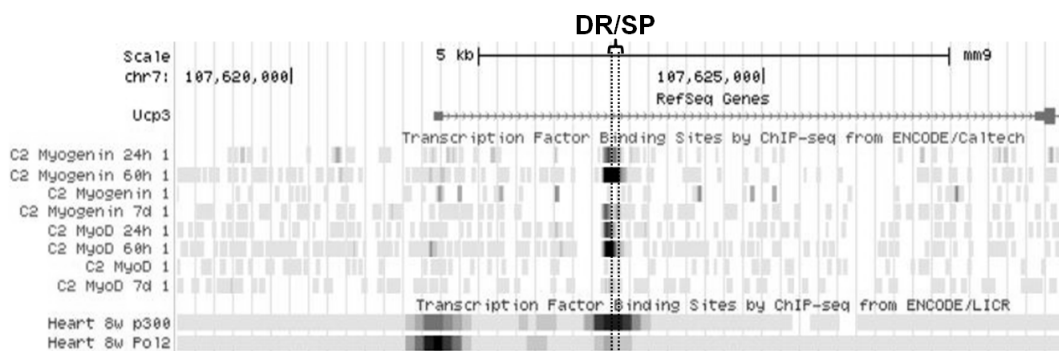
**Figure 31: Stepwise deletion of the first intron uncovers an additional region required for UCP3 expression.** Deletions were generated by incomplete amplification of the IVS1+1505G reporter gene construct. Vectors were transfected into HIB1b cells and treated with Wy14643 and rosiglitazone or vehicle (DMSO). Putative regulatory regions then were screened by Genomatix MatInspector. **A)** Normalised reporter gene activity. "A" and "G" specify the full length reporter gene constructs in the two IVS1+1505 alleles. The Δ-constructs lack different parts of the intron. Open boxes mark non-coding exons, the filled box marks the *Gaussia* luciferase gene. \* p<0.05, 2-way ANOVA for difference to full size IVS1+1505G construct, Holm-Sidak post hoc test. **B)** Transcription factor family abbreviation, full name, strand location and Genomatix family code for the matrices identified in the region distinguishing Δ4b from Δ4a (*Phodopus* sequence). **C)** Region deleted in construct Δ4b. Only binding matrices conserved between mouse, rat and hamster are shown. Asterisks mark nucleotides absent in the Δ4b reporter vector. **D)** Alignment of the putative enhancer elements in mouse, rat and hamster: A STAT or NF1 site, the PPAR $\gamma$ /RXR $\alpha$  binding DR1 site and the and the GC-box. Numbers in brackets indicate nucleotides left out.

### 3.7.1. Data mining reveals binding of MyoD, Myogenin and p300 to the intronic enhancer

The ENCODE (Encyclopaedia of DNA Elements) project is a resource dedicated to collecting information about elements of interest in the human and mouse genome and making them publicly available. Elements of interest include, among others, TFBS, DNase hypersensitive sites and transcripts. The focus lies on whole-genome data, and the project has amassed a large collection of ChIP-seq data. We screened this collection for transcription factors binding within the first intron of the UCP3 gene.

ChIP data for BAT was limited, and yielded no additional insight apart from the PPAR $\gamma$ /RXR binding site described in [73]. In contrast, for C2C12 cells and heart data from several ChIP-seq experiments was available (figure 32). In C2C12 cells ChIP-seq data were available for C/EBP $\beta$ , CTCF, E2F-4, FOSL1, Max, MyoD, myogenin, RNA Polymerase II (Pol2), REST, SRF, TCF12, TCF3 and USF2. Only for three of these factors, binding peaks were found in the region of the UCP3 gene: Pol2 binds near the TSS, and MyoD and myogenin bind in the region of the intronic enhancer. For heart, data were available for CTCF, p300 and Pol2. Here, binding peaks for p300 and Pol2 were detected. Figure 32B shows ChIP-seq data (see appendix 12 for accession numbers) for the UCP3 locus in C2C12 cells and heart. In C2C12 cells, MyoD and myogenin are binding at the intronic enhancer within intron 1. Binding peaks at 60h of differentiation, but is low at day 7 of differentiation and in myoblasts. In contrast, no binding of MyoD and myogenin was found at the core promoter. For heart, both p300 and Polymerase II both show a bimodal binding pattern with one binding peak at the core promoter and a second one at the region of the intronic enhancer. For Polymerase II binding is stronger at the core promoter while for p300 binding is stronger at the intronic enhancer.

Taken together, in the first intron of the UCP3 gene, at least in rodents, several other proteins bind near the DR1/GC-Box module. Binding of MyoD and myogenin was demonstrated for C2C12 cells and binding of p300 binds for heart. Interestingly, binding is strong mid differentiation but greatly reduced late in differentiation, despite the fact that both MyoD and myogenin, at least on mRNA level, are expressed at these states (see 3.1.4)



**Figure 32: Sequence analysis and publicly available ChIP-seq data reveal a MyoD/myogenin binding site directly upstream of the DR1/GC-Box module.** Publicly available ChIP-seq data from C2C12 cells and heart (8 weeks old mice) were used to search for further proteins binding to the intronic enhancer. The image shows the region upstream of UCP3 exon 2 (right) in *Mus musculus*, including intron 1, exon 1 and 2700 bp upstream of the TSS. Chromosome location (chr7 and base positions), scale and gene structures are indicated at the top. ChIP-seq signals for MyoD, myogenin, Pol2 and p300 are displayed below. Dotted vertical lines mark the putative DR1/GC-Box module. For C2C12 cells (C2) different time points during differentiation are shown (24h, 60h, 7d). Tracks without an indicated time point stem from undifferentiated myoblasts.

### 3.7.2. Bioinformatic search for complex modules

Functionally relevant TFBS rarely occur alone. Instead, usually several TFBS are grouped within a regulatory region and need to interact to execute their function. That way several signals can be integrated and complex control systems are generated. Such combinations of elements are referred to as modules. In silico, models can be generated that resemble such modules. These then are a powerful tool to predict meaningful regulatory regions. The initial search for DR1/GC-Box modules revealed such modules in hamster, mouse, rat, man, pig and horse (see 3.6.2). Subsequently, deletion screens, bioinformatics and data mining revealed presence of additional putative TFBS. We combined these pieces of information to obtain a more complete model of the intronic enhancer. Using this model we attempted to identify related enhancers in human, horse and pig. For all predictions, two stringency parameters were applied: For individual TFBS, the matrix similarity determines how many deviations from the consensus sequence are tolerated, with strict allowing for less deviation than relaxed. Secondly, the element order is of importance. We checked whether the TFBS order of a region matches the order found in hamster (strict) or whether permutations are allowed (relaxed). All predictions were based on the Genomatix matrix families.

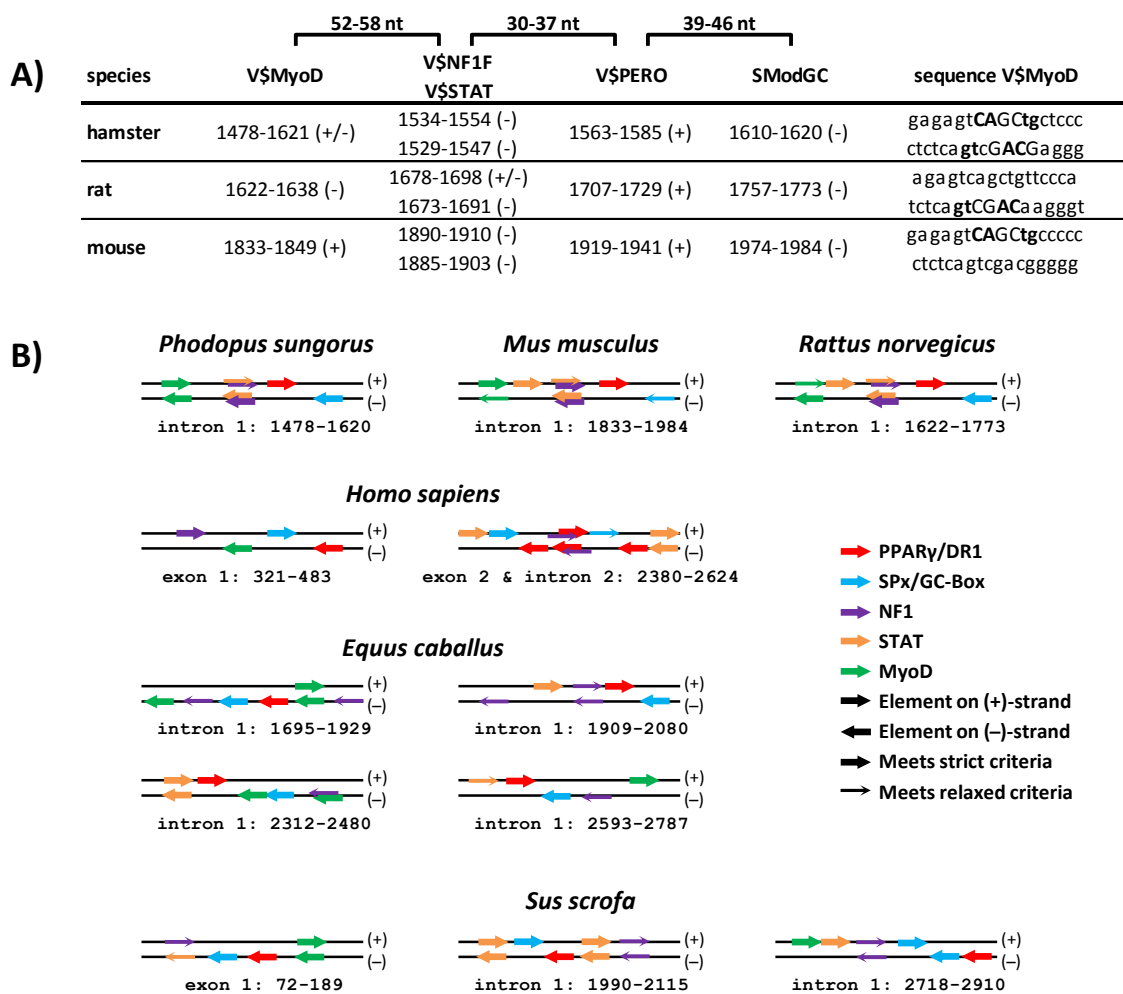
The first step was to search for conserved MyoD/myogenin binding elements within the region covered by the intronic MyoD/myogenin CHIP-Seq peak. MatInspector readily identified a putative binding element of the MyoD matrix family in hamster, mouse and rat, located 55 bp upstream of the NF1/STAT element. The sequence is CAGCTG for all species (figure 33A). While this core sequence is perfectly conserved and reverse complementary for both strands, MatInspector assigns different strand location for the elements due to variances in the flanking region. Notably, a high number of putative individual TFBS is predicted even when using strict sequence similarity: Nine MyoD sites, six GC-Boxes, seven PPAR $\gamma$ - or DR1 elements, seven STAT matrices and five NF1 sites were identified within the 3335 nucleotides of exon 1 and intron 1 of hamster UCP3. To circumvent this problem, we next searched for grouped sets of elements.

We assembled multi-TFBS models (FastM) which fulfilled three requirements. Firstly, they contained a DR1/GC-Box resembling the rodent module. Secondly, they contained a MyoD family binding element. Thirdly, they contained either a NF1 element, or a STAT element, or both. Afterwards, exon 1, intron 1 and exon 2 of the UCP3 gene were screened using these models (ModellInspector). In case of human, furthermore intron 2 was included. The all-strict model readily identified the intronic enhancer region in mouse, rat and hamster (figure 33A), but was unable to do so in pig, human and horse. However, by searching for alternative element orders, one additional region each in mouse, human, pig and horse were identified (see appendix 9). For human and pig, the regions were located within exon 1. For all other species, the regions were located in intron 1. Lastly, regions with strict element order, but relaxed TFBS similarity were searched. This time, five additional regions in hamster, three in rat, seven in horse and one in pig were uncovered (see appendix 9). Screening for models relaxed in both parameters yielded an excess of predictions. Element location, order and orientation of selected regions are shown figure 33B.

Last, we searched for DR1/GC-Box modules with strict matrix similarity and independent of distance to the TSS. This identified a total of 20 regions of interest across all six species, with some regions containing more than one putative module. We then proceeded to manually search for STAT, NF1 and MyoD TFBS in proximity of these regions (MatInspector). Furthermore the regions identified by ModellInspector were inspected. Of all the regions identified by all strategies, ultimately we picked two regions in human, three in pig and four in horse which best resemble the rodent enhancer setup.

Schematic delineations of these regions, as well as the putative rodent enhancer regions, are depicted in figure 33B. Detailed information on the regions as well as the parameters used for prediction can be found in appendix 9.

In summary, we generated a possible model for the setup of the UCP3 intron 1 enhancer in hamster, mouse and rat. Using this information, putative enhancer regions resembling the rodent UCP3 enhancer module were identified. We now have a list containing a manageable number of promising candidate regions awaiting validation. Furthermore, we identified the putative MyoD/myogenin binding element conserved between hamster, mouse and rat. The location of this element co-localizes with the location of the MyoD/myogenin CHIP-seq peak (see 3.7.1).



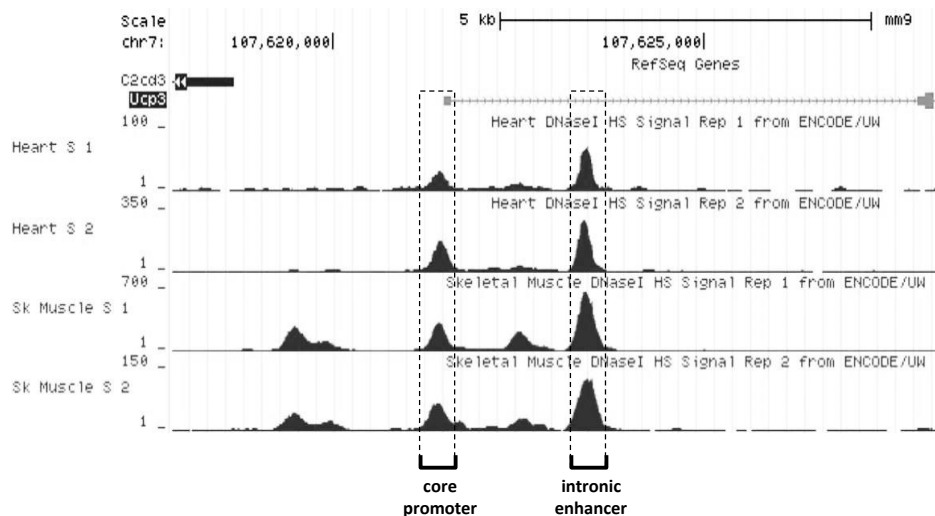
**Figure 33: Putative enhancer regions in different species: A)** Putative setup of the enhancer modules in mouse, rat and hamster. The last column contains the sequence of the putative MyoD element (capital letters: consecutive core sequence, bold letters: highest conserved). Only the TFBS with the highest sequence similarity are included. The numbers mark the location in relation to the TSS. (+) and (-) indicate the DNA strand on which the module resides. The numbers differ from the IVS numbering by the size of the first intron. **B)** Schematic of the enhancer region in hamster, rat and mouse, and putative enhancer regions within the UCP3 genes of human, horse and pig. Parallel black lines indicate the two DNA strands while arrows indicate putative TFBS. Arrow thickness indicates the sequence similarity to the consensus sequence. In contrast to A), putative binding elements of lower sequence similarity are included.

### 3.8. DNase I hypersensitive sites supports relevance of the intronic enhancer

Using data from the ENCODE project, we gathered further evidence about the enhancer module. The whole region of the UCP3 gene was inspected for the presence of activating histone modifications and presence of DNase I hypersensitive (DHS)-sites in mouse. DHS-sites are a marker of ongoing transcriptional regulation.

Screening the data on activating histone modifications yielded little insight. A large set of CHIP data on histone modification was available, but the different signals were broad and sketchy, inconclusive or partly even contradictory (not shown). In contrast, distinct DHS signals were present at both the core promoter and intron 1 enhancer in SKTM and heart (figure 34). No data were available for BAT. In both tissues, the strongest DHS signal was located at the intronic enhancer. In addition to these sites, an additional DHS region was present about 2000 bp upstream of the TSS in SKTM, but not in heart. Furthermore, one smaller peak was present between the core promoter and the intronic enhancer, again specific for SKTM.

In summary, the prominent DHS-site at the intronic enhancer region supports importance of that region for regulation of transcription. The presence both in SKTM and heart indicates that the intronic enhancer is relevant for transcription regulation in these tissues as well. Furthermore, a DHS-site located about 2000 bp upstream of the TSS of UCP3 hints towards an additional upstream enhancer, which might possibly be SKTM specific.



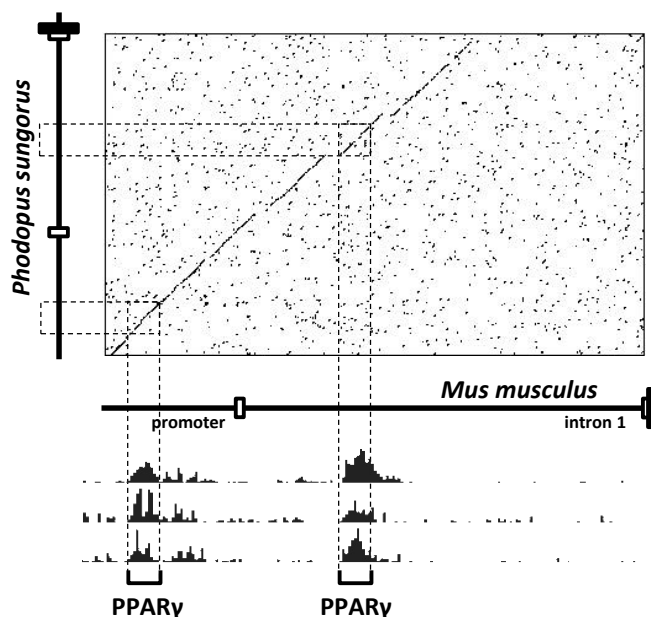
**Figure 34: DNase I hypersensitive (DHS) sites in the vicinity of the UCP3 Gene:** Publicly available DHS-site data (ENCODE) from skeletal muscle and heart were inspected to aid identification of putative enhancer regions. The image shows the region upstream of UCP3 exon 1 (right) up to exon 1 of the neighbouring C2cd3 gene. Chromosome location (chr7 and base positions), scale and gene structures are indicated at the top. DHS-site signals of two skeletal muscle and two heart samples are plotted below. Dashed boxes indicate location of core promoter and the intronic enhancer.

### 3.9. Discovery of an additional upstream DR1 element binding PPAR $\gamma$

In SKTM, three DHS-sites are located within the region of the UCP3 gene. One is located at the intronic enhancer, one at the core promoter and one roughly 2000 bp upstream of the TSS. For the first two sites, a plethora of data supports their importance for regulation of UCP3 transcription. In contrast, little is known about putative upstream regulator regions. Until recently, only a COUP-TF2 element [98] was described to be located upstream of the core promoter. The COUP-TF2 element is not located in the region covered by the DHS site, though. Very recently, colleagues reported a nuclear response factor 2 (Nrf2) binding antioxidant response element (ARE) [47] which integrated ROS signalling in heart and SKTM. Interestingly, this element is located within the region covered by the upstream DHS signal (figure 34).

Screening ChIP-seq data from a recent publication [169], we observed two putative PPAR $\gamma$  binding sites in mouse epididymal and inguinal WAT, and BAT. One was located at the region of the intronic enhancer, the other 2000 bp upstream of the TSS (figure 35). Both peaks cover few hundred nucleotides. While the binding element is known for the intronic peak [73], no distinct element is described for the upstream peak. Notably, this upstream peak co-locates with the upstream DHS-site (figure 34) and the ARE. ENCODE data was screened for other putative binding factors, but none of the other transcription factors for which ChIP-seq data was available bound near the upstream DHS-site. Dotplot (EMBOSS GUI) and alignments (SDSC workbench) of the *Mus musculus* and the *Phodopus sungorus* UCP3 regulatory region demonstrate that the elements marked by ChIP-seq peaks in mouse most likely are also conserved in hamster (figure 35).

In summary, three findings support presence of a regulatory region roughly 2000 bp upstream of the UCP3 TSS. Presence of a Nrf2 binding ARE, a DHS-site and a PPAR $\gamma$  ChIP-seq peak. Currently it is not sure which function the putative enhancer region has in which tissue, since neither of the three pieces of evidence was found in all three tissues, but in each tissue one or another piece has been found.



**Figure 35:** The first intron of *Phodopus sungorus* UCP3 resembles the first 4000 bp of its *Mus musculus* counterpart. The region upstream of the UCP3 promoter, core promoter, exon one, and the first intron of UCP3 in *Mus musculus* and *Phodopus sungorus* were compared in a dotplot. The *Phodopus* gene is found at the Y-axis, the *Mus* gene at the X-axis. Empty boxes denote non-translated exons, filled boxes mark the translated part of exon 2. To the *Mus* gene, PPAR $\gamma$  ChIP-seq reads from BAT, iWAT and eWAT are aligned. Dashed-line boxes project the ChIP signals onto the dotplot and to the *Phodopus* gene. The SP-binding GC-Box is located within the intronic PPAR $\gamma$  peak.

### 3.10. Tissue specific regulation of UCP3 expression

The initial hypothesis, that the IVS1+1505 element binds a BAT specific transcription factor, did not prove true. Instead, the element binds two general transcription factors, SP1 and SP3 (figure 22), which are present in all cell lines and tissues. Moreover, EMSA (figure 9) and reporter gene assays (figure 11) suggested that the IVS1+1505G element is of relevance in SKTM as well. This relevance was further underlined by ChIP-seq data demonstrating binding of MyoD and myogenin near the intronic module (figure 32). We therefor formulated three new hypotheses why the IVS1+1505G/A polymorphism had different effects in *Phodopus sungorus* SKTM and BAT:

- 1) Another enhancer element is active in SKTM that compensates for the loss of function at the intronic enhancer, but this enhancer element is not part of the reporter gene vectors used.
- 2) The agonist cocktail that has been used in cell culture experiments is different from the receptor agonists present in a muscle cell in vivo. Thus, an important stimulating agonist, who would allow for UCP3 transcription in presence of the IVS1+1505A allele in SKTM, is not present in our cell culture experiments.
- 3) The protein-complex binding the intronic enhancer in BAT and SKTM is different in a way that the weakened SP1/3 binding to the IVS1+1505A allele is either stabilised or compensated for in SKTM, but not in BAT. This last hypothesis presupposes that said SKTM protein is absent in C2C12 cells.

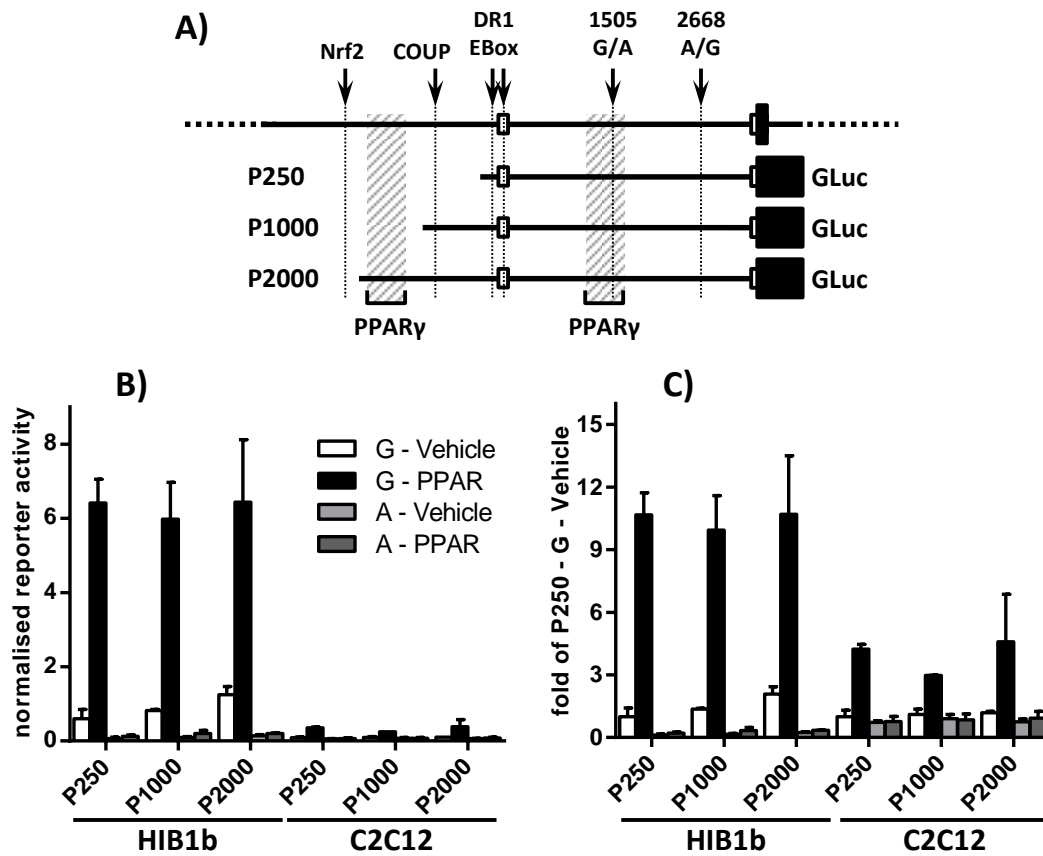
#### 3.10.1. The putative upstream enhancer elements does not influence reporter activity

As a first step we included the upstream COUP-TF2 element and the putative upstream enhancer in our reporter gene constructs. For both the IVS1+1505G and the IVS1+1505A reporter construct the promoter was enlarged twice, generating vectors carrying a 1000 bp (P1000) and a 2000 bp promoter (P2000) fragment. The initial constructs which only contain the core promoter are referred to as "P250". The 1000 bp fragment includes the COUP-TF2 site while the 2000 bp fragment contains the far upstream PPAR $\gamma$  binding region (figure 36A). Since the Nrf2 binding ARE was not known at the time the vectors were constructed, it was not part of any construct.

All six vectors (three promoter sizes and two intronic GC-Box alleles) were transfected into HIB1b and C2C12 cells and 16 hours later cells were stimulated. Reporter gene assays were carried out after 24 hours of stimulation. For both cell lines, stimulation medium contained insulin (20 nM) and T3 plus (1 nM) vehicle or PPAR agonists: Wy14643 (1  $\mu$ M) and rosiglitazone (1  $\mu$ M) for HIB1b, rosiglitazone (1  $\mu$ M) and GW0742 (0.2  $\mu$ M) for C2C12. For all six reporter vectors in both stimulatory conditions, normalised reporter activity in HIB1b cells was greater than in C2C12 cells (figure 36B). When standardised to the vehicle treated P250 IVS1+1505G reporter activity (figure 36C), the reporter constructs exhibited a higher fold induction by PPAR agonists in HIB1b cells than in C2C12 cells. Interestingly, despite the large difference in promoter size between the P250, P1000 and P2000 vector, no significant differences in reporter activity were observed, aside from a trend to a higher non-stimulated activity of the G-allelic P2000 vector in HIB1b. None of the vectors containing the IVS1+1505A allele was responsive to PPAR agonists, neither in C2C12, nor HIB1b cells.

In conclusion neither of the two upstream elements can compensate for the loss of the intronic GC-Box in our experimental setting. The upstream PPAR $\gamma$  binding site furthermore seems to have no effect on PPAR $\gamma$  agonist signalling in our experimental setup.





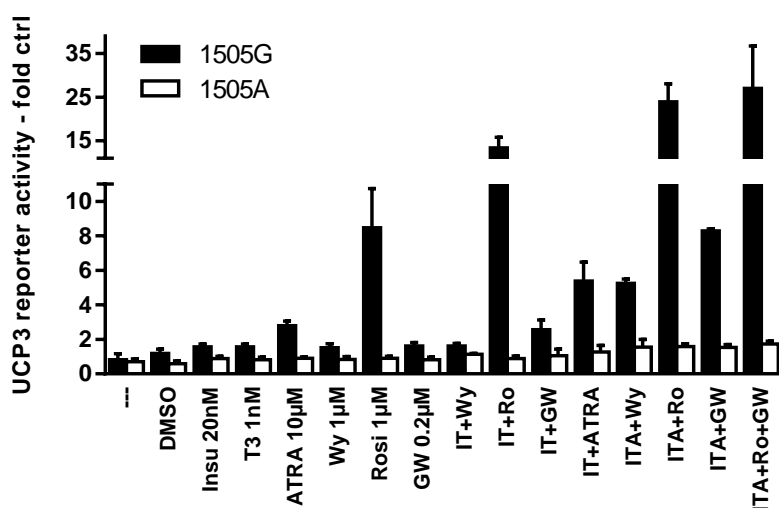
**Figure 36: Neither inclusion of the CoupTF2 nor the upstream PPAR $\gamma$  element can rescue the effect of the IVS1+1505A allele.** Different reporter vectors containing differently sized promoter fragments (250 bp, 1000 bp or 2000 bp) were generated and transfected into HIB1b and C2C12 cells. **A)** Schematic diagram of the three reporter gene constructs and relevant published TFBS. Empty boxes designate non-translated exons, filled boxes mark the translated region. PPAR $\gamma$  ChIP-seq peaks are marked by the hatched areas. These constructs were assayed for their basal activity and their PPAR agonist (1  $\mu$ M Wy14643/1  $\mu$ M rosiglitazone for HIB1b, 1  $\mu$ M rosiglitazone/0.2  $\mu$ M GW0742 for C2C12) responsiveness. For both cell lines medium contained 20 nM insulin and 1 nM T3. P250, P1000 and P2000 indicate the size of promoter fragments. **B)** Normalised reporter gene activity. **C)** Reporter gene activity standardised to the vehicle-treated P250-IVS1+1505G reporter construct. The reporter gene constructs used through the previous experiments carried the 250 bp promoter fragment. A recently discovered Nrf2 element is absent in all constructs as it was not yet discovered at the time the constructs were cloned.

### 3.10.2. Synergism of rosiglitazone, all-trans retinoic acid and T3 in C2C12 cells

The absence of an observable effect of inclusion of the upstream PPAR $\gamma$  element might stem from the absence of the correct stimulus. While rosiglitazone is the suitable agonist for the PPAR component, putative binding partners might need binding of their respective agonist as well. Thus, C2C12 cells were transfected with the P2000 reporter gene vector in both alleles and treated for 24 hours with a wide array of agonists alone and in several combinations.

Figure 37 shows the raw reporter activity standardised to the mean of controls (vehicle and non-treated). Raw reporter activity was used because ATRA treatment influenced activity of the CMV promoter used in the normalisation vector, a fact that has been described previously [170]. Three levels of synergism seemed to exist. Rosiglitazone alone was able to induce reporter activity 8.5-fold. When T3 and insulin were added, the fold induction was 13.4-fold. Further addition of ATRA led to a 23-fold induction compared to vehicle/non-treated. ATRA alone yielded a 2.8-fold increase of reporter activity. Neither T3 nor insulin led to strong activation of reporter activity when used alone but increase reporter activity 5.4-fold when combined with ATRA. The PPAR $\delta$ -agonist GW0742 did not yield obvious induction or synergism. No treatment used in this experiment was able to significantly induce activity of the IVS1+1505A reporter, although a trend to increased activity was visible for the most potent cocktails.

In summary, the UCP3 reporter gene construct requires presence of several different agonists for maximal activation in C2C12 cells. In this cocktail, ATRA, T3 and rosiglitazone act synergistically. Although this cocktail leads to a strong induction of the IVS1+1505G construct, no cocktail was found that can overcome the loss of the intronic GC-Box by the IVS1+1505A allele.



**Figure 37: In C2C12 cells, the P2000 reporter requires a cocktail of several receptor agonists for full activity:** C2C12 cells were transfected with the P2000 IVS1+1505G reporter and, 16 hours later, treated with differentiation medium containing the respective agonists for 24 hours. The graph shows the raw *Gussia* luciferase activity standardised to the mean of no treatment (---) and DMSO. On the x-axis, the different agonists or agonist combinations are designated, along with the respective concentration. In combination treatments, the same concentrations were as in the single treatments. Insu: Insulin, Wy: Wy14643, Rosi/Ro: Rosiglitazone, GW: GW0742, IT: Insulin and T3, ITA: Insulin, T3 and ATRA.

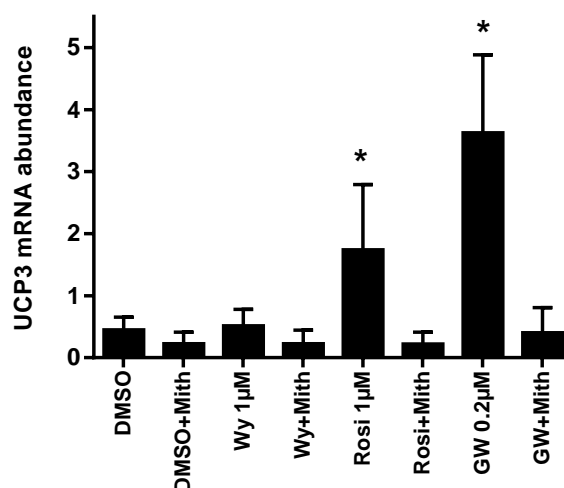
### 3.10.3. PPAR agonist induced UCP3 expression is sensitive to mithramycin in C2C12 cells

Next we wanted to know whether GC-Box binding transcription factors are of importance for expression of the endogenous UCP3 gene in C2C12 cells. We thus treated fully differentiated cells with PPAR agonists in presence or absence of mithramycin. We thus differentiated C2C12 cells for eight days and subsequently stimulated them for 24 hours with Wy14643 (1  $\mu$ M), Rosiglitazone (1  $\mu$ M) or GW0742 (0.2  $\mu$ M), either in absence or in presence of 400 nM mithramycin. The abundance of UCP3 mRNA was measured by quantitative PCR (qPCR).

While Wy14643 had no effect, rosiglitazone led to a 3.9-fold and GW0742 to a 8-fold increased abundance of UCP3 transcript when compared to vehicle (figure 38). Parallel exposure to mithramycin completely blocked both agonist effects. Furthermore, in all four mithramycin treated wells, UCP3 mRNA abundance showed a tendency to be lower than in the vehicle (DMSO) treated well, but this trend did not reach significance.

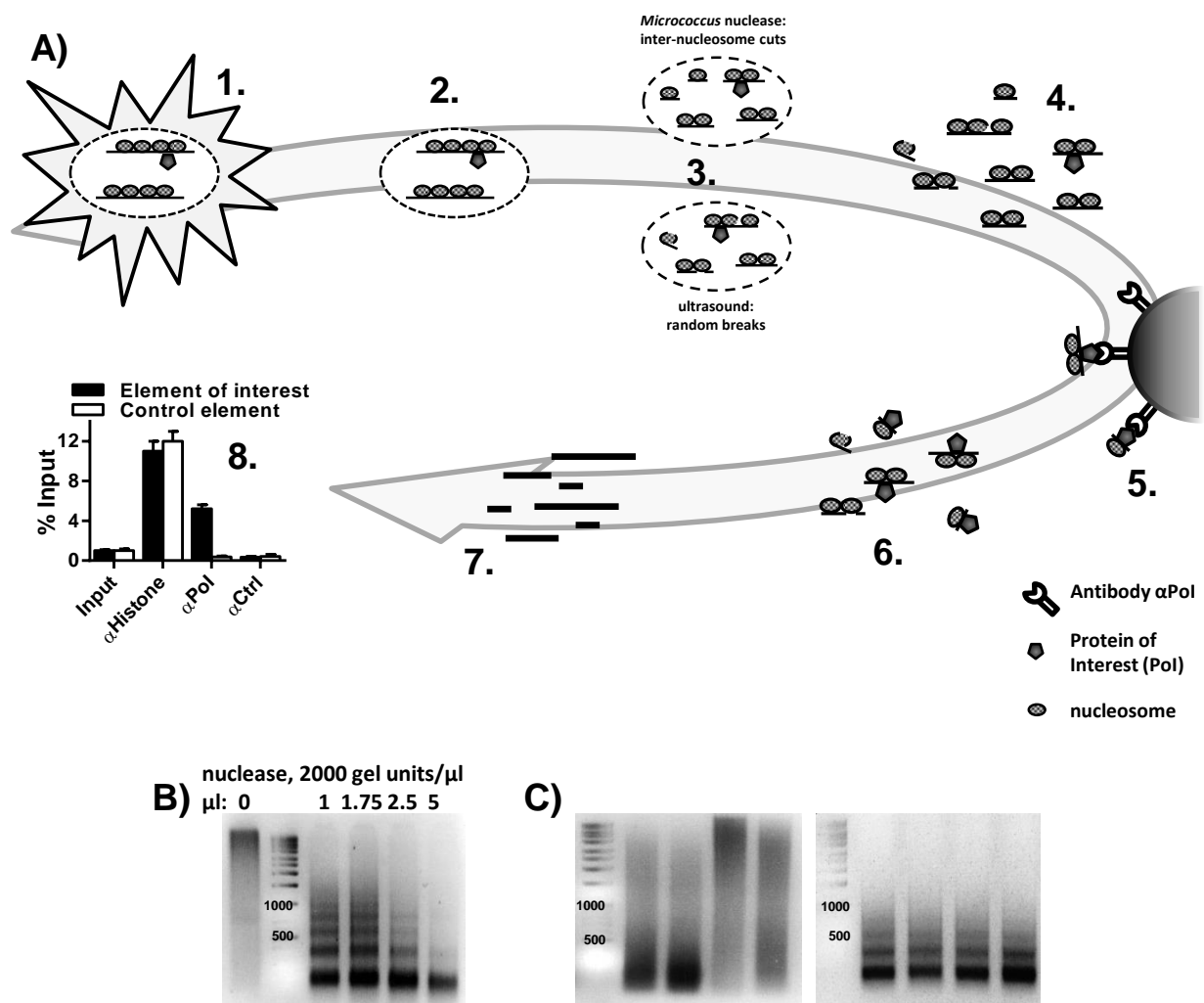
This demonstrates that PPAR-agonist induced induction of UCP3-transcription is dependent on GC-Box binding transcription factors in fully differentiated C2C12 cells. The dependency is true for both PPAR $\gamma$ - and PPAR $\delta$ -agonists.

**Figure 38: Endogenous UCP3 transcript is responsive to PPAR $\gamma$ - and  $\delta$ -agonists in C2C12 cells:** C2C12 cells were differentiated for eight days and afterwards stimulated with PPAR agonists in presence or absence of 400 nM mithramycin for 24 hours. Values were inter-day normalised. Three different agonists were used: Wy14643 (Wy, PPAR $\alpha$ ), rosiglitazone (Rosi, PPAR $\gamma$ ) and GW0742 (GW, PPAR $\delta$ ). DMSO served as vehicle control for both PPAR-agonists and mithramycin. Agonists and agonist concentrations are stated below each bar. \*  $p < 0.05$ , 1-way ANOVA for difference to DMSO, Holm-Sidak post hoc test.



### 3.11. Chromatin Immunoprecipitation

Next we aimed to demonstrate binding of SP1 and SP3 to the IVS1+1505G element in the nuclear genome using Chromatin Immunoprecipitation (ChIP). The basic principle of ChIP is outlined in figure 39A. Briefly, DNA-bound proteins are covalently crosslinked to their binding site. Nuclei are released and the chromatin is fragmented by digestion with micrococcus nuclease or by ultrasound treatment. Crosslinking time and amount of nuclease (figure 39B) or ultrasound treatment needed to be optimised to yield chromatin fragments of 150 to 1000 bp. Fragments too long reduce the resolution while shorter fragments reduce PCR efficiency. Afterwards, antibodies are used to target the protein of interest, thereby enabling precipitation of antibody-chromatin complexes using protein A or protein G covered beads. After elution of complexes, crosslinking is reversed, protein digested and DNA is purified. The purified DNA then serves as a template for PCR where abundance of a region of interest is assayed in each precipitation. For the experiments shown below, micrococcus nuclease was employed for shearing as it yielded more consistent and reproducible shearing results (figure 39C).



**Figure 39: Chromatin shearing and preparation.** **A)** Schematic workflow: **1)** Protein-DNA interactions are covalently cross-linked by treatment with formaldehyde. **2)** Nuclei are released. **3)** Chromatin is fragmented either by *Micrococcus* nuclease digest or using ultrasound. **4)** Fragmented chromatin is released by ultrasound treatment. **5)** Beads loaded with an antibody raised against the protein of interest (PoI) or controls (Histone, neg Control) are incubated with the chromatin fragments. **6)** After washing an elution, fragments bound by the PoI should be enriched. **7)** After removal of protein the abundance of the fragment of interest as well as a control region are measured using quantitative PCR **(8)**. Typically, PCR data are standardised to %Input or fold negative control antibody for each primer pair. **B)** Shearing optimisation: Conditions for shearing are optimised to yield chromatin fragments ranging from 150-1000 bp by varying amounts of micrococcus nuclease. Nuclease treatment yields DNA fragments sized a multiplicity of 150 bp. The distribution achieved with 2.5  $\mu$ l nuclease is optimal. **C)** Reproducibility of ultrasound shearing (left) and nuclease shearing (right). Of four cell batches sheared with ultrasound only two meet the expectations while two show suboptimal shearing. For nuclease shearing, all four cell batches sheared show comparable fragment distribution.

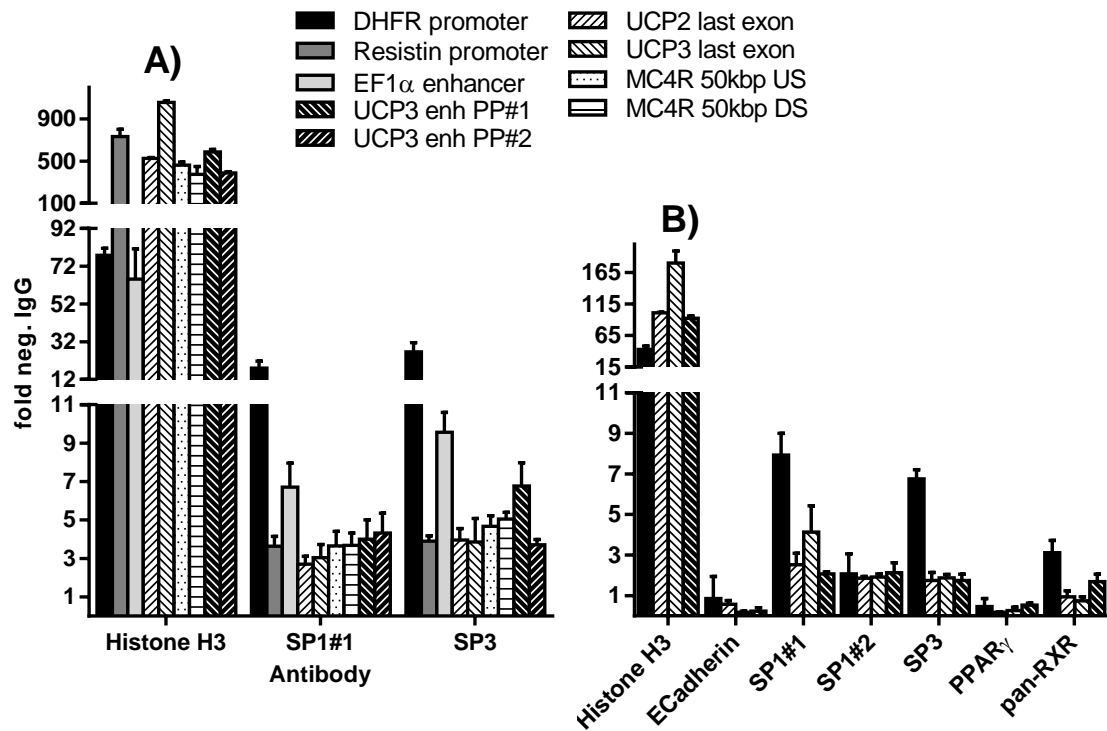
### **3.11.1. Methodical problems: Neither SP1 and SP3 antibodies, nor published positive control antibodies precipitate the UCP3 intron 1 enhancer module**

The classic ChIP approach is to use antibodies targeting the endogenous proteins expressed by a cell. The advantage of this strategy is that the physiological amount of transcription factor in its natural conformation/modification status serves for precipitation. The drawback of this strategy is that ChIP grade antibodies are required. Antibodies suitable for supershift and western blot may not be suitable for ChIP, and ChIP-grade antibodies usually are validated for the human protein in human cells. For ChIP, iBPA cells were used, as previous experiments (figure 29. 3.5.2) demonstrated mithramycin-sensitive expression of UCP3 protein.

Cells were fully differentiated and treated with 5  $\mu$ M rosiglitazone and 5  $\mu$ M ATRA for six hours before crosslinking. 5  $\mu$ g of each antibody were used, except for  $\alpha$ Histone H3. Of H3 antibody 10  $\mu$ l were used but no concentration was stated by the manufacturer. Two independent experiments were carried out: The first experiment aimed to test an array of positive and negative control target regions. Three positive control regions known to bind SP1 (DHFR and resistin promoter, EF1 $\alpha$  intron 1 enhancer) and four negative control regions (50 kDa upstream and downstream of the MC4R gene, the last introns of UCP2 and UCP3) far off known regulatory elements were selected. Lastly, two primer pairs targeting the UCP3 intron 1 enhancer were used. Four antibodies were used:  $\alpha$ Histone H3, normal IgG,  $\alpha$ SP1 and  $\alpha$ SP3, all raised in rabbit. For all IPs, abundance the target regions were quantified using qPCR. Figure 40A shows abundance standardised to normal IgG (negative control antibody). For all regions, the histone antibody purified 50- to 900-fold more DNA than the negative control antibody, validating that in principle assay and readout work. For the SP1 and SP3 antibody, the DHFR promoter achieved the highest enrichment (14- and 23-fold) over negative control IgG. For the EF1 $\alpha$  enhancer, enrichment was 6.7- and 9-fold. For the UCP3 enhancer, only the SP3 antibody led to some enrichment (6.7-fold) above the negative control antibody, and only for one primer pair. All other primer/antibody combinations, including negative control DNA regions, yielded 2.7- to 5-fold enrichment over negative IgG. This demonstrates that an antibody binding a transcription factor will non-specifically enrich any DNA element when compared to a negative control IgG. A meaningful interaction can only be identified by comparing enrichment of different genome regions.

In a second experiment, four selected regions and eight different antibodies were tested. In addition to the four antibodies from the first experiment,  $\alpha$ ECadherin was included as a second negative control while the PPAR $\gamma$  and pan-RXR antibody used in a related publication [73] served as positive controls. Furthermore, another SP1 antibody was tested. All antibodies were raised in rabbit. Again, the histone antibody yielded good enrichment for all regions, although for all antibodies and all regions enrichment was generally lower (figure 40B). Other than enrichment of the DHFR promoter by the SP1 and SP3 antibody used in the first experiment, no enrichment was detected. Neither antibody purified more of the UCP3 intron 1 element than of the negative control regions.

In principle ChIP works in our hands, demonstrated by the fact that the histone antibody precipitates up to 1000-fold more DNA than a negative control antibody. In contrast, enrichment by the antibodies binding SP1 and SP3 is low when comparing positive control regions with negative control regions. Furthermore, no antibody yielded reliable enrichment of the IVS1+1505 element, including the PPAR $\gamma$  and RXR antibodies used as positive controls. Thus, the data gathered does not allow any conclusion about binding of SP1 and SP3 to the intronic enhancer module.



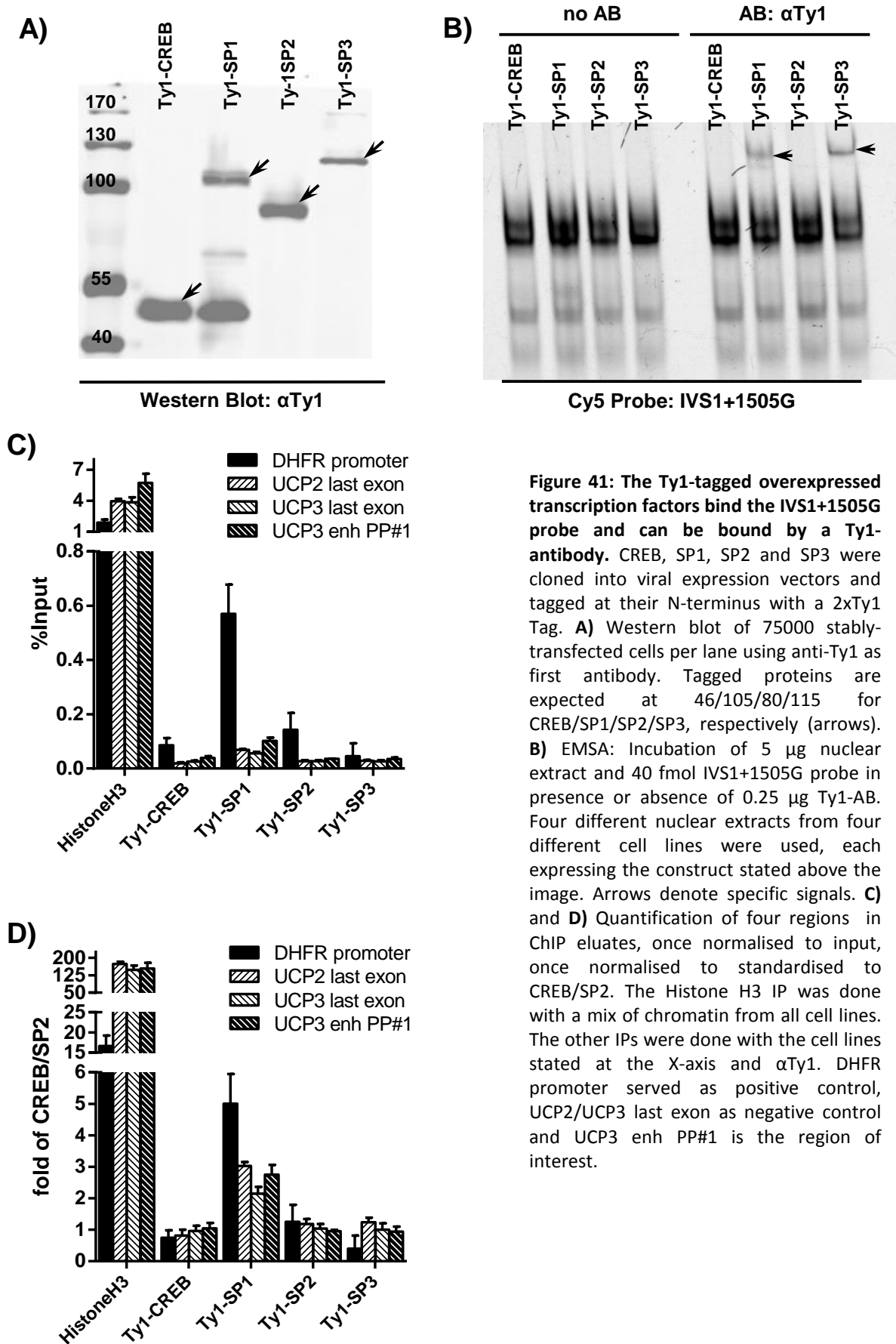
**Figure 40: Technical problems in ChIP: positive control antibodies fail to precipitate the IVS1+1505G enhancer.** ChIP was carried out with *Micrococcus* nuclease sheared chromatin generated from iBPA-L2 cells. **A)** In the first experiments, a wide array of primer pairs was compared in four IPs (negative IgG,  $\alpha$ Histone H3,  $\alpha$ SP1#1 and  $\alpha$ SP3). Three of the primers were amplifying putative positive control regions (DHFR promoter, Resistin promoter and EF1 $\alpha$  intronic enhancer), four were negative control regions (UCP2+3 last Exon, MC4R gene 50 kbp up- and downstream) and two primer pairs targeted the UCP3 intronic enhancer (UCP3 enh PP#1+#2). **B)** Four of the primer pairs from A) were used on a second set of ChIP samples. In addition to the first set, more antibodies were included:  $\alpha$ ECadherin as a second negative control,  $\alpha$ SP1#2 as a second antibody against the prime candidate, and  $\alpha$ PPAR $\gamma$  and  $\alpha$ pan-RXR as positive controls. Error bars stem from technical replicates.

### 3.11.2. ChIP: Overexpression of tagged proteins

To circumvent problems with non-ChIP-grade antibodies, we used antibodies binding epitope tags. These tags are well defined and characterised, and excessively tested antibodies are available. The drawback of this strategy is that the tagged target proteins have to be expressed from vectors provided by the experimenter and thus expression level and splice variants might be abnormal. Furthermore, the protein conformation may be influenced by the tag.

A 2x Ty1-tag was selected for ChIP as a ChIP-grade, monoclonal antibody was available. The tag was inserted into a retroviral vector, upstream of the full *Mus musculus* CDS of CREB, SP1, SP2 and SP3, generating N-terminally tagged proteins. CREB and SP2 served as negative control ChIP targets. After infection and selection of cells, expression of target proteins was assayed by western blot (figure 41A). As the tag just added 22 AAs, the fusion proteins were expected close to their native size of 46/105/80/115 kDa for CREB/SP1/SP2/SP3, respectively. For all four cell lines, Ty1-reactive proteins were detected at these sizes. For the SP1 cell line, a second, 45 kDa Ty1-reactive signal was found, possibly stemming from a shorter isoform. Next, nuclear extracts were generated and used for EMSA incubation with the IVS1+1505G probe (figure 41B). Addition of Ty1-antibody to the four different nuclear extracts in EMSA led to formation of supershifts for the SP1 and SP3 expressing cell lines, but not for the CREB/SP2 cell lines. Lastly, ChIP was performed. Histone H3 was used as a positive control IP as described in 3.11.1. Data was normalised to the respective input samples. Histone H3 IP was carried out with a mixture of all four chromatin samples. Figure 41C and 41D show element enrichments for all five IPs, once normalised to input and once standardised to SP2/CREB. Four primer pairs were used: DHFR promoter, UCP3 enhancer PP#1, UCP2 last exon and UCP3 last exon. For all four regions, Histone H3 IP yielded enrichment above all Ty1 IPs. Histone H3 yields lowest enrichment for the DHFR promoter, an observation also made in the two experiments before (figure 40). Of all Ty1 IPs, only the combination Ty1-SP1 IP/DHFR promoter qPCR exceeded enrichment of the negative control regions. None of the other tagged transcription factors was able to enrich the DHFR promoter or the UCP3 intron element compared to the two negative control regions. Comparing the different tagged proteins, Ty1-SP1 precipitated more of each element than the other three tagged factors.

In summary, none of the epitope-tagged transcription factors allowed reliable enrichment of the UCP3 intron 1 enhancer. As enrichment of the positive control region is poor as well, we conclude that methodical problems still need to be solved before meaningful results can be obtained. These problem were also encountered in 3.11.1 and were not solved by the use of epitope tags. Nevertheless, we still think that using tagged proteins for ChIP is a worthwhile strategy. Our expression system works and all four constructs produce proteins with the expected size. Furthermore, at least in case of SP1 and SP3, these proteins for complexes with their binding site and these complexes can be supershifted using an antibody raised against the tag.



**Figure 41: The Ty1-tagged overexpressed transcription factors bind the IVS1+1505G probe and can be bound by a Ty1-antibody.** CREB, SP1, SP2 and SP3 were cloned into viral expression vectors and tagged at their N-terminus with a 2xTy1 Tag. **A)** Western blot of 75000 stably-transfected cells per lane using anti-Ty1 as first antibody. Tagged proteins are expected at 46/105/80/115 for CREB/SP1/SP2/SP3, respectively (arrows). **B)** EMSA: Incubation of 5 µg nuclear extract and 40 fmol IVS1+1505G probe in presence or absence of 0.25 µg Ty1-AB. Four different nuclear extracts from four different cell lines were used, each expressing the construct stated above the image. Arrows denote specific signals. **C)** and **D)** Quantification of four regions in ChIP eluates, once normalised to input, once normalised to standardised to CREB/SP2. The Histone H3 IP was done with a mix of chromatin from all cell lines. The other IPs were done with the cell lines stated at the X-axis and αTy1. DHFR promoter served as positive control, UCP2/UCP3 last exon as negative control and UCP3 enh PP#1 is the region of interest.



### 3.12. Outlook: Enhancer complex purification

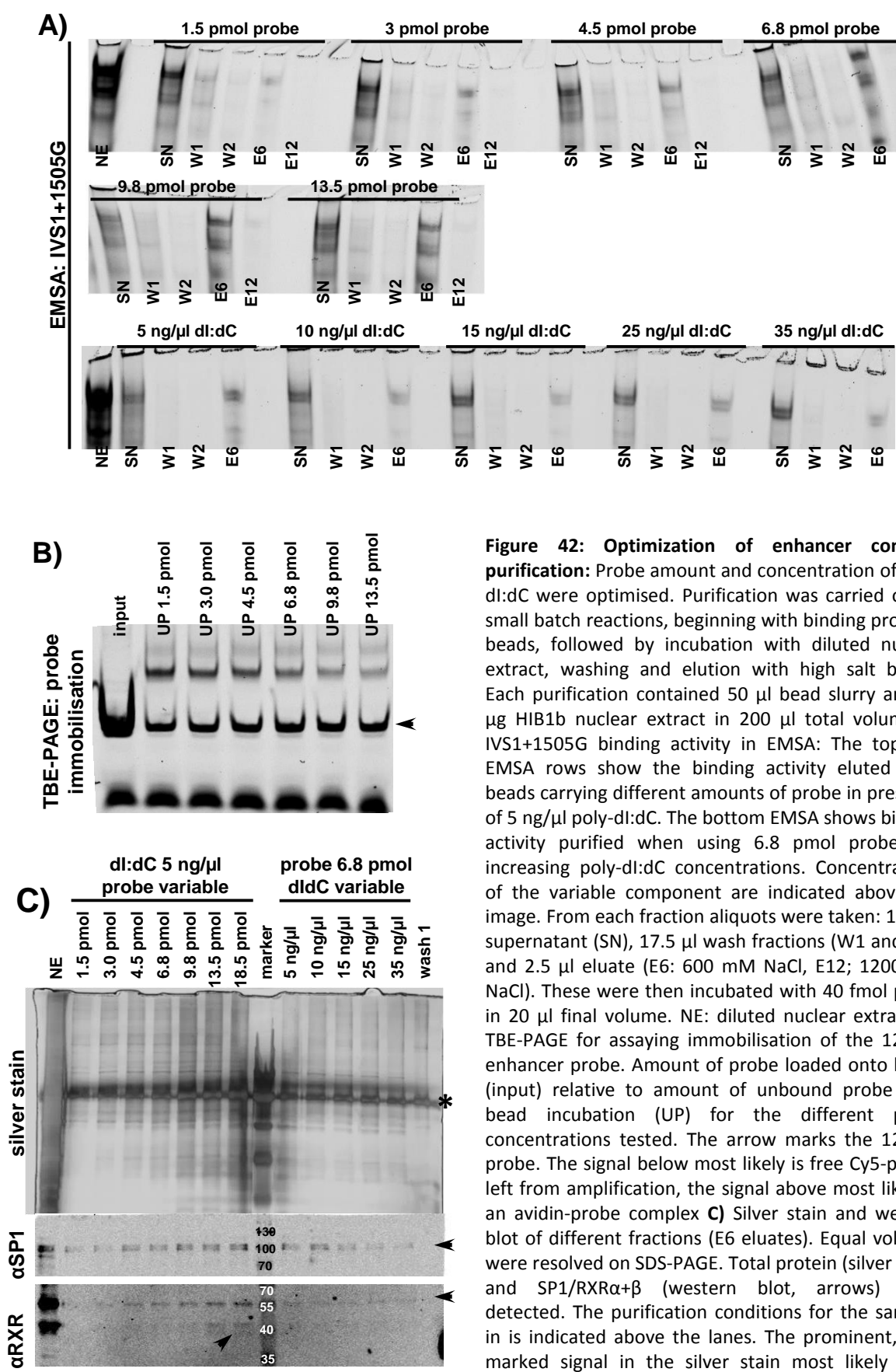
Our current hypothesis is that complex composition at the intronic enhancer makes the difference in regulation of UCP3 expression between BAT and SKTM. Therefore purification of the whole complex, rather than individual TFBS complexes, is the most promising strategy to identify relevant proteins. An approach recently developed is called promoter-trapping [138]. This method relies on PCR-amplifying a module or regulatory element and subsequently using it as a probe for affinity chromatography. Labelling one primer with an anchor suitable for purification, and the other primer with a fluorophore, allows both capturing and detection of the PCR-generated probe. We decided to use biotin and Cy5 as labels. Purification was carried out in small-scale batch reactions using NeutrAvidin agarose.

Primers flanking the NF1/STAT element, the DR1 element and the GC-Box of the *Phodopus sungorus* UCP3 gene were generated. The reporter gene constructs served as template to generate different probes of 120 bp. Using our mutated vectors, different probes carrying different mutations can be generated. In a TBE-PAGE the probe yielded one distinct Cy5-positive signal in absence of nuclear extract, but a non-interpretable smear in presence of nuclear extract (not shown). Optimisation was thus carried out in small batch purifications. For each optimisation step, aliquots of 50  $\mu$ l bead slurry were preloaded with the indicated amount of probe. After washing and equilibration, nuclear extract diluted in 1x EMSA binding was added to the beads. In addition, the buffer was supplemented with poly-dl:dC, igepal CA630 and PPAR $\alpha$ /- $\gamma$ /- $\delta$ /RXR $\alpha$ /TR-agonists. The beads were eluted with high concentrations of NaCl.

We used EMSA for IVS1+1505G binding activity and western blot for SP1 (two splice variants: 95/105 kDa) and RXR $\alpha$ / $\beta$ / $\gamma$  (large variants 50-55 kDa, short variants 38-45 kDa) to track the optimisation and purification process. Binding for these factors was well established by our own work and colleagues [73]. We therefore assumed that conditions favouring binding of these factors would also favour binding of yet unknown proteins. Silver stain was used to estimate total protein amount and complexity. The Cy5 label was used to monitor immobilisation of the 120 bp enhancer probe (figure 42B).

Figure 42 shows the readouts for two optimisation steps in which different amounts of probe and poly-dl:dC were tested. For each purification supernatant, wash fraction and two different eluates, using different salt concentrations, were collected. 17,5  $\mu$ l supernatant or wash fraction, or 2,5  $\mu$ l eluate were diluted with fresh buffer and 31 bp IVS1+1505G probe to assay for binding activity. Figure 43A shows that IVS1+1505G binding activity was mainly found in the supernatant and the first eluate (E6, 600 mM NaCl). Increasing the probe amount loaded onto the beads led to reduction of binding activity in the supernatant and an increase in the eluate, but also increased unspecific smear in the eluate. Increasing the poly-dl:dC concentration present in the binding mix had the inverse effect. Next, equal volumes of all first (E6, 600 mM NaCl) eluates were loaded onto three SDS-PAGEs. One PAGE was used for silver stain, two were used for western blots (figure 43C). With increasing probe amount, also an increase in total protein concentration in the eluates was observed. Again, increasing the poly-dl:dC concentration present in the binding mix had the inverse effect. Western blot for SP1 demonstrates that the SP1 content in the eluates reflects EMSA activity. Western blot for pan-RXR shows the same trend as SP1, although the recovery (eluate compared to input) is very poor.

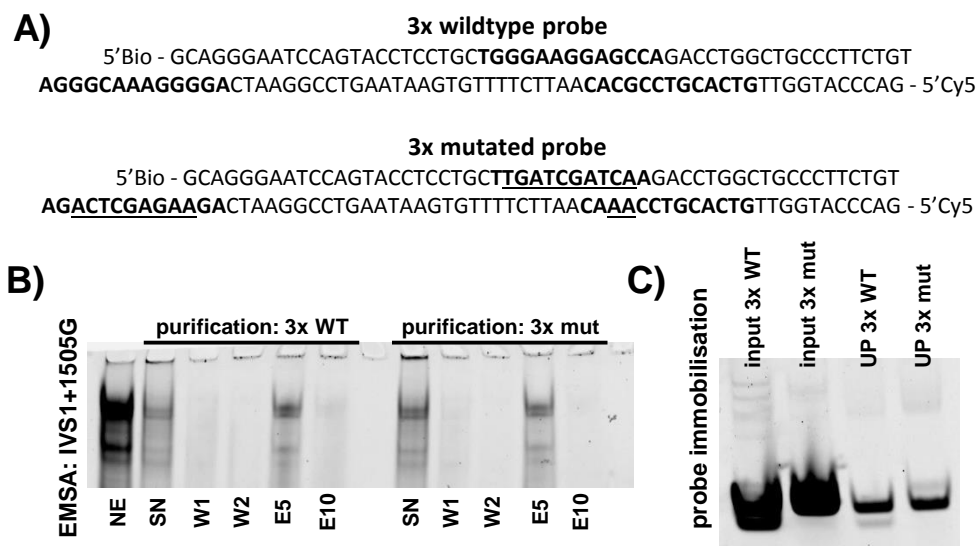
This experiment demonstrates that several components of the enhancer bound protein complex indeed can be purified in parallel using a probe covering all three elements.



**Figure 42: Optimization of enhancer complex purification:** Probe amount and concentration of poly-dl:dC were optimised. Purification was carried out in small batch reactions, beginning with binding probe to beads, followed by incubation with diluted nuclear extract, washing and elution with high salt buffer. Each purification contained 50 μl bead slurry and 75 μg HIB1b nuclear extract in 200 μl total volume **A)** IVS1+1505G binding activity in EMSA: The top two EMSA rows show the binding activity eluted from beads carrying different amounts of probe in presence of 5 ng/μl poly-dl:dC. The bottom EMSA shows binding activity purified when using 6.8 pmol probe and increasing poly-dl:dC concentrations. Concentrations of the variable component are indicated above the image. From each fraction aliquots were taken: 17.5 μl supernatant (SN), 17.5 μl wash fractions (W1 and W2) and 2.5 μl eluate (E6: 600 mM NaCl, E12: 1200 mM NaCl). These were then incubated with 40 fmol probe in 20 μl final volume. NE: diluted nuclear extract. **B)** TBE-PAGE for assaying immobilisation of the 120 bp enhancer probe. Amount of probe loaded onto beads (input) relative to amount of unbound probe after bead incubation (UP) for the different probe concentrations tested. The arrow marks the 120 bp probe. The signal below most likely is free Cy5-primer left from amplification, the signal above most likely is an avidin-probe complex **C)** Silver stain and western blot of different fractions (E6 eluates). Equal volumes were resolved on SDS-PAGE. Total protein (silver stain) and SP1/RXRα+β (western blot, arrows) were detected. The purification conditions for the samples in is indicated above the lanes. The prominent, star-marked signal in the silver stain most likely is an artefact.

In all affinity chromatography experiments some proteins get purified due to binding to the specific ligand, while some proteins are purified via unspecific interactions. We thus aimed to differentiate between specific and unspecific interactions by running a parallel negative control purification. To do so, a probe was generated that carried deleterious mutations in all three elements (figure 43A). Purifications using the optimised poly-dI:dC and probe amounts (10 ng/μl, 6.8 pmol, 75 μg nuclear extract, 200 μl total reaction volume) were carried out with both probes. Immobilisation of the different probes was efficient and comparable (figure 43C). Surprisingly, IVS1+1505G binding activity in the eluates was comparable between both purifications (figure 43B). Sequencing of probe and template vector revealed that all three TFBS present in the mutated probe indeed contained the expected mutations.

While the initial results (figure 42) suggest that enhancer complex purification is a powerful tool to purify enhancer bound protein complexes, the latter experiment raises questions about specificity of the purification. If this strategy shall be pursued further, binding conditions and unspecific interactions will need to be carefully optimised and controlled for.



**Figure 43: Lack of differential binding in enhancer trapping:** 50 μl bead slurry was loaded with either 6.8 pmol 3x WT probe or 3x mutated (3x mut) probe and incubated with 75 μg HIB1b nuclear extract in presence of 10 ng/μl poly-dI:dC in a final volume of 200 μl. Afterwards, beads were washed twice (W1+W2) and eluted with 500 and 1000 mM NaCl (E5+E10) **A)** Sequence of both probes. Mutations in the 3x mut probe was verified by sequencing of probe and template vector. Mutations are underlined, TFBS are bold. **B)** IVS1+1505G binding activity in EMSA. The probe used for purification is stated above the image. Supernatant (SN, 17.5 μl), wash fractions (W1 and W2, 17.5 μl) and eluates (E5: 500 mM NaCl, E10; 1000 mM NaCl, 2.5 μl) were incubated with 40 fmol probe in 20 μl final volume NE: diluted nuclear extract. **C)** Assay of Cy5-labelled probe concentration before (input) and after (UP) probe-bead incubation.

### 3.13. Summary of results

This thesis does answer many questions about the transcriptional regulation of the UCP3 gene, but also brings up new questions. There still is much to understand and uncover, especially regarding commonalities and differences between BAT and SKTM, but several important findings have been made. These findings can be condensed into six central observations:

- The IVS1+1505 element is important in both BAT and SKTM cell culture lines. Complexes are formed with a probe resembling the element in all cell lines (3.1.3). A mutation in the element abolishes PPAR $\gamma$  agonist induced activity of a UCP3 reporter gene construct (3.1.5).
- The IVS1+1505 element contains a GC-Box (3.4.1). SP1 and SP3 bind this GC-Box in EMSA. This was validated by cold competition and supershift experiments (3.4.1). IVS1+150G binding activity furthermore correlates with SP1/SP3 content in heparin affinity chromatography fractions (3.4.4).
- Reduction of SP1 or SP3 abundance in cell culture reduces activity of the IVS1+1505G reporter gene construct (3.4.5.1) in presence of PPAR agonists. Chemical inhibition of GC-Box binding using mithramycin mimics this effect in both reporter gene assays and endogenous UCP3 expression (3.5.2) in BAT. Chemical inhibition of SP binding by mithramycin also affects expression of endogenous UCP3 in a SKTM (3.10.3) cell culture line.
- PPAR $\gamma$  agonist action on UCP3 expression is dependent on presence of the IVS1+1505G element and a PPAR $\gamma$ -binding DR1 element located in juxtaposition. This was demonstrated using mutated reporter gene constructs (3.5.1) and chemical binding inhibition (3.5.2).
- The first intron of the UCP3 gene contains a third binding site which is located directly adjacent to the DR1 and the GC-Box (3.7). This site might bind members of the NF1 or STAT transcription factor families. Furthermore, MyoD and myogenin bind close to the DR1/GC-Box module in C2C12 cells, possibly via a EBox about 55 bp upstream of the putative NF1/STAT element (3.7.2).
- The intronic enhancer is perfectly conserved in the UCP3 genes of *Phodopus sungorus*, *Mus Musculus* and *Rattus norvegicus*. In *Homo sapiens*, *Sus scrofa* and *Equus caballus* we were able to identify regions of comparable TFBS compositio (3.7.2). While these regions still await validation, we are confident that an intronic enhancer is a feature in non-rodent UCP3 genes as well.
- We gathered first evidence for presence of an enhancer region 2000 bp of the UCP3 TSS. While functional data still is scarce the region binds Nrf2 [47] and PPAR $\gamma$  (3.9), and is located within a DNase I hypersensitive site (3.8).

Taken together, we uncovered a complex enhancer module where we expected a simple, single BAT-specific TFBS. We managed to gather plenty of evidence about the composition of enhancer and the basic interplay of the central components. We undertook several steps to further investigate this complex module. Last but not least we gathered initial data on a putative upstream enhancer region.

## 4. Discussion

This thesis originated from the identification of a subpopulation of the djungarian short tailed dwarf hamster (*Phodopus sungorus*) that lacks uncoupling protein (UCP) 3 expression in brown adipose tissue (BAT), but not in skeletal muscle (SKTM) [128]. A base exchange (G→A) at position intervening sequence (IVS)-1+1505 is causal for the tissue specific UCP3 deficiency, most likely by reducing affinity of a transcription factor to an intronic binding element [59]. We expected this factor to be a BAT specific regulator that is not required for UCP3 expression in SKTM. Our main goal was to identify the binding factor and understand its role in regulation of UCP3 expression.

### 4.1. Baseline characterisation of the cell lines used

One important goal was to extend our research on UCP3 regulation on species other than *Phodopus sungorus*. We wanted to uncover whether intronic regulation of UCP3 transcription is just a speciality in hamster or a more general phenomenon. In addition, this step allowed usage of additional model- and assay-systems. As a first step, *Mus musculus* was chosen. Four mouse-derived cell lines were tested as putative model systems: Three BAT-derived cell lines, HIB1b (hibernoma cell line 1b) and iBPA-L1/L2 (immortalised brown preadipocytes), and one SKTM-derived cell line, C2C12. We started by conducting a baseline characterisation to elucidate whether the four cell lines are suitable for research on regulation of UCP3 expression. Four parameters were investigated: Morphological development of the cells, expression of transcription factors required for UCP3 expression, reporter gene activity of *Phodopus sungorus* UCP3 reporter gene constructs and electrophoretic mobility shift assay (EMSA) complex formation.

#### 4.1.1. All four cell lines exhibit basic features required for research on UCP3 expression

All four cell lines exhibited the cell type specific differentiation features (3.1.1) and transcription factor expression patterns (3.1.4) expected for the respective cell type. Transcript abundance for several transcriptional regulators required for UCP3 expression was measured and all were present in both undifferentiated and differentiated cells. Furthermore, the IVS1+1505G reporter gene construct was active in all four cell lines and reporter gene activity increased upon treatment with peroxisome proliferator activated receptor (PPAR) agonists (3.1.5). Lastly, the IVS1+1505G binding complex was present in all cell lines (3.1.2, 3.1.3).

Some features did not match the expectations. UCP3 mRNA was absent in both undifferentiated and differentiated HIB1b cells. One possible explanation is the absence of a required signal molecule. For example, low amount of fat storage inside the cells might mean low intracellular lipid metabolism. Lipid metabolism is a source of PPAR agonists, which are important regulators of UCP3 transcription (1.4.3.1). Another possible explanation is the absence of a regulatory protein required for UCP3 expression. This could be either a transcription factor we did not measure, or a factor whose mRNA does not translate into protein expression. Possible candidates are PPAR $\alpha/\delta$ , thyroid receptor (TR) and retinoid acid receptor (RAR), based on the finding that neither of their agonists (Wy14643/GW0742, triiodothyronine (T3) and all-trans retinoic acid (ATRA), respectively) had effects on reporter gene activity in HIB1b cells (3.1.6). Whether this absence of effect was due to absence of the respective receptor protein or due to lack of a required agonist is unclear. T3 and ATRA, at least in C2C12 cells, were only effective when acting in concert, preferably with PPAR agonists (3.10.2). An unexpected finding was the stimulation of reporter gene activity by indomethacin. This might either be explained by binding and activation of PPAR $\gamma$  at micromolar concentrations [143], or by induction of expression

of PPAR $\gamma$ 2 and CCAAT-enhancer-binding protein (C/EBP)  $\beta$  by a yet unknown mechanism [171]. The latter hypothesis seems less likely, as in the study reporting that mechanism expression PPAR $\gamma$ 2 and C/EBP $\beta$  were increased after two days, while in our experiments effects were observed already after 24 hours.

While some questions about expression on the protein level and agonist induction of UCP3 transcription remain, all four cell lines are suitable for research on the IVS1+1505G element. The main arguments are reporter gene activity, PPAR $\gamma$  agonist responsiveness, expression of key transcription factors and presence of the complex forming proteins.

#### **4.1.2. Initial experiments question tissue specificity of the protein binding the IVS1+1505 element**

The baseline characterisation of the four cell lines demonstrated that all four were suitable for research on UCP3 expression. During this characterisation, some unexpected observations were made that contradict the hypothesis that the IVS1+1505G element was of limited importance in skeletal muscle.

The proteins forming the IVS1+1505G specific EMSA complexes were present in C2C12 cells (3.1.3). While the complex pattern was comparable in all four cell lines, abundance of complexes was not. HIB1b nuclear extract contained the highest amount of complex forming proteins while abundance was lowest in C2C12 extracts. The difference in complex formation is large, but might stem partially from a different purity of the nuclear extracts, mostly due to the different nuclei/cytoplasm ration (HIB1 cells are very small, C2C12 syncytia very large and iBPA cells intermediate). The fact that the complex binding proteins were present in C2C12 cells is in conflict with the initial finding that the complex is absent in SKTM cell lines [59]. The most likely explanation is the refinement of nuclear extract purification during the thesis. This allowed generation from a wider array of cells without the use of potter and pistil, a component which needed to be carefully optimised for every cell line before. Two conclusions can be drawn: Firstly, the complex forming proteins are present in all four cell lines. The binding transcription factors are thus most likely expressed in both BAT and SKTM. Secondly, in SKTM cell lines, and possibly in SKTM tissue, the binding factor is less abundant than in BAT. A very high amount of binding factor would have been a possible explanation why the IVS1+1505A allele has a low effect in SKTM, because reduced binding affinity can be overcome by increasing the abundance of either of the two binding partners (in this case: the transcription factor).

The IVS1+1505A allele abolished responsiveness of the *Phodopus sungorus* reporter gene construct to PPAR agonists in C2C12 cells (3.1.5). This was in contrast to in vivo data from *Phodopus sungorus*. In vivo, the IVS1+1505A allele only mildly reduced UCP3 expression in SKTM. Interestingly, under non-stimulated conditions, no difference in reporter gene activity was observed between the two alleles in C2C12 cells. Two possible explanations arise: Firstly, under baseline conditions indeed both reporter vectors were equally active, which would reflect the initial *Phodopus* finding. Existence of such a non-stimulated condition vivo is doubtful, though. Secondly, in C2C12 cells both reporter vectors were silent in absence of exogenous activators. This would not have been the case in BAT cell lines, either due to intrinsic activity of transcription under basal conditions or due to endogenous ligands always being present. Endogenous PPAR ligands would be prime candidates as they arise from lipid metabolism, which most likely is absent or at least drastically lower in cultured muscle cells compared to fat cells. This is partly conflicting with the low amount of fat storage in HIB1b cells, where the difference between the alleles is largest. Considering the fact that the non-stimulated activity of the IVS1+1505G reporter in C2C12 was comparable to the activity of the inactive IVS1+1505A constructs in all four cell lines, the latter hypothesis seems more likely. Interestingly, reporter gene activity

correlates with complex abundance in EMSA, but whether this is coincidence or causal cannot be decided based on the data available. All comparisons of normalised reporter activities have to be regarded with caution, as the CMV promoter driving the normalisation construct is known exhibit distinct activity in different cell lines [172].

In summary, two findings hint that the IVS1+1505G element and the binding proteins may be relevant for regulation of UCP3 expression in SKTM: Presence of the IVS1+1505G specific EMSA complex and the deleterious effect of the IVS1+1505A allele on reporter gene activity in C2C12 cells. So far we are not able to explain why the cell culture experiments did not reproduce the in vivo findings, but the inconsistencies most likely arise from either of two differences: The experiments do not correctly mimic the in vivo regulation inside the nuclear genome, or the C2C12 cell line does not correctly mimic the situation in *Phodopus sungorus* SKTM. It is important to note that absence of an effect of the IVS1+1505 polymorphism on UCP3 expression in SKTM does not necessarily mean that the element is without regulatory function in SKTM. A more throughout discussion can be found in 4.5.2. Since knowledge about the binding factor will greatly aid understanding, we next focussed on identification of the proteins.

## 4.2. Candidate identification by affinity chromatography

A classic method to identify transcription factors binding to an element of interest is purification by DNA affinity chromatography and subsequent identification by mass spectrometry. While this method is cumbersome and time consuming, it allows unbiased de novo identification of binding factors. The main problems hindering identification are contaminants and non-specifically purified proteins. We attempted to exclude contaminants by comparing purifications with different probe alleles (wildtype versus mutated probe) and by employing affinity strategies which differed in resin, immobilisation chemistry and immobilised anchor.

In our purifications, only few of the identified peptides fit to transcription factors. We could identify several obvious contaminants and thereby omit their validation (see figure 18), but still several possible candidates remained. The tasks were to reduce contaminants in affinity eluates and to tell apart promising candidates from likely false positives. 28 identifications were of proteins directly or indirectly involved in regulation of transcription: seven proteins were part of signalling cascades influencing transcription, nine had functions in chromatin modification, nine either were co-regulators or at least had been proposed to have co-regulator functions, and three were classic transcription factors (figure 18, table 1).

To sort the 28 identified proteins into likely and unlikely candidates, we rated every candidate according to a simple checklist containing the following criteria:

- 1.) The candidate was primarily identified in purifications using wildtype probe.
- 2.) The candidate was identified with both of the different affinity chromatography strategies.
- 3.) The candidate was reproducibility identified, meaning in more than one experiment and with a reasonable number of fragments.
- 4.) The candidate protein is not known to interact with or bind to known obvious contaminants.
- 5.) The literature does not directly contradict a role of the candidate.

In summary, none of the factors proved to be a promising candidate (see 3.2.5). Most were only identified with very few peptides, only with one of the two affinity strategies and independent of the probe allele used in affinity chromatography. Furthermore many of the candidates possibly were purified due to their participation in mRNP complexes, known common contaminants. Initially, several proteins seemed promising due to their role in SKTM and BAT determination, or due to interaction with proteins relevant for regulation of UCP3. After a more detailed analysis most proved to be either anti-adipogenic or linked to UCP3 transcription via very general connections. At that point SP1 and SP3 emanated from our other candidate finding approaches (see 4.3) and we stopped following this approach.

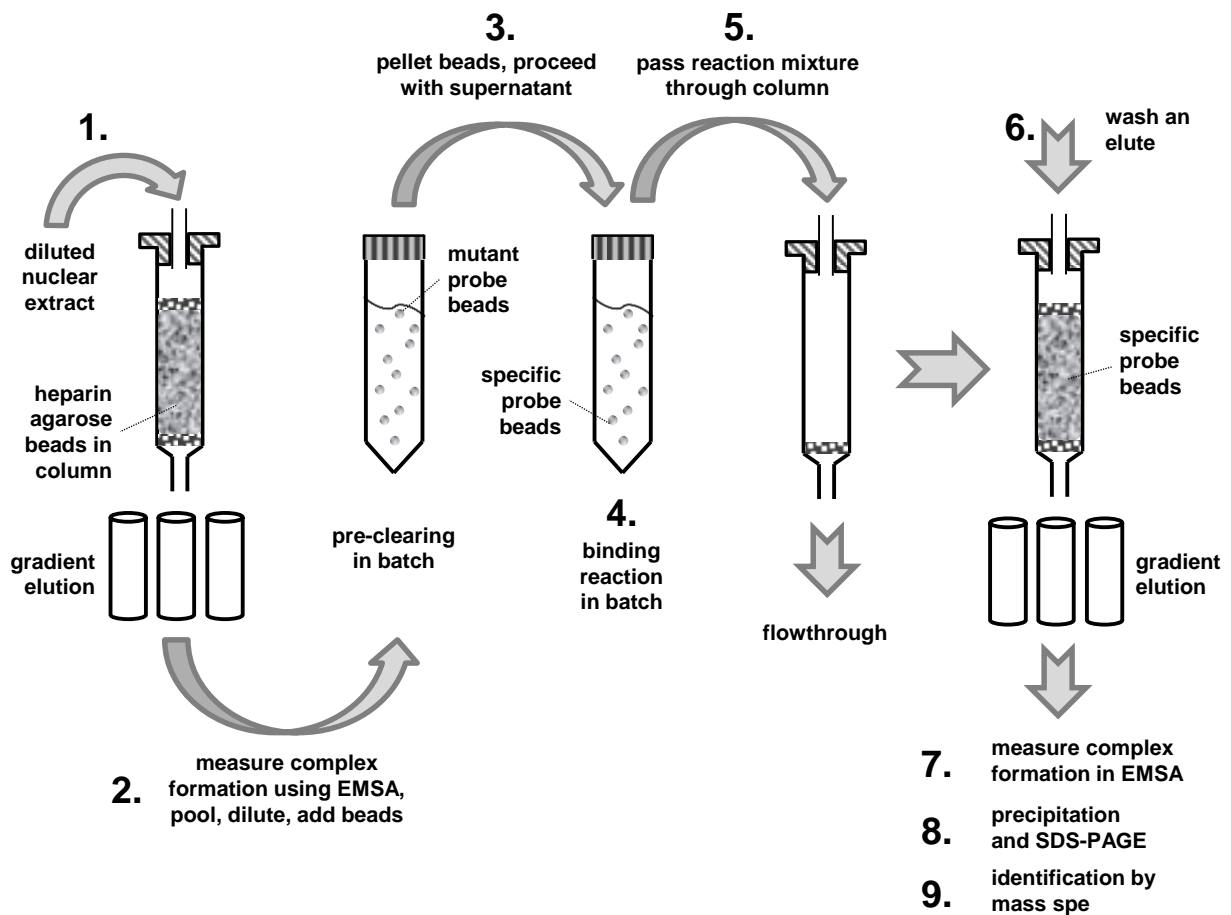
#### **4.2.1. Purification of DNA-binding proteins: An optimised strategy**

Despite not identifying promising candidates, a lot of experience was gained with the general methodology of transcription factor purification. Understanding of the weaknesses and strengths of the different methods allows us to propose an advanced, multi-step workflow for purification of DNA binding transcription factors. While every successful purification step increases purity and reduces complexity of the sample, it also leads to a loss of the protein of interest and has the chance to introduce contaminants. It was thus important to combine steps in a way to maximise yield and purity while keeping contaminants as low as possible.

Figure 44 shows a workflow that allows transcription factor purification in one day. All purification steps use agarose beads as their larger operating volume leads to lower local concentration of immobilised DNA, thereby reducing unspecific binding. Initially, nuclear extract is diluted and pre-purified by heparin affinity chromatography. When carried out in a column setup (figure 6) this step allows processing of large nuclear extract amounts and achieves an increased concentration of the binding activity and a reduction of non-DNA-binding contaminations (see 3.2.6, figure 19 and 23). Eluates are assayed for complex formation using EMSA. Fractions containing the binding activity of interest then are used for an oligonucleotide trapping purification. Incubation is carried out in batch, meaning that the resin is free-floating within the incubation mixture. This reduces loss of activity and allows easier scaling. After an incubation period, the beads are collected in a column to allow efficient washing and easier gradient elution. This step will reduce all contamination besides unspecific DNA-binding contaminations. Fractions containing EMSA-activity can be precipitated and separated by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) to remove salts and other low-mass contaminations. If unspecific DNA-binding contaminants prove to be an issue, such contaminants may be reduced by a pre-clearing step after heparin affinity chromatography. Incubation of the diluted eluates with agarose beads in absence of probe or in presence of mutated probe will deplete proteins binding non-specifically, but not proteins with specific binding activity. The pre-cleared supernatant then can be used for the actual sequence-specific oligonucleotide trapping step.

This strategy omits the preparative EMSA (3.2.3) which we consider a source of contaminants and which also yields only limited recovery. In contrast, it employs two agarose-based chromatography steps which in our hands combined good protein recovery with efficient reduction of complexity and a low chance to introduce contaminations. If gradient mixers and fraction collectors are available, the method might even allow to fractionate different binding proteins.





**Figure 44: Proposed optimized strategy for affinity purification of transcription factors:** For future transcription factor purifications the following strategy should be applied, incorporating experience gained during previous experiments. **1.)** 1-3 ml nuclear extract is diluted and pre-purified by heparin column affinity chromatography. **2.)** All eluate fractions are screened for presence of the binding proteins using EMSA. **3.)** EMSA-active eluates are pooled, diluted and incubated with the resin used for oligonucleotide trapping, either in absence of probe or in presence of mutant probe. The supernatant is removed and transferred to a new tube. **4.)** Specific probe and fresh beads are added to capture specifically binding proteins. **5.)** The beads, loaded with probe-transcription factor complex, are collected in a column. **6.)** The resin is washed and proteins are eluted by a salt gradient. **7-9)** Eluates are assayed for binding activity. Active fractions are pooled and precipitated. Low mass contaminants are removed by SDS-PAGE and proteins are analyzed by mass spectrometry.

### 4.3. Bioinformatics and sequence analysis – identification of the SP transcription factors

Bioinformatics provide a valuable tool to predict transcription factor binding. On the downside, it is very prone to making false positive predictions. This well-known problem was also encountered in our analyses. False-positives originate from pure statistics: If a region is searched for sufficiently many different transcription factor binding sites, and if the algorithm allows for some sequence variance, then there will always be a certain number of matches. Even the most recent algorithms still struggle to tell apart meaningful from random matches due to the failure of many tools to integrate the local sequence information (consensus binding sequences) with surrounding sequences and the local chromatin conformation.

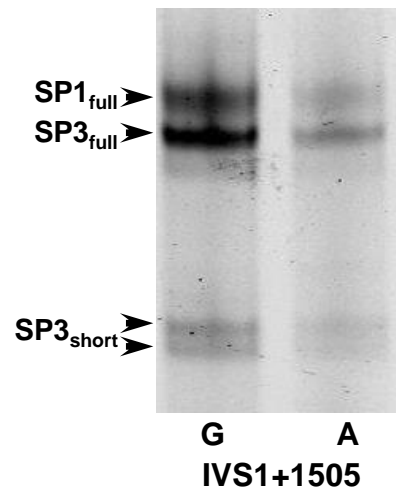
The solution to this problem is to include further information to allow generation of more meaningful hits. For example, this can be accomplished by fine-mapping of the relevant binding site. To do so, mutations are introduced and assayed for their influence on binding activity. The advantage of this method is that it does not depend on knowledge of homologous regions in other species or related sequences within other genes. On the downside, the method is work intense as several mutated probes need to be generated and assayed to reliably pinpoint the binding sequence of interest. Another strategy is based on identification of homologous regions. These regions are then compared to identify conserved sequence blocks and putative multi-TFBS modules, possibly spanning up to few hundreds of basepairs (bp). A recent publication proposes a new standardised method, phylogenetic module complexity analysis (PMCA), that uses module complexity and conservation analysis instead of simple consensus motif search. This algorithm thus provides a promising advancement and may help to discriminate false positive binding predictions from relevant motifs [173].

In this project, we searched for conserved TFBS and also attempted to fine-map the binding site of interest. Initially, we searched for homologous regions to find the relevant conserved TFBS. Putative homologous regions were identified and several TFBS were predicted, but none of the candidates proved to be involved in regulation of UCP3 transcription via the IVS1+1505G element (3.8.1). Retrospectively, this was due to not being able to identify the correct homologous regions in non-rodent species. Instead, incorrect regions were used for conservancy analysis, thereby leading to false TFBS predictions. Nevertheless, sequence alignment and comparison of all competitors used during these experiments (figure 21) subsequently allowed to pinpoint the relevant TFBS and thus ultimately led to identification of the factors binding to the IVS1+1505G element. Grouping all probes and competitors for their ability to bind the same proteins as the IVS1+1505 probe revealed that the sequence CACGCC was crucial for formation of the protein DNA complexes (figure 21). Even minute deviations reduced binding activity in EMSA (figure 20 and 22). In this sequence, the central G corresponds to position IVS1+1505. This sequence closely resembles a GC/GT-Box, an element known to be bound by members of the SP/KLF family of transcription factors.

#### 4.3.1. The SP/KLF family of transcription factors

The SP family of transcription factors contains four transcription factors: SP1, SP2, SP3 and SP4. SP stands both for Specificity Protein and for Sephacryl/Phosphocellulose, the column materials used for the initial purification. The mouse full length proteins consist of 781, 613, 783 and 782 AA, respectively, with shorter splice forms being known or predicted (source: [www.ensembl.org](http://www.ensembl.org)). SP1, SP3 and SP4 contain a comparable set of subdomains and the domain order only shows minor differences. SP2, being shorter, lacks some of the domains common to the other SPs. SP1 and SP3 are expressed ubiquitously, SP2 is expressed in a wide array of tissues, and SP4 is primarily found in brain. All four share the DNA-binding domain, consisting of three Cys<sub>2</sub>His<sub>2</sub> zinc fingers. SP1, SP3 and SP4 bind essentially the same consensus sequence (see 4.3.1), while SP2 exhibits a different binding site preference due to a His to Leu AA variation in the first zinc finger (reviewed in [174]). Because of their ubiquitous expression and their shared binding motif, SP1 and SP3 commonly bind the same promoters and enhancers, although with partly different functions. While both bind the same sequence, only one of both is bound at a single binding a given time. Both proteins are involved in a wide array of different functions, and both are regulated via posttranslational modifications. While SP1 generally acts as an activator of transcription, SP3 can act both as an activator or an inhibitor. Inhibitory activity is induced by SUMOylation and achieved by facilitation of heterochromatin formation ([175], reviewed in [176]).

Closely related to the SP transcription factors are the Krüppel-like factors (KLFs), which also contain three Cys<sub>2</sub>His<sub>2</sub> zinc fingers. They are named after their resemblance to the *Drosophila melanogaster* transcription factor Krüppel, another protein with three Cys<sub>2</sub>His<sub>2</sub> zinc fingers. KLF9, KLF10 and KLF11 share the binding preferences of SP1 and SP3 (reviewed in [177]). In other KLFs, variations within the fingers modify sequence preference. The consensus sequence is 5'-(G/T)GGGCGG(G/A)(G/A)(C/T)-3', but usually presence of either of two core elements is sufficient for binding: A GC-box (GCGGG or CGCCC) or a GT-box (GTGGG or CACCC). Due to their ubiquitous expression and abundance, SP1 and SP3 are the most likely factors binding to the respective elements, and also the most investigated ones. A recent ChIP-seq study for SP1 demonstrated that only 64% of all binding regions contained a GC- or GT-box, meaning that 36% of all binding sites are non-canonical [178]. Binding of SP1 and SP3 to intronic enhancers has been demonstrated before [179,180]. According to the literature, GC/GT-Box probes form four complexes in EMSA [174]: Two fast migrating and two slow migrating ones. The two slow-migrating complexes contain the full length SP1 and SP3 proteins while the fast migrating complexes contain two shorter isoforms of SP3 (figure 45).



**Figure 45: Complex formation pattern for SP1 and SP3:** Two lanes are shown, containing the two alleles of the IVS1+1505 probe and equal amounts of HIB1b nuclear extract. Arrows mark the specific complexes. The binding pattern matches the pattern described in literature where a consensus SP1/SP3 binding probe forms four complexes. The slow migration complexes are formed by the full length versions of SP1 and SP3. The fast migration complexes are formed by short splice forms of SP3.

#### **4.3.2. Validation of SP1/SP3 binding**

Several validation experiments were carried out to demonstrate that indeed SP1 and SP3 are binding to the IVS1+1505G probe in EMSA. Antibodies and miRNAs targeting SP1 and SP3 interfered with complex formation in EMSA in a pattern perfectly in line with literature (figures 23, 26 and 45) [174]. Both knockdown of SP1 and SP3 (3.4.5.1) and binding inhibition using mithramycin (3.5.2) demonstrated involvement of SP1 and SP3 in regulation of UCP3 expression. Lastly, in heparin affinity chromatography, IVS1+1505G binding activity correlated with the concentration of SP1 and SP3 (3.4.4).

ChIP is the most reliable method to demonstrate direct interaction of a transcription factor of interest with a binding site. The main concern about binding of SP1 and SP3 to the intronic GC-Box element *in vivo* is the failure of our ChIP experiments (see 3.9). Neither antibodies binding native SP1 and SP3, nor antibodies for precipitation of tagged SP1 and SP3 were able to enrich the enhancer region above a negative control region. In a worst case scenario this would mean that the interaction of SP1 and SP3 in EMSA is an artefact and that the effects of SP1 and SP3 knockdown on reporter gene activity are indirect. We do not believe this to be the correct interpretation, though. Our controls demonstrated that the ChIP results suffer from technical problems: Antibodies published to purify the same intronic region (PPAR $\gamma$ , pan-RXR [73]) failed to purify the IVS1+1505G element in our hands as well. Furthermore the enrichment we obtained for known SP1-binding positive control regions was low. We thus assume that general technical problems with the ChIP methodology, or with our cell culture system, hindered generation of meaningful results.

A question that remains is why we did not identify SP1 and SP3 in our oligonucleotide trapping eluates. We propose two plausible hypotheses. Both are based on the fact that the oligonucleotide trapping eluates contained binding activity (figure 15), showing a binding pattern matching to the pattern seen in our initial experiments and to the pattern described to be formed by SP1/SP3 in the literature. Firstly, SP1 and SP3 may have been lost during the preparative EMSA, EMSA elution or the subsequent precipitation step. Secondly, SP1 and SP3 were eluted from EMSA, but their identification was hindered by contaminants present. Indeed, the silver stained gels in which the EMSA-eluates were resolved (3.2.4, figure 17) showed bands at 130, 105, 95 and around 70 kDa, which would fit to the size of SP1 and SP3 isoforms seen in western blot (figure 26C) and described in [174]. On the other hand, 105 and 70 kDa perfectly fits to the mass spectrometry identified proteins SND1 (also known as p100) and HSC71 (see 3.2.5). No matter why SP1 and SP3 were not identified in our protein purification followed by mass spectrometry setup, we currently consider the preparative EMSA step to be problematic and thus recommend to omit it for upcoming purifications (see 4.2.1 for an alternative, advanced purification strategy)

In summary, several independent validation experiments prove that the proteins binding to the IVS1+1505G probe *in vitro* are SP1 and SP3. Furthermore we demonstrate that binding of SP1 and SP3 are required for expression of both endogenous UCP3 and the UCP3 mRNA. Lastly, the sequence of the IVS1+1505G element supports involvement of SP1 and SP3. We thus believe that SP1 and SP3 bind to the IVS1+1505G element and regulate transcription of UCP3.

#### **4.4. From a single binding site to a complex enhancer: Additional intronic binding elements**

Our initial hypothesis was that a single, BAT-specific binding site is present within the first intron of the UCP3 gene. Furthermore we expected that this element is related to PPAR signalling. Unexpectedly, the element has proven to be a GC-Box that binds the general transcription factors SP1 and SP3, both proteins with no specific link to PPAR signalling. Furthermore the PPAR binding element was believed to be a DR1 element located within the core promoter. Subsequently, a second DR1 element was identified to be located within the first intron of the mouse UCP3 gene [73]. This element is conserved in hamster and rat, binds PPAR $\gamma$  and RXR, and confers PPAR $\gamma$  agonist action on expression of UCP3. As the GC-Box and this intronic DR1 element are located only 39 or 46 bp apart (in hamster or mouse and rat, respectively), we propose that both elements are functionally interacting.

##### **4.4.1. Intronic DR1/GC-Box interdependence and the promoter DR1 element**

We applied two strategies to investigate the relationship between the DR1 and the GC-Box element: Targeted mutagenesis of the two DR1 elements in our reporter gene constructs (3.5.1) and inhibition by the GC-Box binding inhibitor mithramycin (3.5.2). Three main findings were obtained: Firstly, both intronic elements were indispensable for the effect of PPAR agonists. Secondly, deletion of the promoter DR1 element reduced, but did not ablate PPAR-agonist induction. Thirdly, mutation of either intronic element led to minimal reporter activity. No cumulative effect of multiple mutations was detected.

Both experiments support the same hypothesis: The intronic GC-Box and the intronic DR1 element are absolutely required for activity of the reporter gene construct under PPAR-agonist stimulated conditions. We are convinced that this finding translates to the nuclear UCP3 gene, at least in BAT. So far, little mechanistic insight on the cooperation of the two elements has been gained. Two interpretations are possible: One hypothesis is that binding of SP1 and SP3 to the GC-Box element is needed to mediate the effects of PPAR $\gamma$  and RXR $\alpha$ , which bind to the DR1-element. Alternatively, both the GC-Box and the DR1 element may be indispensable for expression of the UCP3 gene in completely independent, self-autonomous ways. None of the experiments carried out so far allows to differentiate between those two hypotheses. A first step to assay interaction may be to investigate whether distance and orientation of the GC-Box and the DR1-element are of importance. This could be attempted by insertion of a spacer in between, or rotation of either or both of the elements. A second step might be to validate formation of a supercomplex containing both binding elements and their respective binding proteins. This may be achieved by performing an EMSA with two distinct probes, each containing one of the elements of interest (F. Bolze, personal communication). Different fluorescent dyes would allow to identify co-migration of the probes while tagged probes (Biotin/Digoxigenin) would allow supershift/co-shift experiments. Initial experiments in this direction were carried out, but remained unsuccessful (not shown).

Despite the lack of direct experimental evidence, there is circumstantial evidence that favours the hypothesis that the SP-binding GC-Box and the PPAR $\gamma$ /RXR-binding DR1 element act together in one complex to execute a concerted function: The close proximity of the two elements and the fact that PPARs and SP transcription factors have been reported to interact before [181-184]. On a mechanistic level, SP1 and SP3 may be required to aid or recruit binding of PPAR $\gamma$ /RXR $\alpha$ . Apparently, binding of SP1/SP3 to the GC-Box does not require the DR1 element, as complex formation is seen with the 31bp IVS1+1505G probe alone (see 3.4.1). Conversely, we failed to demonstrate binding of PPAR $\gamma$ /RXR $\alpha$  to an EMSA probe only containing the intronic DR1 element (not shown), though there are other

explanations possible for this. Alternatively, only the DR1/GC-Box combined complex may be able to recruit all cofactors required to mediate transcriptional activation. For example, SP transcription factors can recruit the histone acetylase p300 and are also regulated by p300 [185-188]. Another hypothesis is that SP1 and SP3 mediate interaction of the intronic enhancer with the promoter, possibly by DNA-bending. This is in line with the fact that SP1 has been demonstrated to bend DNA [189] and can facilitate enhancer-promoter interaction [190]. At the current state of work we are unable to favour one of these hypotheses, and we are aware that they are not mutually exclusive.

#### **4.4.2. Identification of additional binding elements: A putative NF1/STAT element, and binding of MyoD, myogenin and p300**

Sequence alignments uncovered several short stretches of perfect sequence conservancy between mouse, rat and hamster in the vicinity of the DR1/GC-Box module. These may harbour additional regulatory elements (not shown). To identify such elements, we carried out a deletion screen covering the whole first intron of our UCP3 reporter gene constructs. Furthermore we screened ChIP-seq data deposited in the ENCODE project for transcription factors binding near the intronic DR1/GC-Box.

The deletion screen was carried out in HIB1b cells and identified two regions of interest (see 3.7). One of the regions was directly adjacent to the DR1/GC-Box, and could be narrowed down to about 36 bp (see figure 31D). Bioinformatics identified a Nuclear Factor 1 (NF1)- and a signal transducer and activator of transcription (STAT)-binding element. The ENCODE search revealed binding of MyoD and myogenin in C2C12 cells and binding of RNA-Polymerase II (Pol2) and p300 in heart. p300 is required for expression of UCP3 [79] and preferentially binds to the intronic enhancer module (3.7.1). We believe that it may be recruited by combined action of the DR1 and GC-Box binding factors. A similar mechanism is described for the recruitment of CBP, a protein closely related to p300, by KLF5 and PPAR $\delta$  [91]. MyoD and myogenin exclusively bind to the enhancer, while p300 and Pol2 show a bimodal pattern with one binding peak near the transcriptional start site (TSS) and another near the intronic enhancer (figure 32). Interestingly, binding of MyoD and myogenin is not seen for all differentiation states of C2C12 cells. Binding is absent in proliferating cells, peaks mid differentiation and is low in fully differentiated cells. These findings raise three questions.

The first question is which protein or proteins do bind to the NF1/STAT site. A possible, straightforward answer is: A protein of the NF1 transcription factors family and/or a STAT-factor. While it is tempting to propose this, further validation experiments are required. Proteins of the NF1 family of transcription factors, also known as CTF (CAAT box transcription factor), bind the consensus sequence TGGNNNNNN(N)CCA [191]. The family has four members in vertebrates (reviewed in [192]), which bind as hetero- or homodimers [193]. All four members exist in different splice forms and differ in their spatial and temporal expression pattern. The STAT-family has seven members in mammals, which are the effectors of the Janus-Kinase (JAK)-STAT pathway. This pathway relays cytokine signalling. Notably, the JAK-STAT pathway plays a role in regulation of BAT and WAT function (reviewed in [194]). In summary, both NF1 and STAT proteins are promising candidates for binding upstream of the intronic DR1 element and thus should be a focus of upcoming validation experiments.

Secondly, the binding sites for MyoD and myogenin have not been identified yet. A possible binding site for was identified by bioinformatics (3.7.2): The element is located 55 bp upstream of the NF1/STAT element. The sequence of this element, CAGCTC, fits to the MyoD consensus binding motif: An EBox fitting to the consensus for MyoD binding: CASCTG [195] (Myo-EBox). Examination of the C2C12 ChIP-seq signals reveals that the MyoD and myogenin peaks cover the putative MyoD, NF1 and STAT elements. Alternatively, basic helix-loop-helix (bHLH) TFs can interact with NF1 binding elements

[191]. Consensus NF1 sites can contain other binding sites interjected between the TGG and CCA half sites. Candidates for such an interjection are bHLH factors, including MyoD and myogenin. Comparing the sequence of the putative UCP3 intron 1 NF1 binding site, CTGGGAAGGAGCCA, to the sequence of the proposed NF1/myogenin binding site, TTGGCACGGWGCCA, shows that these sites partly resemble each other. While the hypothesis that myogenin, and possibly MyoD, binds to the NF1 element in cooperation with NF1-factors is tempting, it is too early to draw any reliable conclusion. In summary, to pinpoint the binding site for MyoD and Myogenin, experimental evidence is needed. A suitable first step is to mutate the Myo-EBox in reporter gene construct and to assay whether this mutation influences reporter gene activity in C2C12 cells. If so, binding to this region should be assayed using EMSA.

The third question deals with the bimodal binding pattern of Pol2 and p300. Two possible explanation seem most likely: The bimodal binding pattern observed for Pol2 and p300 could originate either from two independent binding loci for either factor, or from one single complex containing both binding the core promoter region and the intronic enhancer. In the latter hypothesis, the pattern of Pol2 and p300 binding supports the hypothetical formation of one large promoter-enhancer complex, possibly via SP1-mediated DNA bending (see 4.4.1 and 4.9).

In summary, identification of the NF1/STAT element and binding of MyoD, myogenin in C2C12 cells, and p300 in heart, indicate that the intronic enhancer is important for regulation of UCP3 in SKTM and heart as well. This is also supported by intronic DNaseI hypersensitive (DHS) sites in heart and SKTM (3.8). Such a diverse enhancer region does allow for a lot of regulatory complexity. Possible steps towards understanding the mechanism of action are pinpointing the binding motifs for all factors and identification of the factors binding the NF1/STAT element. If indeed NF1 and STAT transcription factors act via the intronic enhancer, this would greatly underline the importance of this region for regulation of UCP3 expression.

#### **4.5. Relevance of the intronic enhancer in skeletal muscle and heart**

During the course of this study it became evident that the UCP3 intron 1 enhancer is likely to be relevant for regulation of UCP3 expression in SKTM. The main arguments for this hypothesis are: 1.) MyoD and myogenin bind to the enhancer region in C2C12 cells (3.7.1). 2.) Reporter gene assays demonstrate an influence of the IVS1+1505 allele on reporter activity in C2C12 cells (3.1.5). 3.) PPAR $\gamma$ , RXR and SP1/SP3 are expressed in C2C12 cells (3.1.4 and 3.4.2), regulate expression of UCP3 in SKTM and C2C12 cells, and act via the intronic enhancer in BAT. 4.) A DNase I hypersensitive (DHS)-site is present at the location of the intronic enhancer (3.8). We furthermore assume that the enhancer is active in heart as well, but so far only binding of p300 (3.7.1) and presence of a DHS site (3.8) support this.

While the enhancer may be equally important in all three tissues, several experiments indicate that mechanism of action and element specific binding vary. Firstly, the IVS1+1505A allele affects UCP3 expression to a different degree in *Phodopus sungorus* BAT and SKTM [59]. Secondly, several tissue specific regulators may act via the intronic enhancer. Thirdly, the physiological regulation of UCP3 is known to be different between SKTM and BAT (discussed in 1.4.1.2). We assume that this differential regulation originates from different usage of the binding elements in the different tissues. Furthermore tissue specific co-regulators may play a role.

#### 4.5.1. Tissue specific binding to the four putative elements of the enhancer

Of all four elements, the GC-Box seems to be the most unlikely candidate to achieve tissue specific regulation. While it was discovered due to a base exchange with a BAT-specific effect (IVS1+1505G→A, [128]), we demonstrated that this element binds SP1 and SP3, two ubiquitous transcription factors. Later, several experiments indicate that binding to this element is important in SKTM as well. The obvious hypothesis is that SP1 and SP3 likewise do bind to the GC-Box in BAT, SKTM and heart.

Even if so, SP1 and SP3 are known targets of a wide array of post-transcriptional modifications. SP1 can be acetylated by p300 [186], phosphorylated by PKA [196], is SUMOylated [197], glycosylated [198] and ubiquitinated [199]. The same is true for SP3, where SUMOylation is believed to be able to switch SP3 from an activator to a repressor [176]. Differences in the modification pattern of SP1 or SP3 in SKTM and BAT might allow tissue specific expression regulation.

While SP1 and SP3 are the most prominent GC-Box binding factors, other proteins are known to share the binding preferences of SP1 and SP3. Possible candidates are the Krüppel-like factors (KLFs) which are closely related to the SP-family [177]. Similar to the SP transcription factors, binding of KLFs is sensitive to mithramycin [200]. For example, one publication reports mithramycin-sensitive regulation of UCP3 expression by KLF5, but in that case the regulation takes place at the core promoter [91]. Of the KLFs, KLF 9, 11 and 15 are of special interest. KLF 9 and KLF11 share their binding site preferences with the SP1 and SP3. KLF11 plays a role in regulation of UCP1 transcription via the UCP1 upstream enhancer [201], is a regulator of hepatic lipid metabolism [202] and important for PPAR $\gamma$  signalling [203]. KLF15 regulates UCP1 expression, adipogenesis [204] and cardiac lipid metabolism [205]. In contrast to SP1 and SP3, KLFs show differential, tissue specific expression patterns. It is thus possible that KLFs can bind to the intronic GC-Box under certain conditions in vivo in a tissue specific manner.

Located directly adjacent to the GC-Box is the intronic DR1 element. It was identified by a ChIP-seq experiment [73,74] in mouse cell cultures and is bound by PPAR $\gamma$  and RXR. PPAR $\gamma$  signalling is a crucial pathway regulating UCP3 expression in SKTM, BAT and heart (see 1.4.13, 3.10.2 and 3.10.3). Generating several mutated reporter gene constructs we demonstrated that this element is sufficient and indispensable for PPAR $\gamma$  signalling (3.5.1). While it has been reliably demonstrated by both colleagues and us that the intronic DR1 element does confer the effect of PPAR $\gamma$  agonists, this does not rule out that other proteins can act via the same element as well. Apart from PPAR $\gamma$ /RXR $\alpha$  dimers, several other transcription factor dimers that can bind DR1 elements, including PPAR $\alpha$ /RXR $\alpha$ - and PPAR $\delta$ /RXR $\alpha$ -dimers [206]. Furthermore, RXR $\alpha$  can form dimers with the nuclear hormone receptors RAR $\alpha$ , TR $\beta$  and VDR [207], which are all known to regulate expression of UCP3 (see 1.4.1.3). Thus, one or several of these different nuclear hormone receptor dimers might be able to replace the PPAR $\gamma$ /RXR dimer at the intronic DR1 element and thereby allow for tissue specific regulation.

Upstream of the DR1 element, another regulatory region is located. This region contains two putative, overlapping binding sites: A NF1 element and a STAT element. In contrast to the other binding elements, we do not have experimental evidence which factors bind the putative this element. In SKTM very limited evidence exists that myogenin might act via the putative NF1 element (see 4.4.2). For other tissues, the only hint towards the binding factor are the matrix identifications by MatInspector. This region thus is the biggest unknown of the enhancer. Reporter gene experiments support that it is functional in BAT, and both NF1 and STAT family transcription factors play a role in regulation of BAT, SKTM and heart gene expression [194,208-211]. A first step in to clarify regulation via this region is to elucidate whether the NF1-matrix, the STAT-matrix or both matrices are functional and relevant. The second step then would be to identify the binding factor or factors.



The last transcription factors known to regulate UCP3 expression via the intronic enhancer are MyoD and myogenin. Both are well established regulators of UCP3 expression [79,99]. MyoD and myogenin act via EBoxes, short elements of the sequence CANNTG. Bioinformatics identified a putative EBox upstream that is located about 55 bp upstream of the NF1/STAT element in mouse, rat and hamster. Deletion of this element did not influence reporter gene activity in the HIB1b deletion screen, fitting to the fact that neither MyoD, nor myogenin is expressed in BAT or HIB1b cells (3.1.4). The element co-localizes with MyoD and myogenin ChIP-seq peaks in C2C12 cells (3.7.1), but so far no hard evidence supports that it indeed binds either or both factors. Apart from MyoD and myogenin, other EBox binding factors are known. E-proteins can either heterodimerize with MyoD and myogenin [195], or can replace them at the end of differentiation [212]. Interestingly, bioinformatics identify the putative EBox-element of mouse and hamster as a Transcription factor E2a (E12/E47) binding site (see appendix 9). In contrast to MyoD and myogenin, not all EBox binding proteins are SKTM specific. Thus, other proteins other than MyoD and myogenin may act via the putative EBox in SKTM, and the element may be bound by other transcription factors in other tissues as well. For two of these EBox factors, Transcription factor E2a (E12/E47) (alias TCF3) and TCF12, ChIP seq data actually are available in ENCODE, but only for timepoints mid differentiation where MyoD and myogenin binding was high. No binding was detected at that time points.

In addition to the transcription factors binding to the four elements mentioned above, the regulatory complex will also contain co-regulators, which add another layer of regulation that can greatly differ between different tissues (see examples in 1.4.4). So far the only co-regulator recruited to the intronic enhancer is p300, a ubiquitous histone acetylase required for UCP3 expression [79], as is demonstrated by ChIP-seq in heart (3.7.1). In addition of to this ubiquitous factor, the enhancer also may recruit tissue specific transcription co-activators. To fully understand the regulatory activity of the UCP3 intron 1 enhancer, it thus might be required to also identify the co-activators recruited in the different tissues.

#### **4.5.2. Different effect of the IVS1+1505A allele in muscle and C2C12 cells**

Our initial hypothesis was that the *Phodopus sungorus* UCP3 intron 1 IVS1+1505G element is bound by a BAT specific transcription factor. This was based on the finding that hamsters homozygous for the IVS1+1505A allele completely lack UCP3 mRNA and protein in BAT. In contrast, abundance in SKTM is only slightly reduced [59]. In this study we investigated this difference in cell culture. Interestingly, experiments in C2C12 cells did not reproduce the resistance of hamster muscle to the effect of the IVS1+1505A allele. Firstly, in reporter gene assays the IVS1+1505A allele completely blocked agonist-induced activation of reporter gene activity (see 3.8.2). Secondly, mithramycin, an inhibitor of GC-Box binding, completely blocked induction of endogenous UCP3 transcript by PPAR $\gamma$  and PPAR $\delta$  agonists (see 3.8.3). Thirdly, complex formation occurred with nuclear extracts from both BAT and SKTM cell lines. To identify the reason for the difference between the initial finding in *Phodopus sungorus* muscle and our experiments in C2C12 cells, it is important to understand the differences between the two different systems.

One difference is that cultured cell lines do not correctly replicate the in vivo situation found in a tissue. In cultured cells both internal and external factors most likely differ from the situation in tissue. It is thus possible that a factor required for the partial suppression of the IVS1+1505A allelic effect on UCP3 expression in SKTM is absent in our C2C12 cell culture. Such a difference could originate both from an intrinsic difference between muscle the C2C12 cells, and from an unsuitable differentiation state of the C2C12 cells. The factor may be a protein, a chromatin modification or an intra- or extracellular signal, which is absent or different in cell culture. So far we only tested small molecule signals for their

ability to overcome the effect of the IVS1+1505A allele. Despite applying several agonists alone and in combination we were not able to find any such condition (see 3.1.6 and 3.10.2). We are aware that this does not rule out that there may be a condition where the IVS1+1505A reporter gene construct is active in C2C12 cells, but due to the countless possible options we were unable to test all.

Furthermore, it is possible that *Mus musculus* is different from *Phodopus sungorus* in some component relevant for regulation of UCP3 expression. For example, slight differences between mouse and hamster proteins may render the mouse complex more susceptible to disturbances. Furthermore, the mouse GC-Box element differs in sequence from the hamster element (figure 31D) and has weaker competition potential (see 3.6.2). This does hint towards a lower affinity of SP1/SP3 for the mouse element, and thus, a more susceptible complex. A combination of both phenomena would explain the inconsistencies encountered.

Apart from the two model systems, also the experiments and readouts differ between the initial finding in hamster and our cell culture experiments. In hamster, the effect of a natural occurring mutation on endogenous UCP3 expression was measured. In the C2C12 cells, reporter gene activity of an extra chromosomal reporter construct and the effect of a chemical inhibitor, mithramycin, were analysed. Both reporter gene experiments and mithramycin treatment experiments do not perfectly reproduce the in vivo regulation of the UCP3 gene in *Phodopus sungorus*.

The reporter gene construct is different from the endogenous UCP3 gene, mithramycin treatment has different effects than a point mutation. Our reporter constructs miss several features that might be relevant for correct regulation of UCP3. Neither long distance enhancers, nor regulatory sites downstream of the start-codon are included. One of these sites may be required for UCP3 expression in SKTM of IVS1+1505A hamsters. Our search for such elements started with the region upstream of the UCP3 core promoter. Neither inclusion of a COUP-TF2 binding site 800 bp upstream, nor inclusion of a PPAR $\gamma$  site 2000 bp upstream of the TSS reduced the effect of the IVS1+1505A allele (see 3.10.1). Recently a nuclear response factor 2 (Nrf2) site was discovered located further upstream [47], which is a promising candidate.

A further problem for our reporter gene assays may be the post-confluent state of cell lines during which they were carried out. Neither of the cell lines contained endogenous UCP3 transcript in this state (see 3.1.4). In the case of C2C12 cells this was necessary as differentiated C2C12 cells evade transfection. In HIB1b cells we worked with this state as we did not see differences between Lipofectamin-transfected, post-confluent cells and Nucleofector-transfected, differentiated cells in our initial experiments (not shown). We attempted to solve problem by stable integration of our reporter gene constructs but so far we did not succeed (not shown).

Mithramycin treatment may not reflect the effect of the IVS1+1505A effect in vivo. While it may completely block binding to the GC-Box, the IVS1+1505A allele might retain some binding activity (see 3.1.2). The residual binding activity may be sufficient to ensure expression of the UCP3 gene in SKTM. This hypothesis would raise the question why residual binding activity is sufficient to sustain UCP3 expression in SKTM, but not in BAT. Furthermore, mithramycin might have additional secondary effects on UCP3 expression, for example via impairing the binding of KLF5 to the core promoter [91]. This effect could either be additive with the impairment of intronic enhancer action or mimic the effect of the IVS1+1505A allele in BAT.

With the data currently available it is difficult to answer why the IVS1+1505A allele only mildly influences UCP3 expression in SKTM and why inhibition of SP transcription factor binding suppresses expression of UCP3 reporter gene constructs and the endogenous gene in C2C12 cells. Currently our preferred model is most that the weakened binding to the GC-Box in *Phodopus sungorus* SKTM is stabilised by, or compensated for by a muscle specific factor. The factor may be absent in our C2C12 cell culture experiments because C2C12 cells do not correctly reproduce the in vivo situation. Initially, MyoD and myogenin were considered to be candidates for this a factor, but they are already expressed in non-differentiated C2C12 cells (3.1.4). Instead, EBox binding proteins may be relevant (discussed in 4.5.1). Another promising candidate is PPAR $\delta$ . PPAR $\delta$  regulates UCP3 expression in SKTM (see 1.4.4) and PPAR $\delta$  agonists do induce endogenous UCP3 expression in fully differentiated C2C12 cells (3.10.3). In contrast, PPAR $\delta$  agonists have no effect on reporter gene activity in post-confluent cells (3.10.2). Induction of the endogenous transcript furthermore is sensitive to mithramycin. As PPAR $\delta$  signalling depends on KLF5 binding [91], inhibition of KLF5 binding [200] might explain the effect of mithramycin on PPAR $\delta$  signalling. This makes KLF5 an additional important candidate to explain resistance of SKTM UCP3 expression to the IVS1+1505A allele.

Ultimately only the combination of two experiments is able to clarify the relevance of binding to the intronic GC-Box in SKTM. The first experiment is ChIP in *Phodopus sungorus* SKTM tissue. This experiment would answer the question whether the IVS1+1505A allele prevents binding of SP1 and SP3 to the intronic enhancer in SKTM. The second experiment is deletion of the complete putative intronic GC-Box element from the nuclear genome in C2C12 cells using CRISPR/Cas9 or TALEN mediated genome editing. Four outcomes are possible: 1.) If SP1 and/or SP3 do bind to the intronic enhancer in SKTM of IVS1+1505A hamsters and removal of the GC-Box leads to loss of UCP3 expression, then binding is stabilised by a muscle specific factor in hamster. 2.) If binding is prevented by the IVS1+1505A allele and removal of the GC-Box leads to loss of UCP3 expression, then another protein binds to the GC-Box in SKTM. 3.) If SP1 and SP3 still bind to the enhancer in SKTM of IVS1+1505A animals and removal of the GC-Box has no influence on UCP3 expression, then SP1 and SP3 bind to an alternative site in SKTM. 4.) If SP1 and SP3 do not bind to the intronic enhancer in IVS1+1505A animals and GC-Box removal has no effect on UCP3 expression, then neither the GC-Box element, nor binding of SP1 or SP3, is required in SKTM.

#### **4.5.3. Summary: The enhancer in brown adipose tissue, skeletal muscle and heart**

The intronic enhancer was detected due to a mutation in a single TFBS binding an unknown protein in BAT. Now we know that the intron harbours a complex regulatory region most likely containing four regulatory elements: A GC-Box, a DR1 element, a NF1 or STAT binding site, and an EBox. Data support seven transcriptional regulators to act via this intronic enhancer: SP1, SP3, PPAR $\gamma$ , RXR, MyoD, myogenin and p300. Furthermore, several experiments demonstrate that the enhancer is relevant for UCP3 expression in BAT, SKTM and heart. This includes EMSA experiments, reporter gene assays, and ChIP and DHS site data. Tissue specific regulation is possible for all of the elements described (see 4.5.1). Furthermore, these factors may recruit tissue specific co-regulators. Even though the exact tissue specific binding patterns are not fully understood to date, it is undeniable that the enhancer is one of the main regions regulating UCP3 expression. Therefore, understanding the tissue specific regulatory enhancer complex will be a very important step towards understanding expression regulation of UCP3.

#### 4.6. Presence of the enhancer module in other species

The intronic enhancer was initially identified due to the ablation of UCP3 expression induced by a mutation within the first intron of the UCP3 gene of *Phodopus sungorus* [128]. Despite the fact that *Phodopus sungorus* is an interesting model organism per se, we quickly wondered whether the first intron of UCP3 was equally important for expression regulation in other species as well. Initially, we started our search in mouse and rat, but later also examined horse, pig and human.

Sequence alignment and dotplot analysis revealed that the intronic DR1/GC-Box module was conserved in mouse, hamster and rat. Interestingly, in all three species the distance between the two elements was comparable and both elements faced each other orientation-wise. Further sequence analysis, using more sophisticated bioinformatics, identified comparable modules in human, pig and horse. We therefore decided to examine intronic DR1/GC-Box modules with this orientation more closely. Using EMSA competition, we were able to validate binding of SP1 and SP3 to the GC-Box component of selected modules (3.6.2) from mouse, rat, human and pig. Interestingly, in all species the elements were identified in roughly comparable distance to the TSS.

To validate these modules, reporter gene constructs were cloned for *Mus musculus* and *Homo sapiens*, but none of the vectors exhibited any reporter gene activity. Despite several attempts, we were unable to solve that problem. A plausible explanation for this phenomenon may be polyadenylation sequences between core promoter and luciferase gene, which are present in *Homo sapiens* and *Mus musculus*, but absent in *Phodopus sungorus*. Whatever the reason for the inactivity of the non-*Phodopus* reporter gene constructs may be, it prevented validation of the putative modules inside our cell culture systems.

We therefore resorted to use bioinformatics once more. After identification of the putative NF1/STAT element and uncovering binding of MyoD and myogenin to an unknown binding site, we refined our bioinformatic models. The new models proved suitable for prediction of the enhancer region in mouse, rat and hamster, and located a putative MyoD binding motif in direct proximity of the NF1/STAT element (3.7.2). We then used several models with relaxed prediction criteria and manually searched for MyoD, NF1 and STAT elements within the vicinity of the aforementioned DR1/GC-Boxes. That way, comparable putative enhancer regions in human, horse and pig (see 3.7.2) were identified. As none of these non-rodent enhancer regions perfectly resembles the enhancer region, and as for each species more than one putative region was detected, further validation experiments are needed.

In contrast, the situation is conclusive for mouse, rat and hamster: The near-perfect conservation of all elements, the ChIP-seq data from mouse heart and C2C12 cells, the EMSA data and the cell culture experiments strongly support the importance of the intronic enhancer in the three rodent species.

#### 4.7. Regulation of UCP3 expression – interaction of promoter and enhancer

Two regions are known to be crucial for regulation of UCP3 expression: The intronic enhancer and the core promoter. Both regions contain several putative binding sites, and the proposed binding proteins and regulatory functions of these regions are partly overlapping. Notably, in the core promoter, a triple EBox and a DR1 element are located while the intronic enhancer contains at least a single EBox and another DR1 element. In both regions, the EBoxes are supposed to bind MyoD and myogenin ([95] and 3.7.1), while binding of different nuclear hormone receptor heterodimers has been suggested for the DR1 elements (see 1.4.1.3 and [73]). The questions arises, which regulatory element binds which transcription factor, and how the different regulatory regions interact.

Now we know that the promoter DR1 element does not bind PPAR $\gamma$ /RXR dimers (3.9 and [73]) and that neither MyoD, nor myogenin bind to the promoter triple EBox (3.7.1). These factors instead bind to the intronic enhancer, which also explains while studies using intronless reporter gene constructs failed to see PPAR $\gamma$  effects on UCP3 expression [78]. This raises the question which transcription factors act via the core promoter. For the triple EBox this question still is unanswered. For the promoter DR1 element, at least in muscle, colleagues have demonstrated binding of RAR and PPAR $\delta$  [91,95]. Less is known about proteins binding to the promoter DR1 in BAT, but our reportergene assays demonstrate that the promoter DR1 element is an important regulatory element (3.5.1) in BAT as well. In addition to these factors, other hormone receptors that can act via a DR1 elements. Both TR and PPAR $\alpha$  bind such elements and are known regulators of UCP3 expression (1.4.1.3). In addition, a GC-Box and a CCAAT-Box have been described adjacent to the promoter DR1 element, binding KLF5 and C/EBP $\beta$ , respectively. Future experiments are needed to elucidate which transcription factors act via promoter elements and which bind at the enhancer.

No matter where regulatory TFBS are located, ultimately all signals need to be integrated to start transcription from the TSS. The most direct mechanism is that elements located outside of the core promoter nucleotide-wise may be in close proximity to the promoter when considering the three dimensional nucleosome structure. This proximity may be achieved by DNA-bending, bridge/scaffold proteins or the packaging architecture by the nucleosomes. In such a case, promoter and enhancer would not form two independent regulatory complexes, but rather one large supercomplex containing both DNA sites. Such interactions may be facilitated by the transcription factors themselves, for example by SP1 or SP3 [189,190], or co-factors interacting with the transcription factors. This mechanism already has been suggested for the first intron of UCP2, UCP3 intron 1, and the UCP3 promoter upon identification of the intronic DR1 element [73]. Apart from direct interactions, enhancers can influence chromatin condensation and thereby allow or prohibit accessibility of other regulatory DNA elements. Both mechanisms might be relevant for regulation of UCP3 expression, supported by the bimodal binding pattern of Pol2 and p300 (3.7.1), where ChIP-seq reports binding of these proteins at both promoter and intronic enhancer.

#### 4.8. Identification of a putative upstream enhancer

Recently evidence for the existence of a putative upstream regulatory regions has accumulated. Firstly, COUP-TF2 binding regulatory element was reported upstream of the core promoter [98]. Secondly, screening CHIP-seq data [169], we identified a PPAR $\gamma$  binding region 2000 bp upstream of the TSS (3.9) of UCP3. Furthermore, a DHS-site at the same region supports a site of ongoing transcriptional regulation in SKTM (3.8).

We assayed a possible influence of the respective regions using reporter gene constructs. Neither inclusion of the COUP-TF2 element, nor inclusion of the putative PPAR $\gamma$  binding/DHS region into our reporter gene constructs had any effect on PPAR-induced reporter activity. In case of COUP-TF2, the absence of an effect may stem from the absence of the respective activating signal pathway in our assays, although COUP-TF2 mRNA was present in all cell lines used (see 3.1.4). Recently, a Nrf2-binding antioxidant response element (ARE) was discovered. This ARE mediates the effect of reactive oxygen species (ROS) signalling on UCP3 expression in SKTM [47]. In addition, the Nrf2 site is functional in heart [47], despite absence of a DHS-site. The ARE is located directly upstream of the PPAR $\gamma$  ChIP-seq peak and also is covered by the SKTM specific DHS-site. Despite being located close to the putative PPAR $\gamma$  binding region, the ARE element was not part of our reporter construct.

In summary, three pieces of evidence support existence of an enhancer located roughly 2000 bp upstream of the UCP3 TSS: Presence of a DHS-site, a PPAR $\gamma$  ChIP-seq peak and a functional, Nrf2 binding, ARE. Being slightly closer to the TSS of the neighbouring C2cd3 gene than to the TSS of UCP3 the enhancer, we wondered whether this enhancer regulates C2cd23 rather than UCP3, but the fact that the Nrf2 element regulates expression of UCP3 [47] argues against that. We therefore conclude that this enhancer regulates UCP3 expression.

#### 4.9. Regulation of UCP3 transcription: A refined model

We have collected sufficient data to propose a greatly refined model of transcriptional regulation of UCP3 expression. Not only does this model add new transcription factors to the array of already known regulators, but it furthermore describes two additional regions which mediate transcriptional regulation.

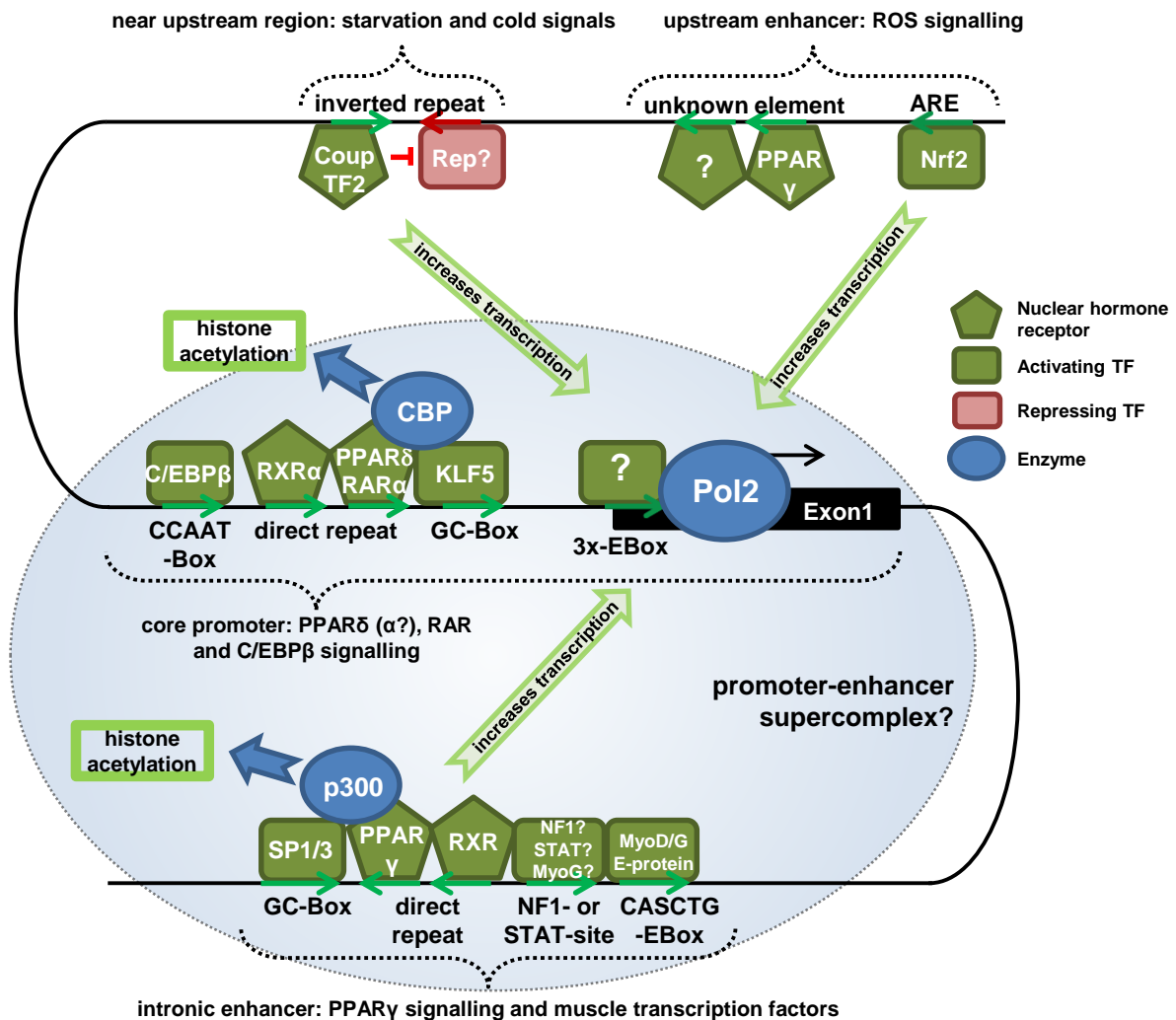
The first major regulatory region is a putative enhancer located about 2000 bp upstream of the TSS of UCP3 in mouse. The region is marked by a DHS-site in SKTM, and contains an Nrf2 binding ARE, and a yet unknown element that binds PPAR $\gamma$ . While the Nrf2 site has been demonstrated to mediate ROS-signalling [47] in SKTM and heart, it is not clear yet if the PPAR $\gamma$  binding site is functionally relevant. Little is known whether this region also contains other regulatory elements, but on the level of reporter gene assays the region seems to be of secondary importance, at least in BAT and in absence of ROS stress.

Between this region and the TSS COUP-TF2 and a yet unknown repressor bind to an inverted repeat. This element integrates cold- and starvation-signals [98]. The element is neither located in a DHS-site, nor are any other known regulatory elements located nearby.

The second major regulatory region is the core promoter, which covers roughly 250 bp upstream of the TSS and exon 1. Most notably, it contains a triple E-Box (3x-Ebox), a CCAAT-Box, a GC-Box and a DR1 element. The latter two elements are both required for PPAR $\delta$  [91] and retinoic acid [95] signalling in SKTM, and their binding proteins are able to recruit the histone acetylase CBP. Furthermore they may be relevant for PPAR $\alpha$  and thyroid receptor (TR) signalling. The CCAAT-Box binds C/EBP $\beta$ , while the proteins binding the EBox are unknown.

The third major regulatory region is the intronic enhancer. In contrast to the putative upstream enhancer, this region is indispensable for expression of UCP3. Covered by a prominent DHS-site, it contains a GC-Box, a DR1 element [69], an EBox and a NF1 and/or STAT element. The DR1 element and the GC-Box are required for PPAR $\gamma$  signalling in BAT, SKTM and possibly in heart as well. They are bound by PPAR $\gamma$ /RXR heterodimers and SP1/SP3, respectively. At least in BAT, both components are absolutely necessary for both basal and PPAR agonist stimulated expression of UCP3. In SKTM, the EBox element binding factors are little investigated, but most likely include MyoD and myogenin. For the NF1/STAT site, so far no binding proteins have been predicted, but its deletion strongly reduces reporter gene construct activity in BAT cell lines. In addition, the intronic enhancer recruits the histone acetylase p300. The enhancer may physically interact with the promoter, thereby forming one large complex.

In summary, transcriptional regulation of UCP3 expression is dominated by an intronic enhancer and the core promoter which both bind important tissue specific and metabolism related transcription factors. Together, these regions integrate PPAR signalling, response to thyroid hormone and retinoids. Both regions are crucial for chromatin decondensation by recruiting histone acetylases. CBP is recruited to the promoter while p300 is recruited to the enhancer, in both cases via a combined GC-Box/DR1 elements binding a SP/KLF transcription factor and a PPAR nuclear hormone receptor, respectively.



**Figure 46: Refined model of regulation of UCP3 transcription:** Schematic graph describing the regions relevant for regulation of UCP3 transcription. The regulatory region of the UCP3 gene is depicted as an S-shaped thick black line with a filled black box indicating the location of exon 1. Transcriptional activators are marked green while repressors are red. Enzymes are indicated blue. Four regions (dotted brackets) are relevant for regulation of UCP3 expression. The core promoter, the intronic enhancer and two upstream enhancer regions. The intronic enhancer and the core promoter are likely to participate in one large regulatory supercomplex. Both the promoter and the intronic enhancer recruit histone acetylases to open the chromatin structure. All regions relay different signals: The upstream regions are relevant for COUP-TF2 and ROS signalling, the core promoter integrates PPAR $\delta$ , RAR $\alpha$  and C/EBP $\beta$  signalling while the intronic enhancer is crucial for PPAR $\gamma$  signalling and binds muscle specific transcription factors. Furthermore the enhancer may bind proteins of the NF1 or STAT family.



## 5. Outlook

This study investigated the regulation of UCP3 expression in brown adipose tissue (BAT). The thesis was triggered by identification of a BAT-specific base exchange that ablates a binding present within the first intron of the UCP3 gene. Surprisingly, three additional transcription factor binding sites (TFBS) were identified within the first intron of the UCP3 gene: A GC-Box element, a direct repeat (DR) 1 element, a nuclear factor (NF) 1 or signal transducer and activator of transcription (STAT) element and a putative EBox. Furthermore, experiments, bioinformatics and data-mining indicate that the enhancer plays is important for regulation of UCP3 expression in skeletal muscle and heart. Lastly, we collected evidence that a putative upstream enhancer may exist that also participates in regulation of UCP3 expression.

So far, only three of the four putative binding sites within the first intron of UCP3 are reliably validated: The GC-Box, the DR1-element and the NF1/STAT element. For the NF1/STAT element it is still unknown which of the two putative binding sites is functional. In addition, a putative MyoD binding element, an EBox with the sequence CAGCTG, was annotated by bioinformatics, but to date no experimental evidence is available whether this element is functional. Lastly, the spacing between the individual elements is large enough to accommodate additional binding elements. The first goal of future experiment can be to validate the EBox and investigate the two components of the NF1/STAT element. This can be achieved by mutagenesis of the elements in our reporter gene vectors and subsequent reporter gene assays. Next, deletion or mutation of the inter-element regions could be used to screen for additional elements.

Once it is known which elements are functional and of importance, the question arises which proteins bind to them. In BAT, the binding proteins for the GC-Box and the DR1-element are known, and we expect the same proteins to bind to these elements in SKTM and heart. We also know of MyoD and myogenin binding in SKTM, although this binding may be temporal. For the NF1/STAT element to date no binding protein is known yet. The identification of proteins binding to the EBox and the NF1/STAT element thus is a second important goal. This could be achieved either by affinity chromatography, or by bioinformatic prediction followed by electrophoretic mobility shift assay (EMSA) and reporter gene assay validation. A possible strategy to aid identification is to use the enhancer-trapping approach described in 3.12 [138]. This approach would factor in protein-protein interactions between the different components of the intronic enhancer and may allow identification of co-factors recruited by the whole complex. That way, new regulatory pathways may be identified, hopefully extending the list of physiological stimuli which regulate UCP3 expression.

In addition, it is interesting to understand which pathways are relevant in which tissue. For example, it is known that changes in UCP3 expression differ between BAT and SKTM in response to fasting and cold (see 1.4.1.2). Some of the aforementioned transcription factors are expressed widely or even ubiquitously, but for several elements of the enhancer, tissue specific interactions are possible (see 4.5.1). It is may thus be worthwhile to assay the importance of regulatory elements and their binding transcription factors in different tissues. For example, the putative MyoD/myogenin binding EBox was irrelevant in HIB1b cells, but may be functional in BAT and heart. The NF1/STAT site might be bound by different proteins in BAT than in SKTM. The experiments described in the paragraph above thus have to be carried out at least in HIB1b and C2C12 cells. Furthermore the deletion screen (3.7) should be repeated in C2C12 cells, as putative muscle specific elements (see 3.7.1) would have evaded

detection in HIB1b cells. Last, but not least, binding candidates need to be validated with ChIP in the respective cell lines or the tissues themselves.

While the study was based on a finding in *Phodopus sungorus*, EMSA competition experiments, sequence alignments and bioinformatics (3.7.2), suggested putative related enhancer regions in the UCP3 genes of other species. These element elements still await experimental validation. Initial attempts to generate human and mouse reporter gene constructs were not successful, possibly due to aberrant polyadenylation. Two possible methods may circumvent this problem. Firstly, for every species of interest a new reporter vector could be generated. Insertion of the luciferase cassette at the end of intron 1, between core promoter and the intronic enhancer, would circumvent problems with cryptic intronic polyadenylation signals. Secondly, the intronic enhancer region could be removed from the already existing functional *Phodopus sungorus* reporter gene constructs. Insertion of the putative enhancer elements from other species would then demonstrate whether they are able to compensate for the loss of the *Phodopus sungorus* enhancer module. No matter how well conserved intronic regulation of UCP3 expression proves to be, it might provide interesting insight in the evolution of the UCP3 gene, especially when UCP3 related genes from the basal *protheria* or even non-mammalian species like birds are included.

Apart from our work on the intronic enhancer, we also uncovered evidence for an additional enhancer region located 2000 bp upstream of the TSS of the mouse UCP3 gene. To date it is uncertain whether a complex upstream enhancer really exists and which functions it has. Published data [47] suggests that it is required for mediating the effect of ROS-signalling on UCP3 expression in SKTM and heart. In addition, ChIP-seq data uncovered a PPAR $\gamma$  binding region nearby. So far we observed no effect of the upstream PPAR binding site on reporter gene activity in C2C12 and HIB1b cells. This may originate from absence of the Nrf2 site in our constructs, result from the lack of relevant signals or signalling proteins, or be a consequence of using unsuitable cell lines or an unsuitable differentiation state of the cell lines. It is thus important to first find a suitable reporter system that can be used to assay the activity of the upstream enhancer. Once such a system is available, further experiments may be planned.

The conclusion of this thesis is a model on regulation of UCP3 transcription that combines all of these findings as well as insight from several studies of colleagues. This model (figure 46) still does not include all transcription factors known to regulate UCP3 expression. The binding sites for VDR, TR and PPAR $\alpha$  are undetermined and the factors binding a triple EBox within the core promoter are still unknown. Possible strategies to identify binding proteins for the promoter EBox are bioinformatics and affinity chromatography, as described above, while binding sites for the nuclear hormone receptors of interest may be identified by ChIP-seq followed by a mutation/deletion screen. Nevertheless we think that our model both advances and combines the current knowledge about regulation of UCP3 expression and clarifies some contradictions and uncertainties that existed at the beginning of this thesis. Some questions still prevail, but we hope that the goals and strategies described in this outlook will help to understand the tissue specific regulation of UCP3 transcription in response to the physiological signals known to influence UCP3 expression, and possible also in response to stimuli yet unknown.

## 6. Summary – PhD thesis Christoph Hoffmann

Uncoupling protein 3 (UCP3) is a protein with proposed functions in lipid metabolism and reactive oxygen species mitigation. UCP3 is expressed in brown adipose tissue (BAT), skeletal muscle (SKTM) and heart. Physiological parameters regulating its expression include local free fatty acid concentration and handling capacity. On the molecular level, this regulation was believed to be mediated via a direct repeat (DR)1 element within the core promoter. Such a DR1 element can bind, beside other nuclear hormone receptors, peroxisome proliferate activated receptors (PPARs). Being considered important regulators of cellular energy metabolism, PPARs are activated by intermediates of lipolysis and other fatty acid derivatives.

Discovery of an intronic base exchange, IVS1+1505G→A, which abolished UCP3 expression on both mRNA and protein level in BAT of *Phodopus sungorus*, shifted our focus to the first intron. Despite being located more than 1500 bp downstream of the promoter DR1 element, the A-allele appeared to act by completely blocking PPAR-agonist activation of UCP3 transcription. Interestingly, the base exchange only led to a minor decrease of UCP3 expression in SKTM. Our goal was to understand the mechanism behind the effect and elucidate the tissue specificity observed.

We discovered a SP1/SP3 binding GC-Box element in the first intron of the UCP3 gene, which was impaired by the IVS1+1505A allele. In parallel, ChIP-seq experiments performed by colleagues uncovered two putative PPAR-binding DR1 elements in mouse: one intronic element, and one element 2000 bp upstream of the transcriptional start site. Employing SP-binding inhibitors and generating mutated reporter gene vectors, we were able to prove that the intronic DR1 element and the GC-Box are interdependent and indispensable for expression of UCP3. Furthermore we demonstrated that both elements together are sufficient and required for PPAR $\gamma$  agonist activation of UCP3 transcription. We also uncovered conservation of both intronic elements in mouse, rat and hamster, and gathered evidence for presence of related elements in horse, pig and human. In contrast, we did not find evidence for an influence of the upstream DR1 element on UCP3 transcription.

We screened for further regulatory elements in the vicinity of the intronic DR1/GC-Box module by carrying out a deletion screen with our reporter gene constructs. Indeed, we uncovered another regulatory region in juxtaposition, which possibly harbours a nuclear factor (NF) 1 and/or signal transducer and activator of transcription (STAT) element. With the help of bioinformatics, we furthermore were able to discover an EBox element in close proximity. Data and publication mining also suggest activity of the intronic enhancer in SKTM and heart and demonstrates binding of MyoD and myogenin in C2C12 cells, a SKTM cell line. Furthermore, this revealed that a histone acetylase, p300, preferably binds to the intronic enhancer region, at least in heart.

Taken together, we uncovered and characterized a complex enhancer region within the first intron of the UCP3 gene. We provide strong evidence that this enhancer binds multiple transcription factors previously believed to act via the core promoter. That way, we were able to greatly refine the previous model on regulation of UCP3 transcription. We are convinced that further research on regulation of UCP3 expression has to focus on the first intron and propose several possible strategies to proceed and solve remaining and arising questions. Finally, we undertook first steps on this path and provide first hypotheses on mechanisms not fully understood to date.

## D. Appendix

### 6.1.1. Appendix 1: Chemicals

<b>Substance</b>	<b>Supplier</b>	<b>Cat. No.</b>
1,4-Dithiothreit	Carl Roth	6908
100 bp-DNA Ladder EQUALIZED	Carl Roth	T833
1kbp-DNA Ladder	Carl Roth	Y014
3,3',5-Triiodo-L-thyronine sodium salt	Sigma Aldrich	T6397
3-Isobutyl-1-methylxanthine	Sigma Aldrich	I5879
70% Ethanol, denatured	Carl Roth	T913
Aceton	Carl Roth	9372
Acrylamide:Bisacrylamide Solution, 30%, 37.5:1	Carl Roth	3029
Acrylamide:Bisacrylamide Solution, 40%, 37.5:1	Carl Roth	T802
Agar Agar	Carl Roth	5210
Agarose GTQ	Carl Roth	6352
Ampicillin	Carl Roth	K029
AMPS	Carl Roth	9592
ATP solution, 100 mM	Fermentas (Thermo)	#R0441
ATRA	Sigma Aldrich	R2625
Biotin	Carl Roth	3822
Blasticidin solution	InvivoGen	ant-bl-1
Boric Acid	Carl Roth	5935
Bromophenol Blue	Carl Roth	A512
BSA Fraction V	Carl Roth	8076
Calcium Chloride	Carl Roth	5291
CHAPS	Carl Roth	1479
Collagen A solution, 1 mg/ml	Biochrom	L7220
Coomassie G250	Carl Roth	3862
cyanogen bromide	Sigma Aldrich	C91492
Dexamethasone	Sigma Aldrich	D1159
di-Sodium hydrogen carbonate	Carl Roth	P030
di-sodium Hydrogene Phosphate	Carl Roth	P030
DMEM high glucose, liquid medium	Sigma Aldrich	D5796
DMEM:F12, powdered medium	Gibco	32500
DMSO	Carl Roth	4720
dNTP set: dATP, dTTP, dCTP, dGTP; 100 mM	Fermentas (Thermo)	#R0181
Donor Horse Serum	Gibco	#16050
EDTA	Carl Roth	8043
EGTA	Carl Roth	3054
Ethanol 96%, denatured	Carl Roth	T171
Ethanol 99.8% p.a	Carl Roth	9065
Fetal Bovine Serum	Biochrom	S0615
Formaldehyde 37% Solution	Carl Roth	4979
Forskolin	Cayman Chemical	11018
Genestein	Cayman Chemical	10005167
Gentamycin	Biochrom	A2712
Glacial acetic acid	Carl Roth	3738
Glucose	Carl Roth	HN06
Glutamine	Carl Roth	HN08
Glycerol	Carl Roth	3783
Glycin	Carl Roth	3908
GW0742	Cayman Chemical	10006798

**Chemicals continued (H-Y)**

<b>Substance</b>	<b>Supplier</b>	<b>Cat. No.</b>
Heparin, sodium salt	Carl Roth	4042
HEPES	Carl Roth	9105
Hexadimethrine bromide	Sigma Aldrich	107689
Hydrogen chloride Solution, 6M	Carl Roth	0281
Igepal CA630	Sigma Aldrich	18896
Indomethacine	Sigma Aldrich	17378
Insulin solution, human	Sigma Aldrich	I9278-5ML
Magnesium chloride hexahydrate	Carl Roth	2189
Magnesium sulfate heptahydrate	Carl Roth	P027
Methanol	Carl Roth	4627
Mithramycin	Cayman Chemical	11434
Nitrocellulose	926-31092	Li-Cor
NP40 substituent	Fluka	74385
Nuclease Free Water	Qiagen	129114
Oil Red O	Sigma Aldrich	O0625
Orange G	Carl Roth	0318
Penicilling Streptomycin premixed	Biochrom	A2212
Phenol:Chlorophorm:Isoamylalkohol	Carl Roth	A156
Phosphatase Inhibitor Cocktail	Sigma Aldrich	P5726-1ML
Phosphate Buffered Saline Tablets	Gibco	18912-014
PMSF	Carl Roth	6367
Poly(deoxyinosinic-deoxycytidylic) acid sodium salt	Sigma Aldrich	P4929
Potassium chloride	Carl Roth	6781
Protease Inhibitor Cocktail	Sigma Aldrich	P8340-1ml
Puromycin	InvivoGen	ant-pr-1
retinal	Sigma Aldrich	R2500
Rosiglitazone	Cayman Chemical	71740
RotiSafe	Carl Roth	3865
Saccharose	Carl Roth	4621
Salmon Sperm DNA	Invitrogen	15.632-011
Silver nitrate solution 1,7% (0.1M)	Merck	1.090.811.000
sodium acetate	Carl Roth	6779
Sodium Carbonate	Carl Roth	A135
Sodium Chloride	Carl Roth	3957
Sodium Deoxycholate	Carl Roth	3484
Sodium Dodecyl Sulfate	Carl Roth	4360
Sodium Fluoride	Sigma Aldrich	S6776
Sodium Hydroxide	Carl Roth	6771
Sodium hydroxide solution, 4mol/l	Carl Roth	T198
Sodium Panthotenate	Carl Roth	3812
Sodium Thiosulfate	Carl Roth	HN08
Spectomycin	Sigma Aldrich	S9007
TEMED	Carl Roth	2367
TRIS	Carl Roth	4855
TRIS-hydrochloride	Carl Roth	9090
TRIsure	Bioline	BIO-38033
Triton X100	Carl Roth	3051
Trypsin/EDTA solution	BioChrom	L2163
Tryptone	Carl Roth	8952
Tween 20	Carl Roth	9127
Whatman paper, 0.34 mm thickness	Carl Roth	4928
Wy14643	Calbiochem	681725
Yeast extract	Carl Roth	2363

## 6.1.2. Appendix 2: Plastic ware

Plastic ware	Supplier	Cat. No.
96 well plates, transparent, for Bradford/BCA	ZEFA	260895
384 well plates, for qPCR	4titude	4ti-0382
96 well plates, white, for luciferase assay	Greiner	655904
adhesive clear sheets, for qPCR	4titude	4ti-0560
cell strainer, 40 µM	BD Biosciences	352340
cellculture dishes 100x20mm	Sarstedt	83.1802
cellculture dishes 150x20mm	Sarstedt	83.1803
cellculture dishes 35x10mm	Sarstedt	83.1800
cellculture dishes 60x15mm	Sarstedt	83.1801
cellculture plates, 12 well	Biochrom/TPP	92412
cellculture plates, 24 well	Biochrom/TPP	83.1836
cellculture plates, 48 well	BD Biosciences	353230
cellculture plates, 6well	Biochrom/TPP	92406
cellculture plates, 96 well	Sarstedt	831.835
gloves, latex, large	ZEFA	GRIP-L
gloves, nitril, large	VWR	112-2756
multipette tips, 1000 µl	Eppendorf	0030 089.430
multipette tips, 200 µl	Eppendorf	0030 089.413
multipette tips, 2500 µl	Eppendorf	0030 089.448
petri dishes, sterile	Sarstedt	821.473.001
pipette tips, boxed, 12.5 µl, for matrix pipette	Thermo Scientific	#7421
pipette tips, boxed, sterile, 10 µl	Sarstedt	701.130.210
pipette tips, boxed, sterile, 100 µl	Sarstedt	70.760.212
pipette tips, boxed, sterile, 1000 µl	Sarstedt	70.762.211
pipette tips, boxed, sterile, 2.5 µl	Sarstedt	70.1130.212
pipette tips, boxed, sterile, 200 µl	Sarstedt	70.760.211
pipette tips, loose, 10 µl	Sarstedt	70 1130
pipette tips, loose, 1000 µl	Sarstedt	70.762
pipette tips, loose, 200 µl	Sarstedt	70.760.002
pipette tips, loose, 5000 µl	Eppendorf	0030 000.978
reaction tubes for PCR, 200 µl	Sarstedt	72.991.992
reaction tubes for screw caps, 2 ml	Sarstedt	72.609
reaction tubes, 1.5 ml	Sarstedt	72.690.001
reaction tubes, 1.5 ml, black	Gilson	1158365
reaction tubes, 15 ml	Sarstedt	62.554.502
reaction tubes, 2 ml	Sarstedt	72.691
reaction tubes, 5 ml	Sarstedt	62.558.201
reaction tubes, 50 ml	Sarstedt	62.547.254
screw caps for 2 ml and 1.5 ml reaction tubes	Sarstedt	65.716.999
sensors for Scepter counter, 60 µm	Millipore	PHCC60050
serological pipettes, sterile, 10 ml	Corning/Costar	4488
serological pipettes, sterile, 25 ml	Corning/Costar	4489
serological pipettes, sterile, 5 ml	Corning/Costar	4487
serological pipettes, sterile, 50 ml	Corning/Costar	4490
syringe filters, sterile, CA, 25mm/0.2 µm	ZEFA	CA2025RS
syringe filters, sterile, CA, 25mm/0.45 µm	ZEFA	CA4525RS

### 6.1.3. Appendix 3: Enzymes

Enzyme	Supplier	Cat. No.
<b>Restriction Endonucleases</b>		
AarI	Fermentas (Thermo Scientific)	# ER1581
BamHI	Fermentas (Thermo Scientific)	# ER0051
BglII	Fermentas (Thermo Scientific)	# ER0081
Esp3I	Fermentas (Thermo Scientific)	# ER0451
BsaI-HF	New England Biolabs	R3535S
EcoRI	Fermentas (Thermo Scientific)	# ER0271
EcoRV	New England Biolabs	R0195S
Esp3I	Fermentas (Thermo Scientific)	# ER0451
HindIII	Fermentas (Thermo Scientific)	# ER0501
MluI	Fermentas (Thermo Scientific)	# ER0561
NcoI	Fermentas (Thermo Scientific)	# ER0571
NotI	Fermentas (Thermo Scientific)	# ER0591
SalI	Fermentas (Thermo Scientific)	# ER0641
XbaI	Fermentas (Thermo Scientific)	# ER0681
XhoI	Fermentas (Thermo Scientific)	# ER0691
<b>Other enzymes</b>		
FastAP	Fermentas (Thermo Scientific)	# EF0654
Herculase II	Agilent Technologies	600675
Proteinase K	Fermentas (Thermo Scientific)	# EO0492
Q5 Polymerase	New England Biolabs	M0491S
T4 DNA Ligase	Fermentas (Thermo Scientific)	# EL0014
T4 DNA Polymerase	Fermentas (Thermo Scientific)	# EP0061
T4 Polynucleotide Kinase	Fermentas (Thermo Scientific)	# EK0031
<b>Enzyme Mixes and Enzyme Kits</b>		
QuantiTect RT Kit	Qiagen	205313
QuikChange II Kit	Agilent Technologies	200523
RNAseA/T1 mix	Fermentas (Thermo Scientific)	# EN0551
SensiMix Sybr no Rox	BioLine	QT650-20
<b>buffers and buffer additional were purchased along with the enzyme</b>		

### 6.1.4. Appendix 4: Cell lines and organisms

Organism	Supplier	Cat. No.
NEB5 $\alpha$ competent E. coli	NEB	C2988J
Platinum E cells	Cell Biolabs	RV-101
C2C12 cells	DSMZ	ACC 565
HEK293 cells	DSMZ	ACC 305
H1B1b cells	provided by Bruce Spiegelman	
iBPA-L1 cells	provided by Patrick Seale	
iBPA-L2 cells	provided by Alexander Pfeiffer	

### 6.1.5. Appendix 5: Antibodies

Primary antibody	Supplier	Cat. No.
Anti-Actin, cloneC4	Millipore	MAB1501
Cox IV	Cell Signaling (NEB)	#4844
E-Cadherin	Santa Cruz Biotechnology	sc-7870
Histone H3 (D2D12)	Cell Signaling (NEB)	#4620
pan-actin	Chemicon (Millipore)	MAB1501
PPAR $\gamma$ (H-100)	Santa Cruz Biotechnology	sc-7196
RXR (N 197)	Santa Cruz	sc-774
Sp1 #1	Merck Millipore	07-645
Sp1 #2 (PEP2)	Santa Cruz	sc-59x
Sp3 (H-225) X	Santa Cruz Biotechnology	sc-13018 X
Ty1 monoclonal - classic	Diagenode	MAB-054-050
UCP3	Pierce Antibodies	PA1-055
normal rabbit IgG	Cell Signaling (NEB)	#2729
Secondary antibody	Supplier	Cat. No.
rabbit anti goat IRDye 680	Li-Cor	926-32221
goat anti rabbit IRDye 800	Li-Cor	926-32211
goat anti rabbit IRDye 680	Li-Cor	926-32222
donkey anti mouse IRDye 800	Li-Cor	926-32212
donkey anti goat IRDye 800	Li-Cor	926-32214

### 6.1.6. Appendix 6: Kit Systems

Kit	Supplier	Cat. No.
Agilent RNA 6000 Nano Kit	Agilent Technologies	5067-1511
Amaxa Cell Line Nucleofector Kit V	Amaxa (Lonza)	VCA-1003
Dual-Luciferase Reporter Assay System	Promega	E1960
HighCell# CHIP Kit, protein G	Diagenode	kch-maghigh-G16
Lipofectamin LTX with PlusReagent	Invitrogen (LifeTechnologies)	15338100
LowCell# CHIP Kit, protein A	Diagenode	kch-maglow-A16
LowCell# CHIP Kit, protein G	Diagenode	kch-maglow-G16
Pierce BCA Protein Assay Kit	Pierce (Thermo Scientific)	PI-23225
PureYield Plasmid Midiprep System	Promega	A2495
PureYield Plasmid Miniprep System	Promega	A1222
RotiQuant Bradford Assay Kit	Carl Roth	K015.1
SE Cell Line 96-well Nucleofector	Amaxa (Lonza)	V4SC-1096
SimpleCHIP enzymatic Kit	Cell Signalling (NEB)	#9003S
SV Total RNA Isolation System	Promega	Z3105
Wizard SV Gel and PCR Clean-Up System	Promega	A9282



### 6.1.7. Appendix 7: Oligonucleotide sequences

#### qPCR primers for measuring transcript abundance

transcript	forward primer	reverse primer
CDX1	GCTAACCTGGGGCTCACA	GGGCATAGACTTCCTAGGGG
CDX2	CATCACCATCAGGAGGAAAAG	GCAAGGAGGTCACAGGACTC
CDX4	CAGTTAACCTGGGCCTTTCT	TCCGTTTCTCTGCTCCTGTT
COUP-TFII	GCCATAGTCCTGTTACCTCA	ATATCCCGGATGAGGGTTTC
CTCT	AACCAGCCCAAACAGAACC	ACAGCCCAATAGTCCTGGTG
GusB	ACTATGGGCATTTGGAGGTG	ACTCCTCACTGAACATGCCA
HPRT	CAGGCCAGACTTTGTTGGAT	TTGCGCTCATCTTAGGCTTT
MyoD	GAAGTCGTCTGCTGTCTCAAAGG	CGCCACTCCGGGACATAG
Myogenin	CCTAAAGTGGAGATCCTGCG	ACGATGGACGTAAGGGAGTG
p300	GGGAAGAGAACACCAGCAAC	ATCAGGTCACAGGGGATGAG
PPAR $\alpha$	TCCCTGTTTGTGGCTGCTAT	TGCCATCTCAGGAAAGATCA
PPAR $\gamma$ (1+2)	TCAGCTCTGTGGACCTCTCC	ACCCTTGCATCCTTCAACAG
PPAR $\delta$	CGAGTTCTTGCGAAGTCTCC	TGTCCTGGATGGCTTCTACC
SP1	GGAGAGGCCATTCATGTGTAA	GGTGGTAATAAGGGCTGAAGG
SP2	TGCAACTGGTTCTTCTGTGG	AGAGCCTTTGGAAGGAGGAG
SP3	CCAATCTTGGGAAAAAGAAGC	CATCTCTTCCAGCTTCCACAG
TFIIB	TGGAGATTTGTCCACCATGA	GAATTGCCAAACTCATCAAAACT
UCP1	TCTCTGCCAGGACAGTACCC	AGAAGCCCAATGATGTTTCAG
UCP2	ACTGGGCAGAGGATGAAGAA	ACTGTGCCCTTACCATGCTC
UCP3	AAGATGGTGGCTCAGGAGG	GGACGAAACACGGAGGACTA
$\beta$ -actin	AGAGGGAAATCGTGCCTGAC	CAATAGTGATGACCTGGCCGT

#### qPCR primers for CHIP

region	forward primer	reverse primer
UCP3 enhancer PP1	CCTGAGCAAATGGAGAGCAT	ACAGTGAGACAAGCCTGAGGA
UCP3 enhancer PP2	AAGCCTTGCTGCTCTTCTGT	GTGAGACAAGCCTGAGGAGG
UCP3 last exon	GAGAGCAGGAGGAAGTGTGG	CACCTTAGGGCAAGAACGAG
UCP2 last exon	TCCTAGTTCGCCAACCTCTCTAC	TCAACCCCTTATTACAGACCT
EF1a enhancer	CAAGGAGCTCAAAATGGAGG	TCAAAGCCCGAGGTGACTA
Resistin promoter	ACCTCTCTTGGGGTCAGATGT	CTGGGTATTAGCTCCTGTCCC
DHFR promoter	CCTTAGCCCTACCCACCATT	GTTCTGTCTGGGACGGGG
MC4R 50kbp upstream	TGTGCTCAAAGCAGGAATTG	GTCCTGCATTTTCCCCTGTA
MC4R 50kbp downstream	GTAAAGCACACTGCCCGATT	TCTTTCTTACCTGGTTGCC

#### Sequencing primers

Region of interest	Sequence
insert in pJET1.2 (for)	CGACTCACTATAGGGAGAGCGGC
insert in pJET1.2 (rev)	AAGAACATCGATTTTCCATGGCAG
insert in pMXs EF1a (for from EF1)	GAGTTTGGATCTTGGTTCAATTCTC
insert in pMXs EF1a (rev from BGH)	TAGAAGGCACAGTCGAGG
insert in pMXs IRES Puro (for)	CTTACACAGTCCTGCTGACCAC
insert in pMXs IRES Puro (rev)	ACATATAGACAAACGCACACCG
inside emGFP primer (for)	GACAACCACTACCTGAGCAC
Luciferase cassette in pGL3 (rev)	GGTTTGTCCAAACTCATCAATG
miRNAs in pcDNA6.2 emGFP miR	GACAACCACTACCTGAGCAC
pGL3 reporter gene constructs (for)	CTAGCAAAATAGGCTGTCCC
shRNAs in pTer	ACTAGAAGGCACAGTCGAGG
UCP3 intronic enhancer (for)	CCTAGCTCTCCAGGCAAATC
UCP3 intronic enhancer (rev)	TCTCCAGGCTCTTCTTTCAG

**shRNA sequences for transient knockdown, top (+) strand according to sequencing**

<b>Pool scrambled shRNAs</b>	
shScram-1	GAGGAGTTCGTGTTACTAATTCAAGAGATTAGTAACACGAACTCCTCTTTTT
shScram-2	GACCTACTATATGTAACTATTCAAGAGATAGTTACATATAGTAGGTCTTTTT
shScram-3	GCATGGACACGTCTAATAATTCAAGAGATTATTAGACGTGTCCATGCTTTTT
shScram-4	GACACCTTGAGCTTATAGATTCAAGAGATCTATAAGCTCAAGGTGTCTTTTT
<b>Pool shRNAs targeting SP1</b>	
shSp1-1	GGATGGATCTGGTCAAATATTCAAGAGATATTTGACCAGAACCATCCTTTTT
shSp1-2	CATCATGCCTTGATAAATATTCAAGAGATATTTATCAAGGCATGATGTTTT
shSp1-3	GATCACTCCATGGATGAAATTCAAGAGATTTTCATCCATGGAGTGATCTTTTT
shSp1-4	GACTCAGTATGTGACCAATTTCAAGAGAATTGGTCACATACTGAGTCTTTTT
<b>Pool shRNAs targeting SP3</b>	
shSp3-1	ATTATATCCAGTCGCCTGTTTCAAGAGAACAGGCGACTGGATATAATTTTT
shSp3-2	GCTATGGATAGTTCAGACATTCAAGAGATGTCTGAACTATCCATAGCTTTTT
shSp3-3	GCAACACATTTGTCATATATTCAAGAGATATATGACAAATGTGTTGCTTTTT
shSp3-4	GCCAGTGGTCAAATATATTTCAAGAGAATATATTTTGACCACTGGCTTTTT

**miRNA sequences for transient knockdown, top (+) strand according to sequencing**

Sp1 miR1	TGCTGTTGAGCAGCATTACAGTGACGTTTTGGCCACTGACTGACGTCATGTATGCTGCTCAA
Sp1 miR2	TGCTGTCAAGCATCAGACTATACTTCGTTTTGGCCACTGACTGACGAAGTATACTGATGCTTGA
Sp3 miR1	TGCTGTTTGATTAGAGCCAGGAATGAGTTTTGGCCACTGACTGACTCATTCTCTAATCAAAA
Sp3 miR2	TGCTGTAATCAAGGCATCATCTCTTCGTTTTGGCCACTGACTGACGAAGAGATTGCCTTGATTA
shBle miR1	TGCTGTGATGAACAGGGTCACGTCGTGTTTTGGCCACTGACTGACACGACGTGCCTGTTTCATCA
LacZ miR1	TGCTGAAATGTACTGCGCGTGGAGACGTTTTGGCCACTGACTGACGTCCTCACGCAGTACATTT
UCP1 miR1	TGCTGAAGAGAAGTACTCTTGACTGGTTTTGGCCACTGACTGACCAGTCCAAGTACTTCTCTT
UCP1 miR2	TGCTGTTATGTGGTACAATCCACTGTGTTTTGGCCACTGACTGACACAGTGGAGTACCACATAA

<b>Virus construct</b>	<b>combination of miRNAs</b>
Sp1 single KD	SP1 miR1, SP1 miR2
Sp3 single KD	SP3 miR1, SP3 miR2
Sp1/3 double KD A	SP1 miR1, SP3 miR1
Sp1/3 double KD B	SP1 miR2, SP3 miR1
Ctrl1	UCP1 miR1, UCP1 miR2
Ctrl2	LacZ miR1, shBle miR1

**Primers used for site directed mutagenesis**

<b>Deletion</b>	<b>Sequence*</b>
promoter DR1 site	GTCAACTAGCTTCTCAGAATTGATATCGCTGGTGCCTAAGGCC
intronic DR1 site	GACCTGGCTCCCTTCTTCTGTAACCTCGAGAACTAAGGCCCTGAATAAGTG
NF1/STAT site	GGCAGGGAATCCAGTACCTCCTGCTCTAGAGGAGCCAGACCTG
IVS1+1504 A allele	CTAAGGCCCTGAATAAGTGTTCCTTAACAAACCTGCACTGTTGGTA

**\*for every primer, a second, reverse-complement primer was ordered in addition**

**oligonucleotides for annealing of the Ty1-tag**

2x Ty1-tag, top	GATCGCCACCATGGGAGAGGTGCACACCAACCAGGACCCCTGGACG CCGAGGTGCACACCAACCAGGACCCCTGGACGCCGGATCCCAAGC
2x Ty1-tag, bot	GGCCGCTTGGGATCCGGCGTCCAGGGGGTCTGGTTGGTGTGCACCT CGGCGTCCAGGGGGTCTGGTTGGTGTGCACCTCTCCCATGGTGGC

### Primers used for generation of reporter gene deletion constructs

Deletion	IVS1 base*	primers used
Δ1(for)	70 - 389	TGTGAGCCTCCATGAAAAGGGAG
Δ1 (rev)		TTGCAAAGGTCTGCTGCCCA
Δ2(for)	431 - 743	ACTGTTGCTTGAGGAAGCCTGG
Δ2 (rev)		GCCCTGAGGTCATGTCTGACCTTA
Δ4a(for)	1110 - 1415	CCAGAGCACTTTCTGCTGGAGG
Δ4a (rev)		CCAGTACCTCCTGCTGGGAAGG
Δ4b(for)	1110 - 1451	CCAGAGCACTTTCTGCTGGAGG
Δ4b (rev)		GCCCTTCTGTAGGGCAAAGGG
Δ5(for)	1459 - 1559	AGAAGTAGTATAGTCCCCTTTGCCCTACAGAAGG
Δ5 (rev)		AGAAGTAGTATTAACACGCCTGCACTGTTGGTAC
Δ6(for)	1559 - 1979	CTGGGTACCAACAGTGCAGGC
Δ6 (rev)		CCAAAGTGAGTTTGAATCTAAAAGAGGGCTG
Δ7(for)	1975 - 2207	GCAGGGTAGTTTTTAAATCTAGTTTACAGATGAGAAA
Δ7 (rev)		CAACCTATATCAGATGTCCTGAATATCAGATATTTGTATT
Δ8(for)	2278 - 2533	CACTGCAATTTTACCCTGTTATGAAGCAT
Δ8 (rev)		TGTCTGTAAGTACCAAGGTTCCGC
Δ9(for)	2533 - 2892	GGAGGAGCTGTCAGGAAGGGAC
Δ9 (rev)		GCCATGAGATGGATGAGAGTGATACTG
Δ10(for)	2891 - 3151	CAGATACCCCTCGACCCTGTCTC
Δ10 (rev)		CCCTGGAAATAGAACAGAGCCAGACA
ΔInt(for)	-2 - 3215	TGTGAGTCTAGCCAAGGTAGGGTATGC
ΔInt (rev)		CTGCCCCCGGAAGTGAAGTA

\*base location in accordance to the IVS1 nomenclature. Position 1 denotes the first base of Intron 1.

### cloning primers and oligonucleotides

amplicon	forward primer	reverse primer
GFP-miRNAs	AGAGGTCTCCGATCAGAGAACCCTACTGCTTACTGGC	AGAGGTCTCGTCGAGGGCCCTCTAGATCAACCCTTT
CDS of CREB	CGTCTCGGATCCATGACCATGGAATCTGGAGC	CGTCTCCGGCCTCAATCTGATTTGTGGCAGTAAAGG
CDS of SP1	CGTCTCGGATCCATGAGCGACCAAGATCACTCC	CGTCTCCGGCCAGATGTCTCTTGGACCCA
CDS of SP2	GGATCCATGAGCGCAGATCCACAGAT	GCGGCCGCCTCAGTTGGCCTTACAAGC
CDS of SP3	CGTCTCGGATCCATGACCGCTCCCGAAAAGC	CGTCTCCGGCCTTACTCCATTGTCTCATTTCAGA
Gluc cassette	GGTCTCCCATGGGAGTCAAAGTTCTGTTTGGCCCTG	GGTCTCTTAGATGCATGCTCGAGCGG
P2000 promoter	CGTCTCAGATCTTTTACAGGCCACCGACCCTCA	CGTCTCAGATCTGGTTTAGCTTCTGGACTTGG

### Primers, probes and competitors for DNA-affinity chromatography

Deletion	Modification	Sequence*
ssDNA trapping anchor	5' amino C6 modifier	ACACACACAC
T18 competitor		TTTTTTTTTTTTTTTTTTTT
primer enhancer trapping probe	5' Biotin modifier	GCAGGGAATCCAGTACCTCCT
primer enhancer trapping probe	5' Cy5 modifier	CTGGGTACCAACAGTGCAGG
probe for MagBeads	5' Biotin modifier	GTGTTTTCTTAAACACGCCTGCACTGTTGGTA
probe for MagBeads with spacer	5' Biotin modifier	A14-GTGTTTTCTTAAACACGCCTGCACTGTTGGTA
probe for Oligonucleotide	3' overhang (GT)x5	GTGTTTTCTTAAACACGCCTGCACTGTTGGTA

\*for every probe, a second, non-labelled, reverse-complement primer was ordered in addition

### Probes and competitors used in EMSA

Element/Probe	Modification	Sequence*
IVS1505G Cy5	5' Cy5 modifier	GTGTTTTCTTAACACGCCTGCACTGTTGGTA
IVS1505A Cy5	5' Cy5 modifier	GTGTTTTCTTAACACACCTGCACTGTTGGTA
SP1 consensus Cy5	5' Cy5 modifier	ATTTCGATCGGGGCGGGGCGAGC
IVS1505G		GTGTTTTCTTAACACGCCTGCACTGTTGGTA
IVS1505A		GTGTTTTCTTAACACACCTGCACTGTTGGTA
SP1 consensus		ATTTCGATCGGGGCGGGGCGAGC
IVS1504A		GTGTTTTCTTAACAAGCCTGCACTGTTGGTA
Pig putative GC Box 1		GTGCCACCTAGGGGCAGGGAGGGAGTTCAG
Pig putative GC Box 2		AGGGAGTTCAGTCCCTCCACCAAGGCTGACC
Human putative GC Box		TGAAGTTCCTGGGGGCAGGCACAGCAGCCTG
Mouse putative GC Box		CATGTCTCTAAACATGCCTACCCTGCTCTTC
Rat putative GC Box		TTCTCTAAACACGCCCACACCGCTA
AP1		TTGTCCTAGTCAGCCAGCTGTG
C/ebp		GTGTGTATTATGCAGGACTGCACTGTTGGTA
Cdx		GCATTTTATTACCACGCCTGCACTGTTGGTA
Cmyb		GTGTTTTCAGCAACCGCCTGCACTGTTGGTA
CTCF		TGCTTGAGTGCCCTCTGGTGGGCAATAGGA
ETF		GTGTTTTCACATTCCCTCCGGCACTGTTGGTA
Msx		GTGTTTTTCTAATTGGCCTGCACTGTTGGTA
MyoD		CCCCAACAGCTGTTGCCTGA
NFκB		TTGGCTGTGGTACTTTCCAACGGAA
p53		GTGTTTCAGGACATGTCCAGGCATGTCTCTA
Pbx		GTGGATGATTGACAGGTCTGCACTGTTGGTA
Sry		GTGTTTTGTAAACAATAGTGCCTGTTGGTA
Tbp		GTGTTTTGGTATAAATCCTGCACTGTTGGTA

\*for every sequence, a second, non-labelled, reverse-complement primer was ordered in addition

## 6.1.8. Appendix 8: Machines and devices

<b>Detection and Measurement Equipment</b>		
AxioVert 40 Microscope	Zeiss	451207-0000-000
AxioVision 4.8 Materialpaket Standard, AxioCam ICc1	Zeiss	410130-0505-001
DMI4000 Fluorescence Microscope	Leica	specific quote
Infinite M200 Microplate reader	Tecan	30016056
modifications: 30015897, 30033861, 30034835, 30015904, 30015883, 30015892, 30015893, 30015894		
LI-COR Odyssey Infrared Imaging System	Li-Cor	Ody-2197
LightCycler 480 (384 well)	Roche Applied Science	5015243001
MyCycler ThermalCycler (96 well)/ Gradient Upgrade	BioRad	170-9701/-9707
NanoQuant Plate	Tecan	30034027
PTC200 PCR-Cycler (2x48 well)	MJResearch	discontinued
Spectronic Genesys 10 Bio	Thermo Scientific	840-208200
Typhoon TRIO+	GE Healthcare	63-0055-86
<b>Cell culture equipment</b>		
BioWizard SilverLine SL-170	Kojair Tech Oy	SL-170
HeraCell 240	Heraeus (Thermo Scientific)	51026331
Nucleofektor 96-well Shuttle System	Amaxa (Lonza)	AAM-1001S
Nucleofektor II	Amaxa (Lonza)	AAB-1001
Scepter handheld coulter counter	Millipore	PHCC20060
Wasserbad WNB45	Memmert	WNB45
<b>Basic equipment</b>		
AquaLine AL12	Lauda	LCB 0725
BioRuptor Plus	Diagenode	UCD-300 TS
CyberScan pH510	EuTech Instruments	EC-PH510/21S
Eppendorf Research plus 10-100, 8-Kanal	Eppendorf	3122000035
Eppendorf Research plus 30-300, 8-Kanal	Eppendorf	3122000051
Eppendorf Research plus 3-pack, 1-Kanal	Eppendorf	3120000909
Kelvitron T	Heraeus (Thermo Scientific)	51015264
Magnetic stirrer RCT basic	IKA	0003810000
Matix Electronic 384 Equalizer Pipette	Thermo Scientific	2139-11
Multipette Stream	Eppendorf	4986000017
Polymax1040	Heidolph	543-42205-00
Reax2 head-over-end mixer	Heidolph	541-21001-00
Sarpette Automatic	Sarstedt (Thermo Scientific)	NC9783498
Thermomixer Comfort	Eppendorf	5355000011
Unimax 2010	Heidolph	542-10020-00
Vibrax VXR basic	IKA	0002819000
Vortexer MS3 basic	IKA	0003617000
WiseCube	Witeg	Wis20
<b>Centrifuges</b>		
Eppendorf 5417R	Eppendorf	22621807
Eppendorf 5417C	Eppendorf	22621807
Eppendorf 5804R	Eppendorf	5805000327
Sorvall Evolution RC	Sorvall (Thermo Scientific)	728611
Mini centrifuge MCF2360	LMS Co. LTD	5944400
<b>Electrophoresis Equipment</b>		
Gel iX Imager Geldokumentationssystem	Intas	specific quote
Mini-PROTEAN Tetra Cell	BioRad	165-8000
MultiSUB Choice Electrophoresis Tank	Cleaver Scientific	MSChoice
MultiSUB Mini Electrophoresis Tank	Cleaver Scientific	MSMini
PerfectBlue Doppelgelsystem Twin ExW S	PeqLab	45-2010-C
PerfectBlue Doppelgelsystem Twin L	PeqLab	45-2020-C
PowerPac basic	BioRad	164-5050
Trans-Blot SD Semi-Dry Transfer Cell	BioRad	170-3940

### 6.1.9. Appendix 9: Bioinformatic resources

<b>A)</b> model	matrix 1	distance	matrix 2	distance	matrix 3	distance	matrix 4
SMod1505	U\$1505G	0-20 nt	V\$SP1F				
MyoD-NF1-PERO-SP1	V\$MyoD	10-90 nt	V\$NF1F	10-70 nt	V\$PERO	10-70 nt	SMod1505
NF1-MyoD-PERO-SP1	V\$NF1F	10-70 nt	V\$MyoD	10-70 nt	V\$PERO	10-70 nt	SMod1505
PERO-SP1-MyoD-NF1	V\$PERO	10-70 nt	SMod1505	10-70 nt	V\$MyoD	10-70 nt	V\$NF1F
PERO-SP1-NF1-MyoD	V\$PERO	10-70 nt	SMod1505	10-70 nt	V\$NF1F	10-90 nt	V\$MyoD
MyoD-PERO-SP1-NF1	V\$MyoD	10-90 nt	V\$PERO	10-70 nt	SMod1505	10-70 nt	V\$NF1F
NF1-PERO-SP1-MyoD	V\$NF1F	10-70 nt	V\$PERO	10-70 nt	SMod1505	10-90 nt	V\$MyoD
PERO-SP1	V\$PERO	10-70 nt	SMod1505				

#### B)

condition	strand	core sim.	matrix sim.
strict	both	>0.75	optimized
relaxed	both	>0.70	>opt.-0.05

#### C)

V\$MyoD	Genomatix:	Myoblast determining factors
V\$NF1F	Genomatix:	Nuclear factor 1
V\$PERO	Genomatix:	PPAR/DR1 elements
V\$SP1F	Genomatix:	GC-Box factors SP1/GC
V\$STAT	Genomatix:	Signal transducer/activator of transcription
U\$1505G	User matrix:	caCGCCymcbg

**Table 3: Criteria and models used for bioinformatics.** **A)** Matrix families and distances between matrices for all models used. The Sub-model (SMod) was integrated into the larger model with matrix occurrence of 50%. It thus was sufficient if only one of the matrices defined in a submodel was detected. For all models incorporating V\$NF1F an alternative version was generated by replacing V\$NF1F by the V\$STAT. **B)** Strict and relaxed criteria used for matrix identification. **Strand:** On which strand the element may be located. **Core sim.:** Required sequence match to the matrix core (1=100%). **Matrix sim.:** Fraction of the whole matrix that needs to match. Optimised (opt.) denotes the recommendation by Genomatix. **C)** Name, origin and description of all matrices used. **Genomatix:** Matrix families defined in the standard Genomatix library. **User matrix:** Matrices generated using MatDefine. Capital letters in user matrices indicate the matrix core.

<b>Matrix families</b>	
U\$IVS1505G	no family, user defined matrix
V\$MYOD	Myoblast determining factors
V\$NF1F	Nuclear factor 1
V\$PERO	Peroxisome proliferator-activated receptor
V\$SP1F	GC-Box factors SP1/GC
V\$STAT	Signal transducer and activator of transcription

<b>Transcription factor matrices</b>		<b>Family</b>
U\$IVS1505G	GC/GT Box based on Phodopus IVS1+1505 element	user matrix
V\$GC.01	GC box elements	V\$SP1F
V\$MYF6.01	Myogenic factor 6 (herculin)	V\$MYOD
V\$MYOD.01	Myogenic regulatory factor MyoD (myf3)	V\$MYOD
V\$MYOGENIN.02	Myogenic bHLH protein myogenin (myf4)	V\$MYOD
V\$NF1.01	Nuclear factor 1	V\$NF1F
V\$NF1.02	Nuclear factor 1 (CTF1)	V\$NF1F
V\$NF1.03	Non-palindromic nuclear factor I binding sites	V\$NF1F
V\$NF1.04	Nuclear factor 1	V\$NF1F
V\$PPAR_RXR.01	PPAR/RXR heterodimers, DR1 sites	V\$PERO
V\$PPAR_RXR.02	PPAR/RXR heterodimers, DR1 sites	V\$PERO
V\$PPARG.02	Peroxisome proliferator-activated receptor gamma	V\$PERO
V\$PPARG.03	Peroxisome proliferator-activated receptor gamma, DR1 sites	V\$PERO
V\$SGN1.01	bHLH factor Sgn-1 (Salivary Glands 1), achaete-scute complex homolog 3 (Ascl3)	V\$MYOD
V\$SP1.02	Stimulating protein 1, ubiquitous zinc finger transcription factor	V\$SP1F
V\$SP1.03	Stimulating protein 1, ubiquitous zinc finger transcription factor	V\$SP1F
V\$SP2.01	Sp2, member of the Sp/XKLF transcription factors with three C2H2 zinc fingers	V\$SP1F
V\$SP4.01	Sp4 transcription factor	V\$SP1F
V\$SP4.02	Sp4 transcription factor (secondary DNA binding preference)	V\$SP1F
V\$STAT.01	Signal transducers and activators of transcription	V\$STAT
V\$STAT1.01	Signal transducer and activator of transcription 1	V\$STAT
V\$STAT1.02	Signal transducer and activator of transcription 1	V\$STAT
V\$STAT3.02	Signal transducer and activator of transcription 3	V\$STAT
V\$STAT5.01	STAT5: signal transducer and activator of transcription 5	V\$STAT
V\$STAT5A.01	Signal transducer and activator of transcription 5A	V\$STAT
V\$STAT5B.01	Signal transducer and activator of transcription 5B	V\$STAT
V\$STAT6.01	STAT6: signal transducer and activator of transcription 6	V\$STAT
V\$TCF12.01	Helix-loop-helix transcription factor 12	V\$MYOD
V\$TCFE2A.02	Transcription factor E2a (E12/E47) (secondary DNA binding preference)	V\$MYOD
V\$TCFE2A.03	Transcription factor E2a (E12/E47)	V\$MYOD
V\$TIEG.01	TGFβ-inducible early gene (TIEG) / Early growth response gene alpha (EGRalpha)	V\$SP1F

**Table 4: Description of the binding sites identified by MatInspector. Top:** Genomatix transcription factor families: Code and full names. **Bottom:** Individual binding matrices: Code, full name, and parental matrix family to which the individual matrix belongs.

Family	Matrix	Element location			Matix sim.	Family	Matrix	Element location			Matix sim.
		start	stop	strand				start	stop	strand	
<b>Equus caballus region 1 - Intron 1</b>					<b>Mus musculus region 1 - Intron 1</b>						
V\$MYOD	V\$MYOGENIN.02	1695	1711	(-)	0.910	V\$MYOD	V\$TCFE2A.03	1833	1849	(+)	0.963
V\$NF1F	V\$NF1.04	1721	1741	(-)	0.858*	V\$MYOD	V\$TCF12.01	1834	1850	(-)	0.956*
V\$SP1F	V\$SP1.02	1767	1783	(-)	0.851	V\$STAT	V\$STAT3.02	1854	1872	(+)	0.956
V\$PERO	V\$PPAR_RXR.01	1819	1841	(-)	0.764	V\$STAT	V\$STAT5A.01	1885	1903	(-)	0.855
V\$MYOD	V\$MYOD.01	1852	1868	(+)	0.974	V\$STAT	V\$STAT.01	1887	1905	(+)	0.858*
V\$MYOD	V\$TCFE2A.03	1853	1869	(-)	0.993	V\$NF1F	V\$NF1.01	1890	1910	(-)	0.929
V\$NF1F	V\$NF1.03	1909	1929	(-)	0.877*	V\$NF1F	V\$NF1.04	1890	1910	(+)	0.888
<b>Equus caballus region 2 - Intron 1</b>					<b>Phodopus sungorus region 1 - Intron 1</b>						
V\$NF1F	V\$NF1.03	1909	1929	(-)	0.877*	V\$PERO	V\$PPARG.02	1919	1941	(+)	0.948
V\$STAT	V\$STAT3.02	1961	1979	(-)	0.941	U\$IVS1505G	U\$IVS1505G	1974	1984	(+)	0.890
V\$NF1F	V\$NF1.02	1989	2009	(+)	0.767*	<b>Rattus norvegicus region 1 - Intron 1</b>					
V\$NF1F	V\$NF1.03	1989	2009	(-)	0.898*	V\$MYOD	V\$TCFE2A.03	1489	1505	(+)	0.965
V\$PERO	V\$PPARG.03	2014	2036	(+)	0.832	V\$MYOD	V\$TCF12.01	1490	1506	(-)	0.962*
V\$SP1F	V\$SP2.01	2064	2080	(-)	0.837	V\$STAT	V\$STAT5A.01	1540	1558	(-)	0.855
<b>Equus caballus region 3 - Intron 1</b>					<b>V\$STAT</b>						
V\$STAT	V\$STAT6.01	2312	2330	(-)	0.927	V\$NF1F	V\$NF1.01	1542	1560	(+)	0.858*
V\$STAT	V\$STAT6.01	2313	2331	(+)	0.886	V\$NF1F	V\$NF1.04	1545	1565	(-)	0.849
V\$PERO	V\$PPARG.02	2351	2373	(+)	0.887	V\$NF1F	V\$NF1.04	1545	1565	(+)	0.868*
V\$MYOD	V\$TCFE2A.03	2380	2396	(+)	0.970	V\$PERO	V\$PPARG.02	1574	1596	(+)	0.949
V\$SP1F	V\$TIEG.01	2393	2409	(-)	0.879	V\$SP1F	V\$SP4.02	1616	1632	(-)	0.871*
V\$NF1F	V\$NF1.03	2459	2479	(-)	0.917*	U\$IVS1505G	U\$IVS1505G	1611	1621	(+)	0.883
V\$MYOD	V\$TCFE2A.03	2460	2476	(+)	0.976	<b>Rattus norvegicus region 1 - Intron 1</b>					
V\$MYOD	V\$MYF5.01	2464	2480	(-)	0.918	V\$MYOD	V\$TCF12.01	1621	1637	(+)	0.931*
<b>Equus caballus region 4 - Intron 1</b>					<b>V\$MYOD</b>						
V\$STAT	V\$STAT5B.01	2593	2611	(+)	0.919*	V\$STAT	V\$MYF6.01	1622	1638	(-)	0.973
V\$PERO	V\$PPARG.02	2660	2682	(+)	0.890	V\$STAT	V\$STAT3.02	1642	1660	(+)	0.943
V\$SP1F	V\$SP4.01	2719	2735	(-)	0.878	V\$STAT	V\$STAT5A.01	1673	1691	(-)	0.858
V\$NF1F	V\$NF1.03	2740	2760	(-)	0.891*	V\$STAT	V\$STAT.01	1675	1693	(+)	0.858*
V\$MYOD	V\$TCFE2A.02	2779	2795	(+)	0.941	V\$NF1F	V\$NF1.01	1678	1698	(-)	0.846
<b>Sus scrofa region 1 - Exon 1</b>					<b>V\$NF1F</b>						
V\$STAT	V\$STAT6.01	72	90	(-)	0.796*	V\$NF1F	V\$NF1.04	1678	1698	(+)	0.863*
V\$NF1F	V\$NF1.01	74	94	(+)	0.796*	V\$PERO	V\$PPARG.02	1707	1729	(+)	0.949
V\$SP1F	V\$SP2.01	94	110	(-)	0.822	V\$SP1F	V\$SP4.01	1757	1773	(-)	0.885
V\$PERO	V\$PPAR_RXR.01	121	143	(-)	0.833	U\$IVS1505G	U\$IVS1505G	1974	1984	(+)	0.890
V\$MYOD	V\$MYOGENIN.02	172	188	(+)	0.905	<b>Homo sapiens region 1 - Exon 1</b>					
V\$MYOD	V\$MYOD.01	173	189	(-)	0.940	V\$NF1F	V\$NF1.04	321	341	(+)	0.882
<b>Sus scrofa region 2 - Intron 1</b>					<b>V\$MYOD</b>						
V\$STAT	V\$STAT5A.01	1990	2008	(-)	0.863	V\$SP1F	V\$SP2.01	439	455	(+)	0.824
V\$STAT	V\$STAT1.02	1992	2010	(+)	0.862	V\$PERO	V\$PPARG.02	461	483	(-)	0.889
V\$SP1F	V\$SP2.01	2029	2045	(+)	0.884	<b>Homo sapiens region 2 - Exon 2/Intron 2</b>					
V\$PERO	V\$PPARG.03	2048	2070	(-)	0.877	V\$STAT	V\$STAT3.02	2380	2398	(+)	0.968
V\$STAT	V\$STAT1.02	2072	2090	(-)	0.915	V\$SP1F	V\$SP1.01	2444	2460	(+)	0.891
V\$STAT	V\$STAT5.01	2074	2092	(+)	0.850*	V\$PERO	V\$PPARG.03	2479	2501	(-)	0.882
V\$NF1F	V\$NF1.01	2095	2115	(-)	0.801*	V\$PERO	V\$PPARG.02	2522	2544	(-)	0.940
V\$NF1F	V\$NF1.04	2095	2115	(+)	0.863*	V\$PERO	V\$PPARG.03	2525	2547	(+)	0.865
<b>Sus scrofa region 3 - Intron 1</b>					<b>V\$NF1F</b>						
V\$MYOD	V\$SGN1.01	2718	2734	(+)	0.942	V\$NF1F	V\$NF1.02	2521	2541	(+)	0.804*
V\$STAT	V\$STAT6.01	2734	2752	(+)	0.856	V\$NF1F	V\$NF1.02	2528	2548	(-)	0.807*
V\$NF1F	V\$NF1.01	2808	2828	(-)	0.789*	V\$SP1F	V\$GC.01	2553	2569	(+)	0.860*
V\$NF1F	V\$NF1.02	2808	2828	(+)	0.773*	V\$PERO	V\$PPAR_RXR.01	2586	2608	(-)	0.781
V\$SP1F	V\$SP1.03	2826	2842	(+)	0.910	V\$STAT	V\$STAT6.01	2604	2622	(-)	0.856
V\$SP1F	V\$SP2.01	2846	2862	(-)	0.813	V\$STAT	V\$STAT5B.01	2606	2624	(+)	0.951
V\$PERO	V\$PPARG.03	2888	2910	(-)	0.845						

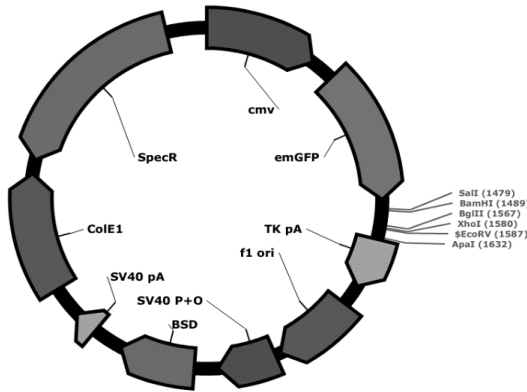
**Table 5: Putative binding sites identified by MatInspector.** The regulatory regions of the UCP3 genes of horse, pig, mouse, hamster, rat and human were screened for DR1/GC-Box modules as described in 3.7.2. Subsequently, MatInspector was used to identify nearby TFBS from the matrix families NF1, STAT and MyoD. For each of the regions depicted in Figure 33, the respective elements are shown. The first column (Family) states the Matrix family to which the identified TFBS (second column: Matrix) belongs. Column 3, 4 and 5 (start, stop, strand) indicate the location of the respective element relative to the TSS and the strand on which the element resides. The last column, Matrix sim., shows how well the detected putative TFBS resembles the consensus sequence. Note that for most of the binding elements only the best match is given. Usually a given binding element also fits to other family members as well. An asterisk indicates an identification that does not meet the strict detection criteria, either due to low matrix similarity, or low core similarity (not shown). Detection criteria can be found in table 3. Matrix families and individual matrix names are explained in table 4.



Species	Module	Location	Strand
<i>Equus caballus</i>	MyoD-NF1F-PERO-GCBox relaxed	293 - 413	(+)
	MyoD-NF1F-PERO-GCBox relaxed	392 - 274	(-)
	MyoD-NF1F-PERO-GCBox relaxed	1067 - 1275	(+)
	MyoD-NF1F-PERO-GCBox relaxed	1853 - 2018	(+)
	STAT-PERO-GCBox-MyoD strict	2291 - 2476	(+)
	STAT-PERO-GCBox-MyoD strict	2312 - 2476	(+)
	MyoD-NF1F-PERO-GCBox relaxed	2381 - 2504	(+)
	MyoD-NF1F-PERO-GCBox relaxed	2381 - 2525	(+)
	MyoD-NF1F-PERO-GCBox relaxed	2381 - 2546	(+)
	MyoD-NF1F-PERO-GCBox relaxed	2381 - 2557	(+)
	MyoD-NF1F-PERO-GCBox relaxed	2454 - 2504	(+)
	MyoD-NF1F-PERO-GCBox relaxed	2454 - 2525	(+)
	MyoD-NF1F-PERO-GCBox relaxed	2454 - 2546	(+)
	MyoD-NF1F-PERO-GCBox relaxed	2454 - 2557	(+)
	MyoD-NF1F-PERO-GCBox relaxed	3975 - 3874	(-)
	MyoD-NF1F-PERO-GCBox relaxed	3975 - 3849	(-)
	MyoD-NF1F-PERO-GCBox relaxed	4359 - 4442	(+)
	MyoD-NF1F-PERO-GCBox relaxed	6674 - 6528	(-)
	MyoD-NF1F-PERO-GCBox relaxed	6689 - 6528	(-)
	<i>Homo sapiens</i>	PERO-GCBox-MyoD-NF1F	483 - 321
<i>Mus musculus</i>	MyoD-NF1F-PERO-GCBox relaxed	1494 - 1583	(+)
	MyoD-NF1F-PERO-GCBox strict	1833 - 1984	(+)
	MyoD-NF1F-PERO-GCBox relaxed	1833 - 1984	(+)
	MyoD-STAT-PERO-GCBox strict	1833 - 1984	(+)
	STAT-MyoD-PERO-GCBox strict	3557 - 3732	(+)
	STAT-MyoD-PERO-GCBox strict	3557 - 3740	(+)
	MyoD-NF1F-PERO-GCBox relaxed	6192 - 6045	(-)
	MyoD-NF1F-PERO-GCBox relaxed	6192 - 6025	(-)
<i>Phodopus sungorus</i>	MyoD-NF1F-PERO-GCBox relaxed	958 - 809	(-)
	MyoD-NF1F-PERO-GCBox relaxed	1343 - 1201	(-)
	MyoD-NF1F-PERO-GCBox strict	1489 - 1631	(+)
	MyoD-NF1F-PERO-GCBox relaxed	1489 - 1632	(+)
	MyoD-STAT-PERO-GCBox strict	1489 - 1631	(+)
	MyoD-PERO-GCBox-NF1F1 strict	1489 - 1651	(+)
	MyoD-PERO-GCBox-STAT strict	1489 - 1698	(+)
	MyoD-NF1F-PERO-GCBox relaxed	1505 - 1390	(-)
	MyoD-NF1F-PERO-GCBox relaxed	2922 - 2788	(-)
<i>Rattus norvegicus</i>	MyoD-NF1F-PERO-GCBox relaxed	659 - 542	(-)
	MyoD-NF1F-PERO-GCBox strict	1622 - 1773	(+)
	MyoD-NF1F-PERO-GCBox relaxed	1622 - 1773	(+)
	MyoD-NF1F-PERO-GCBox relaxed	1622 - 1779	(+)
	MyoD-STAT-PERO-GCBox strict	1622 - 1773	(+)
	MyoD-NF1F-PERO-GCBox relaxed	2551 - 2386	(-)
<i>Sus scrofa</i>	MyoD-NF1F-PERO-GCBox relaxed	410 - 269	(-)

**Table 6: Putative enhancer regions identified by ModelInspector.** The regulatory regions of the UCP3 genes of horse, pig, mouse, hamster, rat and human were screened for presence of complex modules as described in 3.7.2. Models used for the module search are described in table 3. The first column (Species) states the respective species whose UCP3 gene was searched. The second column (Module) notes the Model to which the respective region fits. Column 3 and 4 (Location and Strand) indicate the location of the respective element relative to the TSS and the strand on which the element resides. If a region is identified by multiple modules or by a given module more than once, different combinations of TFBS in that region fit to the respective module, usually originating from presence of multiple candidate elements for a given TFBS. Models used for the module search and relaxed and strict criteria are described in table 3.

### 6.1.10. Appendix 10: Vector maps

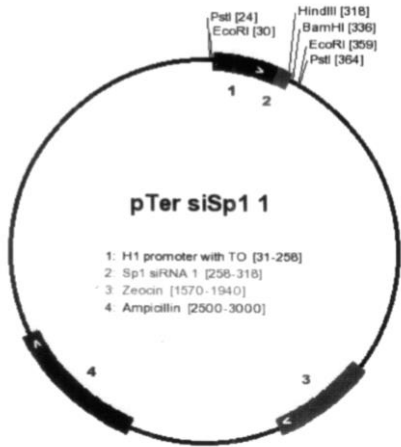


pcDNA6.2-GW/EmGFP  
5699 bp

Vector name:  
**pcDNA6.2-GW/EmGFP empty**

Related/derived vectors:  
**Cloning intermediates for generation of PMXs miRNA vectors**

Origin:  
**Invitrogen; part of BLOCK-iT™ Pol II miRNA RNAi Expression Vector Kit with EmGFP (K4936-00)**

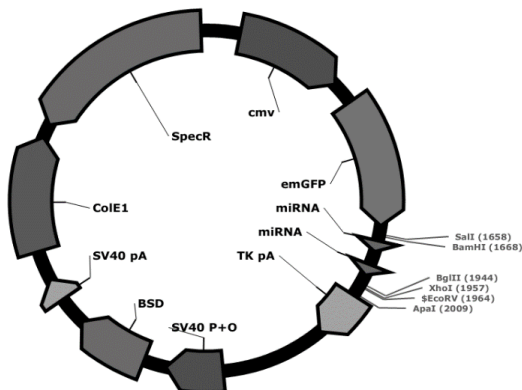


Vector name:  
**pTer siSp1-1**

Related/derived vectors:  
**pTer siSp1-2, -3, -4; siSp3-1, -2, -3, -4, Scram-1, -2, -3, -4**

Origin:  
**Provided by AG Guntram Suske**

Sequences:  
**See appendix 7, page 128: miRNA and shRNA sequences**



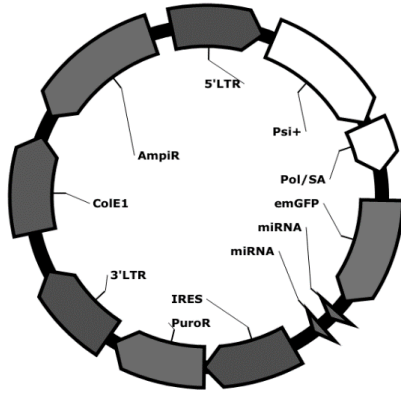
pcDNA6.2-GW/EmGFP 2x miRNA  
5897 bp

Vector name:  
**pcDNA6.2-GW/EmGFP 2x miRNA**

Related/derived vectors:  
**Cloning intermediates for generation of PMXs miRNA vectors**

Origin:  
**Annealing, ligation and concatemerisation of miRNA oligos according to Kit manual**

Sequences:  
**See appendix 7, page 128: miRNA and shRNA sequences**



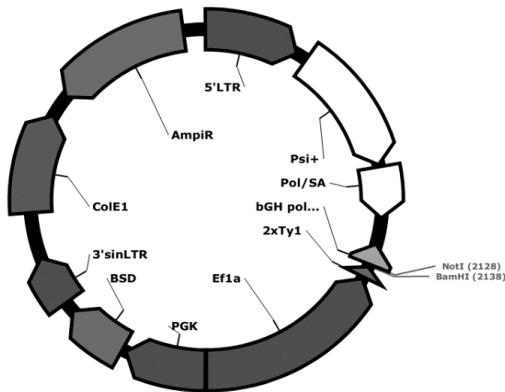
pMXs miRNA IRES Puro  
7057 bp

Vector name:  
**pMXs miRNA IRES Puro**

Related/derived vectors:  
**pMXs miR-SP1, -SP3, -1+3A, -1+3B, -Ctrl1, -Ctrl2**

Origin:  
**Integration of the emGFP-miRNA cassette of pcDNA6.2-GW/EmGFP 2x miRNA (see 2.2.2.1)**

Sequences:  
**See appendix 7, page 128: miRNA and shRNA sequences**



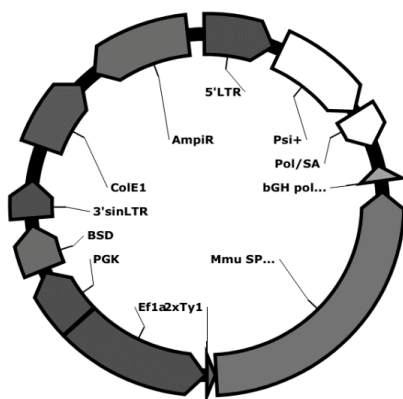
pMXs EF1 2xTy1 BSD  
6818 bp

Vector name:  
**pMXs EF1 2x Ty1 BSD empty**

Related/derived vectors:  
**pMXs EF1 Ty1-CREB, -SP1, -SP2, SP3**

Origin:  
**Insertion of the annealed Ty1 tag oligonucleotides into pMXs EF1 BSD empty (see 2.2.2.2)**

Sequences:  
**See appendix 7, page 129: cloning primers and oligonucleotides**



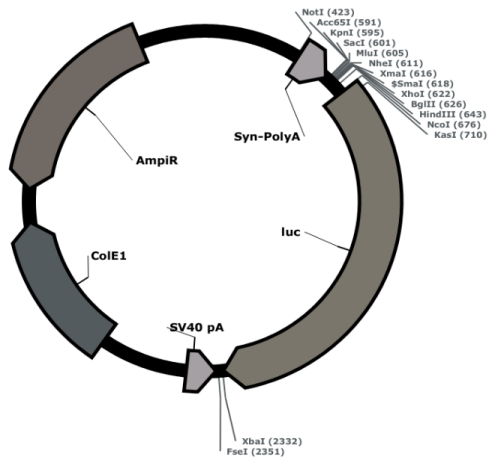
pMXs EF1 2xTy1 SP1 BSD  
9246 bp

Vector name:  
**pMXs EF1 Ty1-SP1 BSD empty**

Related/derived vectors:  
**pMXs EF1 Ty1-CREB, -SP2, SP3**

Origin:  
**Insertion of the amplified SP1 CDS into pMXs EF1 2x Ty1 BSD empty (see 2.2.2.3)**

Sequences:  
**See appendix 7, page 129: cloning primers and oligonucleotides**

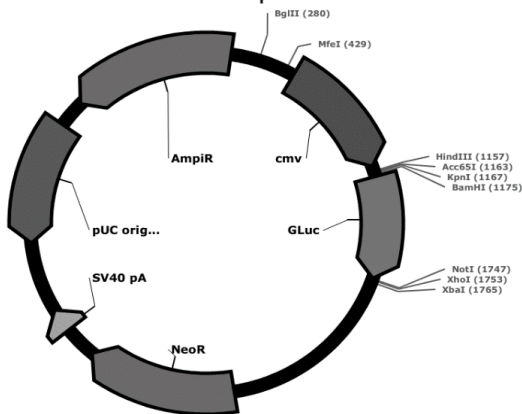


Vector name:  
**pGL3 basic**

Related/derived vectors:  
**All reporter gene constructs used**

Origin:  
**Promega (E1751)**

**pGL3 basic**  
4818 bp

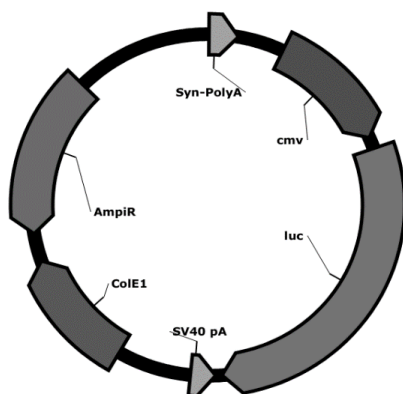


Vector name:  
**pCMV GLuc**

Related/derived vectors:  
**- none -**

Origin:  
**New England Biolabs (N8081S)**

**pCMV GLuc**  
5764 bp

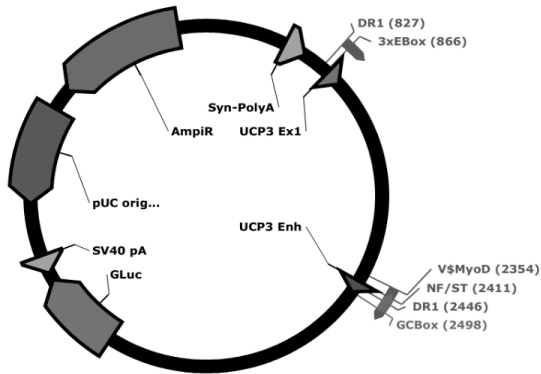


Vector name:  
**pGL3 CMV PLuc**

Related/derived vectors:  
**- none -**

Origin:  
**Insertion of the CMV promoter excised from pCMV GLuc into pGL3 basic (see 2.2.3.1)**

**pGL3 CMV PLuc**  
5678 bp



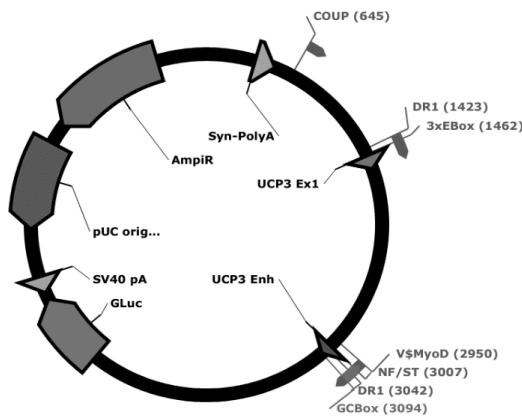
Vector name:  
**pGL3 GLuc *Psu* UCP3 P250**

Related/derived vectors:  
**A large array of mutation and deletion constructs (see 2.2.3)**

Origin:  
**Insertion of an the GLuc cassette amplified from pCMV Gluc into a UCP3 PLuc vector available from [58]**

Sequences:  
**See appendix 7, page 129: primers for generation of deletions and mutations**

**pGL3 GLuc *Psu* UCP3 P250**  
7245 bp

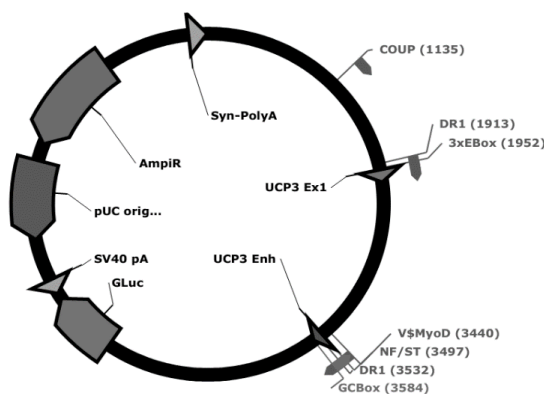


Vector name:  
**pGL3 GLuc *Psu* UCP3 P1000**

Related/derived vectors:  
**pGL3 GLuc *Psu* UCP3 P1000 IVS1+1505G and +1505A**

Origin:  
**Insertion of an upstream promoter fragment into the P250 construct excised from the -2244UCP3luc vector [87] (see 2.2.3.4)**

**pGL3 GLuc *Psu* UCP3 P1000**  
8017 bp



Vector name:  
**pGL3 GLuc *Psu* UCP3 P2000**

Related/derived vectors:  
**pGL3 GLuc *Psu* UCP3 P2000 IVS1+1505G and +1505A**

Origin:  
**Insertion of an upstream promoter fragment into the P250 construct amplified from the -2244UCP3luc vector [87] (see 2.2.3.4)**

Sequences:  
**See appendix 7, page 129: cloning primers an oligonucleotides**

**pGL3 GLuc *Psu* UCP3 P2000**  
9024 bp

### **6.1.11. Appendix 11: Software and internet resources**

#### **Buffer selection for double digests**

<http://www.thermoscientificbio.com/webtools/doubledigest/>

#### **Design of mutagenesis primers**

<http://www.genomics.agilent.com/primerDesignProgram.jsp>

#### **Design of sequencing primers**

<https://ecom.mwgdna.com/services/manage-primers/design-primer.tcl>

#### **Design of qPCR and cloning primers, sequence alignments**

<http://workbench.sdsc.edu/>

#### **Design of miRNA hairpin sequences**

<http://rnaidesigner.lifetechnologies.com/rnaiexpress/>

#### **Dotplot analysis using EMBOSS GUI**

<http://emboss.bioinformatics.nl/>

#### **ENCODE project**

<http://genome.ucsc.edu/encode/>

#### **NEB melting temperature calculator**

<https://www.neb.com/tools-and-resources/interactive-tools/tm-calculator>

#### **Plasma DNA software**

<http://research.med.helsinki.fi/plasmadna/>

#### **TFBS prediction**

<http://www.gene-regulation.com/pub/programs.html>

<http://www.genomatix.de/>

#### **GAP4 Staden package for sequence analysis**

<http://staden.sourceforge.net/>

### **6.1.12. Appendix 12: Accession numbers ChIP-seq data**

#### **Myogenin ChIP-seq in C2C12 cells**

GSM915159, GSM915163, GSM915166, GSM915164

#### **MyoD ChIP-seq in C2C12 cells**

GSM915183, GSM915185, GSM915186, GSM915165

#### **p300 and Pol2 ChIP-seq in heart**

GSM918747, GSM918723

#### **DNaseI hypersensitive sites in heart and muscle**

GSM1014166, GSM1014166, GSM1014189, GSM1014189

#### **PPAR $\gamma$ and RXR ChIP-Seq in 3T3L1**

GSM340794 - GSM340805

#### **PPAR $\gamma$ ChIP-Seq in eWAT, iWAT and BAT**

GSM1018066 - GSM1018068

### 6.1.13. Appendix 13: Complete list of proteins identified by mass spectrometry

Identified Proteins (196)	Accession	#IDs
#IDs: number of peptide hits in a total of 11 samples from 4 experiments, summed up regardless of probe-	and experiment	type
Keratin, type I cytoskeletal 10 Intermediate filament protein;	IPI00755181	809
Keratin, type II cytoskeletal 73 Intermediate filament protein;	IPI00347110	369
Transitional endoplasmic reticulum ATPase ATPase family associated with various cellular activities (AAA);Cell	IPI00622235	253
Keratin, type II cytoskeletal 1 Intermediate filament protein;	IPI00625729	213
Keratin, type II cytoskeletal 2 epidermal Intermediate filament protein;	IPI00622240	178
Keratin, type II cytoskeletal 5 Intermediate filament protein;	IPI00139301	172
Histone H1.2 linker histone H1 and H5 family;	IPI00223713	122
Nucleolin RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IPI00317794	112
Keratin, type I cytoskeletal 14 Intermediate filament protein;	IPI00227140	80
Myosin-9 IQ calmodulin-binding motif;Myosin head (motor domain);Myosin N-terminal SH3-like	IPI00123181	77
Actin, cytoplasmic 1 Actin;	IPI00110850	76
Keratin, type II cytoskeletal 8 Intermediate filament protein;	IPI00322209	75
Serum albumin Serum albumin family;	IPI00131695	71
Putative uncharacterized protein linker histone H1 and H5 family;	IPI00404590	71
IPI00856683-R	IPI00856683-	64
Keratin, type II cuticular Hb5 Intermediate filament protein;	IPI00331459	61
Keratin, type I cytoskeletal 42 Intermediate filament protein;	IPI00468696	60
Staphylococcal nuclease domain-containing protein 1 Staphylococcal nuclease homologue;Tudor domain;	IPI00123129	58
40S ribosomal protein S3 KH domain; Ribosomal protein S3, C-terminal domain;	IPI00134599	56
40S ribosomal protein S4 Ribosomal protein S2;	IPI00123604	53
Keratin, type I cytoskeletal 17 Intermediate filament protein;	IPI00230365	52
Histone H1.5 linker histone H1 and H5 family;	IPI00230133	51
11 kDa protein Core histone H2A/H2B/H3/H4;	IPI00876005	48
Keratin, type II cuticular Hb2 Intermediate filament protein;	IPI00117269	43
40S ribosomal protein S8 Ribosomal protein S8e;	IPI00621229	43
40S ribosomal protein S4, X isoform KOW motif;Ribosomal family S4e;RS4NT (NUC023) domain;S4 domain;	IPI00331092	43
Isoform 2 of Heterogeneous nuclear ribonucleoprotein K KH domain;ROKNT (NUC014) domain;	IPI00224575	39
Heterogeneous nuclear ribonucleoprotein U SAP domain;SPRY domain;Zeta toxin;	IPI00458583	39
Guanine nucleotide-binding protein subunit beta-2-like 1 WD domain, G-beta repeat;	IPI00317740	36
Isoform 1 of Serine/arginine-rich splicing factor 1 RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IPI00420807	36
Tubulin beta-5 chain Tubulin/FtsZ family, GTPase domain;Tubulin C-terminal domain;	IPI00117352	36
IPI00187463-R	IPI00187463	35
40S ribosomal protein S3a Ribosomal S3Ae family;	IPI00331345	34
Isoform 1 of U1 small nuclear ribonucleoprotein 70 kDa RNA recognition motif. (a.k.a. RRM, RBD, or RNP	IPI00625105	33
Histone H1.4 linker histone H1 and H5 family;	IPI00223714	31
60S ribosomal protein L14 Ribosomal protein L14;	IPI00133185	30
Heat shock cognate 71 kDa protein Hsp70 protein;	IPI00323357	30
Uncharacterized protein Ribosomal protein S6e;	IPI00108454	30
FACT complex subunit SPT16 Metallopeptidase family M24;Histone chaperone Rtp106-like;FACT complex	IPI00120344	30
40S ribosomal protein S7 Ribosomal protein S7e;	IPI00136984	29
Histone H2A type 2-C Core histone H2A/H2B/H3/H4;	IPI00272033	28
40S ribosomal protein S18 Ribosomal protein S13/S18;	IPI00317590	28
Isoform 1 of Plasminogen activator inhibitor 1 RNA-binding protein Hyaluronan / mRNA binding family;	IPI00471475	27
60S ribosomal protein L6 Ribosomal protein L6e ;Ribosomal protein L6, N-terminal domain;	IPI00313222	26
60S ribosomal protein L12 Ribosomal protein L11, RNA binding domain;Ribosomal protein L11, N-terminal	IPI00849793	26
Histone H1.1 linker histone H1 and H5 family;	IPI00228616	26
Histone H2B type 1-F/J/L Core histone H2A/H2B/H3/H4;	IPI00114642	25
Nucleophosmin Nucleoplasmin;	IPI00127415	25
Rps16 protein Ribosomal protein S9/S16;	IPI00469918	25
Sorting nexin-9 WASP-binding domain of Sorting nexin protein;PX domain;SH3 domain;	IPI00313275	24
Isoform 3 of Heterogeneous nuclear ribonucleoproteins A2/B1 Nuclear factor hnRNPA1;RNA recognition motif.	IPI00405058	23
60S ribosomal protein L4 Ribosomal protein L4/L1 family;	IPI00111412	23
Isoform B of Methyl-CpG-binding protein 2 Methyl-CpG binding domain;	IPI00775806	23
Isoform 1 of THO complex subunit 4 RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IPI00114407	23
40S ribosomal protein S12 Ribosomal protein L7Ae/L30e/S12e/Gadd45 family;	IPI00225634	23
60S ribosomal protein L23 Ribosomal protein L14p/L23e;	IPI00139780	23
Uncharacterized protein Ribosomal protein S7p/S5e;	IPI00857345	22
Putative uncharacterized protein RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IPI00128441	22
DNA topoisomerase 1 Eukaryotic DNA topoisomerase I, catalytic core;Eukaryotic DNA topoisomerase I, DNA	IPI00109764	22
60S ribosomal protein L11 Ribosomal protein L5;ribosomal L5P family C-terminus;	IPI00331461	22
Putative uncharacterized protein Fibronectin type I domain;Fibronectin type II domain;Fibronectin type III	IPI00652813	21
Putative uncharacterized protein Histidine_kinase_-, DNA gyrase B-, and HSP90-like ATPase;Hsp90 protein;	IPI00229080	21
Uncharacterized protein RBM1CTR (NUC064) family;RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IPI00474144	21
40S ribosomal protein S27 Ribosomal protein S27;	IPI00173160	21
60S ribosomal protein L18 Ribosomal protein L18e/L15;	IPI00555113	20
40S ribosomal protein S21 Ribosomal protein S21e ;	IPI00132950	20
Isoform 2 of Heterogeneous nuclear ribonucleoprotein D0 CBFNT (NUC161) domain;RNA recognition motif.	IPI00230086	18

40S ribosomal protein S10 Plectin/S10 domain;	IP100112448	18
60S ribosomal protein L30 Ribosomal protein L7Ae/L30e/S12e/Gadd45 family;	IP100222549	18
40S ribosomal protein S14 Ribosomal protein S11;	IP100322562	18
Heterogeneous nuclear ribonucleoprotein A/B CBFNT (NUC161) domain;RNA recognition motif. (a.k.a. RRM,	IP100117288	17
Tubulin alpha-1B chain Tubulin/FtsZ family, GTPase domain;Tubulin C-terminal domain;	IP100117348	16
40S ribosomal protein S28 Ribosomal protein S28e;	IP100137736	16
60S ribosomal protein L7 Ribosomal protein L30p/L7e;Ribosomal L30 N-terminal domain;	IP100311236	16
60S ribosomal protein L13 Ribosomal protein L13e;	IP100224505	16
Putative uncharacterized protein BRCA1 C Terminus (BRCT) domain;PADR1 (NUC008) domain;Poly(ADP-ribose)	IP100139168	16
Uncharacterized protein Ribosomal protein S5, N-terminal domain;Ribosomal protein S5, C-terminal domain;	IP100604967	15
Ubiquitin-40S ribosomal protein S27a Ribosomal protein S27a;Ubiquitin family;	IP100470152	15
40S ribosomal protein S20 Ribosomal protein S10p/S20e;	IP100323819	15
Pre-mRNA-processing-splicing factor 8 Mov34/MPN/PAD-1 family;PRO8NT (NUC069), PrP8 N-terminal	IP100121596	14
heterogeneous nuclear ribonucleoprotein A0 RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IP100109813	14
60S acidic ribosomal protein P0 60s Acidic ribosomal protein;Ribosomal protein L10;	IP100314950	14
Elongation factor 2 Elongation factor G C-terminus;Elongation factor G, domain IV;Elongation factor Tu GTP	IP100466069	14
Isoform 2 of 40S ribosomal protein S24 Ribosomal protein S24e;	IP100402981	14
Alpha-actinin-4 Calponin homology (CH) domain;Ca2+ insensitive EF hand;Spectrin repeat;	IP100118899	13
40S ribosomal protein S9 Ribosomal protein S4/S9 N-terminal domain;S4 domain;	IP100420726	13
40S ribosomal protein S19 Ribosomal protein S19e;	IP100113241	13
Proliferation-associated protein 2G4 Metallopeptidase family M24;	IP100119305	13
Isoform 2 of FACT complex subunit SSRP1 HMG (high mobility group) box;Structure-specific recognition protein	IP100407571	13
Isoform 1 of Heterogeneous nuclear ribonucleoprotein A3 RNA recognition motif. (a.k.a. RRM, RBD, or RNP	IP100269661	12
Keratin, type I cuticular Ha1 Intermediate filament protein;	IP100124945	12
Heterogeneous nuclear ribonucleoprotein L RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IP100620362	11
High mobility group protein B1 HMG (high mobility group) box;	IP100420261	11
Putative uncharacterized protein Ribosomal protein L24e;	IP100323806	11
60S ribosomal protein L32 Ribosomal protein L32;	IP100230623	11
60S ribosomal protein L18a Ribosomal L18ae/LX protein domain;	IP100162790	11
Endoplasmic Histidine_kinase_, DNA gyrase B-, and HSP90-like ATPase;Hsp90 protein;	IP100129526	11
Uncharacterized protein Ribosomal protein L7Ae/L30e/S12e/Gadd45 family;	IP100265107	10
Biglycan Leucine Rich Repeat;Leucine rich repeat N-terminal domain;	IP100123194	10
Elongation factor 1-alpha 1 Elongation factor Tu GTP binding domain;Elongation factor Tu domain 2;Elongation	IP100307837	10
Protein DEK DEK C terminal domain;SAP domain;	IP100227720	10
Uncharacterized protein S25 ribosomal protein;	IP100115992	10
Serine/arginine-rich splicing factor 5 RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IP100314709	10
Probable ATP-dependent RNA helicase DDX5 DEAD/DEAH box helicase;Helicase conserved C-terminal	IP100420363	9
U1 small nuclear ribonucleoprotein A RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IP100122350	9
Nucleosome assembly protein 1-like 1 Nucleosome assembly protein (NAP);	IP100123199	9
60S ribosomal protein L3 Ribosomal protein L3;	IP100321170	9
Putative uncharacterized protein RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IP100130343	9
Keratin, type I cuticular Ha5 Intermediate filament protein;	IP100131382	9
Isoform Smooth muscle of Myosin light polypeptide 6	IP100354819	9
60S ribosomal protein L10 Ribosomal protein L16p/L10e;	IP100474637	9
Putative uncharacterized protein Thioredoxin;	IP100122815	8
similar to Protein disulfide isomerase associated 6 Thioredoxin;	IP100854971	8
Polyadenylate-binding protein 1 Poly-adenylate binding protein, unique domain;RNA recognition motif. (a.k.a.	IP100124287	8
Isoform Long of Serine/arginine-rich splicing factor 3 RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IP100129323	8
High mobility group protein B2 HMG (high mobility group) box;	IP100462291	8
60S ribosomal protein L31 Ribosomal protein L31e;	IP100123007	8
Serpin H1 Serpin (serine protease inhibitor);	IP100114733	8
ATP-binding cassette sub-family E member 1 ABC transporter;4Fe-4S binding domain;Possible metal-binding	IP100322869	8
Keratin, type II cytoskeletal 6A Intermediate filament protein;	IP100131368	7
60S ribosomal protein L5 Ribosomal L18p/L5e family;	IP100308706	7
60S ribosomal protein L23a Ribosomal protein L23;Ribosomal protein L23, N-terminal domain;	IP100461456	7
Uncharacterized protein Ribosomal protein L18e/L15;	IP100474407	7
Cytoskeleton-associated protein 4	IP100223047	7
60S ribosomal protein L13a Ribosomal protein L13;	IP100223217	7
40S ribosomal protein S13 Ribosomal S13/S15 N-terminal domain;Ribosomal protein S15;	IP100125901	7
H/ACA ribonucleoprotein complex subunit 2 Ribosomal protein L7Ae/L30e/S12e/Gadd45 family;	IP100133550	7
Interleukin enhancer-binding factor 2 DZF;	IP100318550	7
Type VI collagen alpha 3 subunit Collagen triple helix repeat (20 copies);Kunitz/Bovine pancreatic trypsin	IP100830749	7
Serine hydroxymethyltransferase Serine hydroxymethyltransferase;	IP100454008	6
60S ribosomal protein L8 Ribosomal Proteins L2, RNA binding domain;Ribosomal Proteins L2, C-terminal	IP100137787	6
Vigilin KH domain;	IP100123379	6
60S ribosomal protein L36 Ribosomal protein L36e;	IP100463297	6
40S ribosomal protein S17 Ribosomal S17;	IP100465880	6
60S ribosomal protein L22 Ribosomal L22e protein family;	IP100222546	6
Activated RNA polymerase II transcriptional coactivator p15 Transcriptional Coactivator p15 (PC4);	IP100225633	6
116 kDa U5 small nuclear ribonucleoprotein component Elongation factor G C-terminus;Elongation factor G,	IP100469260	6
H/ACA ribonucleoprotein complex subunit 4 DKCLD (NUC011) domain;PUA domain;TruB family pseudouridylylate	IP100113635	6
Isoform 1 of Heterogeneous nuclear ribonucleoprotein M RNA recognition motif. (a.k.a. RRM, RBD, or RNP	IP100132443	6
Putative uncharacterized protein Ribosomal protein S17;	IP100117569	6



Isoform UBF1 of Nucleolar transcription factor 1 HMG (high mobility group) box;	IPI00114869	6
Cell growth-regulating nucleolar protein LYAR-type C2HC zinc finger ;	IPI00113232	6
Stress-70 protein, mitochondrial Hsp70 protein;	IPI00133903	5
Isoform 2 of ATP-dependent RNA helicase A DEAD/DEAH box helicase;Double-stranded RNA binding	IPI00339468	5
Isoform 3 of Ribosome-binding protein 1 Ribosome receptor lysine/proline rich region;	IPI00121149	5
Putative uncharacterized protein LOC435285 Keratin, high sulfur B2 protein;	IPI00458822	5
Keratin type II Intermediate filament protein;	IPI00124800	5
Histone H3.2 Core histone H2A/H2B/H3/H4;	IPI00230730	5
myosin light chain, regulatory B-like EF hand;	IPI00109044	5
Putative uncharacterized protein RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IPI00553777	4
Poly(rC)-binding protein 1 KH domain;	IPI00128904	4
Ras-related protein Rab-7a Ras family;	IPI00408892	4
60S ribosomal protein L9 Ribosomal protein L6;	IPI00122413	4
Serine/arginine-rich splicing factor 2 RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);	IPI00121135	4
Tubulin beta-2C chain Tubulin/FtsZ family, GTPase domain;Tubulin C-terminal domain;	IPI00169463	4
similar to Rps15a protein Ribosomal protein S8;	IPI00113394	4
D-3-phosphoglycerate dehydrogenase D-isomer specific 2-hydroxyacid dehydrogenase, catalytic domain;D-	IPI00225961	4
60S ribosomal protein L17 Ribosomal protein L22p/L17e;	IPI00453768	4
Small nuclear ribonucleoprotein 200 DEAD/DEAH box helicase;Helicase conserved C-terminal domain;Sec63 Brl	IPI00420329	4
scaffold attachment factor B RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);SAP domain;	IPI00944159	4
Nuclease-sensitive element-binding protein 1 'Cold-shock' DNA-binding domain;	IPI00120886	4
High mobility group protein HMGI-C AT hook motif;	IPI00331612	4
putative ribosomal RNA methyltransferase NOP2 NOL1/NOP2/sun family;P120R (NUC006) repeat;	IPI00311453	3
Casein kinase II subunit alpha Protein_kinase_ domain;	IPI00120162	3
U4/U6.U5 tri-snRNP-associated protein 2 Ubiquitin carboxyl-terminal hydrolase;Zn-finger in ubiquitin-	IPI00457815	3
60S ribosomal protein L10a Ribosomal protein L1p/L10e family;	IPI00127085	3
Eukaryotic initiation factor 4A-I DEAD/DEAH box helicase;Helicase conserved C-terminal domain;	IPI00118676	3
Coatomer subunit alpha Coatomer WD associated region ;Coatomer (COPI) alpha subunit C-terminus;WD	IPI00229834	3
60S ribosomal protein L38 Ribosomal L38e protein family;	IPI00331121	3
X-ray repair cross-complementing protein 6 Ku70/Ku80 beta-barrel domain;Ku70/Ku80 C-terminal	IPI00132424	3
Isoform 1 of H/ACA ribonucleoprotein complex subunit 1 Gar1/Naf1 RNA binding region;	IPI00110931	3
IPI00353560-R	IPI00353560	3
2,4-dienoyl-CoA reductase, mitochondrial short chain dehydrogenase;	IPI00387379	3
60S ribosomal protein L34 Ribosomal protein L34e;	IPI00466153	3
Isoform 1 of Heterogeneous nuclear ribonucleoprotein F RNA recognition motif. (a.k.a. RRM, RBD, or RNP	IPI00226073	3
Heterogeneous nuclear ribonucleoprotein U-like protein 2 SAP domain;SPRY domain;Zeta toxin;	IPI00222208	3
Eukaryotic translation initiation factor 3 subunit A PCI domain;	IPI00129276	3
Lupus La protein homolog La domain;RNA recognition motif. (a.k.a. RRM, RBD, or RNP domain);RNA binding	IPI00134300	2
Putative uncharacterized protein Adaptin N terminal region;Coatomer gamma subunit appendage domain;	IPI00223437	2
Serine protease 23 Trypsin;	IPI00318017	2
Thyroid hormone receptor-associated protein 3	IPI00556768	2
60S ribosomal protein L26 KOW motif;	IPI00132460	2
Protein disulfide-isomerase A3 Thioredoxin;	IPI00230108	2
Isoform Alpha of DNA ligase 3 ATP dependent DNA ligase C terminal region ;ATP dependent DNA ligase	IPI00124272	2
Dolichyl-diphosphooligosaccharide--protein glycosyltransferase 48 kDa subunit Oligosaccharyltransferase 48	IPI00117705	2
Isoform 1 of Heterogeneous nuclear ribonucleoprotein Q RNA recognition motif. (a.k.a. RRM, RBD, or RNP	IPI00406117	2
basement membrane-specific heparan sulfate proteoglycan core protein EGF-like domain;Immunoglobulin I-set	IPI00515360	2
Isoform 1 of Bifunctional polynucleotide phosphatase/kinase Polynucleotide_kinase_3 phosphatase;	IPI00454118	2
Isoform 1 of Heterochromatin protein 1-binding protein 3 linker histone H1 and H5 family;	IPI00342766	2
Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 2OG-Fe(II) oxygenase superfamily;	IPI00127407	2
Eif3b protein Eukaryotic translation initiation factor eIF2A;RNA recognition motif. (a.k.a. RRM, RBD, or RNP	IPI00229859	2
Nucleolar protein 56 Putative snoRNA binding domain;NOP5NT (NUC127) domain;NOSIC (NUC001) domain;	IPI00318048	2
Eukaryotic translation initiation factor 3 subunit C Eukaryotic translation initiation factor 3 subunit 8 N-	IPI00321647	2
Keratin 32 Intermediate filament protein;	IPI00122281	2
RNA polymerase II-associated factor 1 homolog Paf1 ;	IPI00331654	2
Lysine-specific demethylase 5A ARID/BRIGHT DNA binding domain;JmjC domain;jmjN domain;PHD-finger;PLU-	IPI00849089	2

**#IDs: number of peptide hits in a total of 11 samples from 4 experiments, summed up regardless of probe- and experiment type**

## E. Literature

1. Allison DB, Fontaine KR, Manson JE, Stevens J, VanItallie TB (1999) Annual deaths attributable to obesity in the United States. *JAMA* 282: 1530-1538.
2. Haslam DW, James WP (2005) Obesity. *Lancet* 366: 1197-1209.
3. Lean ME, James WP, Jennings G, Trayhurn P (1986) Brown adipose tissue uncoupling protein content in human infants, children and adults. *Clin Sci (Lond)* 71: 291-297.
4. Saito M, Okamatsu-Ogura Y, Matsushita M, Watanabe K, Yoneshiro T, et al. (2009) High incidence of metabolically active brown adipose tissue in healthy adult humans: effects of cold exposure and adiposity. *Diabetes* 58: 1526-1531.
5. van Marken Lichtenbelt WD, Vanhommel JW, Smulders NM, Drossaerts JM, Kemerink GJ, et al. (2009) Cold-activated brown adipose tissue in healthy men. *N Engl J Med* 360: 1500-1508.
6. Virtanen KA, Lidell ME, Orava J, Heglind M, Westergren R, et al. (2009) Functional brown adipose tissue in healthy adults. *N Engl J Med* 360: 1518-1525.
7. Boss O, Samec S, Paoloni-Giacobino A, Rossier C, Dulloo A, et al. (1997) Uncoupling protein-3: a new member of the mitochondrial carrier family with tissue-specific expression. *FEBS Lett* 408: 39-42.
8. Vidal-Puig A, Solanes G, Grujic D, Flier JS, Lowell BB (1997) UCP3: an uncoupling protein homologue expressed preferentially and abundantly in skeletal muscle and brown adipose tissue. *BiochemBiophysResCommun* 235: 79-82.
9. Seale P, Bjork B, Yang W, Kajimura S, Chin S, et al. (2008) PRDM16 controls a brown fat/skeletal muscle switch. *Nature* 454: 961-967.
10. Farmer SR (2008) Brown fat and skeletal muscle: unlikely cousins? *Cell* 134: 726-727.
11. Millet L, Vidal H, Andreelli F, Larrouy D, Riou JP, et al. (1997) Increased uncoupling protein-2 and -3 mRNA expression during fasting in obese and lean humans. *J Clin Invest* 100: 2665-2670.
12. Larkin S, Mull E, Miao W, Pittner R, Albrandt K, et al. (1997) Regulation of the third member of the uncoupling protein family, UCP3, by cold and thyroid hormone. *Biochem Biophys Res Commun* 240: 222-227.
13. Matsuda J, Hosoda K, Itoh H, Son C, Doi K, et al. (1997) Cloning of rat uncoupling protein-3 and uncoupling protein-2 cDNAs: their gene expression in rats fed high-fat diet. *FEBS Lett* 418: 200-204.
14. Weigle DS, Selfridge LE, Schwartz MW, Seeley RJ, Cummings DE, et al. (1998) Elevated free fatty acids induce uncoupling protein 3 expression in muscle: a potential explanation for the effect of fasting. *Diabetes* 47: 298-302.
15. Himms-Hagen J, Harper ME (2001) Physiological role of UCP3 may be export of fatty acids from mitochondria when fatty acid oxidation predominates: an hypothesis. *ExpBiolMed(Maywood)* 226: 78-84.
16. Schrauwen P, Hesselink MK (2004) The role of uncoupling protein 3 in fatty acid metabolism: protection against lipotoxicity? *Proc Nutr Soc* 63: 287-292.
17. Heaton GM, Wagenvoort RJ, Kemp A, Jr., Nicholls DG (1978) Brown-adipose-tissue mitochondria: photoaffinity labelling of the regulatory site of energy dissipation. *Eur J Biochem* 82: 515-521.
18. Enerback S, Jacobsson A, Simpson EM, Guerra C, Yamashita H, et al. (1997) Mice lacking mitochondrial uncoupling protein are cold-sensitive but not obese. *Nature* 387: 90-94.
19. Golozoubova V, Hohtola E, Matthias A, Jacobsson A, Cannon B, et al. (2001) Only UCP1 can mediate adaptive nonshivering thermogenesis in the cold. *FASEB J* 15: 2048-2050.
20. Lafontan M, Berlan M (1993) Fat cell adrenergic receptors and the control of white and brown fat cell function. *J Lipid Res* 34: 1057-1091.
21. Cao W, Daniel KW, Robidoux J, Puigserver P, Medvedev AV, et al. (2004) p38 mitogen-activated protein kinase is the central regulator of cyclic AMP-dependent transcription of the brown fat uncoupling protein 1 gene. *Mol Cell Biol* 24: 3057-3067.
22. de Jesus LA, Carvalho SD, Ribeiro MO, Schneider M, Kim SW, et al. (2001) The type 2 iodothyronine deiodinase is essential for adaptive thermogenesis in brown adipose tissue. *J Clin Invest* 108: 1379-1385.
23. Carvalho SD, Kimura ET, Bianco AC, Silva JE (1991) Central role of brown adipose tissue thyroxine 5'-deiodinase on thyroid hormone-dependent thermogenic response to cold. *Endocrinology* 128: 2149-2159.
24. Bordicchia M, Liu D, Amri EZ, Ailhaud G, Dessi-Fulgheri P, et al. (2012) Cardiac natriuretic peptides act via p38 MAPK to induce the brown fat thermogenic program in mouse and human adipocytes. *J Clin Invest* 122: 1022-1036.
25. Okamura H, Kelly PA, Chabot JG, Morel G, Belles-Isles M, et al. (1988) Atrial natriuretic peptide receptors are present in brown adipose tissue. *Biochem Biophys Res Commun* 156: 1000-1006.
26. Fleury C, Neverova M, Collins S, Raimbault S, Champigny O, et al. (1997) Uncoupling protein-2: a novel gene linked to obesity and hyperinsulinemia. *Nat Genet* 15: 269-272.

27. Mao W, Yu XX, Zhong A, Li W, Brush J, et al. (1999) UCP4, a novel brain-specific mitochondrial protein that reduces membrane potential in mammalian cells. *FEBS Lett* 443: 326-330.
28. Sanchis D, Fleury C, Chomiki N, Goubert M, Huang Q, et al. (1998) BMCP1, a novel mitochondrial carrier with high expression in the central nervous system of humans and rodents, and respiration uncoupling activity in recombinant yeast. *J Biol Chem* 273: 34611-34615.
29. Aquila H, Link TA, Klingenberg M (1987) Solute carriers involved in energy transfer of mitochondria form a homologous protein family. *FEBS Lett* 212: 1-9.
30. Jezek P, Urbankova E (2000) Specific sequence of motifs of mitochondrial uncoupling proteins. *IUBMB Life* 49: 63-70.
31. Locke RM, Rial E, Scott ID, Nicholls DG (1982) Fatty acids as acute regulators of the proton conductance of hamster brown-fat mitochondria. *Eur J Biochem* 129: 373-380.
32. Bouillaud F, Arechaga I, Petit PX, Raimbault S, Levi-Meyrueis C, et al. (1994) A sequence related to a DNA recognition element is essential for the inhibition by nucleotides of proton transport through the mitochondrial uncoupling protein. *EMBO J* 13: 1990-1997.
33. Rafael J, Heldt HW (1976) Binding of guanine nucleotides to the outer surface of the inner membrane of guinea pig brown fat mitochondria in correlation with the thermogenic activity of the tissue. *FEBS Lett* 63: 304-308.
34. Hanak P, Jezek P (2001) Mitochondrial uncoupling proteins and phylogenesis--UCP4 as the ancestral uncoupling protein. *FEBS Lett* 495: 137-141.
35. Jastroch M, Wuertz S, Kloas W, Klingenspor M (2005) Uncoupling protein 1 in fish uncovers an ancient evolutionary history of mammalian nonshivering thermogenesis. *Physiol Genomics* 22: 150-156.
36. Lin CS, Klingenberg M (1982) Characteristics of the isolated purine nucleotide binding protein from brown fat mitochondria. *Biochemistry* 21: 2950-2956.
37. Winkler E, Klingenberg M (1994) Effect of fatty acids on H<sup>+</sup> transport activity of the reconstituted uncoupling protein. *J Biol Chem* 269: 2508-2515.
38. Skulachev VP (1991) Fatty acid circuit as a physiological mechanism of uncoupling of oxidative phosphorylation. *FEBS Lett* 294: 158-162.
39. Fedorenko A, Lishko PV, Kirichok Y (2012) Mechanism of fatty-acid-dependent UCP1 uncoupling in brown fat mitochondria. *Cell* 151: 400-413.
40. Hesselink MK, Keizer HA, Borghouts LB, Schaart G, Kornips CF, et al. (2001) Protein expression of UCP3 differs between human type 1, type 2a, and type 2b fibers. *FASEB J* 15: 1071-1073.
41. Harper JA, Stuart JA, Jekabsons MB, Roussel D, Brindle KM, et al. (2002) Artfactual uncoupling by uncoupling protein 3 in yeast mitochondria at the concentrations found in mouse and rat skeletal-muscle mitochondria. *Biochem J* 361: 49-56.
42. Heidkaemper D, Winkler E, Muller V, Frischmuth K, Liu Q, et al. (2000) The bulk of UCP3 expressed in yeast cells is incompetent for a nucleotide regulated H<sup>+</sup> transport. *FEBS Lett* 480: 265-270.
43. Echtay KS, Roussel D, St-Pierre J, Jekabsons MB, Cadenas S, et al. (2002) Superoxide activates mitochondrial uncoupling proteins. *Nature* 415: 96-99.
44. Vidal-Puig AJ, Grujic D, Zhang CY, Hagen T, Boss O, et al. (2000) Energy metabolism in uncoupling protein 3 gene knockout mice. *J Biol Chem* 275: 16258-16266.
45. Cadenas S, Buckingham JA, Samec S, Seydoux J, Din N, et al. (1999) UCP2 and UCP3 rise in starved rat skeletal muscle but mitochondrial proton conductance is unchanged. *FEBS Lett* 462: 257-260.
46. Cadenas S, Echtay KS, Harper JA, Jekabsons MB, Buckingham JA, et al. (2002) The basal proton conductance of skeletal muscle mitochondria from transgenic mice overexpressing or lacking uncoupling protein-3. *J Biol Chem* 277: 2773-2778.
47. Anedda A, Lopez-Bernardo E, Acosta-Iborra B, Saadeh Suleiman M, Landazuri MO, et al. (2013) The transcription factor Nrf2 promotes survival by enhancing the expression of uncoupling protein 3 under conditions of oxidative stress. *Free Radic Biol Med* 61C: 395-407.
48. Echtay KS, Esteves TC, Pakay JL, Jekabsons MB, Lambert AJ, et al. (2003) A signalling role for 4-hydroxy-2-nonenal in regulation of mitochondrial uncoupling. *EMBO J* 22: 4103-4110.
49. Esterbauer H, Schaur RJ, Zollner H (1991) Chemistry and biochemistry of 4-hydroxynonenal, malonaldehyde and related aldehydes. *Free Radic Biol Med* 11: 81-128.
50. Criscuolo F, Mozo J, Hurtaud C, Nubel T, Bouillaud F (2006) UCP2, UCP3, avUCP, what do they do when proton transport is not stimulated? Possible relevance to pyruvate and glutamine metabolism. *Biochimica Et Biophysica Acta-Bioenergetics* 1757: 1284-1291.
51. Huang SG, Klingenberg M (1996) Chloride channel properties of the uncoupling protein from brown adipose tissue mitochondria: a patch-clamp study. *Biochemistry* 35: 16806-16814.

52. Vozza A, Parisi G, De Leonardis F, Lasorsa FM, Castegna A, et al. (2014) UCP2 transports C4 metabolites out of mitochondria, regulating glucose and glutamine oxidation. *Proc Natl Acad Sci U S A*.
53. Trenker M, Malli R, Fertschai I, Levak-Frank S, Graier WF (2007) Uncoupling proteins 2 and 3 are fundamental for mitochondrial Ca<sup>2+</sup> uniport. *Nat Cell Biol* 9: 445-452.
54. De Marchi U, Castelbou C, Demaurex N (2011) Uncoupling protein 3 (UCP3) modulates the activity of Sarco/endoplasmic reticulum Ca<sup>2+</sup>-ATPase (SERCA) by decreasing mitochondrial ATP production. *J Biol Chem* 286: 32533-32541.
55. Meirhaeghe A, Amouyel P, Helbecque N, Cottel D, Otabe S, et al. (2000) An uncoupling protein 3 gene polymorphism associated with a lower risk of developing Type II diabetes and with atherogenic lipid profile in a French cohort. *Diabetologia* 43: 1424-1428.
56. Fang QC, Jia WP, Yang M, Bao YQ, Chen L, et al. (2005) Effect of polymorphism of uncoupling protein 3 gene -55 (C>T) on the resting energy expenditure, total body fat and regional body fat in Chinese. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi* 22: 485-488.
57. Xu K, Zhang M, Cui D, Fu Y, Qian L, et al. (2011) UCP2 -866G/A and Ala55Val, and UCP3 -55C/T polymorphisms in association with type 2 diabetes susceptibility: a meta-analysis study. *Diabetologia* 54: 2315-2324.
58. Nau K, Fromme T, Meyer CW, von Praun C, Heldmaier G, et al. (2008) Brown adipose tissue specific lack of uncoupling protein 3 is associated with impaired cold tolerance and reduced transcript levels of metabolic genes. *J Comp Physiol B* 178: 269-277.
59. Fromme T, Hoffmann C, Nau K, Rozman J, Reichwald K, et al. (2009) An intronic single base exchange leads to a brown adipose tissue-specific loss of Ucp3 expression and an altered body mass trajectory. *Physiol Genomics* 38: 54-62.
60. von Praun C, Burkert M, Gessner M, Klingenspor M (2001) Tissue-specific expression and cold-induced mRNA levels of uncoupling proteins in the Djungarian hamster. *Physiol Biochem Zool* 74: 203-211.
61. Brun S, Carmona MC, Mampel T, Vinas O, Giralt M, et al. (1999) Uncoupling protein-3 gene expression in skeletal muscle during development is regulated by nutritional factors that alter circulating non-esterified fatty acids. *FEBS Lett* 453: 205-209.
62. Brun S, Carmona MC, Mampel T, Vinas O, Giralt M, et al. (1999) Activators of peroxisome proliferator-activated receptor-alpha induce the expression of the uncoupling protein-3 gene in skeletal muscle: a potential mechanism for the lipid intake-dependent activation of uncoupling protein-3 gene expression at birth. *Diabetes* 48: 1217-1222.
63. Gong DW, He Y, Karas M, Reitman M (1997) Uncoupling protein-3 is a mediator of thermogenesis regulated by thyroid hormone, beta3-adrenergic agonists, and leptin. *J Biol Chem* 272: 24129-24132.
64. Giacobino JP (1999) Effects of dietary deprivation, obesity and exercise on UCP3 mRNA levels. *Int J Obes Relat Metab Disord* 23 Suppl 6: S60-63.
65. Nagase I, Yoshida S, Canas X, Irie Y, Kimura K, et al. (1999) Up-regulation of uncoupling protein 3 by thyroid hormone, peroxisome proliferator-activated receptor ligands and 9-cis retinoic acid in L6 myotubes. *FEBS Lett* 461: 319-322.
66. Lockridge JB, Sailors ML, Durgan DJ, Egbejimi O, Jeong WJ, et al. (2008) Bioinformatic profiling of the transcriptional response of adult rat cardiomyocytes to distinct fatty acids. *J Lipid Res* 49: 1395-1408.
67. Boss O, Samec S, Desplanches D, Mayet MH, Seydoux J, et al. (1998) Effect of endurance training on mRNA expression of uncoupling proteins 1, 2, and 3 in the rat. *FASEB J* 12: 335-339.
68. Cabrero A, Alegret M, Sanchez R, Adzet T, Laguna JC, et al. (1999) Etomoxir, sodium 2-[6-(4-chlorophenoxy)hexyl]oxirane-2-carboxylate, up-regulates uncoupling protein-3 mRNA levels in primary culture of rat preadipocytes. *Biochem Biophys Res Commun* 263: 87-93.
69. Kelly LJ, Vicario PP, Thompson GM, Candelore MR, Doebber TW, et al. (1998) Peroxisome proliferator-activated receptors gamma and alpha mediate in vivo regulation of uncoupling protein (UCP-1, UCP-2, UCP-3) gene expression. *Endocrinology* 139: 4920-4927.
70. Teruel T, Smith SA, Peterson J, Clapham JC (2000) Synergistic activation of UCP-3 expression in cultured fetal rat brown adipocytes by PPARalpha and PPARgamma ligands. *Biochem Biophys Res Commun* 273: 560-564.
71. Cabrero A, Alegret M, Sanchez R, Adzet T, Laguna JC, et al. (2000) Peroxisome proliferator-activated receptor alpha (PPARalpha) activators, bezafibrate and Wy-14,643, increase uncoupling protein-3 mRNA levels without modifying the mitochondrial membrane potential in primary culture of rat preadipocytes. *Arch Biochem Biophys* 380: 353-359.
72. Matsuda J, Hosoda K, Itoh H, Son C, Doi K, et al. (1998) Increased adipose expression of the uncoupling protein-3 gene by thiazolidinediones in Wistar fatty rats and in cultured adipocytes. *Diabetes* 47: 1809-1814.

73. Bugge A, Siersbaek M, Madsen MS, Gondor A, Rougier C, et al. (2010) A novel intronic peroxisome proliferator-activated receptor gamma enhancer in the uncoupling protein (UCP) 3 gene as a regulator of both UCP2 and -3 expression in adipocytes. *J Biol Chem* 285: 17310-17317.
74. Nielsen R, Pedersen TA, Hagenbeek D, Moulos P, Siersbaek R, et al. (2008) Genome-wide profiling of PPARgamma:RXR and RNA polymerase II occupancy reveals temporal activation of distinct metabolic pathways and changes in RXR dimer composition during adipogenesis. *Genes Dev* 22: 2953-2967.
75. Son C, Hosoda K, Matsuda J, Fujikura J, Yonemitsu S, et al. (2001) Up-regulation of uncoupling protein 3 gene expression by fatty acids and agonists for PPARs in L6 myotubes. *Endocrinology* 142: 4189-4194.
76. Hwang CS, Lane MD (1999) Up-regulation of uncoupling protein-3 by fatty acid in C2C12 myotubes. *Biochem Biophys Res Commun* 258: 464-469.
77. Cabrero A, Alegret M, Sanchez RM, Adzet T, Laguna JC, et al. (2000) Down-regulation of uncoupling protein-3 and -2 by thiazolidinediones in C2C12 myotubes. *FEBS Lett* 484: 37-42.
78. Pedraza N, Rosell M, Villarroya J, Iglesias R, Gonzalez FJ, et al. (2006) Developmental and tissue-specific involvement of peroxisome proliferator-activated receptor-alpha in the control of mouse uncoupling protein-3 gene expression. *Endocrinology* 147: 4695-4704.
79. Solanes G, Pedraza N, Iglesias R, Giralt M, Villarroya F (2003) Functional relationship between MyoD and peroxisome proliferator-activated receptor-dependent regulatory pathways in the control of the human uncoupling protein-3 gene transcription. *Mol Endocrinol* 17: 1944-1958.
80. Kodera Y, Takeyama K, Murayama A, Suzawa M, Masuhiro Y, et al. (2000) Ligand type-specific interactions of peroxisome proliferator-activated receptor gamma with transcriptional coactivators. *J Biol Chem* 275: 33201-33204.
81. Dressel U, Allen TL, Pippal JB, Rohde PR, Lau P, et al. (2003) The peroxisome proliferator-activated receptor beta/delta agonist, GW501516, regulates the expression of genes involved in lipid catabolism and energy uncoupling in skeletal muscle cells. *Mol Endocrinol* 17: 2477-2493.
82. Pedraza N, Solanes G, Carmona MC, Iglesias R, Vinas O, et al. (2000) Impaired expression of the uncoupling protein-3 gene in skeletal muscle during lactation: fibrates and troglitazone reverse lactation-induced downregulation of the uncoupling protein-3 gene. *Diabetes* 49: 1224-1230.
83. Forman BM, Tontonoz P, Chen J, Brun RP, Spiegelman BM, et al. (1995) 15-Deoxy-delta 12, 14-prostaglandin J2 is a ligand for the adipocyte determination factor PPAR gamma. *Cell* 83: 803-812.
84. Mottillo EP, Bloch AE, Leff T, Granneman JG (2012) Lipolytic products activate peroxisome proliferator-activated receptor (PPAR) alpha and delta in brown adipocytes to match fatty acid oxidation with supply. *J Biol Chem* 287: 25038-25048.
85. Cabrero A, Llaverias G, Roglans N, Alegret M, Sanchez R, et al. (1999) Uncoupling protein-3 mRNA levels are increased in white adipose tissue and skeletal muscle of bezafibrate-treated rats. *Biochem Biophys Res Commun* 260: 547-556.
86. Murray AJ, Panagia M, Hauton D, Gibbons GF, Clarke K (2005) Plasma free fatty acids and peroxisome proliferator-activated receptor alpha in the control of myocardial uncoupling protein levels. *Diabetes* 54: 3496-3502.
87. Finck BN, Lehman JJ, Leone TC, Welch MJ, Bennett MJ, et al. (2002) The cardiac phenotype induced by PPARalpha overexpression mimics that caused by diabetes mellitus. *J Clin Invest* 109: 121-130.
88. Luquet S, Lopez-Soriano J, Holst D, Fredenrich A, Melki J, et al. (2003) Peroxisome proliferator-activated receptor delta controls muscle development and oxidative capability. *FASEB J* 17: 2299-2301.
89. Wang YX, Zhang CL, Yu RT, Cho HK, Nelson MC, et al. (2004) Regulation of muscle fiber type and running endurance by PPARdelta. *PLoS Biol* 2: e294.
90. Tanaka T, Yamamoto J, Iwasaki S, Asaba H, Hamura H, et al. (2003) Activation of peroxisome proliferator-activated receptor delta induces fatty acid beta-oxidation in skeletal muscle and attenuates metabolic syndrome. *Proc Natl Acad Sci U S A* 100: 15924-15929.
91. Oishi Y, Manabe I, Tobe K, Ohsugi M, Kubota T, et al. (2008) SUMOylation of Kruppel-like transcription factor 5 acts as a molecular switch in transcriptional programs of lipid metabolism involving PPAR-delta. *Nat Med* 14: 656-666.
92. Cheng L, Ding G, Qin Q, Huang Y, Lewis W, et al. (2004) Cardiomyocyte-restricted peroxisome proliferator-activated receptor-delta deletion perturbs myocardial fatty acid oxidation and leads to cardiomyopathy. *Nat Med* 10: 1245-1250.
93. Allenby G, Bocquel MT, Saunders M, Kazmer S, Speck J, et al. (1993) Retinoic acid receptors and retinoid X receptors: interactions with endogenous retinoic acids. *Proc Natl Acad Sci U S A* 90: 30-34.
94. Buroker NE, Young ME, Wei C, Serikawa K, Ge M, et al. (2007) The dominant negative thyroid hormone receptor beta-mutant {Delta}337T alters PPAR{alpha} signaling in heart. *Am J Physiol Endocrinol Metab* 292: E453-460.

95. Solanes G, Pedraza N, Iglesias R, Giralt M, Villarroya F (2000) The human uncoupling protein-3 gene promoter requires MyoD and is induced by retinoic acid in muscle cells. *FASEB J* 14: 2141-2143.
96. Berry DC, Noy N (2009) All-trans-retinoic acid represses obesity and insulin resistance by activating both peroxisome proliferation-activated receptor beta/delta and retinoic acid receptor. *Mol Cell Biol* 29: 3286-3296.
97. Wong KE, Szeto FL, Zhang W, Ye H, Kong J, et al. (2009) Involvement of the vitamin D receptor in energy metabolism: regulation of uncoupling proteins. *Am J Physiol Endocrinol Metab* 296: E820-828.
98. Fromme T, Reichwald K, Platzer M, Li XS, Klingenspor M (2007) Chicken ovalbumin upstream promoter transcription factor II regulates uncoupling protein 3 gene transcription in *Phodopus sungorus*. *BMC Mol Biol* 8: 1.
99. Kim D, Jitrapakdee S, Thompson M (2007) Differential regulation of the promoter activity of the mouse UCP2 and UCP3 genes by MyoD and myogenin. *J Biochem Mol Biol* 40: 921-927.
100. Lin CS, Klingenberg M (1980) Isolation of the uncoupling protein from brown adipose tissue mitochondria. *FEBS Lett* 113: 299-303.
101. Petrovic N, Walden TB, Shabalina IG, Timmons JA, Cannon B, et al. (2010) Chronic peroxisome proliferator-activated receptor gamma (PPARgamma) activation of epididymally derived white adipocyte cultures reveals a population of thermogenically competent, UCP1-containing adipocytes molecularly distinct from classic brown adipocytes. *J Biol Chem* 285: 7153-7164.
102. Carroll AM, Haines LR, Pearson TW, Brennan C, Breen EP, et al. (2004) Immunodetection of UCP1 in rat thymocytes. *BiochemSocTrans* 32: 1066-1067.
103. Patane G, Anello M, Piro S, Vigneri R, Purrello F, et al. (2002) Role of ATP production and uncoupling protein-2 in the insulin secretory defect induced by chronic exposure to high glucose or free fatty acids and effects of peroxisome proliferator-activated receptor-gamma inhibition. *Diabetes* 51: 2749-2756.
104. Waldeck-Weiermair M, Malli R, Naghdi S, Trenker M, Kahn MJ, et al. (2010) The contribution of UCP2 and UCP3 to mitochondrial Ca(2+) uptake is differentially determined by the source of supplied Ca(2+). *Cell Calcium* 47: 433-440.
105. Tseng YH, Kokkotou E, Schulz TJ, Huang TL, Winnay JN, et al. (2008) New role of bone morphogenetic protein 7 in brown adipogenesis and energy expenditure. *Nature* 454: 1000-1004.
106. Tomlinson E, Fu L, John L, Hultgren B, Huang X, et al. (2002) Transgenic mice expressing human fibroblast growth factor-19 display increased metabolic rate and decreased adiposity. *Endocrinology* 143: 1741-1747.
107. Hondares E, Rosell M, Gonzalez FJ, Giralt M, Iglesias R, et al. (2010) Hepatic FGF21 expression is induced at birth via PPARalpha in response to milk intake and contributes to thermogenic activation of neonatal brown fat. *Cell Metab* 11: 206-212.
108. Huang H, Song TJ, Li X, Hu L, He Q, et al. (2009) BMP signaling pathway is required for commitment of C3H10T1/2 pluripotent stem cells to the adipocyte lineage. *Proc Natl Acad Sci U S A* 106: 12670-12675.
109. Reshef R, Maroto M, Lassar AB (1998) Regulation of dorsal somitic cell fates: BMPs and Noggin control the timing and pattern of myogenic regulator expression. *Genes Dev* 12: 290-303.
110. Schulz TJ, Huang P, Huang TL, Xue R, McDougall LE, et al. (2013) Brown-fat paucity due to impaired BMP signalling induces compensatory browning of white fat. *Nature* 495: 379-383.
111. Qian SW, Tang Y, Li X, Liu Y, Zhang YY, et al. (2013) BMP4-mediated brown fat-like changes in white adipose tissue alter glucose and energy homeostasis. *Proc Natl Acad Sci U S A* 110: E798-807.
112. Gesta S, Tseng YH, Kahn CR (2007) Developmental origin of fat: tracking obesity to its source. *Cell* 131: 242-256.
113. Trajkovski M, Ahmed K, Esau CC, Stoffel M (2012) MyomiR-133 regulates brown fat differentiation through Prdm16. *Nat Cell Biol* 14: 1330-1335.
114. Chen Y, Siegel F, Kipschull S, Haas B, Frohlich H, et al. (2013) miR-155 regulates differentiation of brown and beige adipocytes via a bistable circuit. *Nat Commun* 4: 1769.
115. Chau MD, Gao J, Yang Q, Wu Z, Gromada J (2010) Fibroblast growth factor 21 regulates energy metabolism by activating the AMPK-SIRT1-PGC-1alpha pathway. *Proc Natl Acad Sci U S A* 107: 12553-12558.
116. Hondares E, Iglesias R, Giralt A, Gonzalez FJ, Giralt M, et al. (2011) Thermogenic activation induces FGF21 expression and release in brown adipose tissue. *J Biol Chem* 286: 12983-12990.
117. Seale P, Kajimura S, Yang W, Chin S, Rohas LM, et al. (2007) Transcriptional control of brown fat determination by PRDM16. *Cell Metab* 6: 38-54.
118. Sun L, Xie H, Mori MA, Alexander R, Yuan B, et al. (2011) Mir193b-365 is essential for brown fat differentiation. *Nat Cell Biol* 13: 958-965.
119. Puigserver P, Wu Z, Park CW, Graves R, Wright M, et al. (1998) A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* 92: 829-839.

120. Walden TB, Petrovic N, Nedergaard J (2010) PPARalpha does not suppress muscle-associated gene expression in brown adipocytes but does influence expression of factors that fingerprint the brown adipocyte. *Biochem Biophys Res Commun* 397: 146-151.
121. Hondares E, Rosell M, Diaz-Delfin J, Olmos Y, Monsalve M, et al. (2011) Peroxisome proliferator-activated receptor alpha (PPARalpha) induces PPARgamma coactivator 1alpha (PGC-1alpha) gene expression and contributes to thermogenic activation of brown fat: involvement of PRDM16. *J Biol Chem* 286: 43112-43122.
122. Braun T, Gautel M (2011) Transcriptional mechanisms regulating skeletal muscle differentiation, growth and homeostasis. *Nat Rev Mol Cell Biol* 12: 349-361.
123. Timmons JA, Wennmalm K, Larsson O, Walden TB, Lassmann T, et al. (2007) Myogenic gene expression signature establishes that brown and white adipocytes originate from distinct cell lineages. *Proc Natl Acad Sci U S A* 104: 4401-4406.
124. Tiraby C, Tavernier G, Lefort C, Larrouy D, Bouillaud F, et al. (2003) Acquisition of brown fat cell features by human white adipocytes. *J Biol Chem* 278: 33370-33376.
125. Powelka AM, Seth A, Virbasius JV, Kiskinis E, Nicoloso SM, et al. (2006) Suppression of oxidative metabolism and mitochondrial biogenesis by the transcriptional corepressor RIP140 in mouse adipocytes. *J Clin Invest* 116: 125-136.
126. Scime A, Grenier G, Huh MS, Gillespie MA, Bevilacqua L, et al. (2005) Rb and p107 regulate preadipocyte differentiation into white versus brown fat through repression of PGC-1alpha. *Cell Metab* 2: 283-295.
127. Hansen JB, Jorgensen C, Petersen RK, Hallenborg P, De Matteis R, et al. (2004) Retinoblastoma protein functions as a molecular switch determining white versus brown adipocyte differentiation. *Proc Natl Acad Sci U S A* 101: 4112-4117.
128. Liebig M, von Praun C, Heldmaier G, Klingenspor M (2004) Absence of UCP3 in brown adipose tissue does not impair nonshivering thermogenesis. *Physiol Biochem Zool* 77: 116-126.
129. Moxley RA, Jarrett HW (2005) Oligonucleotide trapping method for transcription factor purification systematic optimization using electrophoretic mobility shift assay. *J Chromatogr A* 1070: 23-34.
130. Gadgil H, Jarrett HW (2002) Oligonucleotide trapping method for purification of transcription factors. *J Chromatogr A* 966: 99-110.
131. Towbin H, Staehelin T, Gordon J (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* 76: 4350-4354.
132. Renart J, Reiser J, Stark GR (1979) Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: a method for studying antibody specificity and antigen structure. *Proc Natl Acad Sci U S A* 76: 3116-3120.
133. Klaus S, Choy L, Champigny O, Cassard-Doulier AM, Ross S, et al. (1994) Characterization of the novel brown adipocyte cell line HIB 1B. Adrenergic pathways involved in regulation of uncoupling protein gene expression. *J Cell Sci* 107 ( Pt 1): 313-319.
134. Ross SR, Choy L, Graves RA, Fox N, Soleyjeva V, et al. (1992) Hibernoma formation in transgenic mice and isolation of a brown adipocyte cell line expressing the uncoupling protein gene. *Proc Natl Acad Sci U S A* 89: 7561-7565.
135. Graham FL, van der Eb AJ (1973) A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* 52: 456-467.
136. Felgner PL, Gadek TR, Holm M, Roman R, Chan HW, et al. (1987) Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. *Proc Natl Acad Sci U S A* 84: 7413-7417.
137. Schreiber E, Matthias P, Muller MM, Schaffner W (1989) Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. *Nucleic Acids Res* 17: 6419.
138. Jiang D, Moxley RA, Jarrett HW (2006) Promoter trapping of c-jun promoter-binding transcription factors. *J Chromatogr A* 1133: 83-94.
139. Lehmann JM, Moore LB, Smith-Oliver TA, Wilkison WO, Willson TM, et al. (1995) An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor gamma (PPAR gamma). *J Biol Chem* 270: 12953-12956.
140. Sznajdman ML, Haffner CD, Maloney PR, Fivush A, Chao E, et al. (2003) Novel selective small molecule agonists for peroxisome proliferator-activated receptor delta (PPARdelta)--synthesis and biological activity. *Bioorg Med Chem Lett* 13: 1517-1521.
141. Yen PM (2001) Physiological and molecular basis of thyroid hormone action. *Physiol Rev* 81: 1097-1142.
142. Keller H, Devchand PR, Perroud M, Wahli W (1997) PPAR alpha structure-function relationships derived from species-specific differences in responsiveness to hypolipidemic agents. *Biol Chem* 378: 651-655.

143. Lehmann JM, Lenhard JM, Oliver BB, Ringold GM, Kliewer SA (1997) Peroxisome proliferator-activated receptors alpha and gamma are activated by indomethacin and other non-steroidal anti-inflammatory drugs. *J Biol Chem* 272: 3406-3410.
144. Barnett J, Chow J, Ives D, Chiou M, Mackenzie R, et al. (1994) Purification, characterization and selective inhibition of human prostaglandin G/H synthase 1 and 2 expressed in the baculovirus system. *Biochim Biophys Acta* 1209: 130-139.
145. Johnson JL, Wimsatt J, Buckel SD, Dyer RD, Maddipati KR (1995) Purification and characterization of prostaglandin H synthase-2 from sheep placental cotyledons. *Arch Biochem Biophys* 324: 26-34.
146. Karmakar S, Mahajan MC, Schulz V, Boyapaty G, Weissman SM (2010) A multiprotein complex necessary for both transcription and DNA replication at the beta-globin locus. *EMBO J* 29: 3260-3271.
147. Jarbouli MA, Wynne K, Elia G, Hall WW, Gautier VW (2011) Proteomic profiling of the human T-cell nucleolus. *Mol Immunol* 49: 441-452.
148. Park JH, Kang HJ, Kang SI, Lee JE, Hur J, et al. (2013) A multifunctional protein, EWS, is essential for early brown fat lineage determination. *Dev Cell* 26: 393-404.
149. Dolfini D, Mantovani R (2013) YB-1 (YBX1) does not bind to Y/CCAAT boxes in vivo. *Oncogene* 32: 4189-4190.
150. Kuhn A, Voit R, Stefanovsky V, Evers R, Bianchi M, et al. (1994) Functional differences between the two splice variants of the nucleolar transcription factor UBF: the second HMG box determines specificity of DNA binding and transcriptional activity. *EMBO J* 13: 416-424.
151. Mariappan MM, D'Silva K, Lee MJ, Sataranatarajan K, Barnes JL, et al. (2011) Ribosomal biogenesis induction by high glucose requires activation of upstream binding factor in kidney glomerular epithelial cells. *Am J Physiol Renal Physiol* 300: F219-230.
152. Young DW, Hassan MQ, Pratap J, Galindo M, Zaidi SK, et al. (2007) Mitotic occupancy and lineage-specific transcriptional control of rRNA genes by Runx2. *Nature* 445: 442-446.
153. Sun H, Tu X, Liu M, Baserga R (2007) Dual regulation of upstream binding factor 1 levels by IRS-1 and ERKs in IGF-1-receptor signaling. *J Cell Physiol* 212: 780-786.
154. Drakas R, Tu X, Baserga R (2004) Control of cell size through phosphorylation of upstream binding factor 1 by nuclear phosphatidylinositol 3-kinase. *Proc Natl Acad Sci U S A* 101: 9272-9276.
155. Grueneberg DA, Pablo L, Hu KQ, August P, Weng Z, et al. (2003) A functional screen in human cells identifies UBF2 as an RNA polymerase II transcription factor that enhances the beta-catenin signaling pathway. *Mol Cell Biol* 23: 3936-3950.
156. Valineva T, Yang J, Palovuori R, Silvennoinen O (2005) The transcriptional co-activator protein p100 recruits histone acetyltransferase activity to STAT6 and mediates interaction between the CREB-binding protein and STAT6. *J Biol Chem* 280: 14989-14996.
157. Maloveryan A, Finta C, Osterlund T, Kogerman P (2007) A possible role of mouse Fused (STK36) in Hedgehog signaling and Gli transcription factor regulation. *J Cell Commun Signal* 1: 165-173.
158. McDermott A, Gustafsson M, Elsam T, Hui CC, Emerson CP, Jr., et al. (2005) Gli2 and Gli3 have redundant and context-dependent function in skeletal muscle formation. *Development* 132: 345-357.
159. James AW, Leucht P, Levi B, Carre AL, Xu Y, et al. (2010) Sonic Hedgehog influences the balance of osteogenesis and adipogenesis in mouse adipose-derived stromal cells. *Tissue Eng Part A* 16: 2605-2616.
160. Suh JM, Gao X, McKay J, McKay R, Salo Z, et al. (2006) Hedgehog signaling plays a conserved role in inhibiting fat formation. *Cell Metab* 3: 25-34.
161. Park JH, Lee JJ, Yoon S, Lee JS, Choe SY, et al. (2001) Genomic cloning of the Hsc71 gene in the hermaphroditic teleost *Rivulus marmoratus* and analysis of its expression in skeletal muscle: identification of a novel muscle-preferred regulatory element. *Nucleic Acids Res* 29: 3041-3050.
162. Duffy DJ, Millane RC, Frank U (2012) A heat shock protein and Wnt signaling crosstalk during axial patterning and stem cell proliferation. *Dev Biol* 362: 271-281.
163. Endoh H, Maruyama K, Masuhiro Y, Kobayashi Y, Goto M, et al. (1999) Purification and identification of p68 RNA helicase acting as a transcriptional coactivator specific for the activation function 1 of human estrogen receptor alpha. *Mol Cell Biol* 19: 5363-5372.
164. Jensen ED, Niu L, Caretti G, Nicol SM, Teplyuk N, et al. (2008) p68 (Ddx5) interacts with Runx2 and regulates osteoblast differentiation. *J Cell Biochem* 103: 1438-1451.
165. Caretti G, Schiltz RL, Dilworth FJ, Di Padova M, Zhao P, et al. (2006) The RNA helicases p68/p72 and the noncoding RNA SRA are coregulators of MyoD and skeletal muscle differentiation. *Dev Cell* 11: 547-560.
166. Katano-Toki A, Satoh T, Tomaru T, Yoshino S, Ishizuka T, et al. (2013) THRAP3 interacts with HELZ2 and plays a novel role in adipocyte differentiation. *Mol Endocrinol* 27: 769-780.
167. Hoffmann C, Zimmermann A, Hinney A, Volckmar AL, Jarrett HW, et al. (2013) A Novel SP1/SP3 Dependent Intronic Enhancer Governing Transcription of the UCP3 Gene in Brown Adipocytes. *PLoS One* 8: e83426.



168. Aruga J, Nagai T, Tokuyama T, Hayashizaki Y, Okazaki Y, et al. (1996) The mouse zic gene family. Homologues of the *Drosophila* pair-rule gene odd-paired. *J Biol Chem* 271: 1043-1047.
169. Siersbaek MS, Loft A, Aagaard MM, Nielsen R, Schmidt SF, et al. (2012) Genome-wide profiling of peroxisome proliferator-activated receptor gamma in primary epididymal, inguinal, and brown adipocytes reveals depot-selective binding correlated with gene expression. *Mol Cell Biol* 32: 3452-3463.
170. Nakashima T, Sun SY, Lotan R, Fujiwara T, Yasumatsu R, et al. (2001) All-trans-retinoic acid enhances the effect of adenovirus-mediated wild-type p53 gene transfer in head and neck squamous cell carcinoma. *Laryngoscope* 111: 1459-1464.
171. Styner M, Sen B, Xie Z, Case N, Rubin J (2010) Indomethacin promotes adipogenesis of mesenchymal stem cells through a cyclooxygenase independent mechanism. *J Cell Biochem* 111: 1042-1050.
172. Qin JY, Zhang L, Clift KL, Huler I, Xiang AP, et al. (2010) Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *PLoS One* 5: e10611.
173. Claussnitzer M, Dankel SN, Klocke B, Grallert H, Glunk V, et al. (2014) Leveraging cross-species transcription factor binding site patterns: from diabetes risk Loci to disease mechanisms. *Cell* 156: 343-358.
174. Suske G (1999) The Sp-family of transcription factors. *Gene* 238: 291-300.
175. Stielow B, Sapetschnig A, Wink C, Kruger I, Suske G (2008) SUMO-modified Sp3 represses transcription by provoking local heterochromatic gene silencing. *EMBO Rep* 9: 899-906.
176. Valin A, Gill G (2007) Regulation of the dual-function transcription factor Sp3 by SUMO. *Biochem Soc Trans* 35: 1393-1396.
177. Philipsen S, Suske G (1999) A tale of three fingers: the family of mammalian Sp/XKLF transcription factors. *Nucleic Acids Res* 27: 2991-3000.
178. Raiber EA, Kranaster R, Lam E, Nikan M, Balasubramanian S (2012) A non-canonical DNA structure is a binding motif for the transcription factor SP1 in vitro. *Nucleic Acids Res* 40: 1499-1508.
179. McEwen DG, Ornitz DM (1998) Regulation of the fibroblast growth factor receptor 3 promoter and intron I enhancer by Sp1 family transcription factors. *J Biol Chem* 273: 5349-5357.
180. Bianchi M, Crinelli R, Giacomini E, Carloni E, Magnani M (2009) A potent enhancer element in the 5'-UTR intron is crucial for transcriptional regulation of the human ubiquitin C gene. *Gene* 448: 88-101.
181. Singh AK, Battu A, Mohareer K, Hasnain SE, Ehtesham NZ (2010) Transcription of human resistin gene involves an interaction of Sp1 with peroxisome proliferator-activating receptor gamma (PPARgamma). *PLoS One* 5: e9912.
182. Okazaki M, Iwasaki Y, Nishiyama M, Taguchi T, Tsugita M, et al. (2010) PPARbeta/delta regulates the human SIRT1 gene transcription via Sp1. *Endocr J* 57: 403-413.
183. Aguilo F, Camarero N, Relat J, Marrero PF, Haro D (2010) Transcriptional regulation of the human acetoacetyl-CoA synthetase gene by PPARgamma. *Biochem J* 427: 255-264.
184. Sugawara A, Uruno A, Kudo M, Ikeda Y, Sato K, et al. (2002) Transcription suppression of thromboxane receptor gene by peroxisome proliferator-activated receptor-gamma via an interaction with Sp1 in vascular smooth muscle cells. *J Biol Chem* 277: 9676-9683.
185. Billon N, Carlisi D, Datto MB, van Grunsven LA, Watt A, et al. (1999) Cooperation of Sp1 and p300 in the induction of the CDK inhibitor p21WAF1/CIP1 during NGF-mediated neuronal differentiation. *Oncogene* 18: 2872-2882.
186. Kumar P, Tripathi S, Pandey KN (2014) Histone Deacetylase Inhibitors Modulate the Transcriptional Regulation of Guanylyl Cyclase/Natriuretic Peptide Receptor-A Gene: Interactive roles of Modified Histones, HATS, p300, and Sp1. *J Biol Chem*.
187. Li J, Wang W, Liu C, Wang W, Li W, et al. (2013) Critical role of histone acetylation by p300 in human placental 11beta-HSD2 expression. *J Clin Endocrinol Metab* 98: E1189-1197.
188. Azahri NS, Di Bartolo BA, Khachigian LM, Kavurma MM (2012) Sp1, acetylated histone-3 and p300 regulate TRAIL transcription: mechanisms of PDGF-BB-mediated VSMC proliferation and migration. *J Cell Biochem* 113: 2597-2606.
189. Ikeda K, Nagano K, Kawakami K (1993) Possible implications of Sp1-induced bending of DNA on synergistic activation of transcription. *Gene* 136: 341-343.
190. Deshane J, Kim J, Bolisetty S, Hock TD, Hill-Kapturczak N, et al. (2010) Sp1 regulates chromatin looping between an intronic enhancer and distal promoter of the human heme oxygenase-1 gene in renal cells. *J Biol Chem* 285: 16476-16486.
191. Funk WD, Wright WE (1992) Cyclic amplification and selection of targets for multicomponent complexes: myogenin interacts with factors recognizing binding sites for basic helix-loop-helix, nuclear factor 1, myocyte-specific enhancer-binding factor 2, and COMP1 factor. *Proc Natl Acad Sci U S A* 89: 9484-9488.
192. Gronostajski RM (2000) Roles of the NFI/CTF gene family in transcription and development. *Gene* 249: 31-45.

193. Kruse U, Sippel AE (1994) Transcription factor nuclear factor I proteins form stable homo- and heterodimers. *FEBS Lett* 348: 46-50.
194. Richard AJ, Stephens JM (2014) The role of JAK-STAT signaling in adipose tissue function. *Biochim Biophys Acta* 1842: 431-439.
195. Cao Y, Yao Z, Sarkar D, Lawrence M, Sanchez GJ, et al. (2010) Genome-wide MyoD binding in skeletal muscle cells: a potential for broad cellular reprogramming. *Dev Cell* 18: 662-674.
196. Rohlf C, Ahmad S, Borellini F, Lei J, Glazer RI (1997) Modulation of transcription factor Sp1 by cAMP-dependent protein kinase. *J Biol Chem* 272: 21137-21141.
197. Wang YT, Chuang JY, Shen MR, Yang WB, Chang WC, et al. (2008) Sumoylation of specificity protein 1 augments its degradation by changing the localization and increasing the specificity protein 1 proteolytic process. *J Mol Biol* 380: 869-885.
198. Yang X, Su K, Roos MD, Chang Q, Paterson AJ, et al. (2001) O-linkage of N-acetylglucosamine to Sp1 activation domain inhibits its transcriptional capability. *Proc Natl Acad Sci U S A* 98: 6611-6616.
199. Abdelrahim M, Safe S (2005) Cyclooxygenase-2 inhibitors decrease vascular endothelial growth factor expression in colon cancer cells by enhanced degradation of Sp1 and Sp4 proteins. *Mol Pharmacol* 68: 317-329.
200. Frigo DE, Sherk AB, Wittmann BM, Norris JD, Wang Q, et al. (2009) Induction of Kruppel-like factor 5 expression by androgens results in increased CXCR4-dependent migration of prostate cancer cells in vitro. *Mol Endocrinol* 23: 1385-1396.
201. Yamamoto K, Sakaguchi M, Medina RJ, Niida A, Sakaguchi Y, et al. (2010) Transcriptional regulation of a brown adipocyte-specific gene, UCP1, by KLF11 and KLF15. *Biochem Biophys Res Commun* 400: 175-180.
202. Zhang H, Chen Q, Yang M, Zhu B, Cui Y, et al. (2013) Mouse KLF11 regulates hepatic lipid metabolism. *J Hepatol* 58: 763-770.
203. Yin KJ, Fan Y, Hamblin M, Zhang J, Zhu T, et al. (2013) KLF11 mediates PPARgamma cerebrovascular protection in ischaemic stroke. *Brain* 136: 1274-1287.
204. Mori T, Sakaue H, Iguchi H, Gomi H, Okada Y, et al. (2005) Role of Kruppel-like factor 15 (KLF15) in transcriptional regulation of adipogenesis. *J Biol Chem* 280: 12867-12875.
205. Prosdocimo DA, Anand P, Liao X, Zhu H, Shelkay S, et al. (2014) Kruppel-like factor 15 is a critical regulator of cardiac lipid metabolism. *J Biol Chem*.
206. Heinaniemi M, Uski JO, Degenhardt T, Carlberg C (2007) Meta-analysis of primary target genes of peroxisome proliferator-activated receptors. *Genome Biol* 8: R147.
207. Kliewer SA, Umesono K, Mangelsdorf DJ, Evans RM (1992) Retinoid X receptor interacts with nuclear receptors in retinoic acid, thyroid hormone and vitamin D3 signalling. *Nature* 355: 446-449.
208. Pan J, Fukuda K, Saito M, Matsuzaki J, Kodama H, et al. (1999) Mechanical stretch activates the JAK/STAT pathway in rat cardiomyocytes. *Circ Res* 84: 1127-1136.
209. Graves RA, Tontonoz P, Ross SR, Spiegelman BM (1991) Identification of a potent adipocyte-specific enhancer: involvement of an NF-1-like factor. *Genes Dev* 5: 428-437.
210. Spitz F, Salminen M, Demignon J, Kahn A, Daegelen D, et al. (1997) A combination of MEF3 and NFI proteins activates transcription in a subset of fast-twitch muscles. *Mol Cell Biol* 17: 656-666.
211. Trenerry MK, Della Gatta PA, Cameron-Smith D (2011) JAK/STAT signaling and human in vitro myogenesis. *BMC Physiol* 11: 6.
212. Londhe P, Davie JK (2011) Sequential association of myogenic regulatory factors and E proteins at muscle-specific genes. *Skelet Muscle* 1: 14.

## **F. Acknowledgements**

Zuerst danke ich Professor Dr. Martin Klingenspor und Dr. Tobias Fromme für die Bereitstellung des Themas, die Betreuung und all die andere, vielfältige Unterstützung bei dieser Doktorarbeit.

Besonderer Dank gebührt Anika Zimmermann für ihre Unterstützung im Labor und die lange, gute und angenehme Zusammenarbeit. Dies beinhaltet auch Dank dafür, dass du mir manchmal in den Hintern getreten hast, natürlich immer unverdient.

Ebenfalls danken möchte den Professoren Bernhard Küster und Harry W. Jarrett danken, sowie ihren Teams: Prof. Küster und seinem Team für die Durchführung der MS Identifikationen, und Prof. Jarrett für das sehr lehrreiche und angenehme Praktikum in den USA.

Des Weiteren möchte ich der gesamten Arbeitsgruppe danken, vor allem meinen Bürogenossen, für viele interessante wissenschaftliche Diskussionen, gute Ideen und auch das ein ums andere Mal tatkräftige Hilfe.

Nicht minder dankbar bin ich für das angenehme Arbeitsumfeld, moralische Unterstützung, und in manchen Situationen auch für Geduld und Verständnis. Dies gilt nicht nur für meine Arbeitsgruppe, sondern für alle im Institut.

Zuletzt, möchte ich mich bei den Menschen aus meinem privaten Umfeld bedanken, allen voran bei meinen Eltern, aber natürlich auch meinen vielen, guten Freunden. Ihr habt mir ein Umfeld und Rückhalt gegeben, wie man es sich nur wünschen kann. Auch diejenigen, die kein Wort, keinen Handschlag und keinen Gedanken zu meiner Doktorarbeit beigetragen haben, waren dadurch wichtig für diese.

## Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die bei der promotionsführenden Einrichtung Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der TUM zur Promotionsprüfung vorgelegte Arbeit mit dem Titel:

### **A novel SP1/SP3 dependent intronic enhancer governing transcription of the UCP3 gene**

am Lehrstuhl für Molekulare Ernährungsmedizin unter der Anleitung und Betreuung durch Univ.-Prof. Dr. Martin Klingenspor ohne sonstige Hilfe erstellt und bei der Abfassung nur die gemäß § 6 Abs. 6 und 7 Satz 2 angegebenen Hilfsmittel benutzt habe.

Ich habe keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht, oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.

Ich habe die Dissertation in dieser oder ähnlicher Form in keinem anderen Prüfungsverfahren als Prüfungsleistung vorgelegt.

Die vollständige Dissertation wurde noch nicht veröffentlicht.

Ich habe den angestrebten Doktorgrad **noch nicht** erworben und bin **nicht** in einem früheren Promotionsverfahren für den angestrebten Doktorgrad endgültig gescheitert.

Die öffentlich zugängliche Promotionsordnung der TUM ist mir bekannt, insbesondere habe ich die Bedeutung von § 28 (Nichtigkeit der Promotion) und § 29 (Entzug des Doktorgrades) zur Kenntnis genommen. Ich bin mir der Konsequenzen einer falschen Eidesstattlichen Erklärung bewusst.

Mit der Aufnahme meiner personenbezogenen Daten in die Alumni-Datei bei der TUM bin ich einverstanden.

Freising, den .....

.....

Unterschrift

# Lebenslauf

## Persönliche Angaben

Name: Christoph Hoffmann  
Geburtsdatum: 30. September 1981  
Nationalität: Deutsch  
Beziehungsstatus: Unverheiratet

## Schulbildung

1988 – 1992: Don Bosco Grundschule Wennemen  
1992 – 2001: Städtisches Gymnasium Meschede  
2001 – 2008: Philipps-Universität Marburg, Biologie

## Wissenschaftlicher Werdegang

08/2001 – 06/2003: Grundstudium Biologie (Note: 1,3)  
08/2003 – 07/2006: Hauptstudium: Tierphysiologie, Genetik und Biochemie  
08/2006 – 06/2007: Diplomarbeit – „Protein-DNA Interaktion am ersten Intron des“  
UCP3 Genes (Note: 1,0)  
Philipps-Universität Marburg, Stoffwechselphysiologie, Prof.  
Gerhard Heldmaier  
07/2007 – 03/2008: Wissenschaftlicher Mitarbeiter, Philipps-Universität Marburg,  
Stoffwechselphysiologie, Prof. Gerhard Heldmaier  
04/2007 – 04/2014: Doktorarbeit – “A novel SP1/SP3 dependent intronic enhancer  
governing transcription of the UCP3 gene”  
TU München, EKfZ, Molekulare Ernährungsmedizin, Prof.  
Martin Klingenspor

## Sonstige wissenschaftliche Qualifikationen

- 2007 – 2013: Betreuung von einer Diplomarbeit, einer Masterarbeit und fünf Bachelorarbeiten
- 07/2007 – 10/2007: Verfassung eines DFG-Drittmittelantrages: KL 973/8-1, bewilligt 03/2008
- 08/2008 – 11/2008: Wissenschaftliches Praktikum – Transkriptionsfaktor-DNA Interaktion und DNA-Affinitätschromatographie  
University of Texas at San Antonio, Department of Chemistry,  
Prof. Harry Jarrett

## Konferenzen & Workshops

- 2007: 2D Differential Gel Electrophoresis workshop, Münster
- 2009: Arbeitstagung Mikromethoden der Proteinbiochemie, Martinsried
- 2011: Novel insights in adipose cell functions, Fondation IPSEN, Paris
- 2012: EBEC2012: European Bioenergetics Conference, Freiburg;  
Vortrag – Session S5: UCPs
- 2013: Keystone Symposia: Adipose Tissue Biology, Keystone, Colorado; Poster Präsentation

## Wissenschaftliche Veröffentlichungen

1. **A novel SP1/SP3 dependent intronic enhancer governing transcription of the UCP3 gene in brown adipocytes;** Hoffmann, [...], Klingenspor; *PLOS one*; 2013 Dec 31;8(12):e83426
2. **Leveraging cross-species transcription factor binding site patterns: from diabetes risk loci to disease mechanisms;** Claussnitzer, [...], Hoffmann, [...], Laumen; *Cell*; Manuscript accepted Sept. 11<sup>th</sup>, 2013
3. **An intronic single base exchange leads to a brown adipose tissue-specific loss of Ucp3 expression and an altered body mass trajectory;** Fromme, Hoffmann, [...], Klingenspor; *Physiol Genomics*; 2009 Jun 10;38(1):54-62.