



universität
wien

MASTERARBEIT

Titel der Masterarbeit

„Sedoheptulose Kinase CARKL: Identification of
Alternative mRNA Transcripts and Initial
Characterization of Transgenic Mice“

verfasst von

Csörsz Nagy, BSc

angestrebter akademischer Grad

Master of Science (MSc)

Wien, 2013

Studienkennzahl lt. Studienblatt:

A 066 834

Studienrichtung lt. Studienblatt:

Masterstudium Molekulare Biologie

Betreut von:

Univ.-Prof. Dr. Thomas Decker

Abstract

Outstanding research achievements of the last decades underpin the high impact of metabolic regulation and bioenergetic reprogramming on many diseases, ranging from diabetes and atherosclerosis to cancer and inflammatory disorders. Unbalanced long-term nutrition, dominated by the heavy consumption of refined carbohydrates and fats, is closely linked to such alterations in cellular metabolism.

The recent discovery of the sedoheptulose kinase CARKL revealed that sedoheptulose, a natural ketoheptose present in fruits and vegetables, is an accessible carbon source in humans, as well as animals. CARKL phosphorylates sedoheptulose to sedoheptulose-7-phosphate and makes it available for subsequent cellular reactions. Importantly, CARKL counteracts proinflammatory polarization of macrophages through the modulation of carbon flux and cellular redox state.

In this thesis, I present the theoretical implications of cellular sedoheptulose metabolism, which displays interesting analogies to glucose metabolism. The sedoheptulose kinase CARKL appears to be located at a strategic position to regulate cellular carbon flux at the interface of the glycolysis and the non-oxidative pentose phosphate pathway. In experiments, I started to investigate CARKL mRNA diversity, which resulted in the identification of mRNA transcripts in metabolic-active organs with alternative 3'UTRs. The 3'UTR region of CARKL contains many putative regulatory sites and thereby constitutes a central element to understand CARKL gene expression, mRNA stability and translation. Furthermore, I was involved in the generation of transgenic mice overexpressing CARKL to investigate the function of this kinase under physiological conditions. Four transgenic mouse lines with different CARKL protein levels were established during my thesis. These transgenic mice are now available as a powerful tool to study the function of CARKL in various *in vivo* settings, ranging from nutrition to inflammation and various disease models.

New insights into CARKL regulation and the metabolic networks influenced by this kinase might be beneficial for the flourishing field of heptose biology and might further contribute to a better understanding of diseases linked to changes in carbohydrate metabolism.

Table of contents

Abstract	3
Abbreviations	6
Introduction	8
Background on heptose metabolism	10
Carbohydrate metabolism	10
Sedoheptulose & CARKL	13
Heptoses in primary carbohydrate metabolism	14
CARKL & metabolic regulation	17
The physiological role of CARKL	19
Aims of the master thesis	20
Part 1 – Novel CARKL transcripts	21
Background	21
Regulatory mechanisms of gene expression	21
The CARKL gene	23
The CARKL protein	24
Regulatory mechanisms acting on mRNA	24
Rapid amplification of cDNA ends	28
Aim	30
Results	30
Alternative 3'UTRs of human CARKL	30
Regulator elements of the human CARKL 3'UTR	33
Alternative 3'UTRs of murine CARKL	37
Regulatory elements of the murine CARKL 3'UTR	39
Discussion:	44
Part 2 – Transgenic CARKL mice	49
Background	49
CARKL regulates macrophage polarization in vitro	49
Generation of transgenic CARKL-mice	51
Aim	53
Results	53
Genotyping strategy	53
CARKL protein expression in transgenic mice	57
Discussion	60
Material and Methods	62
Chemicals	62
BMDM isolation	63
Human monocyte isolation	63
Cell culture	64
RNA isolation	64
RNA integrity analysis	64
5' RACE	65
3' RACE	65
Nested RACE-PCR	65
Agarose gel electrophoresis	66

Gel purification	67
Cloning	67
Plasmid preparation & restriction analysis	67
Sequence analysis & processing	68
Mouse tail DNA isolation	68
Mouse genotyping – PCR	68
Mouse genotyping – real time PCR	69
Reverse transcription	70
Real time RT-PCR	71
Protein isolation	72
Western Blotting	72
Supplementary figures	74
Supplement: Zusammenfassung in Deutsch	82
References	83
Acknowledgements	88
Curriculum vitae	89

Abbreviations

ALDO	-	aldolase
ATP	-	adenosine triphosphate
bp	-	base pair
CARKL	-	<u>car</u> bohydrate <u>k</u> inase <u>l</u> ike (Sedoheptulose kinase)
cDNA	-	copy DNA
CIP	-	calf intestine phosphatase
CPE	-	cytoplasmic polyadenylation element
CTNS	-	cystinosin
DNA	-	deoxyribonucleic acid
DHAP	-	dihydroxyacetone phosphate
E4P	-	erythrose 4-phosphate
F1P	-	fructose 1-phosphate
F1,6bP	-	fructose 1,6-bisphosphate
F2,6bP	-	fructose 2,6-bisphosphate
F6P	-	fructose 6-phosphate
FBPase	-	fructose bisphosphatase
G1,6bP	-	glucose 1,6-bisphosphate
G3P	-	glyceraldehyde 3-phosphate
G6P	-	glucose 6-phosphate
Gal1P	-	galactose 1-phosphate
GSH	-	reduced glutathione
GSSG	-	oxidized glutathione
HK	-	hexokinase
IL	-	interleukin
IRES	-	internal ribosome entry site
KHK	-	ketoheokinase
LPS	-	lipopolysaccharide
miRNA	-	micro RNA
MMLV	-	Moloney murine leukemia virus
mRNA	-	messenger RNA
NAD+	-	nicotineamide adenine dinucleotide
NADH	-	reduced nicotineamide adenine dinucleotide
NADP+	-	nicotineamide adenine dinucl. phosphate
NADPH	-	reduced nicotineamide adenine dinucl. Phosphate
nt	-	nucleotide
ORF	-	open reading frame
PCR	-	polymerase chain reaction
PFK	-	phosphofructokinase
PPP	-	pentose phosphate pathway
PTMs	-	post-translational modifications
R5P	-	ribose 5-phosphate
RACE	-	<u>r</u> apid <u>a</u> mplification of <u>c</u> DNA <u>e</u> nds
RIN	-	RNA integrity number
RLM-RACE	-	RNA ligase mediated-RACE
RNA	-	ribonucleic acid
RNA pol	-	RNA polymerase
ROS	-	reactive oxygen species
RT	-	reverse transcriptase

Ru5P	-	ribulose 5-phosphate
S7P	-	sedoheptulose 7-phosphate
S1,7bP	-	sedoheptulose 1,7-bisphosphate
SHPK	-	sedoheptulose kinase (=CARKL)
TAP	-	tobacco acid pyrophosphatase
TCA	-	tricarboxylic acid cycle
TNF	-	tumour necrosis factor (alpha)
UTR	-	untranslated region
X5P	-	xylulose 5-phosphate

Introduction

Heptoses are monosaccharides consisting of seven carbon atoms. They are part of our daily diet and endogenously produced constituents of many organisms. In contrast to the well-known family of hexoses, to which glucose and fructose belong to, the contribution of heptoses to metabolism and health is still poorly investigated.

The importance of heptose metabolites was realized in the 1950's, when Melvin Calvin and colleagues identified sedoheptulose 7-phosphate (S7P) as an essential intermediate of photosynthesis in plants [1]. This was followed by the discovery of the pentose phosphate pathway (PPP) by Bernard Leonard Horecker and co-workers, a metabolic pathway present in most known life-forms, which also utilizes S7P [2-4]. However, since these groundbreaking findings, the biology of heptoses has been largely neglected by scientists over the last century.

The recent discovery of the sedoheptulose kinase CARKL revealed sedoheptulose as an accessible carbon source in humans and various animals [5-7]. Phosphorylation of sedoheptulose by CARKL yields S7P, a carbohydrate phosphate that holds a strategic position to regulate carbon flux at the interface of glycolysis and the non-oxidative pentose phosphate pathway (non-ox PPP) [5, 8]

Both pathways, glycolysis and the pentose phosphate pathway, represent major routes of carbohydrate metabolism and are connected to each other by sharing intermediate molecules. Glucose, used as the primary energy source in most cell types, can be either consumed via glycolysis to produce pyruvate, adenosine triphosphate (ATP) and reduced nicotinamide adenine dinucleotide (NADH) or enter the oxidative branch of the PPP (ox-PPP). The latter generates ribulose 5-phosphate, reduced nicotinamide adenine dinucleotide phosphate (NADPH) and CO₂ through three irreversible reactions [9]. Ribulose 5-phosphate is further used for the synthesis of ribose 5-phosphate (R5P) and nucleotides.

The non-ox PPP consists of fully reversible reactions catalysed by transaldolase and transketolase. These can reconvert C5 bodies into the glycolytic intermediates fructose 6-phosphate (F6P) and glyceraldehyde 3-phosphate (G3P). Alternatively, the non-ox PPP can use F6P and G3P to work in reverse direction and generate C5 bodies without the necessity to meanwhile produce NADPH and CO₂. Depending on momentary systemic needs, co-ordinated carbon scrambling events between these pathways allow to individually adjust cellular redox status as well as precursor formation.

CARKL derived S7P, together with its reaction partner G3P can participate in two distinct reaction of the non-ox PPP. They can be either converted to R5P and xylulose 5-phosphate (X5P) by transketolase or they can be utilized by transaldolase to generate F6P and erythrose 4-phosphate (E4P) [10]. Since S7P might be the rate-limiting substrate in these reactions, its supply regulated by CARKL activity is discussed to modulate carbon scrambling events, non-ox PPP flux and metabolic reprogramming [5, 8, 11].

Bioenergetic adaptation and dynamic metabolic reprogramming are crucial for cellular function and tightly associated with transformation events such as tumorigenesis or immune cell activation [12-14]. Importantly, CARKL was found to counteract pro-inflammatory-type polarization of macrophages, which relies on aerobic glycolysis, while supporting the development of an alternative, anti-inflammatory phenotype, mainly utilizing oxidative phosphorylation [5]. This finding highlights the potential of sedoheptulose kinase to influence cell fate by balancing carbohydrate metabolism and redox regulation.

The background on heptose metabolism, discussed below, is based on a mini-review with the title “Sedoheptulose kinase regulates cellular carbohydrate metabolism by sedoheptulose 7-phosphate supply”, which we wrote during my master thesis. It was published in the journal *Biochemical Society Transactions*, volume 41, part 2 (2013).

In experiments, I started to explore the posttranscriptional regulation of CARKL. Using RACE (rapid amplification of cDNA ends), I could identify novel transcripts with shorter 3'UTRs (untranslated regions) in mouse liver and kidney, as well as in human liver. Mouse bone marrow derived macrophages (BMDMs) and macrophages differentiated from human blood monocytes expressed only the full length CARKL mRNA, containing several regulatory elements within the 3'UTR. These results suggest that metabolically active organs and macrophages may control CARKL expression through different regulatory elements within the 3' untranslated region.

During my master thesis, I also planned and performed the breeding of the first transgenic mouse lines, overexpressing CARKL. I confirmed the transgenic status of these animals on DNA, mRNA and protein levels. In an initial characterization attempt, the transgenic mice didn't show obvious abnormalities, implicating that CARKL overexpression doesn't negatively affect organogenesis. Up to date, our knowledge about the impact of sedoheptulose and CARKL on carbohydrate metabolism, as well as inflammation is very limited. Many questions remain still unanswered, ranging from the exact regulation of CARKL to its physiological role *in vivo*, finally to a possible translation to health, nutrition and medicine.

Background on heptose metabolism

Carbohydrate metabolism

Carbohydrates serve as the main energy source of most cell types, while providing structural building blocks essential for many metabolic processes. Glycolysis and the PPP represent two central pathways of carbohydrate metabolism. Together they utilize different monosaccharides to generate ATP, various reducing equivalents, as well as precursors necessary for the synthesis of critical molecules, such as amino-, nucleic- and fatty-acids.

Beside glucose, which is a very abundant carbohydrate, also several other carbohydrate molecules provide energy and carbon units. The best studied examples are represented by fructose and galactose, which are converted to glycolytic intermediates during metabolization. Our group recently demonstrated that also sedoheptulose is an accessible carbon source in humans [unpublished data]. For convenience, the focus in this overview-chapter is on the metabolic path of glucose.

After cellular uptake by transporter proteins, glucose is phosphorylated by hexokinase (HK) to glucose 6-phosphate (G6P) and mainly consumed via a conserved pathway termed glycolysis (a.k.a. Embden-Meyerhof-Parnas pathway). Glucose is converted to two three carbon (C3) bodies, G3P and dihydroxyacetone phosphate (DHAP), which are finally metabolised to pyruvate, while generating ATP and NADH. Under aerobic conditions, using O₂ as terminal electron acceptor, pyruvate can be completely oxidized to CO₂ and H₂O through the tricarboxylic acid cycle (TCA) and mitochondrial oxidative phosphorylation. Aerobic respiration generates a maximal yield of ATP per molecule of glucose and is the preferred metabolic program of non-proliferating, differentiated cells.

In the case of low oxygen tension or in the absence of oxygen (anaerobic conditions), pyruvate is mainly fermented to lactate (animals) or ethanol (yeast), in order to form ATP and regenerate cellular NAD⁺, which is an essential cofactor of glycolysis. Interestingly, some cells have a preference for this metabolic-route even if ample oxygen is available. This phenomenon is known as aerobic glycolysis or “the Warburg effect”. Although aerobic glycolysis is inefficient in producing ATP, it is an ideal pathway to incorporate nutrients into biomass and therefore favoured by actively proliferating cells, such as cancer cells [13].

The second main route for glucose consumption is the PPP, which can be divided into the oxidative and non-oxidative branches [15]. The oxidative branch enables cells to convert

glucose 6-phosphate into ribose 5-phosphate, the precursor for nucleotide synthesis. Simultaneously NADPH is produced, which is a reducing agent involved in many cellular processes. NADPH is an important coenzyme for anabolic reactions such as triglyceride, phospholipid and cholesterol synthesis. Furthermore, it is used to regenerate reduced glutathione (GSH) from oxidized glutathione (GSSG) pools and thereby essential for proper redox-regulation. In immune cells, NADPH behaves like a double-edged sword, as it can not only protect cells from reactive oxygen species (ROS) by the regeneration of anti-oxidative GSH, but also aid in forming ROS produced by NADPH-dependent oxidases [16].

The non-oxidative arm of the PPP can reconvert C5 bodies, which are not required for biosynthetic reactions, back into glycolytic intermediates. As a result of its reversible nature, the non-ox PPP can also switch to reverse mode, generating R5P from the glycolytic intermediates G3P and F6P, independently of the ox-PPP. This metabolic mode avoids the production of 2 NADPH and 1 CO₂ molecules while converting a C6 into a C5 molecule. Furthermore, the reactions of transaldolase and transketolase produce and consume the C4 and C7 bodies erythrose 4-phosphate (E4P) and S7P, respectively. In plants, fungi and prokaryotes, E4P serves as a precursor molecule for the biosynthesis of the essential amino acids tyrosine, phenylalanine and tryptophan, while it also contributes to vitamin B6 metabolism [17, 18]. The exact role of E4P in animals however remains unknown.

The flux-rate and direction of glycolysis and the ox-PPP are controlled by tightly regulated enzymes that catalyse thermodynamically driven reactions, like hexokinase, phosphofructokinase, pyruvate kinase and glucose 6-phosphate dehydrogenase. In contrast, the non-oxidative arm of the PPP is fully reversible and lacks committed steps. Its regulation is very poorly defined and it's not exactly understood which circumstances decide the pathways mode of operation. The recently discovered sedoheptulose kinase CARKL catalyses a thermodynamically driven phosphorylation reaction, which resides at a strategic position to modulate carbon-flux. Thus, this enzyme may represent a novel regulator at the interface of glycolysis and the non-ox PPP.

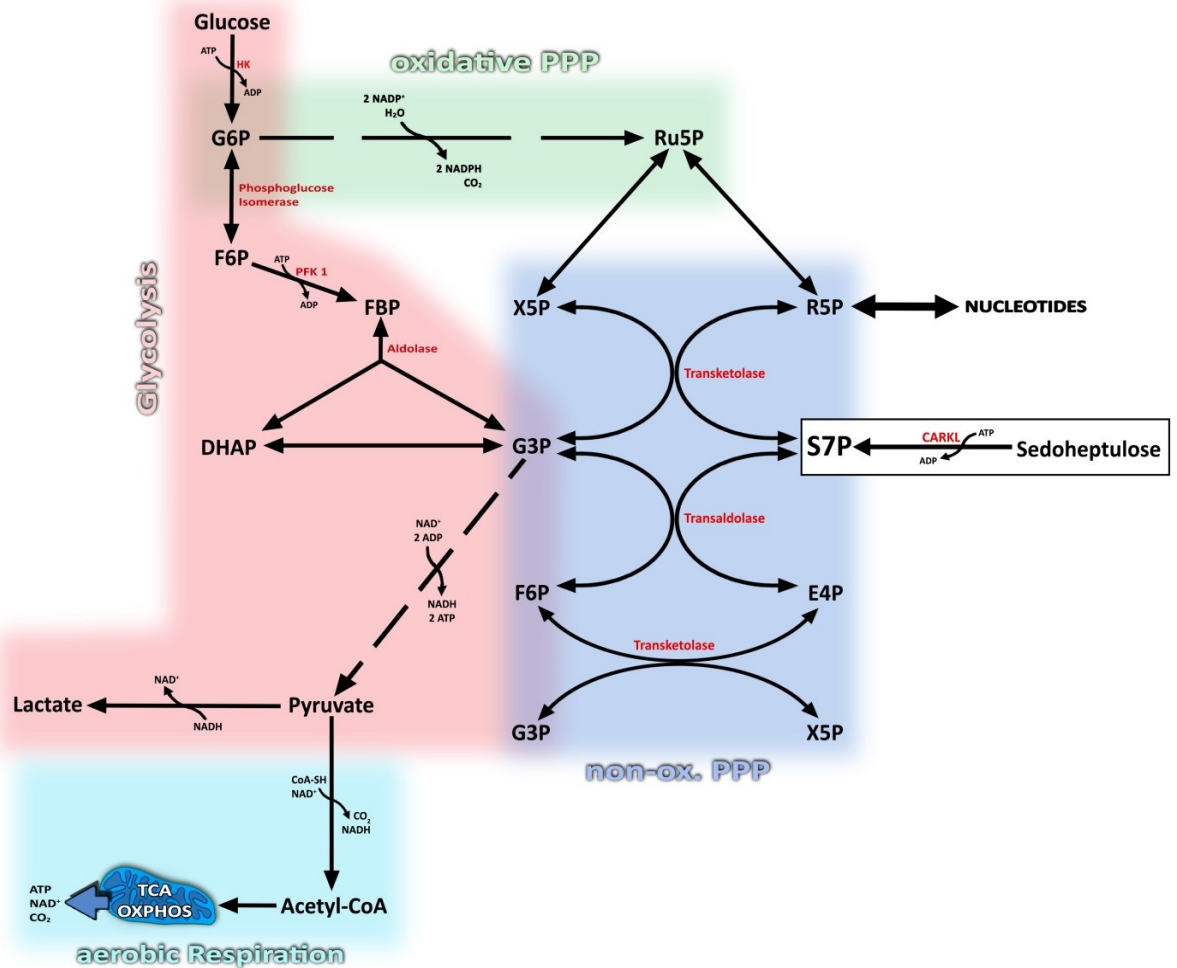


Figure 1: A simplified schematic overview of the main routes of carbohydrate metabolism. Free carbohydrates first need to be phosphorylated before they become available for subsequent biochemical reactions. The hexokinases (HK) phosphorylate D-glucose to glucose 6-phosphate (G6P), while the sedoheptulose kinase CARKL catalyses the phosphorylation of free D-sedoheptulose to sedoheptulose 7-phosphate (S7P). Glycolysis breaks down C6-bodies into two 3-bodies, while generating ATP. The end product pyruvate is fermented to lactate in order to regenerate cellular NAD⁺, an essential cofactor for glycolytic reactions. Under aerobic conditions in many terminally differentiated cell types, pyruvate is used to drive the tricarboxylic acid cycle and the mitochondrial respiratory chain for ATP generation. The oxidative pentose phosphate pathway converts C6-bodies into C5 bodies, necessary for nucleotide synthesis, while generating NADPH, a reducing agent involved in redox homeostasis and anabolic reactions, as well as CO₂. The reactions of the non-ox PPP can recycle unneeded C5 bodies into glycolytic intermediates. Running in reverse mode, the non-ox PPP converts glycolytic intermediates into C5 bodies, without meanwhile generating NADPH and CO₂. CARKL may modulate carbon flux at the interface of the non-ox PPP and glycolysis through thermodynamically driven S7P supply. For the other abbreviations please refer to the list at page 5.

Sedoheptulose & CARKL

Sedoheptulose is present in various fruits and vegetables like carrots, tomatoes, oranges and apples [7]. A specific transporter responsible for the cellular uptake of this ketoheptose has not been identified yet. Beside possible dietary uptake, sedoheptulose can also be formed enzymatically by the dephosphorylation of endogenous S7P [19]. The phosphatase catalysing this reaction is not known to date.

It can be assumed that cellular sedoheptulose is phosphorylated by the sedoheptulose kinase CARKL (a.k.a. SHPK) to S7P, analogously to the phosphorylation of glucose to G6P by hexokinase. This phosphorylation event might retain sedoheptulose inside the cell and make it available as a substrate for subsequent reactions. Interestingly, this metabolically relevant kinase was completely unknown until the 21st century. In the year 2000, Touchman and colleagues reported a novel open reading frame in a region of chromosome 17p13, which is frequently deleted in nephropathic cystinosis patients [20]. The sequence of this putative gene contains putative FGGY carbohydrate kinase domains and was accordingly called CARKL (carbohydrate kinase like gene). Eight years later, Wamelink and co-workers measured high sedoheptulose concentrations in the urine of cystinosis patients with the common 57kb deletion affecting CARKL [6]. Since the deficiency for CARKL in these patients also correlated with decreased sedoheptulose phosphorylation capacity of crude fibroblast protein homogenates, the authors concluded that CARKL may be responsible for their observation. This hypothesis gained support by the results of Kardon et al., who reported that the recombinant CARKL protein showed specificity for sedoheptulose in an enzymatic ADP formation assay [7]. Conclusively, our group recently demonstrated the direct kinase reaction of CARKL, which formed a single sedoheptulose phosphate in vitro [5]. Using multi-NMR analysis, the C-7 atom of sedoheptulose was precisely identified as the phosphate acceptor site.

CARKL is currently restricted to some species and is strongly expressed in metabolically active tissues [unpublished data]. Both, yeast and drosophila don't show sequence homologs to conserved CARKL nucleotide sequences found in vertebrates.

Heptoses in primary carbohydrate metabolism

The phosphorylation event catalysed by CARKL allows cells to directly route free sedoheptulose to carbohydrate metabolism, analogously to the system of hexokinase and glucose. In order to integrate the role of CARKL into established primary carbohydrate metabolism, we compared sedoheptulose with glucose and fructose, as all these molecules exist physiologically in free and diverse phosphorylated forms. **Figure 2** illustrates a comparative, non-linear network based on the participating carbohydrate phosphates. This simplified system is regulated by (i) phosphorylation events, (ii) the interconversion rates of various carbohydrate phosphates, and (iii) the conversion to carbohydrate bisphosphates.

By an enzymatically driven phosphorylation event, free carbohydrates are retained inside the cell and energetically activated for subsequent reactions. The hexokinases 1 to 4 (HK1-4) phosphorylate glucose and are key determinants of cellular carbohydrate flux [21]. The resulting G6P can then directly participate in the biochemical reactions of glycolysis and the oxidative PPP. Fructose phosphorylation by ketohexokinase (KHK) yields fructose 1-phosphate (F1P), a molecule which cannot directly contribute to primary carbohydrate metabolism. First, F1P has to be converted to glyceraldehyde and dihydroxyacetone phosphate (DHAP) [22]. DHAP then immediately enters the carbon cycle, while glyceraldehyde needs another round of phosphorylation. In adipose tissue and muscle, fructose can also be phosphorylated by hexokinase to directly yield the glycolytic intermediate fructose 6-phosphate (F6P) [23]. Since the affinity of hexokinase for fructose is much lower than that of ketohexokinase, this direct path of fructose metabolism may only play a role if fructose is present in great excess. Galactose (not shown in the figure) is phosphorylated to galactose 1-phosphate (Gal1P) by galactokinase. Similarly to F1P, Gal1P has to go through further enzymatic reactions, the so called Leilor pathway, until direct contribution to glycolysis in the form of G6P is achieved [24]. In contrast to fructose and galactose, the formation of S7P from sedoheptulose is analogous to the system of G6P and the hexokinases. A single phosphorylation reaction by CARKL makes this ketoheptose directly available for primary carbohydrate metabolism.

Phosphorylated carbohydrates subsequently become possible substrates for aldolase, transaldolase and transketolase reactions. These reversible enzymes can interconvert carbohydrate phosphates (so called “carbon scrambling”) according to momentary cellular needs. The interface of glycolysis and the non-ox PPP represents a prime example for this. F6P, resulting from the conversion of G6P, can react with E4P to G3P and S7P, which can be subsequently converted to two pentose phosphates. This series of reversible reactions, generating C3, C4, C5, C6 and C7 units, is most likely primarily regulated by the subcellular

localization of the involved enzymes and the concentrations of their substrates (see also **Fig. 4A**). This regulation of carbon scrambling events represents a second level of regulation.

The third regulatory level in our model is represented by carbohydrate bisphosphates, which frequently have additional regulatory functions. Examples include glucose 1,6-bisphosphate (G1,6bP), fructose 1,6-bisphosphate (F1,6bP) and fructose 2,6-bisphosphate (F2,6bP). The existence of sedoheptulose 1,7-bisphosphate (S1,7bP) was also reported, but rarely considered [25-28]. Partly, the formation of S1,7bP can be explained by the condensation reaction of E4P and DHAP, catalysed by aldolase [29]. Furthermore reports show that rabbit liver fractions, enriched for phosphofructokinase (PFK) activity, generated S1,7bP when S7P and ATP were present [28, 30]. However, the identity of the kinase catalysing this reaction was not defined reliably. Karadsheh and co-workers further reported that very high S7P concentrations competitively inhibited F6P phosphorylation [30]. It is interesting that fractions containing fructose 1,6-bisphosphatase (FBPase) activity were also shown to dephosphorylate S1,7bP [31, 32]. These results together indicate that hexose- and heptose-(bis)phosphate shunts may be coexistent (**Fig. 3**). A possible explanation for the incomplete enzymatic discrimination between F6P and S7P may be the structural similarity of these carbohydrates (**Fig 3B**).

It is further notable that perfusion of rat liver with high glucose levels increased S1,7bP formation [27]. The same effect was reported after the incubation of rat liver cytosol with F2,6bP, a glycolytic regulator that allosterically affects PFK-1 activity [28]. In contrast to glucose, glucagon increased S1,7bP dephosphorylation and S7P formation [27]. Taken together, these results suggest that sedoheptulose metabolism is susceptible for hormonal control.

Besides being allosteric regulators of metabolism, carbohydrate bisphosphates are also used as carbon substrates themselves. The enzyme aldolase converts F1,6bP into two triose phosphates, while it converts S1,7bP into one triose phosphates and one tetrose phosphate . These mono-phosphorylated products then re-enter the above discussed carbon-substrate-bridge (**Fig. 2**).

To sum up, sedoheptulose can directly enter primary carbohydrate metabolism after phosphorylation by CARKL as S7P, similarly to the initiation of glucose metabolism by hexokinase. On the other hand, S7P and S1,7bP might be partly utilized by the same enzymes as the fructose metabolites F6P and F1,6bP, namely aldolase, phosphofructokinase and fructose bisphosphatase. The biochemical importance of the two

latter steps however requires further verification. Research about heptoses in carbohydrate metabolism is still at its very beginning and our current view might change within the next years. It is indeed supposable that sedoheptulose phosphorylation through CARKL affects the equilibrium of reactions in the non-ox PPP, an essential pathway present in most organisms, where flux directions and regulatory mechanisms are still largely unexplored.

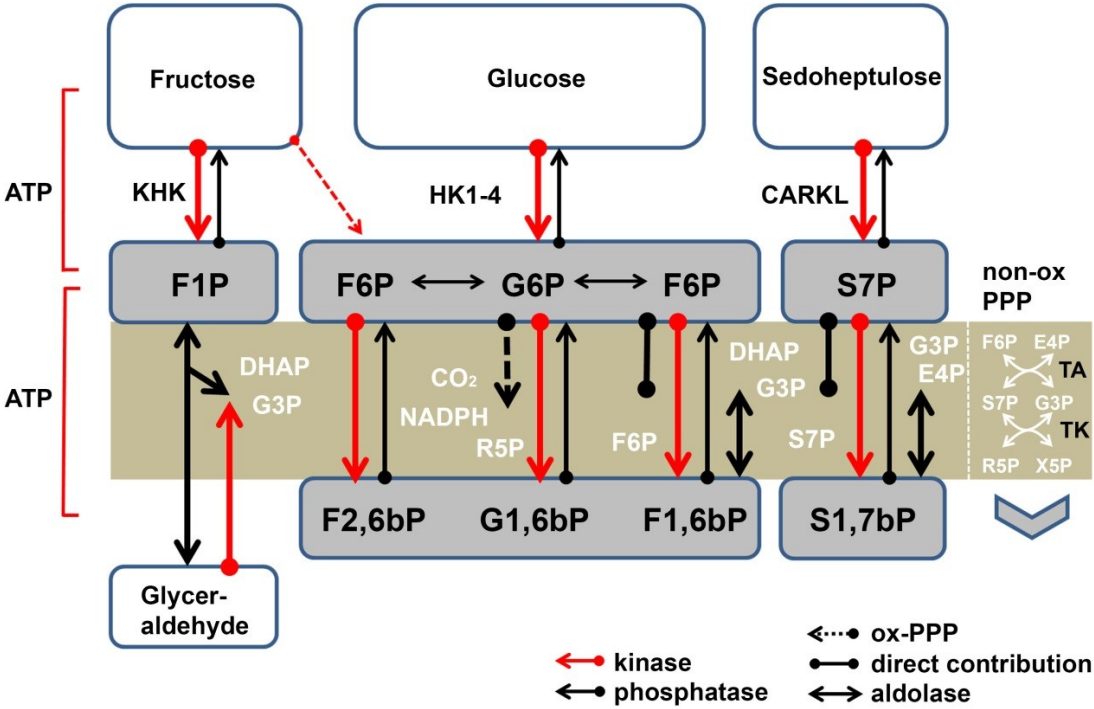


Figure 2: A non-linear model of carbohydrate phosphate interconversions. Glucose, fructose and sedoheptulose catabolism are compared by illustrating reversible and rate limiting kinase/phosphatase reactions. The free carbohydrates are activated by phosphorylation to G6P, F1P or S7P respectively. The carbohydrate phosphates or their metabolites can become substrates for interconversions catalysed by transaldolase (TA), transketolase (TK) and aldolase. G6P can also be consumed by the ox-PPP. F6P, mainly resulting from G6P isomerisation and S7P directly contribute to non-ox PPP reactions. Fructose phosphorylation by ketohexokinase (KHK) yields F1P, which in the presence of aldolase forms DHAP and glyceraldehyde. The latter can be either phosphorylated to G3P or utilized in triglyceride synthesis [33]. Carbohydrate bisphosphates represent an additional level of regulation. PFK-1 phosphorylates F6P to F1,6bP, while PFK-2 forms F2,6bP, a main glycolytic regulator. Reconversion of the bisphosphates into monophosphates is facilitated by phosphofructose biphosphatases. A related pattern of regulation was also reported for S7P and S1,7bP [25-32]. Aldolase splits F1,6bP into G3P and DHAP, S1,7bP into E4P and DHAP respectively. The reversibility of these reactions allows to regulate the bioavailability of carbohydrate monophosphates as substrates for the non-ox PPP, a pathway which scrambles carbon phosphates according to the cellular demand for critical intermediates of different length (C3 to C7). Figure adopted from [8].



Figure 3: Hexose/Heptose (bis) phosphate shunts. (A) A comparison of sedoheptulose- and fructose-(bis)phosphate turnover. S7P can be phosphorylated to S1,7bP and was shown to compete with F6P for phosphorylation, if present in high concentrations[25-32]. While high glucose or F2,6bP stimulated S1,7bP formation, glucagon enhanced S7P accumulation [27, 28]. Aldolase converts the bisphosphates into carbohydrate monophosphates, which feed glycolysis and the non-ox PPP. While F1,6bP is split into two triose phosphates, S1,7bP yields a C4-body (E4P) and DHAP. The ratios of S7P/S1,7bP as well as S1,7bP/F1,6bP might also contribute the regulation of glycolysis and the PPP. **(B)** Fischer projection formulas of D-fructose (ketoheptose) and D-sedoheptulose (ketoheptose). Figure adopted from [8].

CARKL & metabolic regulation

Metabolic pathways are tightly regulated by various mechanisms to ensure proper function. Glycolysis for example, which has been extensively studied in the past, is under influence of allosteric regulation, hormonal control, as well as feedback and amplification loops [10, 34]. Subcellular localization and post-translational modifications of the participating enzymes further contribute to the regulation of carbon flux. Glycolytic carbon usage is adapted to systemic glucose levels, transduced by hormones, as well as to the intracellular energy situation, arranged by factors such as the AMP/ATP ratio.

The reversible nature of the non-ox PPP, allowing various operating modes, requires sophisticated regulatory mechanisms [35]. CARKL represents a likely candidate to regulate these fluxes, as it provides S7P to the system, independent of glucose and other carbohydrates. Its overexpression in RAW-264.7 cells, a mouse macrophage cell line, resulted in the reduction of G3P, R5P and X5P steady-state levels [5]. Knockdown mediated by an RNA interference approach showed the reverse effect. Interestingly, basal sedoheptulose levels were not affected by these CARKL perturbation experiments, indicating a continuous endogenous supply of the ketoheptose. S7P levels were strongly decreased by CARKL loss, but didn't change upon overexpression. These results indicate S7P as a rate-limiting cofactor for glycolysis derived G3P via the non-ox PPP (**Fig. 4**). Supportingly, it was reported that incubation of Krebs ascites tumour cells with glucose resulted in approximately

5-fold higher G3P than S7P concentrations [11]. Since G3P is constantly produced by glycolysis, and the PPP generates equal amounts of G3P and S7P even when flux rates increase, we hypothesize that S7P supply by CARKL is rate limiting for carbon scrambling events at the interface of glycolysis and the non-ox PPP. Regulation of S7P availability might be the mechanism by which CARKL exerts its control over non-ox PPP flux. Additionally to the S7P/G3P ratio, the S7P/S1,7bP and S1,7bP/F1,6bP ratios also seem to be of biological relevance (**Fig. 3**).

Flux through this metabolic interface may be of relevance in various cellular settings, such as bioenergetic adaptations to different stimuli as well as metabolic reprogramming, well known from immune-, stem- and tumour cells.

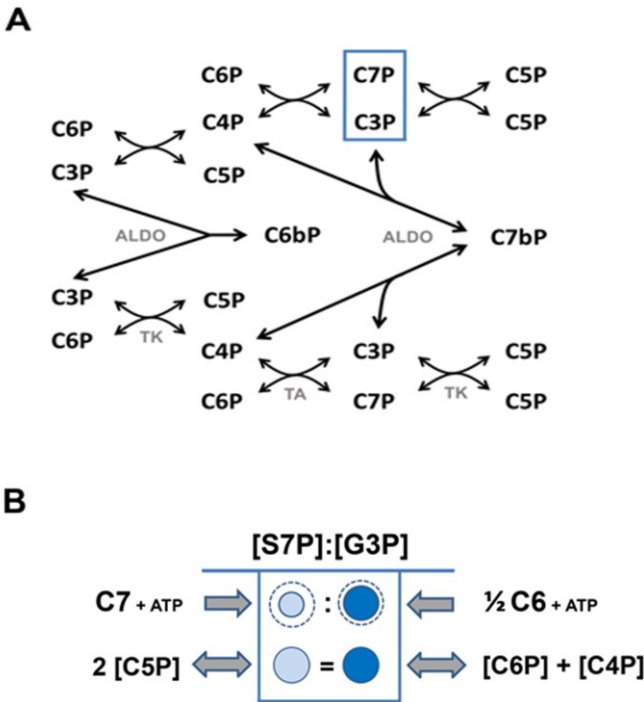


Figure 4: S7P as a rate limiting substrate. (A) A simplified illustration of the classical non-oxidative PPP depicting the reversible reactions of carbohydrate (bis)phosphates, catalysed by transaldolase, transketolase and aldolase (ALDO). (B) A representation of the flux ratios of S7P and G3P, two substrates that are reactions partners in both transaldolase and transketolase reactions. Equal amounts of S7P and G3P are produced by forward and reverse PPP flux, either from two C5 bodies or a hexose- plus tetrose- phosphate respectively. The stoichiometric imbalance may result from glycolytic G3P and CARKL derived S7P. Increased G3P and S7P levels favour the formation of C5 bodies or glycolytic intermediates. Flow direction depends on momentary cellular needs. If equilibrium shifts towards S7P formation (e.g. after CARKL inactivation or knockdown), non-ox PPP flux will increase. The direction of carbon flux is not indicated here, since it appears to be multifactorial and to depend on the momentary cellular demand for various precursors. Figure adopted from [8].

The physiological role of CARKL

As mentioned above, the CARKL gene was first identified in a commonly deleted region of nephropathic cystinosis patients [20]. Cystinosis, an autosomal recessive lysosomal storage disorder, is characterized by a defect of the membrane transport protein cystinosin (CTNS) [36]. Mutations in CTNS result in the accumulation of the amino acid cystine in lysosomes, which leads to crystal formation in different tissues. The introduction of cystine depleting agents and kidney transplantations made this once fatal disease treatable.

Beside CTNS and CARKL, the 57kb deletion, associated with the severe infantile nephropathic type of the disease, also affects the gene transient receptor potential vanilloid 1 (TRPV1) [37]. The evaluation of CARKL deletion separate from CTNS and TRPV1 may be mandatory to understand the clinical heterogeneity of cystinosis [6].

The recent discovery, that CARKL regulates macrophage polarization by modulating cellular carbon flux and redox homeostasis might be further helpful to decode the physiological function of sedoheptulose kinase [5]. CARKL was shown to be rapidly downregulated upon classical macrophage polarization, while enhanced sedoheptulose kinase activity supported alternative polarization. The role of CARKL in macrophage polarization is discussed in more detail in the background section of Part 2 of the thesis.

Aims of the master thesis

The theoretical section of my master thesis, discussed above, is largely based on our review “Sedoheptulose kinase regulates cellular carbohydrate metabolism by sedoheptulose 7-phosphate supply”, which was published in the journal *Biochemical Society Transactions*, volume 41, part 2 (2013).

The aims of the experimental part of my master thesis were divided in two parts: Part one deals with the post-transcriptional regulation of CARKL and focuses on the discovery of CARKL transcripts with novel 3'UTRs. Part two describes the initial characterization of recently generated transgenic mice overexpressing CARKL. Detailed descriptions of the master thesis objectives can be found in the “aim” sections of the respective parts.

Part 1 – Novel CARKL transcripts

As we start to understand the function of sedoheptulose kinase and the impact of heptoses on metabolism, novel questions keep arising about the biological regulation of cellular sedoheptulose utilization. Thus, we started to investigate several levels of regulatory mechanism likely acting on CARKL expression. This part of my master thesis deals with the post-transcriptional regulation of CARKL, while focusing on the identification of CARKL transcripts with novel 3' untranslated regions (3'UTRs).

Background

Regulatory mechanisms of gene expression

Regulation of gene expression in eukaryotes has multiple facets ranging from transcriptional to post-transcriptional and translational regulation, finally to various post-translational modifications on the protein level. Whole arrays of interacting regulatory elements decide on the induction and repression of specific genes in a precisely timed manner [38].

Regulation of gene expression starts with the initiation of the transcriptional process. Protein coding genes, such as CARKL, are transcribed by RNA polymerase II (RNA pol II). The cis-acting regulatory regions of RNA pol II promoters consist of the core promoter and various enhancer elements [39]. These enhancer regions, which can be quite distant away from promoters, are conglomerates of various regulatory elements, which can be specifically bound by proteins called transcription factors. Together these cis- and trans-acting elements serve to fine-tune transcriptional regulation in a cell type- and situation-dependent fashion.

Transcription factors can act as transcriptional activators or repressors, depending on the accessory proteins, which they are associated with [40]. These proteins frequently act as modulators of chromatin structure. An open chromatin structure is associated with transcriptional activation, as it provides access to RNA polymerase and further transcription factors, while condensed chromatin structure correlates with transcriptional repression.

Generally, chromatin is the name for the complex of DNA with proteins, as it is present in the nucleus of cells [41]. The DNA is wrapped around octamers of histone proteins, which themselves carry a huge variety of possible posttranslational modifications such as acetylation, methylation and phosphorylation. Together these are frequently referred as the “histone code”. Histone modifications participate in the regulation of chromatin accessibility as well as the recruitment of remodelling factors. The methylation of DNA represents a further chromatin modification and is associated with transcriptional silencing. Also various alternative histone isoforms can be incorporated into histone octamers, which have special regulatory functions.

The protein coding transcripts produced by RNA pol II undergo further processing events to generate the mature mRNA, commonly summarized as “posttranscriptional modifications”. They include splicing, capping, polyadenylation and RNA editing. Splicing events remove introns and can also lead to the generation of alternatively spliced mRNAs, which may show differences in stability, localization or may encode for alternative protein isoforms. Posttranscriptional addition of the 5'-cap structure and the 3' poly(A) tail is important for mRNA stability and translational efficiency. RNA editing is an emerging modification which deserves more and more attention nowadays and may alter various mRNA features depending on the position of editing [42].

The mature mRNA consists of the protein coding region as, a 5'CAP structure, a poly(A) tail as well as 5' and 3' UTRs. These UTRs are of special importance for mRNA localisation, stability and their translational efficiency. They carry special binding motives for protein and small-RNA species, which can regulate mRNA fate. The latter are mainly known as micro RNAs (miRNAs) which influence mRNA through a phenomenon called RNA interference [43]. The efficiency of protein synthesis from a particular mRNA is controlled by translational regulation. The mechanisms here include ribosome recruitment to the initiation codon and the modulation of translational elongation and termination [44]. Translational regulation is frequently associated with secondary mRNA structure motives as hairpin loops.

The final level of control is provided by posttranslational modifications (PTM) of proteins. Various modifications such as phosphorylation, ubiquitination or nitrosylation, to name but few, are often decisive for protein function, activity and turnover [10]. Many of these appear and disappear very rapidly compared to the time which is necessary to drive changes in transcriptional and translational machineries.

The CARKL gene

The presence of CARKL is currently restricted to some species. It is present in vertebrates, while yeast and drosophila show no sequence homologues to conserved CARKL nucleotide sequences. In humans, the CARKL gene can be found on the (-) strand of chromosome 17, whereas the mouse variant is on the (+) strand of chromosome 11. In both organisms, CARKL has a shared promoter region with the CTNS gene. As publicly available DNA methylation data suggests, parts of the promoter seem to be specific for one or the other gene in humans (<http://genome.ucsc.edu/>). General tissue expression is likely to be regulated by changes on chromatin accessibility level, such as DNA methylation and histone modifications. Promoter bashing experiments in our laboratory revealed functionally important transcription factor binding sites, which were verified by site directed mutagenesis [unpublished data].

In both humans and mice, seven exons are spliced to form the mature mRNA (~3.8 and 2.8kb in length respectively). The full length transcripts contain short 5' and long 3'UTRs. In humans, CARKL mRNA expression is high in liver, kidney, pancreas and heart tissue [20]. Similarly, the mRNA expression in mice is also elevated in liver, kidney and further in adipose tissue and some glands [unpublished data]. CARKL mRNA levels were found to be highly regulated during macrophage polarization: after LPS stimulation sedoheptulose kinase was rapidly downregulated, while IL-4 stimulation led to an increase in message steady-state levels [5]. More details about CARKL and macrophage polarization can be found in part 2 of the master thesis. Touchman and colleagues showed by northern blot analysis that alternative CARKL transcripts may exist, but their sequence and function were not further investigated [20].

Interestingly, CARKL expression in mice also oscillates within the circadian rhythm as shown in a publically available database (<http://bioinf.itmat.upenn.edu/circa/query>) (**Fig. S1**). Its expression is elevated during the day, while it drops during night. It is well known that the light-dark cycle plays an essential role in energy metabolism and body weight control, as well as the development of metabolic disorders.

The CARKL protein

The mature CARKL mRNA encodes a ~52 kDa protein, which was shown to co-localize with G6P dehydrogenase in the cytoplasm [5]. In humans, the CARKL protein consists of 478 amino acids (AA), while the mouse variant contains 476 AA. They share 92.5% interspecies similarity and 87% identity [8]. The sedoheptulose kinase protein is also receptive to post translational modifications. Two independent proteomics screens identified this kinase as an ubiquitin-modified protein [45, 46] These PTMs might be able to regulate the enzymatic activity of the protein as a rapid response to various stimuli.

In silico comparative modelling of the protein suggests that CARKL has a putative catalytic cleft in the centre of the protein [5]. This model further revealed two potential sedoheptulose binding sites and a conserved ATPase domain. Site directed mutagenesis confirmed the functional relevance of the ATPase domain. To understand the mechanisms of this kinase, precise structural data is necessary. The protein structure may be solved by methods such as X-ray crystallography or nuclear magnetic resonance (NMR).

Regulatory mechanisms acting on mRNA

As the following experiments presented in this part of the master thesis deal with modifications in the regulatory part of mRNAs, we switch gears from the protein level back to the world of messenger-RNAs. In order to give rise to the mature mRNA, the primary transcript (pre-mRNA) has to undergo post-transcriptional processing. This consists of capping at the 5' end, polyadenylation at the 3' end, pre-mRNA splicing and sometimes RNA editing. Together these mechanisms ensure the correct translation of mRNAs into their respective protein products.

The 5' end of the transcript receives co-transcriptionally a so called CAP-structure, a 7-methylguanosine (5'm⁷GTP) connected by a 5' to 5' triphosphate linkage, which protects the transcript against degradation by 5' exonucleases. At the 3' end, pre-mRNAs are sequence specifically cleaved by ribonucleases. Subsequently poly(A) polymerase adds a long tail of adenines. This poly(A) tail, associated with poly(A)-binding proteins (PABP), has a length of ~250nt in mammals and is a necessary structure for the export from the nucleus as well as for the escape from nuclear surveillance mechanisms [47, 48]. Poly(A) tails of actively

translated mRNAs are progressively shortened. Messages which are chosen for silencing, for example by the binding of miRNAs, are actively decapped, deadenylated and subjected to RNA decay [49, 50].

During the formation of the mature mRNA, the splicing machinery removes introns, while the exons are joined together. Many splicing events are tissue specific and are able to produce multiple different mRNAs from a single gene. Most of these events are catalysed by the spliceosome, a large complex of small nuclear ribonucleoprotein particles (snRNPs) and proteins, which recognizes pre-mRNA cis-acting sequences such as the 5' splice site, 3' splice site and the branchpoint sequence [51]. Further, auxiliary regulatory mechanisms have an influence on the sequence specific recruitment of the spliceosome [52]. Cis-acting elements, including exon splicing-enhancers (ESEs) and –silencers (ESSs), as well as intron splicing-enhancers (ISEs) and –silencers (ISSs), are bound by proteins which modulate splicing pattern. Among these, the best studied examples are the SR proteins, which usually promote splicing, and hnRNPs, which primarily act as splicing inhibitors. It is further known that pre-mRNAs adopt complex secondary and tertiary structures, which can modulate alternative splicing events [53].

After successful post-transcriptional modification, mature eukaryotic mRNAs share common elements which include the coding region, the 5'CAP, the poly(A) tail as well as regulatory 5' and 3' untranslated regions (UTRs) (**Fig. 5**). These untranslated regions play important roles in the regulation of mRNA fate. Arrays of signals in mainly noncoding, but also coding regions, serve as platforms for the binding of protein and RNA species, which modulate the destiny of a particular mRNA in a cell type and situation dependent manner. The 5'UTR is mainly involved in translational control, while the 3'UTR is a crucial regulator of mRNA localization, translational efficiency and mRNA half-life [49].

It is important to mention that RNA-regulatory motifs are frequently determined by a combination of primary and secondary structures (e.g. various stem-loop structures) [54, 55]. Structural features of the 5'UTR for example are known to have an important role in the regulation of mRNA translation. Translation of the mRNA coding sequence is usually initiated in a CAP-dependent fashion at an AUG codon [56]. The selection of the initiation codon depends on its sequence context. Some mRNAs have long 5'UTR containing upstream initiation codons or open reading frames (uORFs), at which translation is less efficiently initiated. Upstream ORFs might further be capable of regulating the expression of the main open reading frame [57]. CAP-independent translational initiation through RNA secondary structures called internal ribosome entry sites (IRES), was originally discovered in viruses

and subsequently also found in the 5'UTR of many eukaryotic mRNAs, upstream of an initiation codon [58-60].

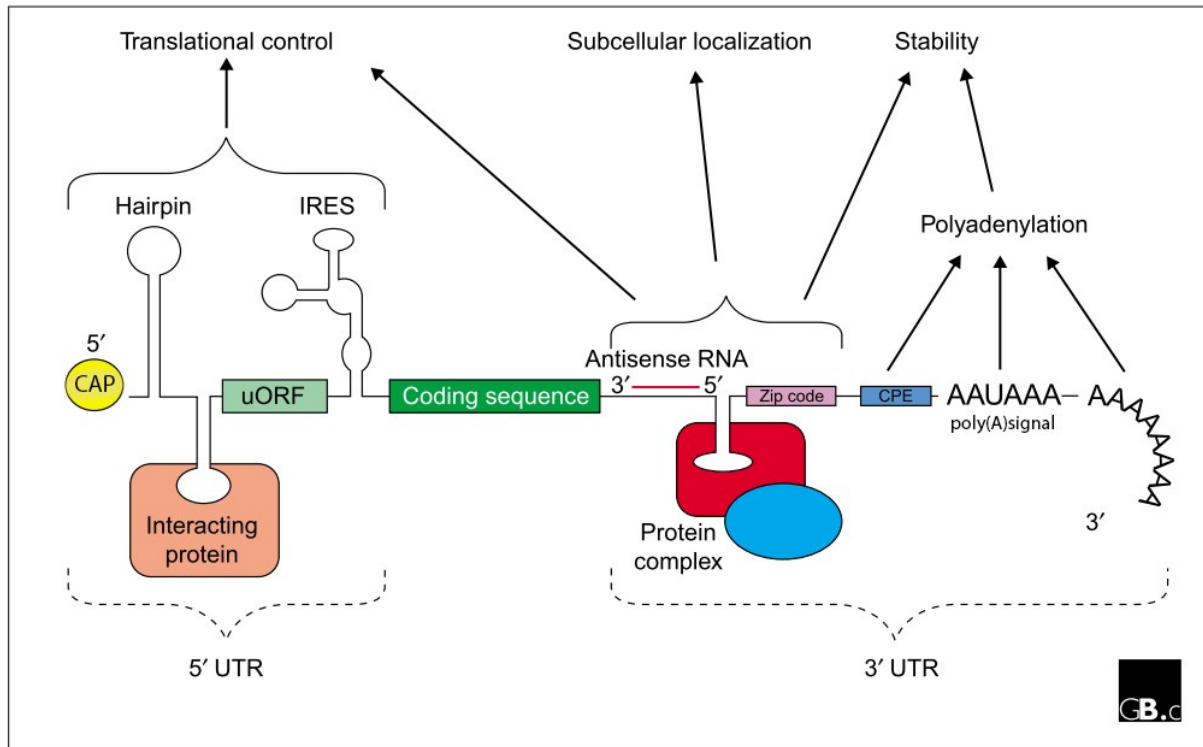


Figure 5: Regulatory elements present in eukaryotic mRNA. This simplified illustration depicts the general structure of eukaryotic mRNAs, which consist of the coding sequence, untranslated regions (UTRs) a 5'CAP. Manifold regulatory elements, mainly residing in the UTRs influence mRNA fate through the modulation of translation, stability and subcellular localization. Please see main text for the description of other regulatory elements. This illustration was adopted from [54].

The 5'UTR controls main aspects of translational initiation, while the 3'UTR is first of all involved in the modulation of transcript turnover, localization and its translational efficiency. A large part of mRNA regulation is carried out by the sequence specific binding of miRNAs to their targets. Micro RNAs are evolutionary conserved non-coding RNAs, with a length 18-25 nucleotides, which are involved in the regulation of fundamental cellular processes [61]. Micro RNA-precursors (pri-miRNAs) are usually transcribed by RNA pol II and processed by endoribonucleases to form the intermediary pre-miRNAs and finally the mature miRNAs. These mature miRNAs are incorporated into the RNA-induced silencing complex (RISC) [62]. Their specificity is defined by the so called “seed sequence”, including positions 2 to 8 from the 5'end of the miRNA, which binds to the target sequences, usually in the 3'UTR of mRNAs. Perfect complementarity between a miRNA and its target leads to mRNA degradation, whereas imperfect pairing results in translational blockade [63]. “Non-canonical” miRNA-mediated mechanisms include the binding of miRNAs to the 5'UTR and coding

sequence, as well as gene expression activation instead of repression [64]. Many miRNAs are able to target a large number of genes (~ 500), while most genes are targeted by at least one miRNA [65]. This implies that by controlling mRNA degradation and translation, miRNAs are able to regulate multiple pathways and orchestrate integrated cellular responses.

Context dependent regulation of transcript stability is not just controlled by small RNA species. Many regulatory pathways for mRNA stability also involve proteins as trans-acting factors, which interact with cis-elements residing in the 3'UTR. A frequently encountered group amongst these is represented by motifs rich in adenine and uridine (AU-rich elements or ARE) [66]. ARE-regulated genes frequently involve highly active proteins such as cell cycle regulators, growth factors and cytokines. Transcripts of immune-regulatory proteins represent a large class in the ARE database [67]. AU-rich elements have a higher complexity than most other cis-elements, as they are not made up of a single consensus sequence. Amongst others, the sequence components AUUUA and UUAUUUAUU have been shown to play a role [66, 68]. AREs are considered to be primarily destabilizing signal sequences, which mediate their effects by the interaction with intracellular proteins [69]. The AU-rich element binding proteins (AUBPs) belong to different protein families. Most of them promote mRNA destabilization or modulate translation, while some others can delay or overcome destabilizing effects [70-72]. It has been shown that ARE-mediated mRNA decay mechanisms are primarily associated with exosomal 3' to 5' degradation [73, 74]. There are also various other classes of proteins modulating mRNA turnover, which are beyond the scope of this master thesis [75].

A further mechanism of transcript regulation is represented by the phenomenon of RNA editing [42, 76]. Generally, RNA editing leads to post-transcriptional base substitutions, which can alter protein binding sites, miRNA binding sites as well as splice sites. Further, RNA editing events can even lead to nuclear sequestration and endonucleolytic cleavage. The substitution of adenine to inosine is the most frequent editing event in humans. Inosine base pairs with cytosine and is therefore read as guanine. RNA editing takes place in double stranded RNA regions and can be predominantly found in Alu repetitive elements. These editing events, carried out by adenosine deaminases acting on RNA (ADAR) are not random. They are predetermined by the sequence and especially RNA structural motifs. A to I editing appears to be tissue specific and greatly contributes to transcriptome diversity. It represents a further mechanism to fine tune gene expression, depending on tissue function and developmental stage.

Targeting of transcripts to specific subcellular compartments also frequently depends on 3'UTR elements [77]. Thus, the remodelling of 3'UTRs in response to extracellular cues contributes to mRNA localization and local protein synthesis. These events are especially in strongly polarizes cells of high importance, such as neurons. The involved 3'UTR elements are known as mRNA zip-codes and are bound by zip-code-binding proteins, which facilitate subcellular mRNA targeting [78]. Another class of 3'UTR regulatory sequences is represented by cytoplasmic polyadenylation elements (CPE) [54]. CPEs, bound by the CPEB-protein family (cytoplasmic polyadenylation element binding), regulate the reactivation of translationally inactive mRNAs via the elongation of their poly(A) tails [79].

Many of the above mentioned regulatory elements can be mainly found in mRNA 3' UTRs. Mammalian cells frequently contain multiple polyadenylation signals, which can lead to alternative 3' end processing [80]. Importantly, 3'UTRs are also subjected to splicing events [81]. The resulting, usually shorter 3'UTR have changed compositions in cis-acting elements and thus different regulation profiles.

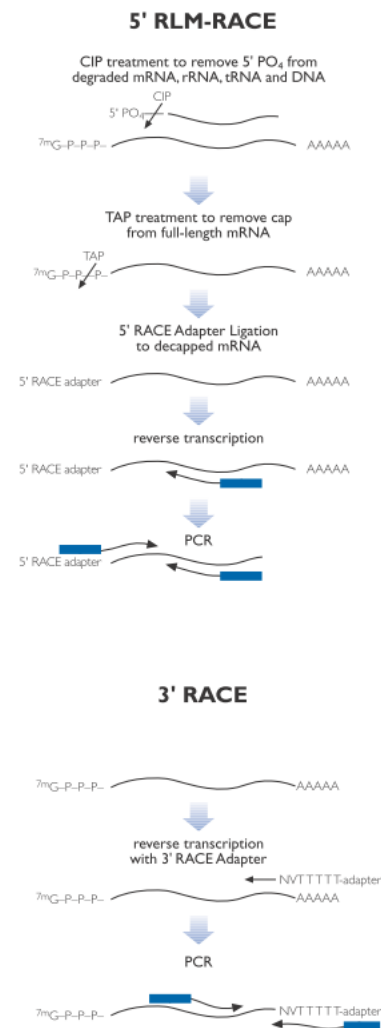
Rapid amplification of cDNA ends

In order to identify alternative CARKL transcripts we chose a technique known as rapid amplification of cDNA ends (RACE). Of the various available RACE-methods, we decided to use RLM-RACE (RNA-ligase mediated RACE), as it provides benefits in terms of specificity [82]. During RLM-RACE, 5' and 3' transcript ends are reverse transcribed in different reactions using 5' and 3' specific adapters. In the second phase, specific ends are amplified in tandem PCR-reactions with sequence specific primers binding in the centre of the desired mRNA and RACE-primers binding the respective RACE adapters. Specific products can be subsequently identified by cloning and sequencing.

5'RACE libraries are enriched for undegraded mRNA 5' ends containing the CAP structure. First, calf intestine phosphatase (CIP) incubation is used to dephosphorylate RNAs containing 5' phosphates, as degraded mRNA and various non-mRNA species. After CIP removal by phenol-chloroform extraction and RNA precipitation, the enzyme tobacco acid pyrophosphatase (TAP) is used to remove the 5'CAP structures of full length mRNAs. This reaction leaves a 5' monophosphate which is utilized for the subsequent ligation of an RNA anchor (5' RACE adapter) catalysed by RNA ligase. Theoretically only undegraded 5'RNA

ends receive the anchor, as previously dephosphorylated RNAs lack the 5' monophosphate. The following reverse transcription reaction converts the RNAs into a more stable cDNA library, which can be used in for nested PCR reactions to amplify the transcripts ends of interest.

The generation of **3' RACE** libraries is much less demanding. Here, we start directly with the reverse transcription step, using the 3'RACE adapter, a modified anchored poly-d(T) oligo, to prime the reaction. The following nested PCR reactions are performed analogously for both 5' and 3' RACE. In the outer PCR reaction, one primer (RACE primer) binds the adapter sequence while the other (gene specific primer) binds at a chosen site in the transcript of interest. A second set of nested primers is used to increase sensitivity and specificity of the amplification. The obtained PCR products can be gel purified, cloned and finally sequenced.



i
Figure 6: Overview RACE procedure. See main text for a more detailed description. Figure was adopted from the manual of the First-choice RLM-RACE kit (Ambion).

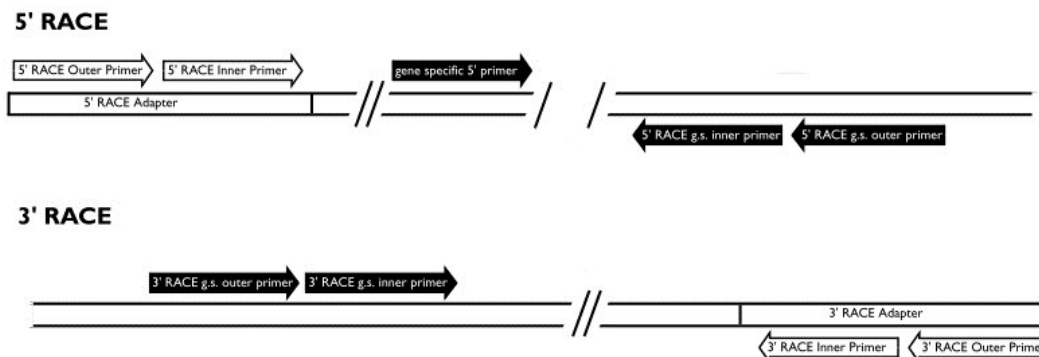


Figure 7: Nested RACE-PCR primer design. See main text for description. Figure was adopted from the manual of the First-choice RLM-RACE kit (Ambion).

Aim

The aim of the first part of my master thesis was to investigate the diversity of CARKL transcripts in various tissues and macrophages derived from humans as well as mice. We applied the above described RACE technique to amplify specific CARKL transcript ends. These were cloned, sequenced and initially characterised using various databases and *in silico* tools

Results

Alternative 3'UTRs of human CARKL

Since alternative transcripts of CARKL have not been further characterised, I started to investigate the diversity of CARKL mRNAs using RLM-RACE. The focus was on differences in the 3'UTR, the region which is of high importance for the regulation of mRNA stability and translation.

We chose human liver tissue and macrophages to perform the RACE-experiments. CARKL was shown to be involved in macrophage polarization, while liver represents a metabolically active organ with high CARKL expression [5, 20]. After RNA purification, we tested total RNA quality using microchannel based electrophoresis (**Fig. S2**). High quality RNA, with an RNA-integrity number (RIN) of >8 was used to generate RACE libraries, containing 3' adapter-bound cDNA. These libraries were used to amplify CARKL transcripts with nested PCRs. 3'RACE reverse primers, which bind the 3'RACE adapter fused to putative poly(A) tails, were used in combination with CARKL forward primers. The CARKL primer used in the outer PCR binds at the 3'end of exon 5, while the primer used in the inner PCR binds exon 7. Specific products obtained from the inner PCR were gel-purified, cloned and sequenced. In order to discriminate between specific and unspecific products, I aligned the obtained sequences to the CARKL full length sequence (GenBank: BC020543.1). Some of the products generated through the 3'RACE nested PCRs either turned out to be unspecific products of the primers, or to result from internal priming of the 3'RACE adapter during the reverse transcription procedure. The latter could be frequently observed in regions with internal poly(A) stretches, mimicking the terminal poly(A) tail.

3'RACE with RNA previously isolated from a human liver sample led the amplification of the full length 3'UTR, approximately equal in length to the published full length sequence. The inner and outer RACE products are shown in **figure 8A**. The full length 3'UTR in human liver had a length of 1950±1bp. The adapter based 3'RACE method doesn't allow predict the exact starting point of the poly(A) tail. Please consider that all indicated 3'UTR ends and therefore also lengths might drift by ±1bp. The full length 3'UTR sequence could also be obtained from a 3'RACE library of macrophages, which were differentiated from human peripheral blood monocytes of two healthy volunteers. The full length 3'UTR in macrophages had a total length of 1959bp.

In human liver, I could furthermore identify two novel CARKL transcripts with shorter 3'UTRs (**Fig. 8B and C**). The first, with a 3'UTR length of 212bp could be amplified from a liver 3'RACE library, reverse transcribed at 48°C (**Fig. 8B**). Sequence analysis revealed that this transcript is most likely terminated by a yet undescribed, alternative polyadenylation (AP) signal. Therefore, this alternative 3'UTR is termed "hu3'UTR_AP" here.

I could identify another novel CARKL 3'UTR in a liver 3'RACE library, reverse transcribed at 45°C (**Fig. 8C**). This product, with a net 3'UTR length of 185bp, results presumably from an alternative splicing (SPL) event, which removes 1645bp of the CARKL 3'UTR (**Fig. 9**). The second novel CARKL 3'UTR is referred here as "hu3'UTR_SPL". The hu3'UTR_SPL product was cloned and sequenced from two independent RACE-PCRs, while hu3'UTR_AP was cloned and sequenced from a single RACE-PCR approach. Generally, at least two clones were sequenced of each PCR product. In contrast to human liver, we could not find transcripts with alternative 3'ends in the 3'RACE library of human macrophages.

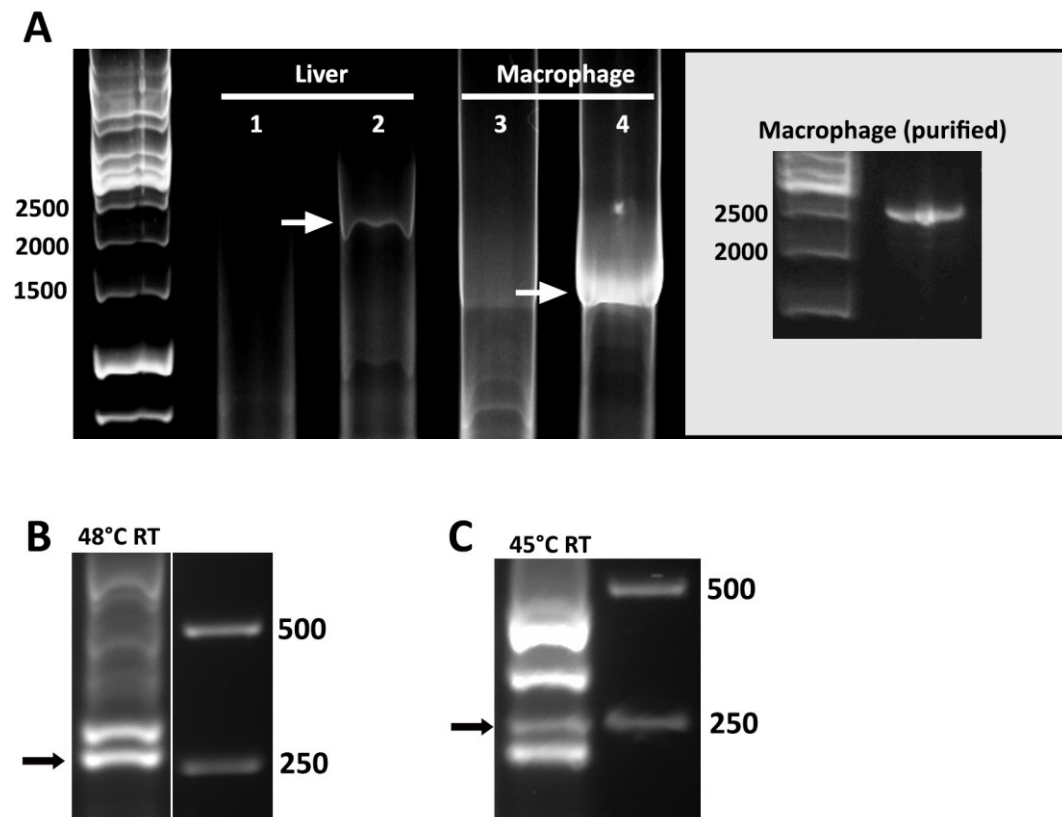


Figure 8: CARL 3' RACE products of human liver tissue and macrophages differentiated from peripheral blood monocytes: Arrows indicate specific products. Other bands visible here either represent unspecific products of the primers or result from internal priming of the 3'RACE adapter during reverse transcription. All human 3'RACE products amplified starting with gene specific primers in exon 5 (outer PCR) and exon 7 (inner PCR). **(A)** Full length CARL transcripts present in liver and macrophages. Lanes 1 and 3 represent the outer RACE-PCR product, where specific bands are usually not observed. Lanes 2 and 4 contain the inner RACE-PCR product, which were subsequently gel purified, cloned and sequenced. The small picture in the box represents the actual size of the product indicated in lane 4 after gel purification and loading of equal amounts of marker DNA. **(B)** Inner 3'RACE indicating a novel CARL 3'UTR product, identified in a human 3'RACE library reverse transcribed at 48°C. The indicated band results presumably from an alternative polyadenylation (hu3'UTR_AP). **(C)** Inner 3' RACE product of another novel CARL 3'UTR detected in human liver, resulting from an alternative splicing event (hu3'UTR_SPL). The reverse transcription step of the 3' RACE library used here was performed at 45°C. In contrast to human liver tissue, the macrophage 3'RACE libraries didn't reveal alternative CARL 3'ends.

Regulator elements of the human CARKL 3'UTR

Publicly available tools predicted multitude of regulatory sites in the full length 3'UTR of CARKL (**Fig. 9A**). Target scan micro RNA (miRNA) regulatory site prediction (<http://www.targetscan.org/>) shows that the 3'UTR might be targeted by numerous miRNAs, for example miR-124/506, miR138 and miR-324-5p (**Fig. 9B**). In addition to putative miRNA binding sites, an AU-rich element (ARE) could be detected by using an online tool of the Vienna TBI-website (<http://rna.tbi.univie.ac.at/cgi-bin/AREsite>).

In the UCSC genome browser (<http://genome.ucsc.edu/>), the CARKL 3'UTR is furthermore listed to undergo A to I RNA-editing events at two distinct sites. One editing position resides in the middle section of the 3'UTR, whereas a whole region is present downstream, harbouring multiple editing sites. The RNA editing sites in the downstream region could also be confirmed in deep sequencing data derived from RNA of human adipocytes differentiated from SVF (stromal vascular fraction) cells (**Fig. S4**) [unpublished data]. The whole downstream RNA editing region resides within an AluSx element. This primate-specific class of short interspersed elements is frequently observed to harbour RNA editing sites [83].

I further analysed the CARKL 3'UTR for common regulatory pattern using RegRNA2.0 (<http://regrna2.mbc.nctu.edu.tw/>) and UTR scan (<http://itbtools.ba.itb.cnr.it/utrscan>). They both predicted four Musashi binding elements (MBE) in the 3'UTR. MBEs might be bound by the Musashi protein, which mediates mRNA translational control [84]. As the used programs predicted MBEs in all human 3'UTRs which were submitted (HSP70, β -actin, β -tubulin) and further two MBEs in a 1500bp random nucleotide sequence, these predictions should be treated sceptically. RegRNA2.0 also predicted various non-coding RNA hybridization regions which all reside within the Alu-Sx element. The submission of various primate Alu-elements led to similar results. As a common role of Alu-elements residing in the 3'UTR has not been reported, the relevance of these predictions is questionable [85].

Interestingly, both novel 3'UTRs lack most the above mentioned regulatory regions: the ARE-site, both RNA-editing sites as well as the majority of the predicted miRNA binding sites. These results suggest that the full length CARKL 3'UTR can be more strictly controlled through the binding of regulatory factors than the novel 3'UTRs hu3'UTR_AP and hu3'UTR_SPL, which could be identified in human liver. Human macrophages expressed solely the full length-variant, which therefore might be receptive to a exclusive range of putative trans-acting factors.

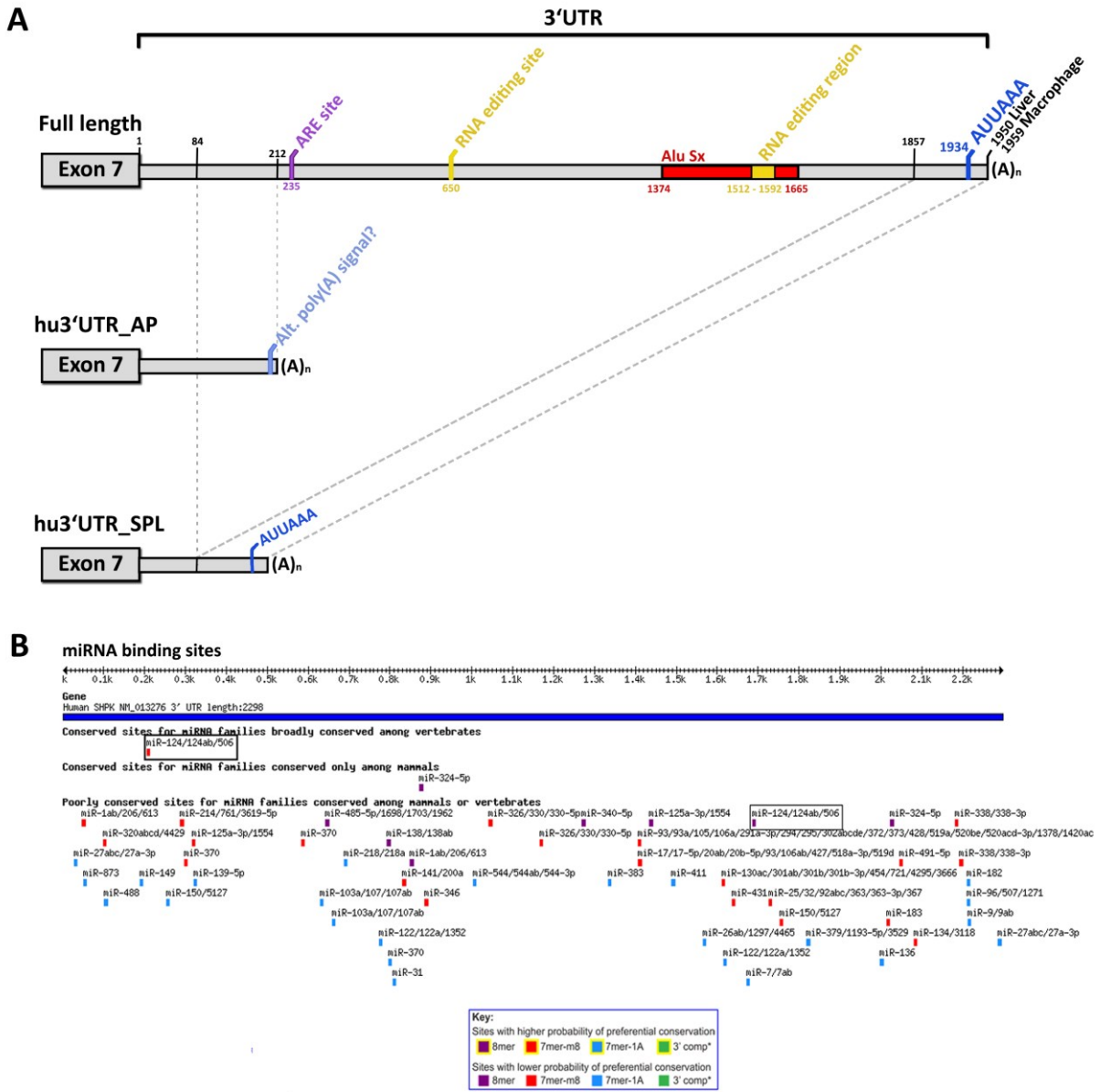


Figure 9: Schematic representation of human CARKL transcript 3'UTRs. (A) The full length 3'UTR contains a variety of regulatory elements, such as RNA editing sites, an ARE site (AUUUA motif) and miRNA binding sites. The full length sequences are terminated by a canonical AUUAAA polyadenylation sequence. The two novel 3'UTRs are significantly shorter and lack most of the regulatory sequences depicted above. hu3'UTR_AP identified in human liver, ends after 212bp and is likely to be terminated by a yet unknown, alternative polyadenylation signal. Another alternative 3'UTR found in liver, hu3'UTR_SPL, results from splicing of the 5'- and 3'terminal ends of the 3'UTR. This leads to a short, 185bp long transcript containing the canonical polyadenylation signal. The numbering in this figure represents the distance in base pairs, while 1 marks the beginning of the 3'UTR. **(B)** miRNA binding site prediction of the full length human-CARKL 3'UTR (TargetScan <http://www.targetscan.org/>). The majority of these putative binding sites are absent in the short 3'UTRs, abrogating possible control mechanisms carried out by the respective miRNAs.

Figure 10 depicts some interesting features of the alternative 3'UTR sequences. The novel transcript hu3'UTR_AP, does not terminate with a canonical- or published alternative-polyadenylation signal (**Fig. 10A**). Neither popular poly(A) prediction tools (Dragon PolyA spotter: www.cbrc.kaust.edu.sa/dps/; PolyA-pred: www.imtech.res.in/raghava/polyapred/), nor manual search could locate one of the known polyadenylation signals (listed in **Fig. S5**). Polyadenylation however cannot be excluded at this site, since ~10% of the human mRNAs do not contain a yet characterized polyadenylation signal [86]. This study further showed that the alternative polyadenylation signals, generally rich in adenosines, tend to reside within 5 to 45 nucleotides upstream of the poly(A)-tail. I tried to locate possible sites of novel alternative polyadenylation signals in hu3'UTR_AP according to these criteria.

In contrast to hu3'UTR_AP, hu3'UTR_SPL ends with exactly the same polyadenylation signal as the full length transcript (**Fig. 9**). In latter case, this is accomplished by a splicing event that removes 1645bp of the 3'UTR, containing the majority of annotated regulatory elements. Supporting our experimental observation, the prediction programme SplicePort (<http://spliceport.cbcb.umd.edu/>) also preestimated possible donor and acceptor splice sites at the putative 3'UTR exon/intron boundaries (**Fig. 10B**). Moreover, prediction of splicing enhancer and silencer elements (<http://www.umd.be/HSF>) showed putative motifs in the proximity of the hu3'UTR_SPL splice sites (**Fig. S6**). Further analysis is necessary to verify the splice site and to identify the polyadenylation signal in the alternative transcripts of human liver.

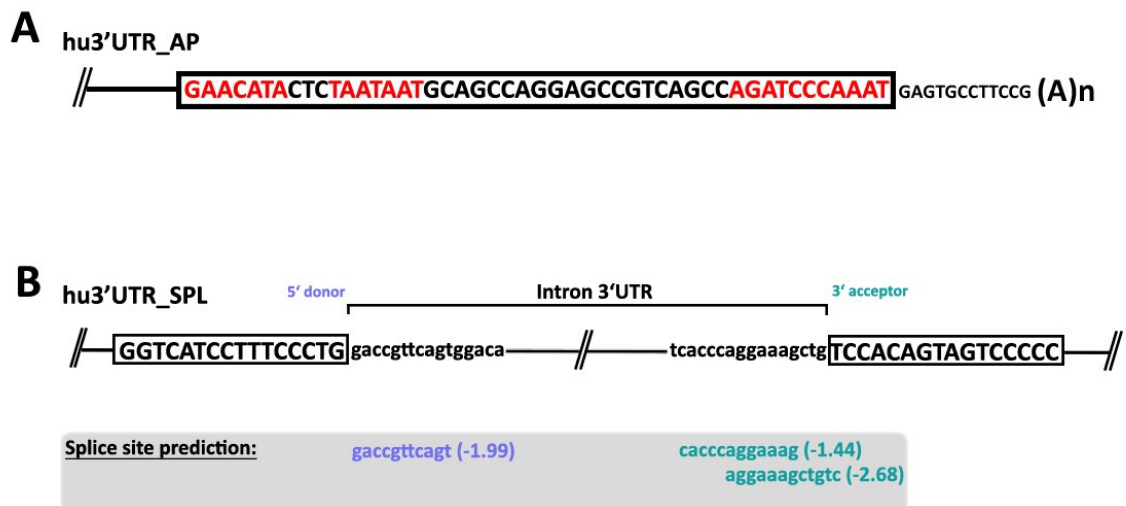


Figure 10: Putative polyadenylation and splice sites of the novel human CARKL 3'UTRs. (A) hu3'UTR_AP doesn't contain a yet described polyadenylation signal. The putative poly(A) signal may reside within the highlighted box at the 3'end of the transcript. Especially interesting parts of the sequence, with high adenosine content are coloured in red. For a list of alternative polyadenylation signal frequencies see **Fig. S5** (B) Donor and acceptor splice sites of hu3'UTR_SPL. The grey box represents predicted spliced sites, generated by the programme SplicePort (<http://spliceport.cbcb.umd.edu/>). A score of -1 produces a false positive rate of 10.74% for AG locations and 8.82% for GT locations, while a score of -2 produces a false positive rate of 48% for AG locations, 32.43% for GT locations. Note that solely based on the sequence information, the donor/acceptor sites of hu3'UTR_SPL could also be shifted two nucleotides to the left (upstream) side: a TG dinucleotide is present both in the 5'end of the exon near the donor splice site as well as in the 5'end of the intron. Please note that sequence prediction programs as the DNA code as input format. All T nucleotides depicted in this figure correspond to U nucleotides in the RNA.

To sum up, I could identify two novel CARKL 3'UTRs in the 3'RACE cDNA libraries derived from human liver RNA. These shorter 3'UTRs lack the majority of predicted regulatory sequences present in the full length CARKL 3'UTR. The elimination of regulatory sequences may allow CARKL expression independent of the presence of the corresponding trans-acting factors. In contrast to liver, human macrophages contained solely CARKL transcripts with full length 3'UTR. This may indicate that CARKL expression in macrophages is more strictly controlled through factors interacting with 3'UTR elements.

Alternative 3'UTRs of murine CARKL

Beside the human transcripts, we were also interested in the diversity of CARKL mRNAs in mouse cells and tissues. As in humans, alternative murine CARKL transcripts have not been characterized so far. I used the same method as for the human transcripts: the generation of RACE libraries, followed by nested PCR, cloning and sequencing. Again, the focus was on the 3'UTR, but we were also able to obtain data from 5'RACE, my method of choice for exploring mRNA variety in the coding region and 5'UTR. For the RACE experiments in mouse, I used RNA from liver and kidney, as well as from bone marrow derived macrophages (BMDMs).

RNA integrity analysis, 3'RACE library generation and the nested RACE-PCRs were performed as described above for the human samples. In mouse, the outer PCR CARKL primer binds exon 6, while the inner primer binds exon 7. After gel purification, cloning and sequencing of the products, I aligned the results to the full length CARKL sequence (NM_029031.3). Analogously to the procedure with human samples, some of the products of the nested PCRs either turned out to be unspecific products of the primers, or to result from internal priming of the 3'RACE adapter during the reverse transcription step. The latter could be observed in regions with internal poly(A) stretches, mimicking the terminal poly(A) tail.

Using the mouse liver and kidney 3'RACE libraries, derived from previously isolated RNA samples, I was able to amplify the full length CARKL 3'UTRs (**Fig. 11A**). These had a length of 1311bp in liver and 1316bp in kidney. Also the library generated from BMDM-mRNA contained the full length transcript (**Fig. 11C**), with a length of 1312bp. The latter RNA was derived from BMDMs isolated from totally 5 wild type C57BL/6N mice.

The murine BMDMs expressed exclusively the CARKL transcript with the full length 3'UTR, analogously to human macrophage. Similarly to human liver, I could identify novel alternative 3'UTRs in murine liver and kidney (**Fig. 11B**). Sequence analysis showed that the amplified transcript ends correspond to net 3'UTR lengths of 334bp and 62bp respectively. Both result most likely from alternative polyadenylation, therefore we refer these transcript ends as "m3'UTR_AP1"(liver) and "m3'UTR_AP2"(kidney) here. It may be possible that m3'UTR_AP1 is also present in kidney, however it could not be cloned so far. An unspecific product of nearly the same size complicates the purification and cloning process in this case. The RACE product in mouse were cloned and sequenced from on independent RACE-PCRs approaches so far. PCR bands in the size of the alternative 3'UTRs could however be

observed in multiple PCR reactions with different polymerases. Generally, at least two clones were sequenced of each cloned PCR product.

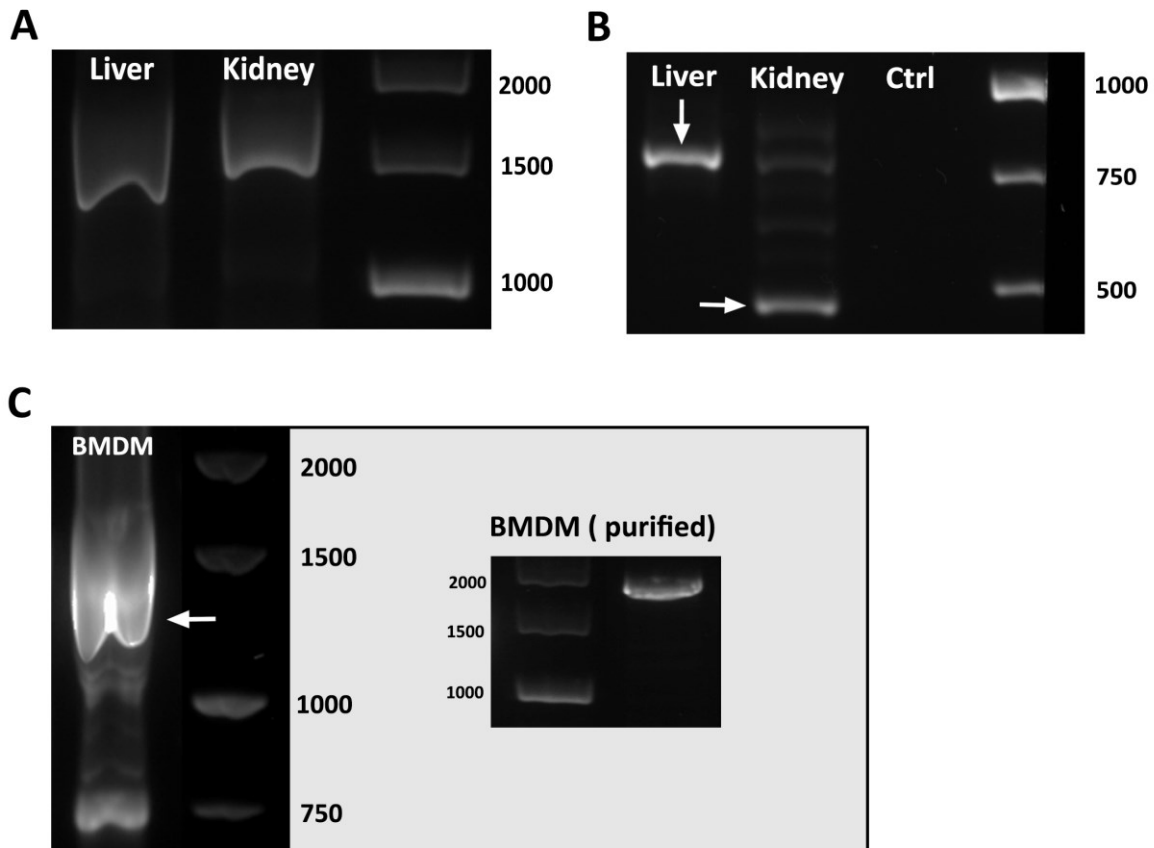


Figure 11: CARKL 3' RACE products of mouse liver, kidney and bone marrow derived macrophages: All 3'RACE nested-PCRs products were amplified with forward primers in exon 6 (inner PCR) and exon 7 (outer PCR) **(A)** Full length CARKL 3'UTR products of mouse liver and kidney (inner PCR). The unequal bandheight may result from gel electrophoresis inhomogeneities and too high content of DNA, as analytical digestion of the respective PCR-product single clones showed approximately equal insert size **(Figures S2)**. **(B)** Novel shorter, alternative CARKL 3'UTRs amplified from mouse liver (m3'UTR_AP1) and kidney (m3'UTR_AP2). Both products result likely from alternative polyadenylation. It is possible that m3'UTR_AP1 is also present in kidney, however it could not be cloned so far. The PCR-control reaction using a plasmid containing the mouse CARKL full length 3'UTR as template did not result in a PCR product, indicating the sequence specificity of the RACE primers. **(C)** 3'RACE product of the full length 3'UTR in BMDMs. The bandshift may result from too high DNA-content on the gel. The small picture in the box shows the same product after gel purification and loading of equal amounts of marker DNA.

Regulatory elements of the murine CARKL 3'UTR

Next I performed a predictive characterisation of the mouse CARKL 3'UTR. Similarly to the human CARKL 3'UTR, several regulatory elements can be found in the full length version of the 3'UTR (**Fig 12**). Target scan miRNA regulatory site prediction prognosticates that various miRNAs might target this region (<http://www.targetscan.org/>), like miR-124/506, miR-128 and miR-149. The miR-124/506 binding site is a conserved element which is also present in the human CARKL 3'UTR. Furthermore two ARE-sites could be detected, located at both ends of the 3' upstream region (<http://rna.tbi.univie.ac.at/cgi-bin/AREsite>).

Next, I analysed the murine CARKL 3'UTR for common regulatory patterns using the publicly available tools RegRNA2.0 (<http://regrna2.mbc.nctu.edu.tw/>) and UTR-scan (<http://itbtools.ba.itb.cnr.it/utrscan>). Both tools predicted an UNR-binding site near the putative poly(A)-signal of m3'UTR_AP1 (**Fig. 12A**). UNR-binding sites might be bound by the UNR-protein, which is known to interact with poly-A binding protein [87]. Chang and colleagues further reported that this protein has a key role in translationally coupled mRNA turnover. RegRNA2.0 and UTR-scan further predicted a K-box, a Brd-box and a GY-box element in the full length CARKL 3'UTR. These boxes, frequently found in Notch signalling target genes, are bound by miRNAs which negatively regulate transcript stability and translational efficiency [88].

The early termination of the short transcripts results in the loss of the bigger part of regulatory elements. The m3'UTR_AP1 retains the UNR-binding site and the ARE-site located in the 5'end of the 3'UTR, while m3'UTR_AP2 lacks both ARE sites, the UNR-binding site and almost all putative miRNA binding sites. The K-box, Brd-box and GY-box elements are also absent in the shorter 3'UTRs.

Analogously to the results in human, I could only identify shorter CARKL 3'UTRs in RACE libraries derived from metabolically active tissues, but not in those derived from macrophages. The full length murine CARKL 3'UTR contains also many regulatory elements, similarly to the human variant. These results suggest that alternative regulation of CARKL mRNA through factors acting on the 3'UTR is a conserved mechanism. Both species contain putative ARE sites, binding sites for miRNA 124/506, as well as non-conserved, species-specific regulatory elements.

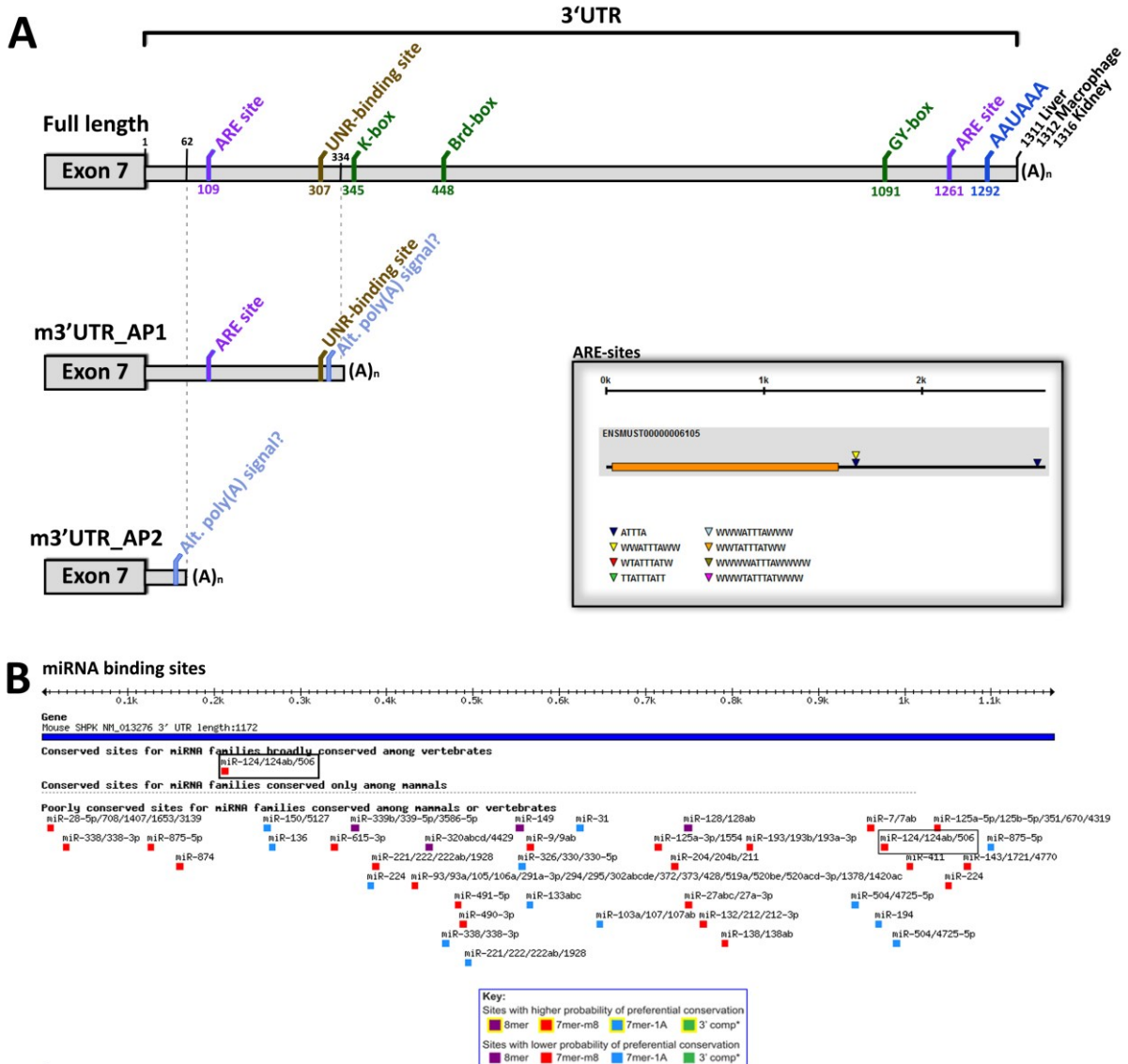


Figure 12: Schematic representation of mouse CARKL transcript 3'UTRs. (A) The full length 3'UTR contains multiple regulatory elements, such as two ARE sites and several miRNA binding sites. The box depicts details about the detected ARE sites. The full length 3'UTRs found in mouse liver, kidney and BMDMs have similar lengths and terminate with a canonical AAUAAA polyadenylation signal. Two novel 3'UTRs could be identified in mouse liver (m3'UTR_AP1) and kidney (m3'UTR_AP2), both terminating without a known polyadenylation signal. m3'UTR_AP1 has a length of 334bp and retains the first ARE site, as well as the putative UNR-binding site. Both ARE-sites and the UNR-binding site are absent in the short 3'UTR found in kidney (m3'UTR_AP2), which has a net length of 62bp. The regulatory elements K-box, Brd-box and GY-box are only present in the full length CARKL 3'UTR. The numbering represents the distance in base pairs, while "1" marks the beginning of the 3'UTR. **(B)** Results of miRNA binding site prediction in the published full length mouse-CARKL 3'UTR (TargetScan <http://www.targetscan.org/>). The majority of these putative binding sites are absent in the short 3'UTRs, abrogating possible control mechanisms carried out by the respective miRNAs.

Figure 13 provides insight into the terminal ends of the alternative 3'UTRs. PolyA-signal prediction tool as Dragon PolyA-spotter and PolyA-pred could not locate a characterized polyadenylation sequence the short 3'UTRs. Manual search for published alternative polyadenylation sequences (listed in **Fig. S5**) also didn't result in hits. As mentioned earlier, polyadenylation events are however still possible at these sites, as many known mRNAs do not contain verified polyadenylation signals [86].

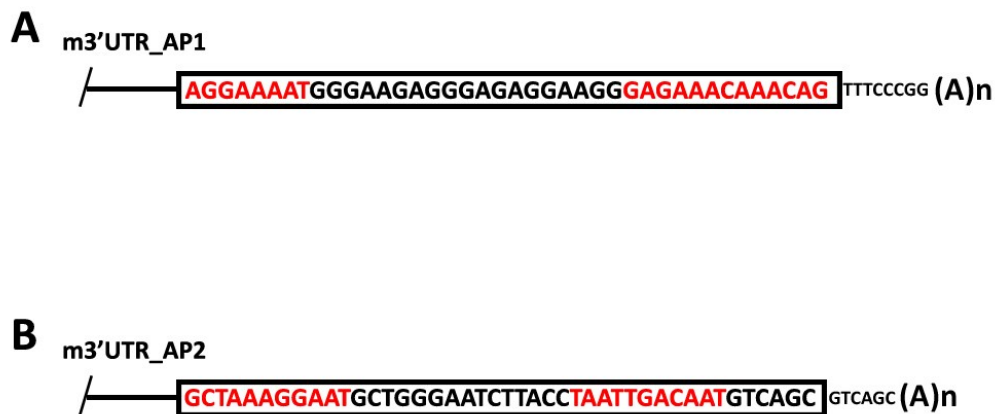


Figure 13: Novel mouse CARKL 3'UTR polyadenylation regions. Both short CARKL 3'UTRs, “liver short AP” as well as “kidney short AP” do not contain yet described polyadenylation signals. Putative polyA-signals may reside within the highlighted box at the 3'end of the transcript. Especially interesting parts of the sequence, with high adenosine content are coloured in red. For a list of alternative polyadenylation signal frequencies see **Fig. S5**. Please note that the sequencing information is provided in the DNA code. All T nucleotides depicted in this figure correspond to U nucleotides in the RNA.

Beside the above described 3'RACE experiments, we also started to investigate CARKL transcript variety in the coding region and 5'UTR using 5'RACE. The workflow of 5'RACE is analogous to 3'RACE, except of the steps required for library generation (described in the background section). For the nested PCRs we used 5'RACE primers (forward), which bind to the 5'RACE adapter, ligated to undegraded mRNAs containing the 5'CAP structure. These were combined with gene specific CARKL primers (reverse), binding in exon 7 (outer PCR) and exon 6 (inner PCR).

We could successfully amplify a full length CARKL 5' fragment from the mouse kidney 5'RACE library (**Fig. 14A**). This fragment ranges from the 5'RACE adapter binding site, which marks the former position of the 5'CAP structure, to the gene specific-primer binding site in exon 6. Sequence analysis of the corresponding clones revealed that the transcriptional start site is initiated 4bp earlier than previously published (ENSMUST0000006105). **Figure 14B** shows the sequence of the mouse kidney 5'UTR. By applying common pattern prediction (RegRNA 2.0 and UTRscan), I could not find regulatory elements residing in this 4bp extension of the 5'UTR extended by these 4 nucleotides.

Further, the 5'RACE experiment didn't produce other specific products than the full length 5'CARKL fragment containing the 5'UTR and exons 1 to 6. Therefore the existence of alternatively spliced CARKL coding regions which include exon 6 and exon 7 is unlikely in mouse kidney. The product was only sequenced between the translational start and the inner RACE primer (exon 6). I used a reverse primer complementary to exon 7 in the outer nested PCR-reaction of the sequenced fragment. As specific 5'RACE products are usually not observed after a single PCR round, primer binding in exon 7 was very likely necessary to generate the observed CARKL 5'fragment.

However, with this approach, I can't exclude the possible existence of alternative splice variants which lack exon 6 or exon 7. This may be investigated in future by applying PCRs with different primer sets. A first approach could be to use a forward primer in the 5'UTR in combination with a reverse primer in the proximal part of the 3'UTR. Up to now, attempts to amplify specific CARKL 5'RACE products in libraries of human and mouse liver as well as macrophages were not successful

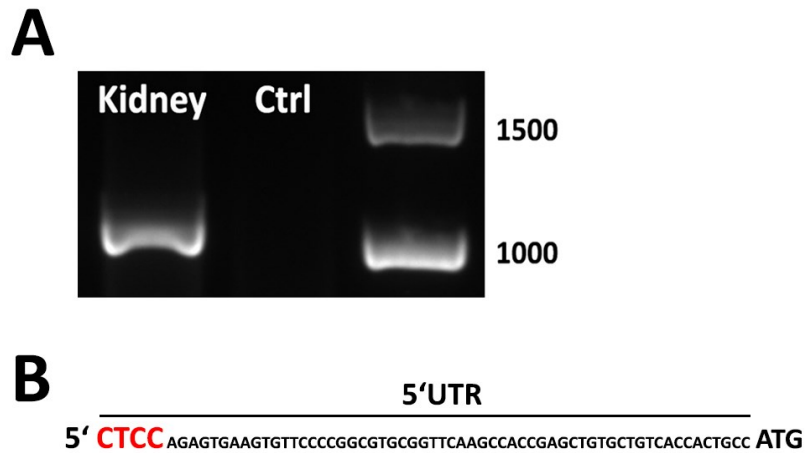


Figure 14: CARKL 5'UTR in mouse kidney: 5'RACE-nested PCR was performed with reverse gene-specific primers in exon 7(outer PCR) and exon 6 (inner PCR) **(A)** The mouse kidney 5'RACE library contained exclusively the full length CARKL 5' fragment. Alternative splicing variants of the 5'UTR as well as the protein coding region between 5'UTR and exon 6 were not observed. The PCR control sample, containing a plasmid carrying the full length mCARKL cDNA didn't produce a product in the nested PCRs, indicating the sequence specificity of the RACE primers. **(B)** Sequence analysis of the 5'UTR revealed that the transcriptional start is 4 bases earlier than previously published (red nucleotides). Please note that the sequencing information is provided in the DNA code. All T nucleotides depicted in this figure correspond to U nucleotides in the RNA.

In summary, the experimental results revealed the existence of alternative CARKL 3'UTRs in metabolically active organs, which partly lack putative regulatory elements present in the full length 3'UTR found in macrophages. The trend to partially use short CARKL 3'UTRs in metabolically active tissues versus the exclusive utilization of full length 3'UTRs in macrophages seems to be conserved in mice and men. I further started to characterize the CARKL 5'UTR and coding region. 5'RACE results of murine kidney showed that alternative splicing of CARKL could not be detected between the transcriptional start site and exon 6, and that transcription starts 4bp earlier than previously reported

Discussion:

In the course of this master thesis I was able to amplify the full length CARKL transcripts in human liver and macrophages, as well as in liver, kidney and BMDMs from *mus musculus*. I could further identify shorter 3'UTRs in human and mouse liver, and also in mouse kidney. In contrast, I couldn't find alternative CARKL transcripts in macrophages, neither in humans, nor in mouse. *In silico* tools and online databases revealed that CARKL has a long, highly regulated 3'UTR, with many potential binding sites for modulatory trans-elements, which may define mRNA fate in a cell- and situation dependent fashion. In summary, our findings lead to the hypothesis that CARKL regulation on the mRNA level is likely to be regulated more stringently in macrophages, where CARKL activity needs to be rapidly changed in response to extracellular stimuli, than in metabolically active tissues, where tonus CARKL expression may be necessary to maintain the balance in primary carbohydrate metabolism.

Different *in silico* tools revealed ARE sites and various miRNA binding sites in the CARKL 3'UTRs of human and mice. Of these putative miRNA binding sites only one miR-124/506 was evolutionary conserved. Further I could also find species specific regulatory sequences. The human CARKL 3' contains RNA editing sites, which mainly reside in an AluSx repetitive element. Alu-repeats are a primate-specific class of short interspersed elements, which frequently contain RNA editing sites [83].

Contrary, common pattern finder programs predicted K-box, Brd-box and GY-box elements in the 3'UTR of mice. These boxes can be frequently found in the target genes of Notch signalling and it was shown that mRNAs containing these elements are negatively regulated by miRNAs [88]. Notch signalling is evolutionarily conserved in metazoan development and has an important role in the regulation of morphogenesis, differentiation and also cellular metabolism [89, 90]. Interestingly, CARKL has indeed been shown to be involved in Notch signalling [91]. The authors identified CARKL as an element that may support Notch activation by using an siRNA-library screen approach in a human mammary carcinoma cell line, which was engineered to report Notch. Curiously, the K-box, Brd-box and GY-box regulatory elements are absent in the CARKL 3'UTR of humans. It remains to be determined how Notch signalling and the sedoheptulose kinase are functionally linked to each other. Further it may be interesting to investigate whether CARKL mRNA turnover is influenced by Notch signalling through a 3'UTR mediated mechanism. Generally, the CARKL 3'UTR seems to represent a very versatile regulatory region, with multiple putative elements involved in turnover modulation.

A key determinant of cellular protein amount is represented by the quantity of the corresponding mRNA available for ribosomal translation. The complex regulation of mRNA levels is defined by a balance of synthesis, processing and degradation. All these events are regulated by numerous intra- and extracellular factors to adapt the steady-state concentration of a given mRNA to the momentary cellular and biochemical conditions. Messenger RNA half-life in mammals can vary between less than 20 minutes and more than 50 hours, while the stability of the same mRNA can significantly vary depending on extracellular stimuli [92]. Changes in mRNA stability, largely determined by cis-acting 3'UTR elements and the availability of corresponding binding factors at a given time, can increase or decrease the steady-state level of transcripts relatively rapidly [93].

Sequence analysis of the shorter CARKL 3'UTRs in liver and kidney showed that they might result from alternative splicing and alternative polyadenylation events. Transcriptome analysis showed that more than 90% of human genes undergo alternative splicing, while the highest isoform variability can be found in the 3'UTR [94]. It has been further estimated that more than half of all mammalian genes undergo alternative polyadenylation events [95]. The selection of alternative poly (A) sites is likely to be a developmentally regulated, tissue-specific process. Importantly, the expression of transcripts with longer 3'UTRs was shown to correlate with lower protein expression levels [96].

According to these findings, the short CARKL isoforms found in liver and kidney tissue may escape from destabilizing protein- and miRNA interactions, and thus ensure constantly sufficient CARKL levels. On the other hand, the full length transcript, with its manifold regulatory elements in the 3'UTR, might be bound by a variety of trans-acting factors, depending on the cellular requirements for sedoheptulose kinase activity. This may provide enough flexibility to rapidly modulate CARKL mRNA steady-state levels as well as translational efficiency in situations such as macrophage activation.

Although I could not find alternative transcripts with short 3'UTRs in resting macrophages, I cannot exclude their existence in an activated status. It has been reported that LPS-stimulated macrophages increase levels of the 64-kDa Cleavage Stimulatory Factor (CstF-64), which induces alternative, proximal poly(A)-site selection [97]. It remains to be tested if CARKL 3'UTR is also affected by stimulation. Further alternative 3'ends may be present in other tissues not tested during my thesis. 3'RACE represents an optimal method to identify such novel 3'UTR variants. However this technique also has its drawbacks. The anchored oligo d(T) primer used for reverse transcription can also bind at internal adenine rich stretches and thus generates a quite high frequency of truncated cDNAs [98]. We could

reduce the occurrence of internal priming by raising the reaction temperature and adding less of the oligo d(T) primer during reverse transcription. Northern blotting combined with primer extension represents an alternative method to explore transcript diversity. As this method is more laborious, does not generate direct sequence information and further requires the use of radionuclides, I preferred RACE.

As a next step, the biological role of the alternative 3'UTRs should be further investigated by various other methods. This includes the quantification of alternative transcript levels by real time PCR and the establishment of expression pattern in various tissues, primary cells as well as cell lines. As liver and kidney are organs with mixed cell populations, homogenous primary cell isolates need to be tested to assign the various CARKL 3'UTRs to specific cell types.

Further, it would be interesting to investigate the effect of shorter 3'UTRs on mRNA stability and translational efficiency. The stability and turnover of the various transcripts with different 3'UTR can be compared by applying transcriptional inhibitors, as actinomycin D to cells and follow mRNA decay over time (e.g. by RT-PCR). The translational efficiency of the respective UTRs could be compared in a representative cell line using luciferase-reporter assays, where the various 3'UTRs are fused downstream of the luciferase gene. The functional elements such as splice sites and novel polyadenylation signals also need further verification by alternative techniques. Additionally, functional characterisation of regulatory sites such as ARE, miRNA binding sites and RNA editing sites should be performed by site directed mutagenesis. Possible changes in mRNA turnover or translational efficiency could be monitored by the methods described above.

Beside functional characterization of 3'UTR elements, also the investigation of the CARKL 5'UTR and coding region should be continued in different cells and tissues. I started to investigate CARKL 5' diversity by applying the 5'RACE technique. As 5'RACE turned out to be much more demanding than 3'RACE and it was only successful in mouse kidney so far (**Fig. 14**). Here, transcription starts 4bp earlier than previously published. The published CARKL 5'UTRs have only a length of 103bp in humans and 59bp in mice, which is very short compared to the length of the 3'UTR (1960bp and 1318bp, respectively). In the murine CARKL 5'UTR, common pattern finder tools didn't predict regulatory sequences related the translational control or mRNA turnover. In contrast, RNAreg2.0 predicted a putative mi-RNA binding site (hsa-miR-12b-2-3p) in the 5'UTR of the human CARKL mRNA. As the CARKL 5'UTR is shorter and contains clearly less putative regulatory elements than the 3'UTR, I also expect it to play a less important role in the dynamic regulation of mRNA turnover.

Nevertheless 5'RACE should be continued to investigate possible sequence diversities, which may occur in the CARKL 5' untranslated regions of different cell types and tissues. 5'RACE with the CARKL transcript is especially difficult because the CARKL RNA has a quite high GC-content in the 5'end, a feature that frequently inhibits the advance of reverse transcriptase. The use of reverse transcriptases with a very high thermostability may help to improve 5'RLM-RACE performance in this case.

The 5'RACE method can also be used to study sequence differences in the coding region, especially if the gene specific primers are designed to bind 3' terminal exons or the 3'UTR. I amplified the murine kidney 5'-CARKL-fragment between the former 5'CAP site and a gene specific primers binding in exon 6. Within this setting, I could not find other transcript variants than the full length coding sequence. However I cannot exclude the existence of transcripts with variations in exon 6 or exon 7. Further, alternative splice variants of the CARKL coding region may be present in other tissues.

Alternatively to 5'RACE, the coding sequence diversity should be further investigated by end-point PCR with different primer settings. The utilization of sequence specific primers binding the 5'- as well as 3'UTR may especially be useful. They could lead to the amplification of all kinds of coding-sequence splice variants, as far as the primer binding sites are included. The very first experiment with this end-point PCR approach, tested in a human liver cDNA library, indicates that CARKL transcripts with alternative coding regions may exist. While the control template, a plasmid containing the full length CARKL sequence, led only to the amplification of products which correspond to the calculated length of the full-length sequence, some additional, shorter products could be observed with the liver cDNA as template. However as cDNA libraries are complex mixtures, derived from a huge number of cellular RNAs, the primers may also amplify unspecific products. This preliminary experiment has to be repeated, followed by the purification, cloning and sequencing of the putative splice variants.

Supporting these observations, various alternative CARKL transcripts with changes in the coding region are predicted in mice, as well as humans (**Fig. S3**). Further, deep sequencing data (Illumina HiSeq) derived from the RNA of adipocytes before and after differentiation from human SVF cells indicates that CARKL may exist in different transcript variants (**Fig. S4**).

During my thesis I started to investigate the diversity of CARKL transcripts and could identify novel 3'UTRs which lack many regulatory elements present in the conventional, full length 3'UTR. Generally, the CARKL mRNA seems to be highly regulated through its 3' upstream region. The relevant cellular factors involved in the modulation of CARKL mRNA turnover still need to be elucidated. It would be also interesting to find out the biological settings which change the fate of CARKL mRNA. Generally a long 3'UTR with manifold regulatory sequences may allow a more stringent regulation of the message. This may be of importance in macrophages, which rapidly downregulate CARKL mRNA upon stimulation with LPS [5]. The role of the regulatory sequences in the CARKL 3'UTR in such biological contexts could be further studied by fusing the 3'UTR to the coding sequence of an instable reporter construct, for example luciferase or GFP. The importance of certain regulatory sequences might further be investigated by site directed mutagenesis. These reporter system approaches might also be combined with various stimuli, which influence CARKL expression and may lead to a better understanding of the cellular and physiological role of 3'UTR-mediated regulatory mechanisms

Part 2 – Transgenic CARKL mice

In the last years, research on heptose metabolism and CARKL made significant progress. Most functional data available was however derived from *in vitro* experiments and needs to be verified *in vivo*. We recently generated transgenic mice, which overexpress CARKL. They represent an appropriate tool for the investigation of CARKL's function *in vivo*, for example its role macrophage polarisation. The second part of my master thesis deals with the breeding and initial characterization of these transgenic mice.

Background

CARKL regulates macrophage polarization in vitro

Inflammation, triggered by infection or injury, is an energy-costly endeavour to provide protection against pathogens, to remove damaged cells and further to initiate essential repair mechanisms. Immune cells participating in inflammation, like classically activated (M1) macrophages, undergo a metabolic shift towards increased aerobic glycolysis and PPP activity, while repressing mitochondrial respiration [14]. These bioenergetic adaptations allow the cells to switch to a highly active state, marked by enhanced phagocytosis, antigen presentation and synthesis of defense factors. The robust metabolic reprogramming towards aerobic glycolysis and increased glutamine use is also frequently observed in tumour cells [13].

Historically, classically activated (M1) macrophages are termed after the stimulation of resting macrophages with lipopolysaccharide (LPS) and interferon- γ (IFN- γ), that results in an innate, inflammatory phenotype, characterized by high microbicidal activity, as well as increased ROS and proinflammatory cytokines (e.g. TNF) production [14, 99]. Contrary, an alternative macrophage phenotype, resulting from activation with IL-4 and IL-13, is associated with anti-inflammatory cytokine production (e.g. IL-10), reduced antigen presentation, tissue repair and humoral immunity [100]. In contrast to M1 macrophages, these so called M2 macrophages rely on oxidative metabolism [99]. In a functional screen for kinases involved in macrophage polarization, CARKL was shown to repress LPS-induced tumour necrosis-alpha (TNF- α) secretion [5].

The ectopic expression of CARKL constrained LPS-induced macrophage activation, resulting in insufficient redox power and blunted intracellular superoxide formation [5]. These cells, with high glycolysis and CARKL-derived S7P levels, accumulated S7P, ribose phosphate and oxidized redox couples, most likely by feedback inhibition acting on the oxidative arm of the PPP. LPS stimulation of human and murine macrophages leads to the rapid downregulation of endogenous CARKL expression human and murine macrophages.

The knockdown of endogenous CARKL in resting macrophages decreased mitochondrial oxygen consumption rates, while increasing extracellular acidification rates, indicating increased lactate production and reduced mitochondrial respiration [5]. Elsewhere, it was reported that high S7P levels inhibit hexose phosphate isomerase and competitively obstruct F6P phosphorylation by PFK [30]. Thus a decrease in S7P levels, resulting from low CARKL activity, might favour glycolysis [8]. Ectopic CARKL overexpression in resting macrophages, associated with increased oxygen consumption rates and decreased lactate production compared to control cells, triggered a mild M2-like, anti-inflammatory phenotype, characterised by decreased TNF and increased MRC1 (mannose receptor 1) expression [5]. CARKL thus appears to be directly related to the decision of macrophage polarization, most likely by modulating cellular bioenergetics and redox state.

The recently generated CARKL overexpressing mice represent a powerful tool for the investigation of CARKL's effect on macrophage polarization and inflammation. They may help us to verify the immune-related aspect of CARKL under physiological conditions.

Generation of transgenic CARKL-mice

In order to investigate sedoheptulose metabolism and the role of CARKL *in vivo*, transgenic mice were created, ubiquitously expressing sedoheptulose kinase. The transgenic mice were generated by pronuclear injection of a construct for CARKL expression.

The used construct, based on the pCAGGS vector is frequently used for generation of transgenic mice, is depicted in **figure 15** [101-103]. Driven by the strong, hybrid AG promoter, it expresses the mouse CARKL coding sequence fused to the 3'UTR of β -globin, which has a stabilizing effect on the transgene mRNA, ubiquitously in all tissues. Further, the CMV immediate early gene enhancer is a pan-active control element in transgenic mice. As it was shown to be activated by LPS stimulation, this construct represents an ideal tool to interfere with LPS-dependent CARKL downregulation [5, 104].



Figure 15: Construct used for the generation of the transgenic CARKL mice. See text for description.

The transgenic mice were generated in a C57BL/6N background. Inbred C57BL/6 mouse strains are commonly used in transgenic-expression and knockout knock-out studies. The two major substrains are C57BL/6J and C57BL/6N [105]. Most substrains derived from the original C57BL/6J strain carry a mutation in the nicotianamide nucleotide transhydrogenase (Nnt) gene resulting from genetic drift [106]. The C57BL/6N substrain was separated in 1951 and lacks this mutation [107]. As the deletion of the Nnt might have effects on bioenergetics and metabolism, the C57BL/6N background is more suitable for the study of CARKL.

The generation of the transgenic mice was performed by pronuclear injection of the CARKL-construct, in cooperation with the University of Veterinary Medicine in Vienna. In our untargeted approach, the transgene integrates into the genome at random sites. Therefore expression is subjected to position effects depending on the integration site. The injected

DNA can further integrate in multiple copy numbers and may lead to expression level differences in the transgenic founders.

Pronuclear injection is widely used for the generation of transgenic mice [108]. **Figure 16** illustrates the simplified workflow of this method. It starts with the preparation of pronuclear-stage embryos from superovulated female mice. The transgene is then microinjected into one of the pronuclei of the fertilized mouse oocytes. Groups of the resulting embryos are reimplanted into the oviduct of pseudopregnant female mice. The resulting pups are screened for transgene integration. This can be for example achieved by PCR detection of the transgenic construct derived from tail DNA.

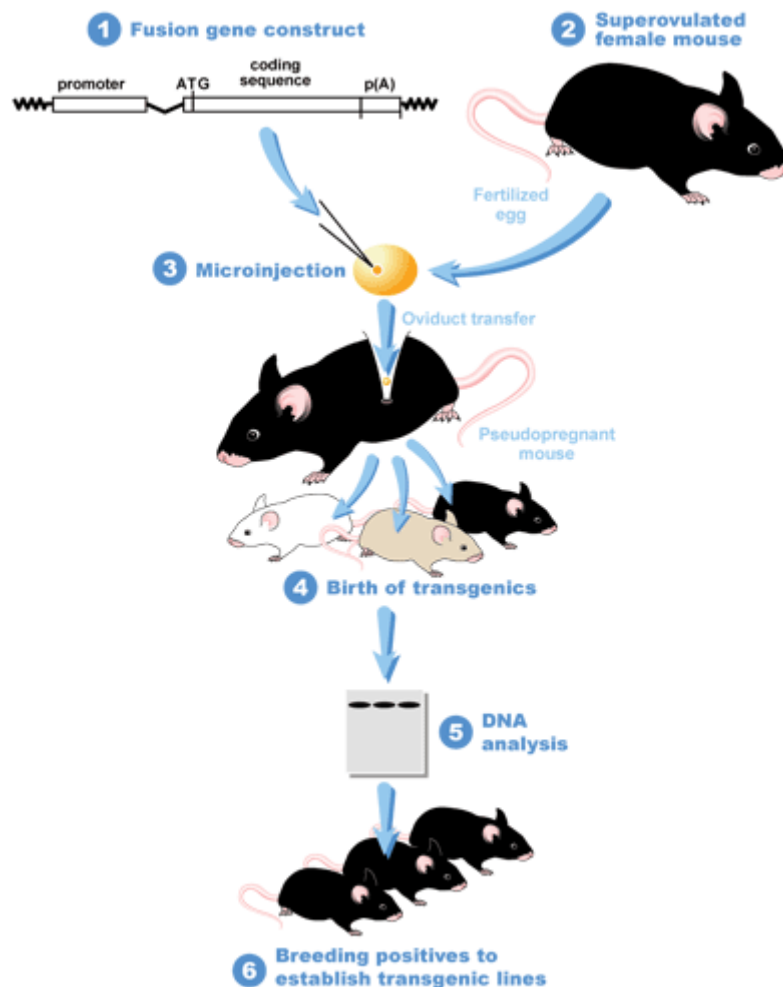


Figure 16: Transgenic mouse generation by pronuclear injection. A more detail procedure can be found in the main text. Image adapted from: Moores Cancer Center (UC San Diego) (<http://cancer.ucsd.edu/research-training/shared-resources/transgenic-core/services/Pages/pronuclear-injection.aspx>)

Aim

The aim of the second part of my master thesis was to plan the breeding and pretesting of four CARKL overexpressing mouse lines. This also included the initial characterization of transgene expression status in various organs.

Results

Genotyping strategy

After the founder of transgenic mouse lines are born, it is essential to establish reliable genotyping methods, as these represent the fundament for successful breeding. Pronuclear injection is not a very effective method, therefore it's important to screen the pups for transgene integration.

Conventional PCR of mouse tail-DNA extracts with transgene specific primers is a sufficient method to distinguish the transgenic founders from their wild type littermates. I used primers amplifying the CARKL transgene construct (pCAGGS-mCARKL). The forward primer is vector specific, while the reverse primer binds CARKL in exon 1.

As a positive control for the PCRs of the founder mice (F0), I used the pCAGGS-mCARKL plasmid as template because DNA of transgene-positive mice was not available yet. **Figure 17A** shows a representative example of a genotyping PCR.

The foster mothers, which received the oocytes from pronuclear injection, gave birth to totally 41 mice. Of these, only six carried the transgenic CARKL construct (14,6%). Five out of these were mated with wild type C57BL/6N mice. The genotyping of the resulting F1 generation (wild type mice and transgenic heterozygotes) was also performed with the above described conventional PCR approach. Four out of five breeding pairs gave rise to transgenic offspring in the F1 generation, which ultimately led to the establishment of four transgenic CARKL-mouse lines: A, B, D and E.

Next, cross-breeding of heterozygous F1 mice was performed. Heterozygous breedings (het-het) usually results in offspring of all 3 genotypes: wild type (wt), heterozygotes (het) and homozygotes (homo). At this point it became necessary to distinguish between CARKL het and homo mice. This is first of all important to maintain successful het-het and homo-homo breedings. As we also started to perform the first experiments with the transgenic mice, we wanted to make sure that we don't mix heterozygotes and homozygotes. The transgene could show gene dosage effects, which might define experimental results.

The untargeted pronuclear injection approach does not allow us to predict the insertion site of the CARKL-construct. Also, frequently multiple transgene copies might be integrated into the genome. Thus it is not possible to distinguish het from homo mice by conventional PCR. I decided to use a real time PCR based approach, which can be performed within a few hours, even with a large number of samples. I hypothesized that genomic DNA of homozygotes should contain roughly twice as much CARKL-construct as DNA of heterozygotes, relatively to total DNA content. For real-time PCR I used equal amounts of genomic mouse tail-DNA as template. These templates were amplified with CARKL-construct specific primers as well as genomic β -actin primers in parallel.

Analysis of real-time PCR data is frequently performed via the threshold cycle (Ct) value, which is determined from a log-linear plot of the PCR signal versus the cycle number [109]. If more specific template is present in a sample, the threshold is reached at an earlier PCR cycle. The Ct (also known as Cq or quantification cycle) equals to the number of cycles at which the fluorescence reaches the threshold. I applied a relative quantification method, known as the $2^{-\Delta\Delta Ct}$ (for convenience also called $\Delta\Delta$ -method) [109]. First I normalized the Ct results of each sample to the Ct of genomic β -actin by calculating the ΔCt -value ($\Delta Ct = Ct^{\text{Construct}} - Ct^{\beta\text{-actin}}$). Then relative CARKL-construct content was estimated by the calculation of the $\Delta\Delta Ct$ -value ($\Delta\Delta Ct = 2^{-(\Delta Ct^{\text{Sample}} - \Delta Ct^{\text{Control}})}$). In each litter, the sample with the lowest ΔCt -value (highest normalized CARKL-construct content) was chosen as the homozygote reference sample. Animals with $\Delta\Delta Ct$ values ~ 1 were assumed to be homozygotes, while those with $\Delta\Delta Ct$ values ~ 0.5 (lower relative construct content) were assumed to be heterozygotes. This method has limitations in terms of accuracy and works with the assumption, that the ΔCt remains constant when the PCR efficiency of both primer-pairs are the same. The results of a representative genomic real time PCR are shown in **Figure 17B-C**. Melting curve analysis illustrates the specificity of the used primers, as they produced a single peak indicating a single PCR product.

Specific CARKL-construct and genomic β -actin products tend to reach the threshold at PCR cycles between 23 and 28. Samples of wild type CARKL mice might also reach the CARKL construct threshold, but usually very late, in the range of the non-template controls.

In some wild type-samples, the CARKL-construct Ct values were observable some cycles earlier, but still significantly later than in transgenic mice. These products showed partially distinct melting curves, suggesting the low-efficiency amplification of an unspecific product. This could be partially explained by the binding of the reverse primer to the endogenous CARKL sequence in wild type mice. Non-template controls were generally observable at late threshold cycles (>35). The reliability of this method is depicted in **figure 17D**. Usually I was able to clearly distinguish between heterozygous and homozygous littermates. Nevertheless, in some cases the $\Delta\Delta$ Ct value of a sample was between 0.70 and 0.80, prohibiting reliable quantification with this method.

To sum up, I successfully applied two methods for genotyping. I used conventional PCR to distinguish between transgenic and wild type mice. Further I used a real time PCR approach to distinguish heterozygotes from homozygotes. These genotyping methods proved to be reliable, as all breedings, which were set up based on the genotyping results, led to the expected offspring genotypes. The het-het breedings produced wild type, heterozygous and wild type mice, while all homozygous breedings pairs had to solely homozygous progeny.

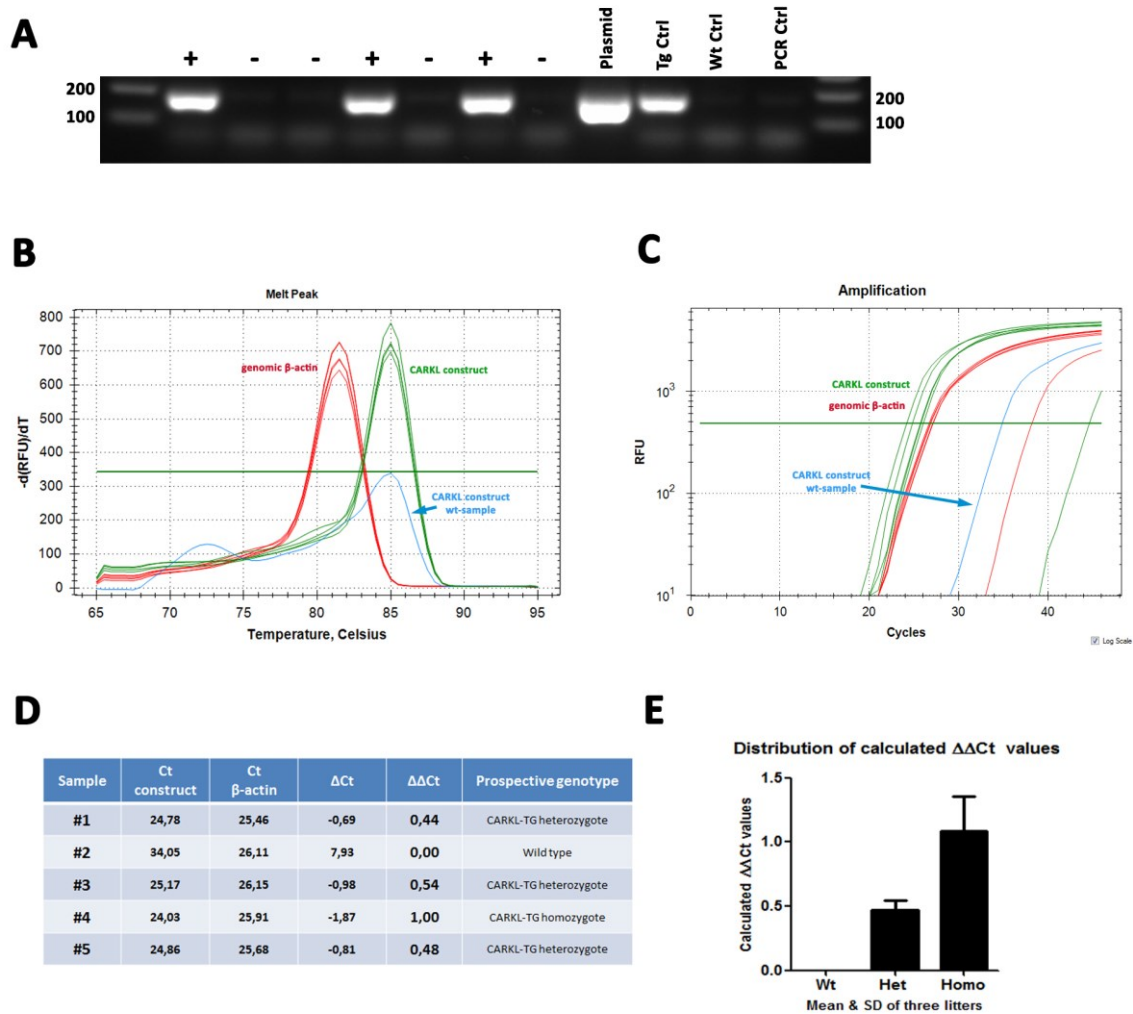


Figure 17: Genotyping of transgenic CARKL mice. (A) A representative example of a conventional genotyping PCR used to distinguish wild type (Wt; -) from transgenic (Tg; +) littermates. Tail-DNA extracts were PCR-amplified with primers specifically binding the CARKL-transgenic construct. I used a plasmid containing the CARKL-transgene construct and DNA of previously genotyped transgenic mice as controls. Further, H₂O was used as a PCR-control. I used a real time PCR approach to be able to distinguish between homozygous and heterozygous transgenic mice (B-E). I hypothesized that genomic DNA of homozygotes should contain roughly twice as much CARKL-construct as DNA of heterozygotes. As template I used equal amounts of genomic DNA and further normalized to genomic β -actin content. Results of representative genomic real time PCR of one litter are shown (B-C). (B) Melting curve analysis indicates that the used primers specifically amplify one product. (C) Specific CARKL-construct and genomic β -actin products tend to reach the threshold at PCR cycles between 23 and 28. Samples of wild type CARKL mice (blue line) might also reach the threshold, usually in the range of the non-template controls, which reach. The amplified product in wild type mice also shows a partially distinct melting curve, indicating the low-efficiency amplification of an unspecific product. Non-template controls (on the far right) might reach the threshold at PCR cycles >35. (D) A representative example of relative CARKL-Tg construct content calculation. The $\Delta\Delta$ Ct values were calculated in relation to the sample with the highest CARKL-Tg construct content (i.e. Δ Ct values). Animals with $\Delta\Delta$ Ct values \sim 1 were assumed to be homozygotes (Homo), while those with $\Delta\Delta$ Ct values \sim 0.5 were assumed to be heterozygotes (Het). (E) A bar graph shows the reliability of this method. It illustrates the distribution of calculated $\Delta\Delta$ Ct values of three litters from het-het breedings, which produce offspring of all possible genotypes (Wt, Het, Homo). Of totally 24 animals in these three litters, only 2 could not be clearly assigned to be hetero- or homozygous, as their calculated $\Delta\Delta$ Ct values were between 0.70 and 0.80.

CARKL protein expression in transgenic mice

Nowadays, genetically engineered mice represent the gold standard tool to study gene function *in vivo* and serve as important models to understand the mechanism of human diseases. To investigate the physiological consequences of sedoheptulose kinase on the whole organism, we generated CARKL-overexpressor mice as described above. Before these animals can be employed for experiments, a general characterization of their transgenic status is indispensable. Mating of the founder animals to wild type C57BL/6N and subsequent crossbreeding led to the establishment of four different transgenic mouse lines (A, B, D, E). The following experiments were performed with mice of the F2 generation.

As protein expression is decisive for transgene activity, we performed western blot analysis for the presence of CARKL. Our aim was to compare CARKL expression of the transgenic mice with their wild type littermates in different organs and further to compare the expression of the different mouse lines with each other.

Therefore we isolated liver, kidney and spleen from 4 animals of each line (2 transgenic vs. 2 wild type mice). We chose liver and kidney as these organs have high CARKL expression. Further we analysed CARKL protein expression in spleen. On all blots, equal protein amounts were loaded and β -actin was used as a loading control. We performed standard SDS-Page with reducing conditions, followed by wet blot transfer. The proteins on the membrane were detected with primary antibodies against specific epitopes and subsequently measured by HRP-coupled secondary antibodies.

Our results show that the transgenic mouse lines A, B and D strongly overexpress the CARKL protein (**Fig. 18**). Transgenic mice of line E express approximately equal CARKL protein amounts as their wild type littermates. Solely the expression in spleen seems to be slightly increased in the transgenic animals of line E. In wild type mice, endogenous CARKL expression is clearly detectable in liver and kidney, while the expression in spleen is lower. In some mouse lines (B and D) it was hardly detectable with our experimental setup. The spleen showed higher expression of β -actin, as I observed a stronger signal here than in the other organs, although equal total protein amounts were loaded.

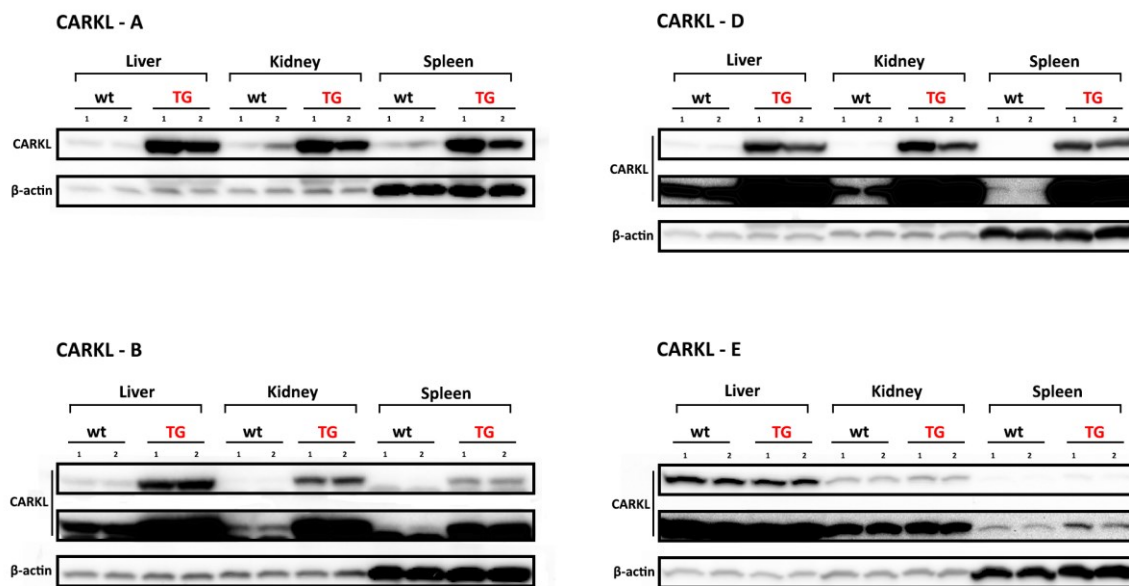


Figure 18: Characterisation of transgenic mouse line sedoheptulose kinase expression on the protein level. Western blot analysis comparing the expression of wild type and transgenic mice of the four CARKL-mouse lines. Protein expression was quantified in liver, kidney and spleen of four mice (2 wild type versus 2 transgenic) per line. Our results show that the lines A, B and D clearly overexpress the CARKL protein. Transgenic mice of the line E express approximately equal CARKL protein amounts as their wild type littermates. Solely the expression in spleen seems to be slightly increased in the transgenic animals of line E. Generally, the basal CARKL expression in wild type mice is higher in liver and kidney than in spleen, which is hardly detectable in some lines (see blots of line B and D). In the lines B, D and E, I used two exposure times to visualize low CARKL expression levels. Double bands in the mid panel of line B result from incomplete stripping and subsequent overexposure during development.

Figure 19 shows the blots which compare the protein expression among the different transgenic lines. The mouse lines A and D show the highest CARKL expression, followed by line B with an intermediate overexpression and line E, with rather low levels. The same pattern can be found in all tested organs. Animal #2 of line A shows lower CARKL levels than #1. The beta-actin levels of this sample suggest that lower amounts of total protein were loaded here.

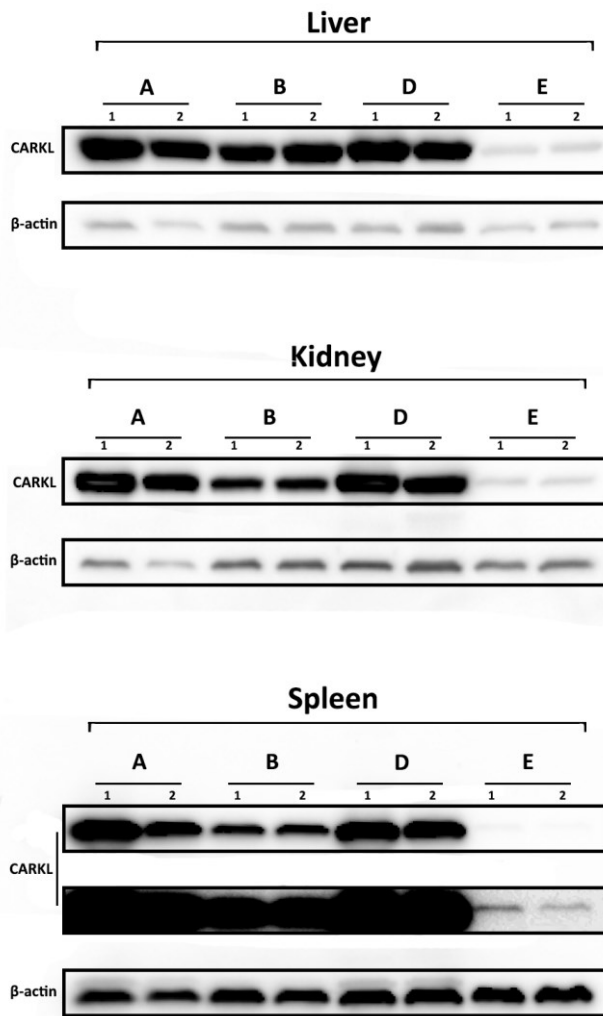


Figure 19: Characterisation of transgenic mouse line sedoheptulose kinase expression on the protein level. Western blot analysis comparing CARKL expression of the four CARKL-mouse lines. β -actin was used as loading control. Protein expression was quantified in liver, kidney and spleen of two transgenic mice per line. Our results show that the mouse line A has the highest CARKL expression, followed by line D. Line B show an intermediate overexpression and line E rather low CARKL levels. The same pattern repeats in all tested tissues. The CARKL expression in animal #2 of line A shows slightly lower CARKL levels than #1. This might be partly explained by lower total protein content in the sample, as indicated by the reduced β -actin levels. In spleen, I used two exposure times, in order to visualize expression in line E.

To sum up, my results show that the transgenic mouse line A expresses the highest level of the CARKL protein, closely followed by line D. Mouse line B shows an intermediate overexpression of the CARKL protein. I could observe similar expression patterns in liver, kidney and spleen. Organs derived from transgenic mice of these lines show clearly an overexpression of the CARKL protein. In contrast, transgenic mice of line E express relatively low amounts of CARKL, comparable to their wild type littermates. CARKL protein expression in other tissues remains to be established. Further it will be important to test the protein expression of primary cells, such as macrophages.

Discussion

In the course of my master thesis, we could establish four transgenic CARKL mouse lines. I tested the integration of the transgenic construct in the genome and further CARKL overexpression on the protein level. The four transgenic mouse lines show differences on the level of protein expression. This may result from transgene integration into regions of different transcriptional activity. It is also very likely that some of the lines harbour multiple copies of the transgenic construct in their genomes.

One of the next important steps in the characterization of the transgenic animals is to determine the absolute transgene copy number and their integration sites. This may be established by a combination of various methods, for example southern blotting, various PCR-techniques, FISH and next generation sequencing [110]. We need to exclude undesired integration of the constructs, which may cause systemic side effects. The random integration of transgenes may for example alter regulatory elements of other genes or even their coding region [111, 112].

I tried to compare the relative CARKL-construct content of the different mouse lines based on the genomic real time PCR data obtained from genotyping. My calculations, based on the $\Delta\Delta C_t$ method, estimated that the mouse lines A and D might harbour more copies of the transgene than B and E. The higher protein expression of the lines A and D supports this observation. It would be interesting to see if exact methods, as southern blotting or whole genome sequencing lead to similar results.

Preliminary CARKL mRNA-expression data, derived from the first experiment with bone marrow derived macrophages of three transgenic mouse lines is shown in supplementary figure 9 (S9). Similarly to overexpression on the protein level, I could measure the highest levels of CARKL mRNA in line A, followed by B with an intermediate overexpression and finally line E. However, these preliminary experiments need further verification. Further it would be interesting to investigate whether CARKL influences macrophage polarization in a similar way as it was shown in cell lines ectopically overexpressing CARKL [5]. Beside further characterisation of primary immune cells, first animal experiments may help to study the role of CARKL in systemic inflammation.

Generally, CARKL overexpression seems not to negatively affect development or health of the transgenic mice. Phenotypically, they are healthy and not distinguishable from their wild type littermates. In a preliminary experiment with a small number of mice, I isolated different organs and could neither observe macroscopic- nor weight-differences between the organs of transgenic animals and their wild type littermates (data not shown). These early results have to be verified with a greater number of experimental animals in the near future.

Nowadays, our knowledge about the physiological relevance of CARKL and heptose metabolism is very limited. Upcoming experiments with the transgenic mice might help to us explore immune-metabolism and lead to a better understanding of chronic disease development and treatment.

Material and Methods

Chemicals

Acrylamide/Bis: 30%, 29:1, BioRad (Hercules, CA)

Acid-phenol:chloroform:isomylalcohol (125:24:1): pH 4.5, Ambion, Life Technologies (Carlsbad, CA)

Agar: select series, Invitrogen (Carlsbad, CA)

Agarose: LE Agarose for electrophoresis, Biozym (Hessisch Oldendorf, D)

APS: Sigma-Aldrich (St. Lois, MO)

Ampicillin sodium salt, Sigma-Aldrich (St. Lois, MO)

Chloroform: MERCK (Darmstadt, D)

Complete inhibitor tablets: Roche (Basel, CH)

D-(+)-Glucose (45% solution), Sigma-Aldrich (St. Lois, MO)

EDTA: Sigma-Aldrich (St. Lois, MO)

Gel-Red: Biotium (Hayward, CA)

HCl: 37%, Sigma-Aldrich (St. Lois, MO)

Isopropanol: molecular biology grade, Applichem (Darmstadt, D)

LB-broth, Miller: Novagen (Madison, WI)

M-CSF: human recombinant macrophage colony stimulating factor, Krackeler Scientific (Albany, NY)

NaCl: Fluka, Sigma-Aldrich (St. Lois, MO)

Na-deoxycholate: Sigma-Aldrich (St. Lois, MO)

Penicillin/streptomycin: 100x: 10.000 units penicillin and 100mg/mL streptomycin, Sigma-Aldrich (St. Lois, MO)

Proteinase K: Sigma-Aldrich (St. Lois, MO)

SDS: Gerbu (Heidelberg, D)

Tris: Sigma-Aldrich (St. Lois, MO)

Triton X-100: Sigma-Aldrich (St. Lois, MO)

Tween 20: Sigma-Aldrich (St. Lois, MO)

X-Gal: Sigma-Aldrich (St. Lois, MO)

BMDM isolation

Bone marrow derived macrophages were isolated from the femur bones of CARKL-transgenic and wild type mice. After the complete removal of flesh from the bones, bone marrow was flushed from both sides using a needle. The obtained bone marrow was filtered through 40µm filters. Afterwards the filtrate was centrifuged (1.200 rpm, 7 min., RT). Erythrocytes were lysed by incubation with erythrocyte lysis buffer (Qiagen) for 4 minutes. After rapid 5-fold dilution with full medium (RPMI 1640 + 10% FBS + 1x penicillin/streptomycin), the remaining cells were centrifuged. The cell pellet was washed once more with full medium, followed by centrifugation. Cell numbers were counted with an automated CASY cell counter (Innovatis). The bone marrow precursor cells were seeded at a density of 1×10^{10} in full medium with 10% L929 conditioned-medium, containing M-CSF for macrophage differentiation. The cells were differentiated for 5 days before they were used for experiments.

Human monocyte isolation

Human monocytes were isolated from peripheral blood of 2 healthy volunteers (50mL each). The blood was diluted 1:1 with PBS, overlaid with lymphocyte separation medium LSM 1077 (PAA) and centrifuged (1,200g, 20 min., slow acceleration, no break). After separation, the disc containing lymphocytes was aspirated and centrifuged (300g, 15min). Cells were washed with PBS and centrifuged once more (300g, 10min). The remaining erythrocytes were incubated with erythrocyte lysis buffer (Qiagen) for 4 minutes. Following erythrocyte lysis, the cells were diluted with FBS containing medium, centrifuged (300g, 10min.), washed 1x with PBS and centrifuged (200g 10min). Platelets can be efficiently removed by two repeated washing steps with PBS + 2% FBS and slow centrifugation (120g, 10 min., brake off), as they remain in the supernatant. Finally the purified cells were resuspended in full medium (RPMI + 20% FBS + 1x penicillin/streptomycin) and seeded into 6-well plates. After 3 hours, when most monocytes had attached, the supernatant containing lymphocytes was removed, followed by a washing step with PBS. The monocytes were cultured in full medium containing 100ng/mL M-CSF. After 5 days, we isolated the RNA of the differentiated macrophages for RACE library generation.

Cell culture

Primary cells were cultured in RPMI 1640 medium (PAA), supplemented with heat inactivated fetal bovine serum (FBS, PAA) and penicillin/streptomycin (100x: 10.000 units penicillin and 100mg/mL streptomycin; Sigma), in an atmosphere with 5% CO₂, 100% humidity and 37°C.

Mouse bone marrow derived macrophages were generally cultured in RPMI 1640 + 10% FBS + 1x penicillin/streptomycin. For plate transfer, the BMDMs were incubated for ~60 min. in 0.0025% trypsin, 0.01% EDTA (GE healthcare) on an orbital shaker.

Human monocytes/macrophages were supplemented with 20% FBS. Apart from that they were cultured under the same conditions as mouse BMDMs.

RNA isolation

We isolated RNA with the RNeasy kit (Qiagen) in combination with QIAshredder cell/tissue homogenisation columns (Qiagen). Before harvest, the cells were washed 2x with cold PBS. Then buffer RLT + 1/100 β-mercaptoethanol were added to the cells to lyse them. The cell culture plates were sealed with parafilm and frozen at -80°C until RNA isolation.

The thawed lysates were first loaded onto the QIAshredder homogenisation columns, which were then centrifuged (max. speed, 2 min.). The flowthrough containing RNA was mixed 1:1 with 70% EtOH, and loaded onto RNeasy columns. The rest of the procedure was performed according to RNeasy protocol instructions.

RNA integrity analysis

The concentration and integrity of RNA used for RACE-library generation was analysed using microchannel based electrophoresis with the Agilent RNA 6000 Nano Assay kit and an Agilent 2100 Bioanalyzer, according to the manufacturer's protocol.

5' RACE

5' RACE was performed using the Ambion FirstChoice® RLM-RACE kit (Life Technologies, Carlsbad, CA) according to the manufacturer's instructions, except of minor adaptations. 10 µg of total RNA were used for the generation of each 5'RACE cDNA library.

RNAs containing 5' phosphates (degraded mRNA & non-mRNA species) were first dephosphorylated using CIP (1h at 37°C). CIP was removed by extraction with acid phenol:chloroform, followed by extraction with chloroform and RNA precipitation with isopropanol. Next TAP was used to remove the 5'CAP structures of full length mRNAs (1h at 37°C), leaving a 5' monophosphate for the subsequent ligation of an RNA anchor with RNA-ligase (1h at 37°C). Reverse transcription of the mRNA, 5' fragments was performed with MMLV-RT and random decamers for 1h at temperatures between 42 and 50°C.

3' RACE

3' RACE was performed using the Ambion FirstChoice® RLM-RACE kit (Life Technologies, Carlsbad, CA) according to the manufacturer's protocol, except of minor adaptations. For the generation of each 3'RACE cDNA library 1 µg of total RNA was reverse transcribed using MMLV-RT and the 3'RACE adapters (anchored oligo d(T) primers with an attached adapter sequence). In some reaction, the amount of used 3'RACE adapters was decreased (1.5 or 2 fold), in order to lower the chance of internal priming. Reverse transcription was carried out for 1h at temperatures between 42 and 50°C.

Nested RACE-PCR

The nested PCRs for the amplification of specific CARKL transcript ends were performed with (Promega), Long-PCR Master Mix (Promega), Phusion (Finnzymes) and Phusion HS-II (Finnzymes) polymerases. The outer RACE-PCR was performed using a RACE cDNA library as a template. The product of this first PCR was used as a template for the inner, nested RACE-PCR.

The 5' and 3' outer and inner PCR primers were used in combination with the below listed gene specific CARKL primers. General guidelines for the PCR-reaction setups are also indicated. All PCRs were performed with vapo-protect Mastercycler ProS (Eppendorf)

thermocyclers. 5'PCR with Phusion polymerases was performed with GC-buffer (buffer recommended for GC-rich templates), 3'RACE with HF-buffer (high fidelity buffer).

CARL RACE primer (5' – 3')	Human	Mouse
5' outer	TCATGCGTGAATACACAGTGG	GGTCAGAGGAGGAGATTCTAGTCA
5' inner	AACCGAGGTGCTGATGTTG	CCTAGGTAGGTCCTGTCTGAAGTAA
3' outer	GAGTGGCCTTGGGTGATTTAC	CCTTACTTCGACAGGACCTACCTA
3' inner	GACACCTCAACCAAGGAATC	CAGCAGAAAGACACCCATCTC

PCR reactions	GoTaq	Long MM	Phusion 5'	Phusion 3'
Template	1-2µL	1-2µL	1-2µL	1-2µL
Buffer	5µL*	25 µL 2xMM	10µL (5x GC)	10 µL (5x HF)
dNTPs	0.2mM	(in 2xMM)	0.2mM	0.2mM
RACE primer	0.4mM	0.4mM	0.4mM	0.4mM
CARL primer	0.4mM	0.4mM	0.4mM	0.4mM
Polymerase	0.012-0.024U/µL	(in 2xMM)	0.01-0.02U/µL	0.01-0.02U/µL
DMSO	-	-	1.5µL	-
Total volume	50µL	50µL	50µL	50µL

*In combination with GoTaq Polymerase, the PCR buffer included in the Ambion RLM-RACE kit was used, instead of the standard Promega 5x buffer.

Cycling conditions	GoTaq	Long MM	Phusion
Initial denaturation	95°C (3min)	94°C (2min)	98°C (1.5 min)
Denaturation	95°C (30s)	94°C (25 s)	98°C (10s)
Annealing	60°C (30s)	58° C (30s)	** °C (30s)
Elongation	72°C (1min/kb)	68° C (1min/kb)	72°C (30-40s/kb)
Final elongation	72° C (10min)	72° C (10min)	72° C (10min)

** The annealing temperature for PCRs with Phusion polymerases is calculated according to a different formula (<http://www.thermoscientificbio.com/webtools/tmc/>). Therefore most PCR reactions with Phusion polymerases were performed at different annealing temperatures (58-65°C), depending on the used primers.

Agarose gel electrophoresis

I used gels with 0.8, 1, 1.2 and 1.5% agarose content in a combination with a standard TAE buffer system. The intercalating agent Gel Red was added 1:50,000 before polymerisation. The DNA samples were mixed with 6x loading dye (Fermentas), if not already containing a loading reagent in the polymerase master mix. 100bp and 1kb gene ruler ladders (Thermo Scientific) were used for product length estimation. Product separation was performed at constant voltage (approx. 10V / cm). DNA was visualized with a Chemidoc XRS 170 imaging system (Biorad).

Gel purification

The gel purification of PCR products was performed using the QIAquick gel extraction kit (Qiagen), according to the manufacturer's protocol. After gel purification, DNA concentration was measured using a Nanodrop 2000c spectrophotometer (Peqlab).

Cloning

Cloning of gel-purified RACE-PCR products was performed with TOPO T/A and TOPO Blunt cloning kits (Life Technologies) according to the manufacturers recommendation. The gel-purified PCR products were mixed with the TOPO vectors and the provided salt solution (4:1:1) and incubated for 5-10 minutes (RT), depending on PCR product size. Until transformation the reaction was kept on ice. I used DH-5 α and TOP10 competent E. coli (both Life Technologies) for transformation. The cloning reaction was added to competent bacteria and incubated on ice for 30 minutes. After a 42°C heatshock for 30 seconds, we chilled the bacteria on ice. For regeneration, the bacteria were incubated in SOC-medium (37°C, 2h, 300rpm) and finally plated on LB-agar plates containing ampicillin (100 μ g/mL). For blue-white screening of colonies, we spread X-gal (stock 40mg/mL) on the plates before adding the bacteria. Plates with transformed bacteria were incubated at 37°C o/n.

Plasmid preparation & restriction analysis

Single colonies of positive clones (white or faintly blue), were inoculated into LB-Amp and incubated o/n at 37°C, 160rpm. Plasmids of the overnight-cultures were purified using the QIAprep spin miniprep kit (Qiagen), according to the manufacturer's protocol.

The plasmids were further analysed by restriction digestions. I performed the digestions with Fast Digest EcoRI (Thermo Scientific), which cuts out the whole PCR-product insert of the respective vectors. The reaction mixture given below was incubated at 37°C for 30 minutes and immediately analysed by agarose gel electrophoresis.

<i>Restriction digestion</i>	
Purified plasmid sample	2.0 μ L
10x Fast Digest green buffer	2.0 μ L
EcoRI (Fast Digest)	0.5 μ L
Total volume	20 μ L

Sequence analysis & processing

Cloned RACE sequences which had approximately the expected insert size after restriction digestion were sent to sequencing (Eurofins MWG Operon). M13 forward and reverse primers (provided by Eurofins MWG operon) were used for the sequencing reactions.

The obtained sequence information was first blasted (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) to filter out unspecific products. Real CARKL RACE products were assembled to contigs (Vector NTI contig express, Life Technologies). The assembled sequences were aligned to the corresponding CARKL full length sequences using AlignX (Vector NTI, Life Technologies). 3'RACE products, which had a very high content of internal adenines at the position of 3'RACE oligo d(T) adapter annealing, were considered as transcripts very likely resulting from an internal priming event during reverse transcription.

Mouse tail DNA isolation

Tails (~0.5cm) of >10 day old mice were incubated o/n with tail buffer (50mM Tris-HCl pH8, 100mM EDTA; 100mM NaCl; 1% SDS; 0.5mg/mL proteinase K) at 55°C. After 5min shaking on a thermomixer 0.35 volumes of saturated (6M) NaCl was added, mixed 5min (RT). The mixture was centrifuged (full speed, 10 min.). The supernatant (w/o top phase and pellet) was transferred into a new tube, 1 volume isopropanol was added, gently mixed and centrifuged (full speed, 10 min.). The pellet was washed with 2 volumes of 70% EtOH, followed by centrifugation (full speed, 15min). The resulting pellets were resuspended in 1 volume of TE-buffer and incubated the mixture at 37°C for 2h.

Mouse genotyping – PCR

Genotyping by standard PCR was performed with primers specifically amplifying the CARKL transgene construct. PCRs were performed with GoTaq polymerase (Promega) and vapo-protect Mastercycler ProS thermocyclers (Eppendorf). Primers, PCR reaction setup and cycling conditions are listed below. The PCR products were analysed by agarose gel electrophoresis.

<i>Genotyping primer</i>	Forward	Reverse
CARKL construct	GCTAACCATGTTTCATGCCTTC	GGGTGCAGCCTCCAATAG

<i>Genotyping PCR</i>	per sample
Template	1 μ L
5x Green buffer	5 μ L
dNTPs	0.2mM
Forward primer	0.4mM
Reverse primer	0.4mM
Polymerase (GoTaq)	0.024U/ μ L
Total volume	25 μ L

<i>Cycling conditions</i>		
Initial denaturation	94°C (3 min)	
Denaturation	94°C (30s)	35x
Annealing	60 °C (30s)	
Elongation	72°C (30s)	
Final elongation	72° C (10min)	

Mouse genotyping – real time PCR

To be able to distinguish between homo- and heterozygous transgenic animals, quantitative real time PCR was performed. Prior to PCR reaction setup, the concentration of the DNA samples were measured using a Nanodrop 2000c spectrophotometer (Pepqab). The samples were subsequently adjusted to the same concentration. I used the same primers for CARKL-transgene amplification as for the standard genotyping PCR described above. As loading control, I amplified a genomic β -actin sequence. The real time PCR reaction was performed with the Promega GoTaq qPCR mastermix, containing BRYT green, a fluorescent DNA binding dye similar to SYBR green. I used BioRad C1000 Thermocyclers combined with the Biorad CFX96 Real time system.

<i>Genotyping primer</i>	Forward	Reverse
CARKL construct	GCTAACCATGTTTCATGCCTTC	GGGTGCAGCCTCCAATAG
Genomic β-actin	TACAGTTTCACCTGCCCTGAG	GGTACGACCAGAGGCATACAG

Genotyping qPCR	per sample
Template	3µL
Forward primer	0.33mM
Reverse primer	0.33mM
GoTaq qPCR Mastermix	7.5µL
Total volume	15µL

Cycling conditions		
Initial denaturation	95°C (2 min)	
Denaturation	95°C (15s)	45x
Annealing & elongation	60 °C (45s)	
Melting curve analysis	95°C (15s), 65°→95°C	

Relative quantification analysis was done by using the delta-delta method. The obtained CARKL-construct Ct values were first normalized to the corresponding genomic β-actin loading control. Relative delta-delta values were calculated by using the formula below to compare CARKL-construct load within one litter.

$$\text{Relative CARKL construct load} = 2^{-(Ct^{\text{CARKL}} - Ct^{\beta\text{-Actin}})_{\text{Sample}} - (Ct^{\text{CARKL}} - Ct^{\beta\text{-Actin}})_{\text{Ref}}}$$

I used the sample with the highest normalized CARKL construct content as reference, corresponding to a homozygous animal. Samples of a relative CARKL construct load of ~1 were assumed to be homozygotes, those with a relative load of ~0.5 heterozygotes.

Reverse transcription

Reverse transcription was performed using the High-Capacity cDNA Reverse Transcription Kit with RNase Inhibitor (Applied Biosystems). Before reverse transcription, the RNA concentration was measured using a Nanodrop 2000c spectrophotometer (Peqlab). The samples of one experiment were adjusted to the same concentration with H₂O (total volume 10µL/sample) and 2µL random hexamer mix were added. These mixtures were heated to 65°C for 10 minutes and afterwards immediately chilled on ice. The reverse transcription master mix (see below) was prepared and 8µL were added to each sample. The reverse transcription cycling was performed with vapo-protect Mastercycler ProS thermocyclers (Eppendorf) according to following program: 25°C for 10min., 37°C for 120min., 85°C 5 sec., 4°C forever. After the completion of reverse transcription, the samples were diluted with 50µL TE-buffer.

<i>Reverse Transcription Mastermix</i>	per sample
10x RT buffer	2.0 µL
25x dNTP mix (100mM)	0.8 µL
Multiscribe reverse transcriptase	1.0 µL
RNase Inhibitor	0.5 µL
Nuclease free H ₂ O	3.7 µL

Real time RT-PCR

The real time PCR reaction was performed with the Promega GoTaq qPCR mastermix, containing BRYT green, a fluorescent DNA binding dye similar to SYBR green. We used BioRad C1000 Thermocyclers combined with the Biorad CFX-96 Real time system. Used primers, reaction mixture and cycling conditions are listed below. β -actin was used as a loading control for each reverse transcription real time PCR. Results were analysed using the delta/delta method.

<i>RT-qPCR primer</i>	Forward	Reverse
CARKL	CAGGCCAAGGCTGTAAT	GCC AGCTGCATCATAGGACT
β -actin	TAGACTTCGAGCAGGAGATGGC	CCACAGGATTCCACCCAAGA

<i>Genotyping qPCR</i>	per sample
Template	3 µL
Forward primer	0.33 mM
Reverse primer	0.33 mM
GoTaq qPCR Mastermix	7.5 µL
Total volume	15 µL

<i>Cycling conditions</i>		
Initial denaturation	95°C (2 min)	
Denaturation	95°C (15 s)	45x
Annealing & elongation	60 °C (45 s)	
Melting curve analysis	95°C (15 s), 65°→95°C	

Protein isolation

The organs for protein isolation (liver, kidney, spleen) were isolated from transgenic and wild type mice, shock frozen in liquid N₂ and kept at -80°C. RIPA buffer (50mM Tris-HCl pH 7.6, 150mM NaCl, 1mM EDTA, 1x Complete) was added and the tissues were homogenized with a Micra D-1 homogeniser. Afterwards detergens mix was added to a final concentration of 1% Triton X-100, 1 % Na-deoxycholate, 0.1% SDS. The mixture was incubated for 15min. on ice on an orbital shaker. Tissue debris were centrifuged (max speed – 16.000rcf, 10 min., 4°C) and the supernatant containing proteins was transferred into fresh tubes.

The protein concentration was measured using the BCA protein assay kit (Thermo Scientific), according to the manufacturer's instruction. As the used organs had very high protein content I diluted liver and kidney 1:25, spleen 1:12.5 for the measurements. For western blot analysis I adjusted the samples to the same concentration with H₂O.

Western Blotting

Discontinuous SDS-PAGE for the separation of protein by size was performed using 12% separating gels and 6% stacking gels. The recipes for the gels are listed below. For SDS-PAGE 15µg of protein were loaded per lane in 2x SDS-PAGE loading buffer (Novex) and 2.5% β-mercaptoethanol. The samples were heated to 95°C for 10 minutes, chilled on ice and spun before loading. For size estimation a prestained Page Ruler Plus ladder (Fermentas) was used. Gels were run at constant voltage (stacking gel 60V, separating gel 100V) in a Tris-Glycine-SDS running buffer.

SDS-PAGE gels	12% separating gel	6% stacking gel
H ₂ O	4 mL	3.3 mL
30% AA/Bis	4 mL	1 mL
Tris buffer	1.9 mL	625 µL
20% SDS	50 µL	25 µL
10% APS	100 µL	50 µL
TEMED	10 µL	5 µL

After the gel-run, the proteins were transferred to PVDF membranes (GE-Healthcare) using tank-blot at a constant voltage of 12V o/n. The blots were blocked with 5% milk in TBST (w. 0.05% Tween) o/n at 4°C. Primary CARL antibodies (1:2000 in 5% milk, polyclonal rabbit, Abcam) were incubated o/n at 4°C. The blots were rinsed and 2x washed with TBST. Secondary antibodies (1:2000 in 5% milk, anti-Rb-HRP, Cell Signaling Technology) were incubated 2h at room temperature. Again, the blots were rinsed and 2x washed with TBST. I developed the blots using mixtures of Super Signal West Pico Chemiluminescent substrate

and Super Signal West Femto Maximum Sensitivity substrate (both Thermo Scientific) according to the manufacturer's protocol with a Chemidoc XRS 170 imaging system (Biorad).

After development, the blots were stripped with stripping buffer (2% SDS, 100mM β -mercaptoethanol, 65.2mM Tris pH6.8) for 10min. at 50°C and intensely washed with TBST. The blots were blocked with 5% milk in TBST (2h, room temperature) and incubated with primary β -actin antibodies (1:5000 in 5% milk, monoclonal mouse, Sigma) for 1h at room temperature. The blots were rinsed and 2x washed with TBST. Afterwards I incubated the blots with secondary antibodies (1:2000 in 5% milk, anti-mouse-HRP, Cell Signaling Technology). Again, the blots were rinsed and 2x washed with TBST. β -actin on the blots was visualized using Super Signal West Pico Chemiluminescent substrate (Thermo Scientific) according to the manufacturer's recommendation.

Supplementary figures

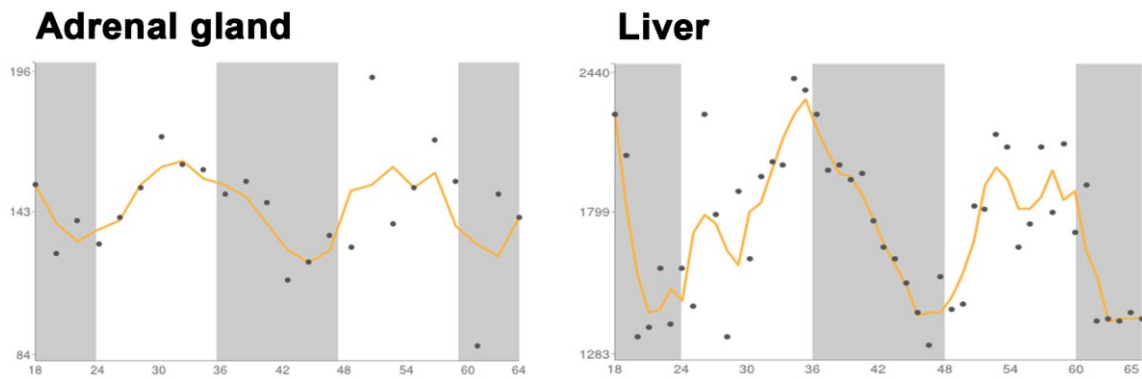


Figure S1: CARKL mRNA expression oscillates within the circadian rhythm. Mouse CARKL mRNA expression data from the database CircaDB (<http://bioinf.itmat.upenn.edu/circa/query>). CARKL expression is elevated during the day, while it drops during night. A similar pattern is observable in adrenal gland and liver. The x-axis shows the time (hours), while the y-axis represents the expression level. The background shaded areas represent the cycle of light and dark conditions.

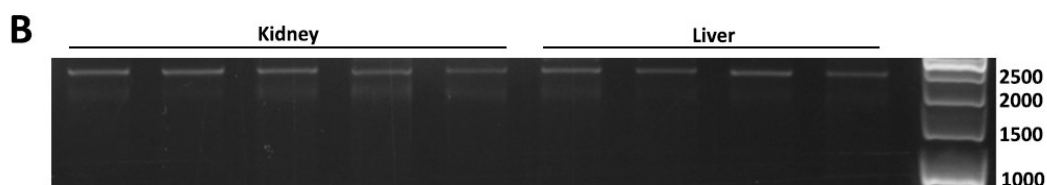
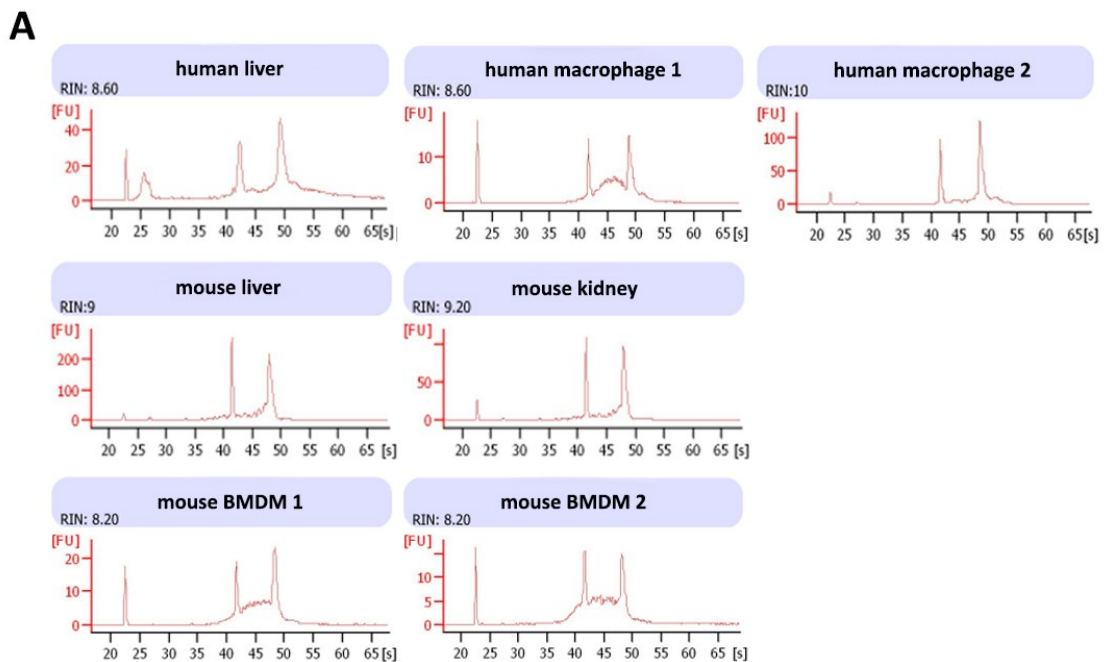


Figure S2: (A) RNA quality graphs of samples used for RACE library generation. Eukaryote total RNA measurement results from an Agilent 2100 bioanalyzer. The first peak from the left represents an internal control, while the two peaks on the right show 18S and 28S ribosomal RNA. Peaks in between result from other RNA

species and their degradation products. The RNA integrity number (RIN), calculated by an algorithm partly based on 18S & 28S rRNA ratio, classifies RNA quality between 1 and 10. A value of 10 represents total RNA of highest quality, while RIN=1 would be totally degraded RNA. **(B) Analytical digestion of 3'RACE CARKL full length single clones from mouse kidney and liver.** These clones correspond to the 3'RACE products shown in Figure 9A. The full length CARKL 3'UTRs in mouse kidney and liver have approximately the same size. The marker on this gel overestimates calculated product size (~1800bp), which was verified by sequencing. Explanations could be that marker and digestion products contained different gel loading dyes. Further it could be that glycerol from the restriction enzyme changed the migration of the DNA-mixture, as I directly loaded the restriction mixture onto the agarose gel.

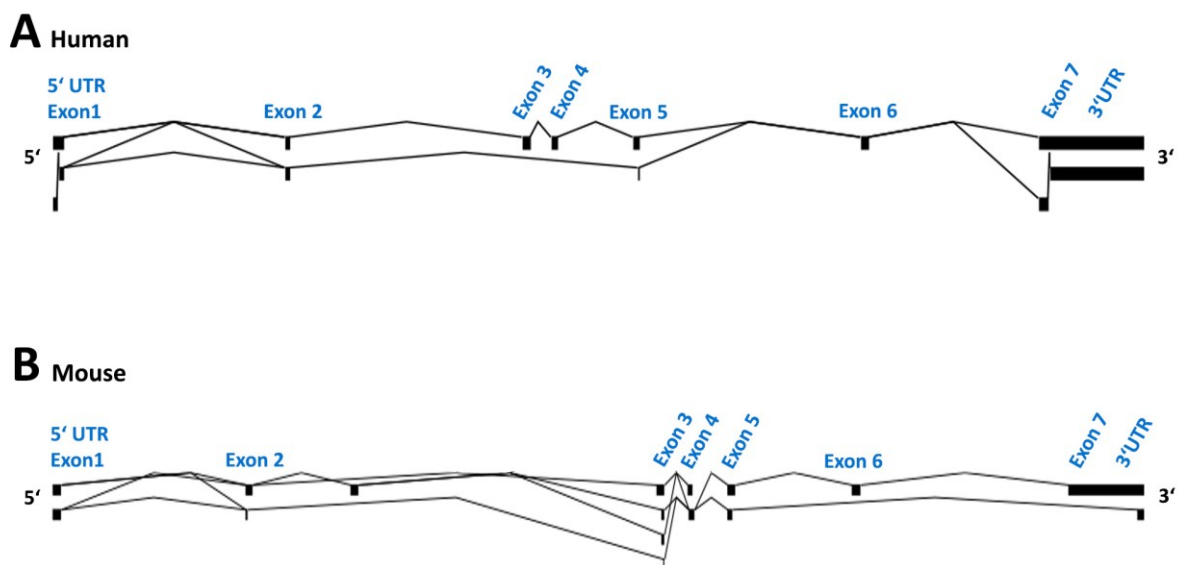


Fig S3: (A) CARKL alternative splicing prediction. A schematic representation of alternative splicing prediction in human and mouse CARKL transcripts, generated by the Swiss Institute of Bioinformatics (SIB) prediction tool in the UCSC genome browser (genome.ucsc.edu/). The blocks indicate exons, while the lines show the introns. These graphs are constructed by analysing experimental RNA transcripts, which serve as basis for the predicted alternative transcripts. **(A)** For the human CARKL mRNA, the predicted alternative features include a splicing event in exon 1, an alternative transcript without exon 3, exon 4 and the 5' part of exon 5, as well as splicing in exon 7. **(B)** For the mouse mRNA, the SIB tool predicts the inclusion of an alternative exon between exon 2 and 3, different lengths of exon 3, a short transcript that ends with exon 4 and a version lacking exon 6, 7 as well as a great part of the 3'UTR.

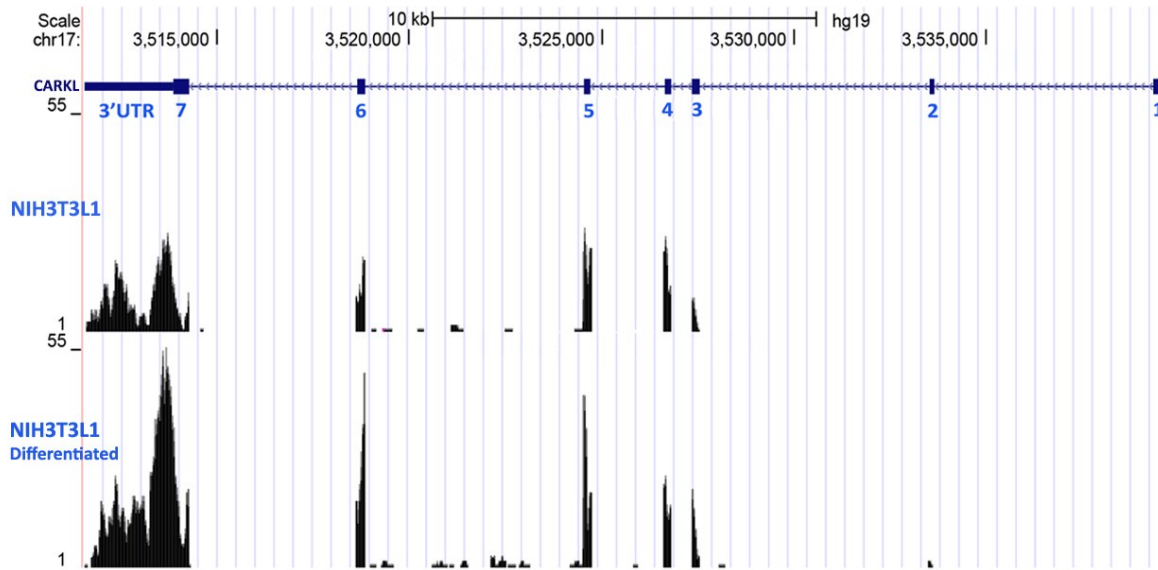
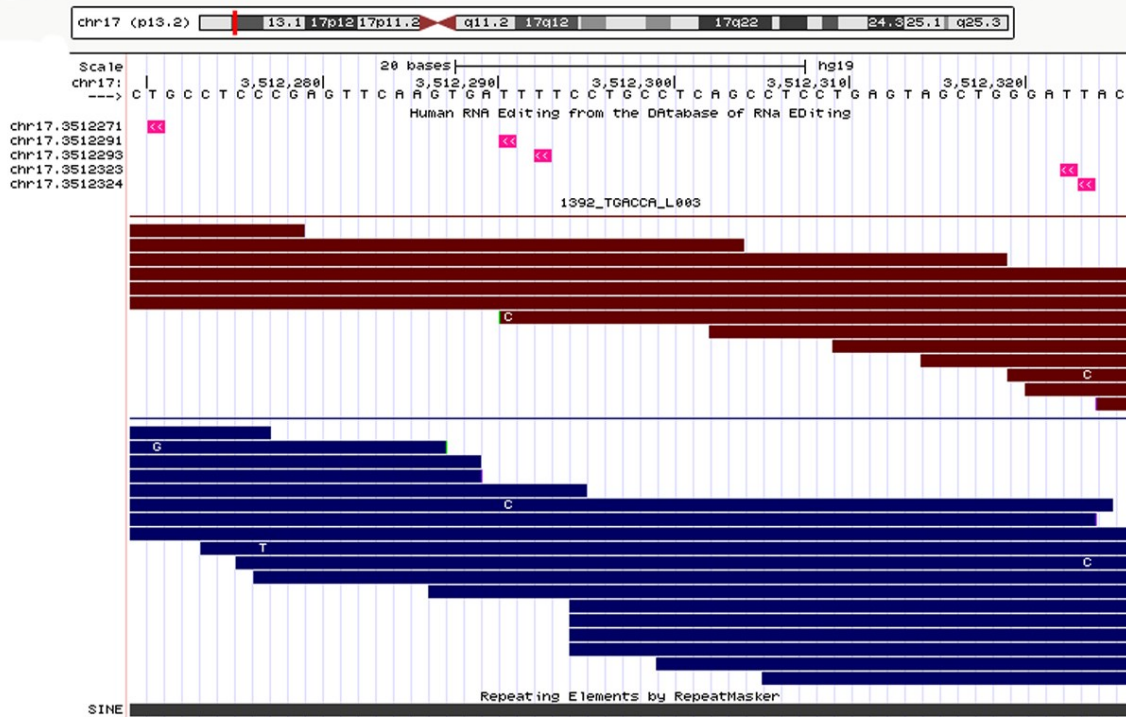
A**B**

Figure S4: Deep sequencing data of CARKL mRNA expression. Deep sequencing data (Illumina HiSeq) derived from the RNA of human SVF (stromal vascular fraction) cells, before and after differentiation to adipocytes, mapped to the human genome (Feb. 2009, GRCh37/hg19) in the UCSC genome browser [unpublished data]. (A) A representation of the total sequencing coverage of RNA mapped to the CARKL gene on the human chromosome 17. The coverage is represented by peak height. Please note that the exons (dark blue squares) are illustrated in an ascending order from the right to the left side, as the CARKL gene resides on the (-) strand of the short arm of chromosome 17. Therefore the 5' end of the gene can be found on the right side, while the 3' end is on the left. Region specific changes in the total coverage upon differentiation might indicate the induction of alternative CARKL transcripts. Especially interesting is the strong increase of coverage in the 5' portion of the 3'UTR, while the 3' end, while there is only a moderate increase in the 3' end. This might result from the induction of a transcript with short 3'UTR, similar to those identified in metabolically active organs. This

increase of the 3'UTR-5'portion in differentiated cells is accompanied by a similar rise of coverage in the exon 7, 6, 3 and the 3'portion of exon 5. Interestingly there is no increase of coverage in exon 4 and the 5'-portion of exon 5. Together this might indicate a splicing event in the coding region. The changes in coverage also depend on other factors, such as base composition and sequencing efficiency. Therefore deep sequencing coverage data alone has to be handled with care and does not allow precise interpretation. Exon 1 and 2 may be absent in the graph due to various difficulties during the sequencing runs or data procession, which cannot be discussed here in detail. (B) RNA editing region in the human CARKL 3'UTR. The red and blue bars represent single reads of deep sequencing data derived from the RNA of adipocytes differentiated from human SVF cells. Blue bars represent reads of the (+) strand, while red bars are reads of the (-) strand. As CARKL resides on the (-) strand, the given sequence has to be red in complement. The pink boxes represent annotated RNA editing sites. Three of these sites could be confirmed in the RNA of human adipocytes. Two are adenosine to inosine editing events (shown here as T→C transition), while another adenosine is changed to cytosine (T→G). The latter may represent a sequencing error, as A-to-C editing is not known to occur. The whole RNA editing region is located inside an AluSx element. These primate-specific short interspersed elements are frequently sites of A-to-I RNA editing [RNA edit 46]. There was no RNA-editing observable in single reads of the cells before differentiation. This indicates that RNA editing in the CARKL 3' UTR might be induced upon differentiation to adipocytes. Please note that the unpublished experimental data in this figure was made available to me in order to get an insight into RNA expression.

<u>PolyA-signal</u>	<u>Percentage</u>	<u>PolyA-signal</u>	<u>Percentage</u>
AAUAAA	58.2%	AAUAGA	0.7%
AUUAAA	14.9%	AAAAAG	0.8%
AGUAAA	2.7%	ACUAAA	0.6%
UAUAAA	3.2%	AAGAAA	1.1%
CAUAAA	1.3%	AAUGAA	0.8%
GAUAAA	1.3%	UUUAAA	1.2%
AAUAUA	1.7%	AAAACA	0.5%
AAUACA	1.2%	GGGGCU	0.3%
		unknown	9.5%

Figure S5: Frequencies of canonical and alternative poly(A) signals. This table was adapted from [86]. Almost 10% of the mRNAs in this study didn't terminate with one of the above listed poly-A signals.

Enhancer motifs

Sequence Position	cDNA Position	Linked SR protein	Enhancer motif	Value 0-100 (variation)
16	+16	SRp40	CCAAACG	86.59 (+38.81 %)
17	+17	SF2/ASF (IgM-BCRA1)	CAAACGA	84.23 (+46.53 %)
17	+17	SF2/ASF	CAAACGA	86.55 (+50.21 %)
21	+21	SRp40	CGACTGC	86.35 (+37.72 %)
22	+22	SC35	GACTGCTG	90.90 (+63.54 %)
24	+24	SF2/ASF (IgM-BCRA1)	CTGCTGT	72.69 (+7.40 %)
42	+42	SC35	GATTACAA	75.48 (+1.71 %)
48	+48	SC35	CATTCTGT	80.39 (+21.42 %)
54	+54	SRp40	TGACACC	88.14 (+45.91 %)
55	+55	SF2/ASF	GACACCA	75.83 (+10.55 %)
62	+62	SRp40	TCTGTGG	81.92 (+17.50 %)
63	+63	SF2/ASF (IgM-BCRA1)	CTGTGGG	73.85 (+11.31 %)
66	+66	SRp55	TGGGTC	76.48 (+10.04 %)
76	+76	SC35	TTCCCTG	79.90 (+19.45 %)
78	+78	SRp40	TCCTGG	84.13 (+27.61 %)
79	+79	SF2/ASF (IgM-BCRA1)	CCCTGga	78.15 (+25.92 %)
88	+88	SC35	gttcagtg	84.02 (+35.95 %)
90	+90	SRp40	tcagtgga	88.20 (+46.18 %)
91	+91	SF2/ASF (IgM-BCRA1)	cagtgga	80.46 (+33.75 %)
91	+91	SF2/ASF	cagtgga	80.72 (+28.65 %)
102	+102	SRp40	ttccaag	88.44 (+47.28 %)

Donor region

Figure S6: Enhancer and silencer motifs in the hu3'UTR_SPL transcript donor and acceptor regions.

Capital letters represent the regions included in the short SPL transcript, while lower case characters represent the 3'UTR intron. The red boxes highlight the exon/intron boundary region.

Sequence position refers to the distance from the beginning of the 3'UTR. The predictions were performed with a publicly available tool (www.umd.be/HSF).

Silencer motifs

Sequence Position	cDNA Position	Sironi's motif	Silencer motif	Value 0-100 (variation)
11	+11	Motif 3 - TCTCCAA	TTCCGCA	64.40 (+11.01 %)
12	+12	Motif 3 - TCTCCAA	TTCCGCA	60.19 (+0.47 %)
13	+13	Motif 3 - TCTCCAA	TGCCCAA	72.75 (+31.86 %)
25	+25	Motif 2 - [T/G]G[T/A]GSGG	TGCTGTA	62.75 (+6.88 %)
35	+35	Motif 3 - TCTCCAA	TTTACTGT	62.39 (+5.97 %)
52	+52	Motif 3 - TCTCCAA	CTGACAC	61.79 (+4.47 %)
62	+62	Motif 2 - [T/G]G[T/A]GSGG	TCGTGGG	66.35 (+15.88 %)
64	+64	Motif 2 - [T/G]G[T/A]GSGG	TGTGGTC	88.10 (+70.25 %)
76	+76	Motif 3 - TCTCCAA	TTCCCTG	79.57 (+48.93 %)
90	+90	Motif 1 - CTAGAGGT	tcagtgga	62.94 (+7.34 %)
90	+90	Motif 2 - [T/G]G[T/A]GSGG	tcagtgga	63.30 (+8.25 %)
92	+92	Motif 2 - [T/G]G[T/A]GSGG	agtgga	66.40 (+16.00 %)
95	+95	Motif 1 - CTAGAGGT	ggacagct	61.77 (+4.41 %)
100	+100	Motif 3 - TCTCCAA	gcttccaa	63.50 (+8.74 %)
105	+105	Motif 1 - CTAGAGGT	caagcagt	66.37 (+15.91 %)
108	+108	Motif 1 - CTAGAGGT	gcagtgct	62.65 (+6.62 %)
116	+116	Motif 2 - [T/G]G[T/A]GSGG	tgtgtgta	75.30 (+38.24 %)
119	+119	Motif 1 - CTAGAGGT	tgtgaggt	71.83 (+29.58 %)
119	+119	Motif 2 - [T/G]G[T/A]GSGG	tgtgaggt	83.25 (+58.13 %)
121	+121	Motif 2 - [T/G]G[T/A]GSGG	tgaggtcc	70.18 (+25.45 %)
124	+124	Motif 3 - TCTCCAA	ggtcccat	76.64 (+41.61 %)
133	+133	Motif 3 - TCTCCAA	tcggccaa	72.75 (+31.86 %)
138	+138	Motif 1 - CTAGAGGT	caagaact	73.01 (+32.54 %)

Enhancer motifs

Sequence Position	cDNA Position	Linked SR protein	Enhancer motif	Value 0-100 (variation)
1832	+1832	SC35	gtcccca	87.89 (+51.47 %)
1834	+1834	SF2/ASF (IgM-BCRA1)	ctcccaa	77.62 (+24.09 %)
1842	+1842	SF2/ASF (IgM-BCRA1)	cacccag	77.62 (+24.09 %)
1842	+1842	SF2/ASF	cacccag	75.01 (+7.53 %)
1843	+1843	SRp40	accagg	79.34 (+5.75 %)
1844	+1844	SF2/ASF (IgM-BCRA1)	cccagga	90.92 (+69.22 %)
1844	+1844	SF2/ASF	cccagga	86.43 (+49.78 %)
1856	+1856	SC35	gtccacag	86.85 (+47.28 %)
1858	+1858	SRp40	CCACAGT	83.59 (+25.15 %)
1859	+1859	SF2/ASF (IgM-BCRA1)	CACAGTA	84.23 (+46.53 %)
1859	+1859	SF2/ASF	CACAGTA	84.10 (+41.16 %)
1860	+1860	SRp40	ACAGTAG	78.50 (+1.93 %)
1866	+1866	SC35	GTCCCCC	78.18 (+12.55 %)
1868	+1868	SF2/ASF (IgM-BCRA1)	CCCCCTT	74.92 (+14.97 %)
1869	+1869	SF2/ASF (IgM-BCRA1)	CCCCCTT	75.69 (+17.57 %)
1878	+1878	SF2/ASF (IgM-BCRA1)	CCACGGT	72.46 (+6.62 %)
1909	+1909	SF2/ASF (IgM-BCRA1)	CTCGGT	85.00 (+49.14 %)
1909	+1909	SF2/ASF	CTCGGT	80.66 (+28.44 %)
1910	+1910	SRp55	TGCGGT	74.50 (+2.45 %)
1912	+1912	SF2/ASF (IgM-BCRA1)	CGGTCAA	70.61 (+0.35 %)
1919	+1919	SF2/ASF (IgM-BCRA1)	CCACGGT	72.46 (+6.62 %)
1926	+1926	SF2/ASF (IgM-BCRA1)	CTGACAA	77.85 (+24.88 %)
1926	+1926	SF2/ASF	CTGACAA	77.05 (+15.07 %)

Acceptor region

Silencer motifs

Sequence Position	cDNA Position	Sironi's motif	Silencer motif	Value 0-100 (variation)
1770	+1770	Motif 2 - [T/G]G[T/A]GSGG	tctgggga	83.25 (+58.13 %)
1772	+1772	Motif 2 - [T/G]G[T/A]GSGG	tggggaga	70.71 (+26.78 %)
1789	+1789	Motif 3 - TCTCCAA	tcctccctc	60.05 (+0.11 %)
1793	+1793	Motif 3 - TCTCCAA	cctcagaa	65.08 (+12.69 %)
1806	+1806	Motif 3 - TCTCCAA	ctgccttg	64.25 (+10.62 %)
1818	+1818	Motif 3 - TCTCCAA	tcctaccag	84.36 (+60.90 %)
1833	+1833	Motif 3 - TCTCCAA	tcctccaa	105.26 (+113.14 %)
1841	+1841	Motif 3 - TCTCCAA	tcaccag	86.22 (+65.54 %)
1848	+1848	Motif 1 - CTAGAGGT	ggaagct	61.77 (+4.41 %)
1855	+1855	Motif 3 - TCTCCAA	gtccaca	62.63 (+6.58 %)
1865	+1865	Motif 3 - TCTCCAA	AGTCCCC	67.87 (+19.67 %)
1867	+1867	Motif 3 - TCTCCAA	TCCCCCTT	71.93 (+29.81 %)
1868	+1868	Motif 3 - TCTCCAA	CCCCCTT	67.81 (+19.54 %)
1889	+1889	Motif 3 - TCTCCAA	CTTCCAT	62.81 (+7.02 %)
1890	+1890	Motif 3 - TCTCCAA	TTCCATG	62.39 (+5.97 %)
1894	+1894	Motif 1 - CTAGAGGT	CATGAGT	69.74 (+24.34 %)
1908	+1908	Motif 2 - [T/G]G[T/A]GSGG	TCTCGGT	66.35 (+15.88 %)
1910	+1910	Motif 2 - [T/G]G[T/A]GSGG	TGCGTCA	60.69 (+1.72 %)
1915	+1915	Motif 3 - TCTCCAA	TCAACCAC	69.46 (+23.64 %)
1925	+1925	Motif 3 - TCTCCAA	TCTGACAA	70.89 (+27.22 %)

Human liver - Short SPL

5' GACACCTCAACCAGAAGGAATCTTAGACAGCAAACTCTTTCGCCAAACGACTGCTGTGAATTTACCTGATTAACATTCTGACACCATCTGTG
GGTCATCCTTCCCTGTCCACAGTAGTCCCCCTTATCCACGGTGTCACTTTCCATGAGTTCAGTTATCTGCGGTCAACCACGGTCTGACAATATT
AAATGGAAAATTCTTCAAACAGCTAAAAAAAAAAAAACCTATAGTGAGTCGTATTAATTCGGATCCGCG 3'

Human liver - Short AP

5' GACACCTCAACCAGAAGGAATCTTAGACAGCAAACTCTTTCGCCAAACGACTGCTGTGAATTTACCTGATTAACATTCTGACACCATCTGT
GGGTCATCCTTCCCTGGACCGTTCAGTGGACAGCTTTCAAGCAGTGTCTGTTGTGAGGTCCCATCTCGGCCAAGAACTTACCTTCAGAACAT
ACTCTAATAATGCAGCCAGGAGCCGTCAGCCAGATCCCAAATGAGTGCCTTCGAAAAAAAAAAAAACCTATAGTGAGTCGTATTAATTCGG
ATCCGCG 3'

Mouse liver - Short AP

5' CAGCAGAAAGACACCCATCTCACCATCACCCCAACAGTGTGGGGGAGAGGCACCTGCCAGACCAGCTGGCCTCTGTGACTAGAATCTCCTC
CTCTGACCTCTCCCTGGGCCATGTGACCCGGGCCCTCTGCAGAGGCATCGTTTCAGAACCTGCACTCCATGCTTCTTTCCAAACAGCTCAAAGA
GTGGGGTGTGGCCCGGTGGTGGGGAGTGGGAGTGCCTGTCCAGGAACGAGGTGCTGAAGCAGGAGGTCCAGAGGGCGTTCTTTCC
CAGTGTGCTTTGGGCAGGATGTAGATGCAGCCTTTGGGGCAGCTCTGGTCATGCTCCAGAGAGACCTTAGCCAGAAGGAGCCTTAGTTCATC
AAGCTCCTTGGCTAAAGGAATGCTGGGAATCTTACCTAATTGACAATGTCAGCATGATCTCTGGGTTGCTTTTCTGGGACGGCTCTGTGGG
TGGACTGTAATTTATACTTGATCTTGAGGTATGATCTAGACTGATGAAGATCTTACCTTCAGGGCATATGATAATCATAAGTCAAGGAATCTCTA
ACCAGACCCCAAACCTAGTGCCTTCCAAGAATAGGATCTTCTGAGAGCTGACTGAAAAATGCAGATGTGGGAGATGGAGAGAATGGGAGGG
AGGAAAATGGGAAGAGGGAGAGGAAGGGAGAAACAACAGTTCCTCCGGTCAAAAAAAAAAAAAACCTATAGTGAGTCGTATTAATTCGGA
TCCGCG 3'

Mouse kidney - Short AP

5' CAGCAGAAAGACACCCATCTCACCATCACCCCAACAGTGTGGGGGAGAGGCACCTGCCAGACCAGCTGGCCTCTGTGACTAGAATCTCCTC
CTCTGACCTCTCCCTGGGCCATGTGACCCGGGCCCTCTGCAGAGGCATCGTTTCAGAACCTGCACTCCATGCTTCTTTCCAAACAGCTCAAAGA
GTGGGGTGTGGCCCGGTGGTGGGGAGTGGGAGTGCCTGTCCAGGAACGAGGTGCTGAAGCAGGAGGTCCAGAGGGCGTTCTTTCC
CAGTGTGCTTTGGGCAGGATGTAGATGCAGCCTTTGGGGCAGCTCTGGTCATGCTCCAGAGAGACCTTAGCCAGAAGGAGCCTTAGTTCATC
AAGCTCCTTGGCTAAAGGAATGCTGGGAATCTTACCTAATTGACAATGTCAGCATAAAAAAAAAAAAAACCTATAGTGAGTCGTATTAATTCGG
ATCCGCG 3'

Figure S7: Full length sequence information of the novel CARKL 3'UTRs. The underlined sequences represent the inner RACE CARKL-primer (5') and the 3'RACE adapter (3') respectively.

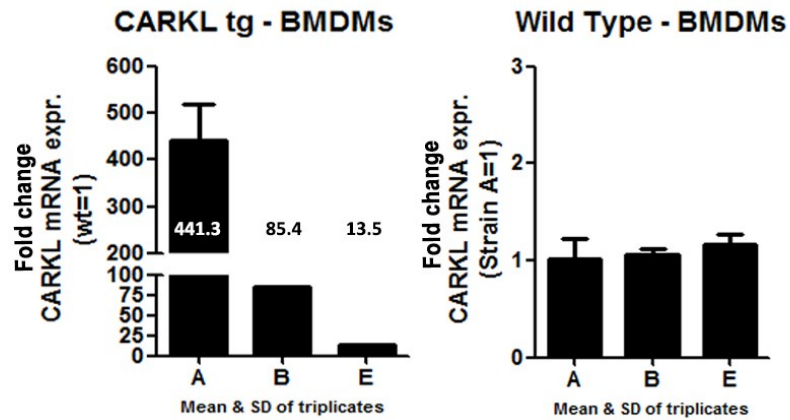


Figure S9: CARKL mRNA expression in BMDMs obtained from transgenic heterozygotes and wild type mice of lines A, B and E. The CARKL mRNA expression was compared by RT-qPCR, using β -actin as a control. Overexpression in the transgenic BMDMs is indicated relatively to the corresponding wild type cells. The numbers show mean values of triplicates, while errors bars depict the standard deviation. Line A BMDMs have more than 400-fold CARKL-mRNA overexpression, followed by line B (~85-fold) and E (~13-fold). The graph on the right shows that wild type littermate BMDMs of the different lines have an approximately equal CARKL-mRNA expression. Here, all values were normalized to the expression of line A.

Supplement: Zusammenfassung in Deutsch

Wichtige Forschungsergebnisse der letzten Jahrzehnte untermauern die Bedeutung von metabolischer Regulation und bioenergetischer Umprogrammierung für viele Erkrankungen wie Diabetes, Atherosklerose, und verschiedener, mit Entzündungen assoziierter Leiden. Solche Veränderungen im zellulären Stoffwechsel sind eng verknüpft mit einer unausgewogenen Ernährungsweise, dominiert durch den hohen Verzehr von raffinierten Kohlenhydraten und Fettsäuren.

Die kürzliche Entdeckung der Sedoheptulose-Kinase CARKL zeigte, dass Sedoheptulose, eine in Obst und Gemüse vorkommende Ketoheptose, als Kohlenstoffquelle für den zellulären Stoffwechsel in Menschen und Tieren genutzt werden kann. CARKL katalysiert die Phosphorylierung von Sedoheptulose zu Sedoheptulose-7-Phosphat und macht sie hiermit für zelluläre Reaktionen verfügbar. Bedeutsamerweise wirkt CARKL der pro-inflammatorische Polarisierung von Makrophagen entgegen, indem sie den zellulären Kohlenstoff-Fluss und Redox-Status moduliert.

In dieser Arbeit präsentiere ich theoretische Überlegungen über den Sedoheptulose-Stoffwechsel, der interessante Analogien zum Stoffwechsel von Glukose aufweist. Die Sedoheptulose-Kinase CARKL scheint in der Lage zu sein den zellulären Kohlenstoff-Fluss zu regulieren, indem sie eine Reaktion an einer strategisch wichtigen Position, der Schnittstelle der Glykolyse und dem nicht-oxidativen Pentose-Phosphat Weg, katalysiert. Im experimentellen Teil der Arbeit habe ich die Diversität von CARKL mRNA untersucht und konnte in metabolisch aktiven Organen Transkripte mit alternativen 3' untranslatierten Bereichen (3'UTR) identifizieren. Die 3'UTR beinhaltet zahlreiche regulatorische Sequenzen und scheint ein zentrales Element in der Regulation der Sedoheptulose Kinase-Genexpression darzustellen. Darüber hinaus habe ich an der Erzeugung von transgenen CARKL Mäusen, welche die Kinase überexprimieren, mitgewirkt. Während meiner Arbeit konnten vier Zuchtlinien etabliert werden, die unterschiedliche Mengen des CARKL-Proteins erzeugen. Die nun verfügbaren transgenen Mäuse stellen ein ideales Werkzeug dar, um die Funktion von CARKL *in vivo* zu untersuchen, zum Beispiel in verschiedenen Ernährungs- und Entzündungsmodellen.

Diese neuen Erkenntnisse über die Regulation von CARKL und den zugrundeliegenden metabolischen Netzwerken könnten zu einem besseren Verständnis von Erkrankungen beitragen, die mit Veränderungen im Kohlenstoff-Metabolismus assoziiert sind.

References

- 1 Bassham, J. B., A.; Kay, L.; Harris, A.; Wilson, A.; Calvin M. (1954) The path of carbon in photosynthesis, XXI. The cyclic regeneration of carbon dioxide acceptor. . J. Amer. Chem. Soc. **76**, 1760 - 1770
- 2 Horecker, B. L., Smyrniotis, P. Z. and Klenow, H. (1953) The formation of sedoheptulose phosphate. The Journal of biological chemistry. **205**, 661-682
- 3 Horecker, B. L. and Smyrniotis, P. Z. (1955) Purification and properties of yeast transaldolase. The Journal of biological chemistry. **212**, 811-825
- 4 Gunsalus, I. C., Horecker, B. L. and Wood, W. A. (1955) Pathways of carbohydrate metabolism in microorganisms. Bacteriological reviews. **19**, 79-128
- 5 Haschemi, A., Kosma, P., Gille, L., Evans, C. R., Burant, C. F., Starkl, P., Knapp, B., Haas, R., Schmid, J. A., Jandl, C., Amir, S., Lubec, G., Park, J., Esterbauer, H., Bilban, M., Brizuela, L., Pospisilik, J. A., Otterbein, L. E. and Wagner, O. (2012) The sedoheptulose kinase CARKL directs macrophage polarization through control of glucose metabolism. Cell metabolism. **15**, 813-826
- 6 Wamelink, M. M., Struys, E. A., Jansen, E. E., Levtschenko, E. N., Zijlstra, F. S., Engelke, U., Blom, H. J., Jakobs, C. and Wevers, R. A. (2008) Sedoheptulokinase deficiency due to a 57-kb deletion in cystinosis patients causes urinary accumulation of sedoheptulose: elucidation of the CARKL gene. Human mutation. **29**, 532-536
- 7 Kardon, T., Stroobant, V., Veiga-da-Cunha, M. and Schaftingen, E. V. (2008) Characterization of mammalian sedoheptulokinase and mechanism of formation of erythritol in sedoheptulokinase deficiency. FEBS letters. **582**, 3330-3334
- 8 Nagy, C. and Haschemi, A. (2013) Sedoheptulose kinase regulates cellular carbohydrate metabolism by sedoheptulose 7-phosphate supply. Biochemical Transactions **41**, 674-680
- 9 Kruger, N. J. and von Schaewen, A. (2003) The oxidative pentose phosphate pathway: structure and organisation. Current opinion in plant biology. **6**, 236-246
- 10 Berg, J. M., Tymoczko, J. L. and Stryer, L. (2011) Biochemistry, Seventh Edition. W.H. Freeman Publishers
- 11 Gumaa, K. A. and McLean, P. (1969) The pentose phosphate pathway of glucose metabolism. Enzyme profiles and transient and steady-state content of intermediates of alternative pathways of glucose metabolism in Krebs ascites cells. The Biochemical journal. **115**, 1009-1029
- 12 Hanahan, D. and Weinberg, R. A. (2011) Hallmarks of cancer: the next generation. Cell. **144**, 646-674
- 13 Vander Heiden, M. G., Cantley, L. C. and Thompson, C. B. (2009) Understanding the Warburg effect: the metabolic requirements of cell proliferation. Science. **324**, 1029-1033
- 14 O'Neill, L. A. and Hardie, D. G. (2013) Metabolism of inflammation limited by AMPK and pseudo-starvation. Nature. **493**, 346-355
- 15 Wamelink, M. M., Struys, E. A. and Jakobs, C. (2008) The biochemistry, metabolism and inherited defects of the pentose phosphate pathway: a review. Journal of inherited metabolic disease. **31**, 703-717
- 16 Park, J. B. (2003) Phagocytosis induces superoxide formation and apoptosis in macrophages. Experimental & molecular medicine. **35**, 325-335
- 17 Herrmann, K. M. and Weaver, L. M. (1999) The Shikimate Pathway. Annual review of plant physiology and plant molecular biology. **50**, 473-503
- 18 Zhao, G., Pease, A. J., Bharani, N. and Winkler, M. E. (1995) Biochemical characterization of gapB-encoded erythrose 4-phosphate dehydrogenase of Escherichia coli K-12 and its possible role in pyridoxal 5'-phosphate biosynthesis. Journal of bacteriology. **177**, 2804-2812
- 19 Hipps, P. P., Ackermann, K., Holland, W. H. and Sherman, W. R. (1981) Synthesis of sedoheptulose from non-dialyzable, endogenous substrates in mammalian tissue extracts. Carbohydrate research. **96**, 1-6
- 20 Touchman, J. W., Anikster, Y., Dietrich, N. L., Maduro, V. V., McDowell, G., Shotelersuk, V., Bouffard, G. G., Beckstrom-Sternberg, S. M., Gahl, W. A. and Green, E. D. (2000) The genomic region encompassing the nephropathic cystinosis gene (CTNS): complete sequencing of a 200-kb segment and discovery of a novel gene within the common cystinosis-causing deletion. Genome research. **10**, 165-173
- 21 Wilson, J. E. (2003) Isozymes of mammalian hexokinase: structure, subcellular localization and metabolic function. The Journal of experimental biology. **206**, 2049-2057
- 22 Diggle, C. P., Shires, M., Leitch, D., Brooke, D., Carr, I. M., Markham, A. F., Hayward, B. E., Asipu, A. and Bonthron, D. T. (2009) Ketohexokinase: expression and localization of the principal fructose-

- metabolizing enzyme. *The journal of histochemistry and cytochemistry : official journal of the Histochemistry Society.* **57**, 763-774
- 23 McKee, T. M., J.R. (2013) *Biochemistry - The Molecular Basis of Life*, updated fifth edition. Oxford University Press
- 24 Holden, H. M., Rayment, I. and Thoden, J. B. (2003) Structure and function of enzymes of the Leloir pathway for galactose metabolism. *The Journal of biological chemistry.* **278**, 43885-43888
- 25 Vanderheiden, B. S. (1965) Phosphate esters of human erythrocytes. IV. Sedoheptulose-1,7-diphosphate, octulose-1,8-diphosphate, inosine triphosphate and uridine diphosphate. *Biochemical and biophysical research communications.* **21**, 265-270
- 26 Paoletti, F., Williams, J. F. and Horecker, B. L. (1979) Detection and estimation of sedoheptulose and octulose mono- and bisphosphates in extracts of rat liver. *Archives of biochemistry and biophysics.* **198**, 620-626
- 27 Blackmore, P. F. and E.A., S. (1982) Regulation of hepatic althro heptulose 1,7-bisphosphate levels and control of flux through the pentose pathway by fructose 2,6-bisphosphate. *FEBS letters.* **142**, 255-259
- 28 Belyaeva, N. F., Golubev, M. A., Grigorovich, J. A., Dubinsky, Z. V., Semenova, N. A., Pitkanen, E. and Korovkin, B. F. (1994) The involvement of fructose 2,6-bisphosphate in substrate cycle control in the nonoxidative stage of the pentose phosphate pathway. A phosphorus magnetic resonance spectroscopy study. *Experientia.* **50**, 780-784
- 29 Horecker, B. L., Smyrniotis, P. Z., Hiatt, H. H. and Marks, P. A. (1955) Tetrose phosphate and the formation of sedoheptulose diphosphate. *The Journal of biological chemistry.* **212**, 827-836
- 30 Karadsheh, N. S., Tejwani, G. A. and Ramaiah, A. (1973) Sedoheptulose-7-phosphate kinase activity of phosphofructokinase from the different tissues of rabbit. *Biochimica et biophysica acta.* **327**, 66-81
- 31 Bonsignore, A., Mangiarotti, G., Mangiarotti, M. A., Deflora, A. and Pontremoli, S. (1963) Cleavage of Sedoheptulose 1,7-Diphosphate by a Purified Rat Liver Diphosphatase. *The Journal of biological chemistry.* **238**, 3151-3154
- 32 Mizunuma, H. and Tashima, Y. (1978) Fructose-1,6-bisphosphatase of the small intestine. Purification and comparison with liver and muscle fructose-1,6-bisphosphatases. *Journal of biochemistry.* **84**, 327-336
- 33 Maswoswe, S. M., Daneshmand, F. and Davies, D. R. (1986) Metabolic effects of D-glyceraldehyde in isolated hepatocytes. *The Biochemical journal.* **240**, 771-776
- 34 Locasale, J. W. and Cantley, L. C. (2011) Metabolic flux and the regulation of mammalian cell growth. *Cell metabolism.* **14**, 443-451
- 35 Montero, F., Nuño, J. C., Sánchez Valdenebro, I., Pérez-Iratxeta, C. and Meléndez-Hevia, E. (1997) Stoichiometric properties of the non oxidative phase of the pentose phosphate cycle. *Nonlinear Analysis, Theory, Methods & Implications.* **30**, 1865-1874
- 36 Nesterova, G. and Gahl, W. A. (2013) Cystinosis: the evolution of a treatable disease. *Pediatr Nephrol.* **28**, 51-59
- 37 Freed, K. A., Blangero, J., Howard, T., Johnson, M. P., Curran, J. E., Garcia, Y. R., Lan, H. C., Abboud, H. E. and Moses, E. K. (2011) The 57 kb deletion in cystinosis patients extends into TRPV1 causing dysregulation of transcription in peripheral blood mononuclear cells. *Journal of medical genetics.* **48**, 563-566
- 38 Hartwell, L. H. H., L.; Goldber, M.L. (2010) *Genetics: From Genes to Genomes*, fourth edition. McGraw-Hill Higher Education
- 39 Shandilya, J. and Roberts, S. G. (2012) The transcription cycle in eukaryotes: from productive initiation to RNA polymerase II recycling. *Biochimica et biophysica acta.* **1819**, 391-400
- 40 Spitz, F. and Furlong, E. E. (2012) Transcription factors: from enhancer binding to developmental control. *Nature reviews. Genetics.* **13**, 613-626
- 41 Davis, C. A., E.; Jenuwein, T.; Reinberg, D. (2009) *Epigenetics*. Cold Spring Harbor Laboratory
- 42 Tang, W., Fei, Y. and Page, M. (2012) Biological significance of RNA editing in cells. *Molecular biotechnology.* **52**, 91-100
- 43 Wilson, R. C. and Doudna, J. A. (2013) Molecular mechanisms of RNA interference. *Annual review of biophysics.* **42**, 217-239
- 44 Kozak, M. (1999) Initiation of translation in prokaryotes and eukaryotes. *Gene.* **234**, 187-208
- 45 Kim, W., Bennett, E. J., Huttlin, E. L., Guo, A., Li, J., Possemato, A., Sowa, M. E., Rad, R., Rush, J., Comb, M. J., Harper, J. W. and Gygi, S. P. (2011) Systematic and quantitative assessment of the ubiquitin-modified proteome. *Molecular cell.* **44**, 325-340
- 46 Danielsen, J. M., Sylvestersen, K. B., Bekker-Jensen, S., Szklarczyk, D., Poulsen, J. W., Horn, H., Jensen, L. J., Mailand, N. and Nielsen, M. L. (2011) Mass spectrometric analysis of lysine ubiquitylation reveals promiscuity at site level. *Molecular & cellular proteomics : MCP.* **10**, M110 003590

- 47 Bernstein, P. P., S.W.; Ross, J. (1989) Poly(A), poly(A) binding protein and the regulation of mRNA
stability. *Trends in Biochemical Sciences*. **14**, 373-377
- 48 Saguez, C., Olesen, J. R. and Jensen, T. H. (2005) Formation of export-competent mRNP: escaping
nuclear destruction. *Current opinion in cell biology*. **17**, 287-293
- 49 Moore, M. J. (2005) From birth to death: the complex lives of eukaryotic mRNAs. *Science*. **309**, 1514-
1518
- 50 Kislauskis, E. H. and Singer, R. H. (1992) Determinants of mRNA localization. *Current opinion in cell
biology*. **4**, 975-978
- 51 Wahl, M. C., Will, C. L. and Luhrmann, R. (2009) The spliceosome: design principles of a dynamic RNP
machine. *Cell*. **136**, 701-718
- 52 Wang, Z. and Burge, C. B. (2008) Splicing regulation: from a parts list of regulatory elements to an
integrated splicing code. *RNA*. **14**, 802-813
- 53 McManus, C. J. and Graveley, B. R. (2011) RNA structure and the mechanisms of alternative splicing.
Current opinion in genetics & development. **21**, 373-379
- 54 Mignone, F., Gissi, C., Liuni, S. and Pesole, G. (2002) Untranslated regions of mRNAs. *Genome biology*.
3, REVIEWS0004
- 55 Meyer, I. M. and Miklos, I. (2005) Statistical evidence for conserved, local secondary structure in the
coding regions of eukaryotic mRNAs and pre-mRNAs. *Nucleic acids research*. **33**, 6338-6348
- 56 Hinnebusch, A. G. and Lorsch, J. R. (2012) The mechanism of eukaryotic translation initiation: new
insights and challenges. *Cold Spring Harbor perspectives in biology*. **4**
- 57 Meijer, H. A. and Thomas, A. A. (2002) Control of eukaryotic protein synthesis by upstream open reading
frames in the 5'-untranslated region of an mRNA. *The Biochemical journal*. **367**, 1-11
- 58 Pelletier, J., Kaplan, G., Racaniello, V. R. and Sonenberg, N. (1988) Cap-independent translation of
poliovirus mRNA is conferred by sequence elements within the 5' noncoding region. *Molecular and
cellular biology*. **8**, 1103-1112
- 59 Kozak, M. (2001) New ways of initiating translation in eukaryotes? *Molecular and cellular biology*. **21**,
1899-1907
- 60 Le, S. Y. and Maizel, J. V., Jr. (1997) A common RNA structural motif involved in the internal initiation of
translation of cellular mRNAs. *Nucleic acids research*. **25**, 362-369
- 61 Ambros, V. (2004) The functions of animal microRNAs. *Nature*. **431**, 350-355
- 62 Hutvagner, G. and Zamore, P. D. (2002) A microRNA in a multiple-turnover RNAi enzyme complex.
Science. **297**, 2056-2060
- 63 Bartel, D. P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*. **136**, 215-233
- 64 Orom, U. A., Nielsen, F. C. and Lund, A. H. (2008) MicroRNA-10a binds the 5'UTR of ribosomal protein
mRNAs and enhances their translation. *Molecular cell*. **30**, 460-471
- 65 Friedmann, R. C. (2009) Most mammalian mRNAs are conserved targets of microRNAs. *Genome
research*. **19**, 92-105
- 66 Gingerich, T. J., Feige, J. J. and LaMarre, J. (2004) AU-rich elements and the control of gene expression
through regulated mRNA stability. *Animal health research reviews / Conference of Research Workers in
Animal Diseases*. **5**, 49-63
- 67 Bakheet, T., Williams, B. R. and Khabar, K. S. (2006) ARED 3.0: the large and diverse AU-rich
transcriptome. *Nucleic acids research*. **34**, D111-114
- 68 Zubiaga, A. M., Belasco, J. G. and Greenberg, M. E. (1995) The nonamer UUAUUUAUU is the key AU-
rich sequence motif that mediates mRNA degradation. *Molecular and cellular biology*. **15**, 2219-2230
- 69 Bevilacqua, A., Ceriani, M. C., Capaccioli, S. and Nicolini, A. (2003) Post-transcriptional regulation of
gene expression by degradation of messenger RNAs. *Journal of cellular physiology*. **195**, 356-372
- 70 Blackshear, P. J. (2002) Tristetraprolin and other CCCH tandem zinc-finger proteins in the regulation of
mRNA turnover. *Biochemical Society transactions*. **30**, 945-952
- 71 Brennan, C. M. and Steitz, J. A. (2001) HuR and mRNA stability. *Cellular and molecular life sciences :*
CMLS. **58**, 266-277
- 72 Zhang, T., Kruys, V., Huez, G. and Gueydan, C. (2002) AU-rich element-mediated translational control:
complexity and multiple activities of trans-activating factors. *Biochemical Society transactions*. **30**, 952-
958
- 73 Chen, C. Y., Gherzi, R., Ong, S. E., Chan, E. L., Rajmakers, R., Pruijn, G. J., Stoecklin, G., Moroni, C.,
Mann, M. and Karin, M. (2001) AU binding proteins recruit the exosome to degrade ARE-containing
mRNAs. *Cell*. **107**, 451-464
- 74 Mukherjee, D., Gao, M., O'Connor, J. P., Rajmakers, R., Pruijn, G., Lutz, C. S. and Wilusz, J. (2002)
The mammalian exosome mediates the efficient degradation of mRNAs that contain AU-rich elements.
The EMBO journal. **21**, 165-174

- 75 Garneau, N. L., Wilusz, J. and Wilusz, C. J. (2007) The highways and byways of mRNA decay. *Nature reviews. Molecular cell biology.* **8**, 113-126
- 76 Bass, B. L. (2002) RNA editing by adenosine deaminases that act on RNA. *Annual review of biochemistry.* **71**, 817-846
- 77 Andreassi, C. and Riccio, A. (2009) To localize or not to localize: mRNA fate is in 3'UTR ends. *Trends in cell biology.* **19**, 465-474
- 78 Doyle, M. and Kiebler, M. A. (2012) A zipcode unzipped. *Genes & development.* **26**, 110-113
- 79 Villalba, A., Coll, O. and Gebauer, F. (2011) Cytoplasmic polyadenylation and translational control. *Current opinion in genetics & development.* **21**, 452-457
- 80 Lianoglou, S., Garg, V., Yang, J. L., Leslie, C. S. and Mayr, C. (2013) Ubiquitously transcribed genes use alternative polyadenylation to achieve tissue-specific expression. *Genes & development.* **27**, 2380-2396
- 81 Hughes, T. A. (2006) Regulation of gene expression by alternative untranslated regions. *Trends in genetics : TIG.* **22**, 119-122
- 82 Schaefer, B. C. (1995) Revolutions in rapid amplification of cDNA ends: new strategies for polymerase chain reaction cloning of full-length cDNA ends. *Analytical biochemistry.* **227**, 255-273
- 83 Levanon, K., Eisenberg, E., Rechavi, G. and Levanon, E. Y. (2005) Letter from the editor: Adenosine-to-inosine RNA editing in Alu repeats in the human genome. *EMBO reports.* **6**, 831-835
- 84 MacNicol, M. C., Cragle, C. E. and MacNicol, A. M. (2011) Context-dependent regulation of Musashi-mediated mRNA translation and cell cycle regulation. *Cell Cycle.* **10**, 39-44
- 85 An, H. J., Lee, D., Lee, K. H. and Bhak, J. (2004) The association of Alu repeats with the generation of potential AU-rich elements (ARE) at 3' untranslated regions. *BMC genomics.* **5**, 97
- 86 Beaudoin, E., Freier, S., Wyatt, J. R., Claverie, J. M. and Gautheret, D. (2000) Patterns of variant polyadenylation signal usage in human genes. *Genome research.* **10**, 1001-1010
- 87 Chang, T. C., Yamashita, A., Chen, C. Y., Yamashita, Y., Zhu, W., Durdan, S., Kahvejian, A., Sonenberg, N. and Shyu, A. B. (2004) UNR, a new partner of poly(A)-binding protein, plays a key role in translationally coupled mRNA turnover mediated by the c-fos major coding-region determinant. *Genes & development.* **18**, 2010-2023
- 88 Lai, E. C., Tam, B. and Rubin, G. M. (2005) Pervasive regulation of Drosophila Notch target genes by GY-box-, Brd-box-, and K-box-class microRNAs. *Genes & development.* **19**, 1067-1080
- 89 Kopan, R. and Ilgan, M. X. (2009) The canonical Notch signaling pathway: unfolding the activation mechanism. *Cell.* **137**, 216-233
- 90 Landor, S. K., Mutvei, A. P., Mamaeva, V., Jin, S., Busk, M., Borra, R., Gronroos, T. J., Kronqvist, P., Lendahl, U. and Sahlgren, C. M. (2011) Hypo- and hyperactivated Notch signaling induce a glycolytic switch through distinct mechanisms. *Proceedings of the National Academy of Sciences of the United States of America.* **108**, 18814-18819
- 91 Izrailit, J., Berman, H. K., Datti, A., Wrana, J. L. and Reedijk, M. (2013) High throughput kinase inhibitor screens reveal TRB3 and MAPK-ERK/TGFbeta pathways as fundamental Notch regulators in breast cancer. *Proceedings of the National Academy of Sciences of the United States of America.* **110**, 1714-1719
- 92 Stolle, C. A. and Benz, E. J., Jr. (1988) Cellular factor affecting the stability of beta-globin mRNA. *Gene.* **62**, 65-74
- 93 Ross, J. (1995) mRNA stability in mammalian cells. *Microbiological reviews.* **59**, 423-450
- 94 Wang, E. T., Sandberg, R., Luo, S., Khrebtkova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P. and Burge, C. B. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature.* **456**, 470-476
- 95 Muro, E. M., Herrington, R., Janmohamed, S., Frelin, C., Andrade-Navarro, M. A. and Iscove, N. N. (2008) Identification of gene 3' ends by automated EST cluster analysis. *Proceedings of the National Academy of Sciences of the United States of America.* **105**, 20286-20290
- 96 Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A. and Burge, C. B. (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science.* **320**, 1643-1647
- 97 Shell, S. A., Hesse, C., Morris, S. M., Jr. and Milcarek, C. (2005) Elevated levels of the 64-kDa cleavage stimulatory factor (CstF-64) in lipopolysaccharide-stimulated macrophages influence gene expression and induce alternative poly(A) site selection. *The Journal of biological chemistry.* **280**, 39950-39961
- 98 Nam, D. K., Lee, S., Zhou, G., Cao, X., Wang, C., Clark, T., Chen, J., Rowley, J. D. and Wang, S. M. (2002) Oligo(dT) primer generates a high frequency of truncated cDNAs through internal poly(A) priming during reverse transcription. *Proceedings of the National Academy of Sciences of the United States of America.* **99**, 6152-6156

- 99 Rodriguez-Prados, J. C., Traves, P. G., Cuenca, J., Rico, D., Aragonés, J., Martín-Sanz, P., Cascante, M. and Bosca, L. (2010) Substrate fate in activated macrophages: a comparison between innate, classic, and alternative activation. *J Immunol.* **185**, 605-614
- 100 Gordon, S. and Martinez, F. O. (2010) Alternative activation of macrophages: mechanism and functions. *Immunity.* **32**, 593-604
- 101 Niwa, H., Yamamura, K. and Miyazaki, J. (1991) Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene.* **108**, 193-199
- 102 Vogler, C., Galvin, N., Levy, B., Grubb, J., Jiang, J., Zhou, X. Y. and Sly, W. S. (2003) Transgene produces massive overexpression of human beta -glucuronidase in mice, lysosomal storage of enzyme, and strain-dependent tumors. *Proceedings of the National Academy of Sciences of the United States of America.* **100**, 2669-2673
- 103 Galbiati, F., Volonte, D., Chu, J. B., Li, M., Fine, S. W., Fu, M., Bermudez, J., Pedemonte, M., Weidenheim, K. M., Pestell, R. G., Minetti, C. and Lisanti, M. P. (2000) Transgenic overexpression of caveolin-3 in skeletal muscle fibers induces a Duchenne-like muscular dystrophy phenotype. *Proceedings of the National Academy of Sciences of the United States of America.* **97**, 9689-9694
- 104 Loser, P., Jennings, G. S., Strauss, M. and Sandig, V. (1998) Reactivation of the previously silenced cytomegalovirus major immediate-early promoter in the mouse liver: involvement of NFkappaB. *Journal of virology.* **72**, 180-190
- 105 Zurita, E., Chagoyen, M., Cantero, M., Alonso, R., Gonzalez-Neira, A., Lopez-Jimenez, A., Lopez-Moreno, J. A., Landel, C. P., Benitez, J., Pazos, F. and Montoliu, L. (2011) Genetic polymorphisms among C57BL/6 mouse inbred strains. *Transgenic research.* **20**, 481-489
- 106 Freeman, H. C., Hugill, A., Dear, N. T., Ashcroft, F. M. and Cox, R. D. (2006) Deletion of nicotinamide nucleotide transhydrogenase: a new quantitative trait locus accounting for glucose intolerance in C57BL/6J mice. *Diabetes.* **55**, 2153-2156
- 107 Mekada, K., Abe, K., Murakami, A., Nakamura, S., Nakata, H., Moriwaki, K., Obata, Y. and Yoshiki, A. (2009) Genetic differences among C57BL/6 substrains. *Experimental animals / Japanese Association for Laboratory Animal Science.* **58**, 141-149
- 108 Ittner, L. M. and Gotz, J. (2007) Pronuclear injection for the production of transgenic mice. *Nature protocols.* **2**, 1206-1215
- 109 Livak, K. J. and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* **25**, 402-408
- 110 Zhang, R., Yin, Y., Zhang, Y., Li, K., Zhu, H., Gong, Q., Wang, J., Hu, X. and Li, N. (2012) Molecular characterization of transgene integration by next-generation sequencing in transgenic cattle. *PloS one.* **7**, e50348
- 111 Yan, B. W., Zhao, Y. F., Cao, W. G., Li, N. and Gou, K. M. (2013) Mechanism of random integration of foreign DNA in transgenic mice. *Transgenic research.* **22**, 983-992
- 112 Durkin, M. E., Keck-Waggoner, C. L., Popescu, N. C. and Thorgeirsson, S. S. (2001) Integration of a c-myc transgene results in disruption of the mouse Gtf2ird1 gene, the homologue of the human GTF2IRD1 gene hemizygotously deleted in Williams-Beuren syndrome. *Genomics.* **73**, 20-27

Acknowledgements

I'm very glad that I got the chance to work on the fascinating topic of heptose biology in the course of my master thesis. Here, I would like to acknowledge all those people who supported and inspired me during the last year and beyond. Especially, I would like to thank:

Dr. Arvand Haschemi, my supervisor at the Medical University of Vienna, for providing me inspiration and excellent mentorship throughout my master thesis, for his great ideas and visions and for supporting my scientific career in every way possible.

Prof. Oswald Wagner, chairman of Clinical Institute of Medical and Chemical Laboratory Diagnosis at the Medical University of Vienna, who enabled me to work on this fascinating topic and provided great scientific as well as financial support.

Prof. Thomas Decker, my supervisor at the University of Vienna, who supported me to perform my thesis at an other University.

Robert Haas, who was my advisor in the first months of the thesis, for teaching me essential laboratory techniques.

Karin Hagenbichler, for helping me a lot in administrative tasks.

Prof. Thomas Rühlicke, Prof. Maria Sibilia, Martina Hammer and Astrid Fabry for the cooperation and generous support with transgenic mice.

Hannes Kessler, Anne Elisabeth Miller and Claudia Quehenberger, my lab members who helped me a lot with my experiments.

The research groups of Prof. Martin Bilban and Prof. Harald Esterbauer, especially Alex, Elisa, Gerfried, Gabriel, Markus and Josefine for useful advices and experimental support.

The research groups of Prof. Friedrich Propst, Dr. Hedwig Suetterluety-Fall and Prof. Wolfgang Miller, for introducing me to techniques in molecular biology during my first internships.

Further I would like to thank my parents and grandparents, who enabled me to study at university and granted me financial as well as mental support. Also I want to thank my girlfriend Denise, my bandmates and my friends who encouraged me and provided me the power to finish my studies.

Curriculum vitae

Name Csörsz NAGY
Address Austria, 1190 Vienna, Döblinger Hauptstraße 6/13
E-mail csörsz.nagy@gmx.at
Birth date 22nd May 1986
Place of birth Dunajská Streda (SK)
Citizenship Austria
Marital Status single

Education

2012 – 2013 Masters Degree Student; Molecular Biology, Immunology & Biochemistry (Supervisors: Prof. Thomas Decker, University of Vienna, Austria; Prof. Oswald Wagner and Arvand Haschemi, PhD, Medical University of Vienna, Austria)

2007 – 2013 Study of Molecular Biology; specialization: Molecular Medicine (University of Vienna, Austria)

2002 – 2006 Secondary school (Bludenz, Austria)

1992 – 2002 Primary school

Employment History

2012 – 2013 Masters Degree Student; Molecular Biology, Immunology & Biochemistry; Center for Translational Research, Dept. for Laboratory Medicine, Haschemi-Lab, Medical University of Vienna (Austria)

4/12 – 7/12 Internship: Molecular Biology & Drosophila techniques; Prof. Wolfgang Miller lab, Medical University of Vienna (Austria)

- 9/11 – 11/11 Internship: Cancer research, Dr. Hedwig Sutterluety-Fall lab, Medical University of Vienna (Austria)
- 12/10 – 2/12 Bachelor Degree Student; Cell Biology, Prof. Friedrich Propst lab, MFPL, University of Vienna (Austria)
- June 2010 Internship: Molecular Diagnostics; Supervisors: Prof. Felix Offner. & Andreas Boesl, Pathology department LKH Feldkirch (Austria)

Additional Experience

- June 2012 Immunological techniques
- May 2012 Techniques in Neuroscience
- Feb. 2012 Epigenetic techniques
- Oct. 2010 Confocal laser scanning microscopy
- Sept. 2010 Light & video microscopy

Publications

Nagy, C. and Haschemi, A. (2013) Sedoheptulose kinase regulates cellular carbohydrate metabolism by sedoheptulose 7-phosphate supply. *Biochemical Transactions* 41(2), 674-680