Inaugural dissertation

for

obtaining the doctoral degree

of the

Combined Faculty of Mathematics, Engineering and Natural Sciences

of the

Ruprecht – Karls – University

Heidelberg

Presented by

Ellen Malovrh M. Sc.

Born in Ljubljana, Slovenia

Oral examination: 16 June 2023

Translated small open reading frames shape the cardiac response to injury

Referees:
Prof. Dr. Dr. Georg Stoecklin
Dr. Mirko Völkers

Summary

Recent studies of cardiac translatomes have revealed widespread translation of small open reading frames into microproteins. Once described as translational noise, these peptides are now recognized as intriguing new players in diverse biological processes. Motivated by their prevalence, I adopted a ribosome profiling-based approach and applied it to mouse myocardium and human heart biopsies to search for previously uncharacterized cardiac microproteins. In addition, heart translatomes were compared with proteomics data and combined with conservation analysis. Integration of these results revealed that novel cardiac microproteins are rare occurrences with little sign of conservation across mammals.

Instead, focusing on annotated but poorly characterized small proteins in the heart led to the identification of cardiac-enriched microprotein termed small integral membrane protein 4 (SMIM4). This promising candidate was used for initial functional characterization, which demonstrated the role of SMIM4 in mitochondrial metabolism, cardiomyocyte viability and growth. Further work will be conducted *in vivo* to understand the contribution of this mitochondrial peptide to heart physiology and cardiac response to stress.

Another recently identified microprotein is small regulatory polypeptide of amino acid response (SPAAR), which features particularly high abundance in murine and human hearts. Its restricted expression pattern and inhibitory effect on mechanistic target of rapamycin complex 1 (mTORC1) suggest that SPAAR plays an important role in regulating cardiac function. Therefore, I studied the role of SPAAR in the heart by using a CRISPR/Cas9-generated knockout mouse model (SPAAR KO), which carries a deletion of the start codon to prevent SPAAR translation, but maintains constant transcript levels. Inducing myocardial ischemia-reperfusion injury in these animals revealed that SPAAR KO mice are protected from myocardial damage early on, as they demonstrated decreased levels of biomarkers for cardiac injury. This protective effect continued at later time points and resulted in sustained left ventricular ejection fraction in SPAAR KO mice, accompanied by reduced infarct sizes.

Strikingly, published single-cell RNA-sequencing data of mammalian hearts indicate enrichment of *Spaar* in cardiac endothelial cells. To understand the molecular mechanisms of SPAAR function, I then focused on the characterization of cardiac endothelial cells. Taking various approaches, I identified several potential targets for mediating SPAAR function. These cell type-specific observations suggest the modulation of cardiac endothelial function by microprotein SPAAR. Combined with the cardioprotective effect after ischemia-reperfusion injury *in vivo*, these findings could prove clinically relevant in the context of ischemic cardiac disease.

Zusammenfassung

Die neusten Studien von Herztranslatomen haben eine weitverbreitete Translation von Mikroproteinen nachgewiesen. Zuvor wurden diese als nicht-funktionell beschrieben. Heute werden diese Peptide als faszinierende neue Regulatoren verschiedener biologischer Prozesse anerkannt. Ich führte Ribosomen-Profiling an murinen Herzen und menschlichen Herzbiopsien durch, um nach neuartigen Mikroproteinen zu suchen. Darüber hinaus wurden diese Translatomanalysen mit Proteomikdaten verglichen und mit Analysen der evolutionären Entwicklungbiologie kombiniert. Die Integration dieser Ergebnisse zeigte, dass neuartige kardiale Mikroproteine selten vorkommen und wenig Anzeichen für eine Erhaltung bei Säugetieren aufweisen.

Stattdessen führte die Analyse von annotierten, aber nicht charakterisierten kleinen Proteinen im Herzen zur Identifizierung von einem kardial angereicherten Mikroprotein, das als "small integral membrane protein 4" (SMIM4) bezeichnet wird. Die funktionelle Charakterisierung dieses Peptids zeigte, dass SMIM4 eine Rolle bei dem mitochondrialen Metabolismus, der Überlebensfähigkeit und Wachstum von Kardiomyozyten spielt. Weitere Arbeiten werden in Mausmodellen durchgeführt werden, um den Beitrag dieses mitochondrialen Peptids zur Herzphysiologie und Herzreaktion auf Stress zu verstehen.

Ein weiteres kürzlich identifiziertes Mikroprotein ist das "small regulatory polypeptide of amino acid response" (SPAAR), das in murinen und menschlichen Herzen besonders häufig vorkommt. Sein eingeschränktes Expressionsmuster und seine hemmende Wirkung auf das "mechanistic target of rapamycin complex 1" (mTORC1) legen nahe, dass SPAAR eine wichtige Rolle bei der Regulierung der Herzfunktion spielt. Um die Rolle von SPAAR im Herzen zu untersuchen, verwendete ich das CRISPR/Cas9-generierte Knockout-Mausmodell (SPAAR KO), das eine Deletion des Startcodons trägt, um die SPAAR Translation zu verhindern, aber ein konstantes Transkript-Niveau beizubehalten. Die Induktion einer myokardialen Ischämie - Reperfusionsschädigung bei diesen Tieren zeigte, dass SPAAR KO Mäuse frühzeitig vor Myokardschäden geschützt waren und ein verringertes Biomarker-Niveau für Herzschädigungen aufwiesen. Diese Schutzwirkung hielt zu späteren Zeitpunkten an und führte zu einer stabilen linksventrikulären Ejektionsfraktion bei SPAAR KO Mäusen, begleitet von reduzierten Infarktgrößen.

Erstaunlicherweise deuten "single-cell RNA"-Sequenzierungsdaten von Säugetierherzen auf eine Anreicherung von *Spaar* in Herzendothelzellen hin. Um die molekularen Mechanismen der SPAAR-Funktion zu verstehen, konzentrierte ich mich dann auf die Charakterisierung von kardialen Endothelzellen. Unter Verwendung verschiedener Methoden identifizierte ich mehrere potenzielle Proteine für die Vermittlung der SPAAR Funktion. Diese Beobachtungen legen eine zelltypspezifischen Veränderung der kardialen Endothelfunktion durch das Mikroprotein SPAAR nahe. Kombiniert mit der kardioprotektiven Wirkung nach einer Ischämie - Reperfusionsverletzung *in vivo* könnten sich diese Erkenntnisse für ischämische Herzerkrankungen als klinisch relevant erweisen.

Contents

List of Figures

List of Tables

Abbreviations

4E-BP	eukaryotic initiation factor 4E-binding protein
AHA	azidohomoalanine
AMPK	adenosine monophosphate-activated protein kinase
B2m	beta-2-microglobulin
BAX	Bcl-2-associated X protein
Btub	beta tubulin
cDNA	complementary DNA
CDS	coding sequence
CITED-4	Cbp/P300 interacting transactivator with Glu/Asp-rich carboxy-terminal domain 4
Col18a1	collagen type 18 alpha 1 chain
Col1a1	collagen type 1 alpha 1 chain
COL5A1	collagen type 1 alpha 1 chain
CRISPR	clustered regularly interspaced short palindromic repeat
cTnT	cardiac troponin T
CYTB	cytochrome b
DNA	deoxyribonucleic acid
DWORF	dwarf open reading frame
ECM	extracellular matrix
EF	ejection fraction
eIF	eukaryotic translation initiation factor
ELISA	enzyme-linked immunosorbent assay
EndMT	endothelial-mesenchymal transition
ERK	extracellular signal-regulated kinase
ETC	electron transport chain
FC	fold change
Fkbp10	FK506-binding protein 10
FKBP1A	FK506 binding protein 1a
FLCN	folliculin
GNB1	guanine nucleotide binding protein, beta 1
GO	gene ontology
GSK3	glycogen synthase kinase-3
h	hour(s)
HA	hemagglutinin
Hprt	hypoxanthine phosphoribosyltransferase

I/R	ischemia-reperfusion
IDE	insulin degrading enzyme
iPSC	induced pluripotent stem cell
KO	knockout
LAD	left anterior descending artery
LFC	lysosomal folliculin complex
lncORF	lncRNA-encoded ORF
lncRNA	long noncoding RNA
MI	myocardial infarction
MiT/TFE	microphthalmia/transcription factor E
MITF	microphthalmia transcription factor
MOXI	micropeptide regulator of beta-oxidation
mRNA	messenger RNA
MS	mass spectrometry
mTOR	mechanistic target of rapamycin
mTORC1	mechanistic target of rapamycin complex 1
mTORC2	mechanistic target of rapamycin complex 2
N	sample size
NRCM	neonatal rat cardiomyocyte
nt	nucleotide(s)
ORF	open reading frame
PE	phenylephrine
PLN	phospholamban
PLOD3	procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3
PRISMA	protein interaction screen on peptide matrix
qRT-PCR	quantitative reverse transcription polymerase chain reaction
Rags	Ras-related GTPases
RHEB	Ras homolog enriched in brain
Ribo-seq	ribosome profiling
RNA	ribonucleic acid
RNA-seq	RNA sequencing
RPF	ribosome-protected footprints
RPL22	60S ribosomal protein L22
RPS6	ribosomal protein S6
rRNA	ribosomal RNA
S6K	ribosomal protein S6 kinase
SAMTOR	S-adenosylmethionine sensor upstream of mTORC1
scRNA-seq	single-cell RNA sequencing
Sdha	succinate dehydrogenase complex flavoprotein subunit A
SDS-PAGE	sodium dodecylsulfate polyacrylamide gel electrophoresis
Ser	serine
SERCA	sarco(endo)plasmic reticulum calcium ATPase
SERPINH1	serpin family H member 1

siRNA	small interfering RNA
SLC38A9	solute carrier family 38 member 9
SLN	sarcolipin
SMCO1	single-pass membrane protein with coiled-coil domains 1
SMIM4	small integral membrane protein 4
sORF	short open reading frame
SPAAR	small regulatory polypeptide of amino acid response
TAC	transverse aortic constriction
TFE3	transcription factor E3
TFEB	transcription factor EB
TFEC	transcription factor EC
ULK1	Unc-51 like autophagy activating kinase 1
uORF	upstream open reading frame
UTR	untranslated region
v-ATPase	vacuolar ATPase
WGA	wheat germ agglutinin
WT	wild-type

1 Introduction

1.1 Small open reading frames encode microproteins

1.1.1 The problem with being small

An open reading frame (ORF) is defined as a span of DNA beginning with a start codon and ending with a downstream in-frame stop codon. ORFs are highly abundant features of all genomes, but only some of them are transcribed into RNA transcripts and an even smaller fraction are eventually translated into proteins. Most eukaryotic messanger RNA molecules encode a single protein, whose sequence is determined by protein-coding sequence (CDS). CDS usually corresponds to the longest ORF in a given transcript, but there can be other, shorter ORFs present as well.

Especially in the last decade, genome-wide studies of translation have revealed thousands of small or short open reading frames (sORFs) with the potential to be translated [1, 2, 3]. sORFs are most commonly defined as **ORFs smaller than 100 codons**, but their upper size limit varies and few studies describe sORFs with up to 200 codons [4, 5].

Traditionally, these sORFs have been **excluded from reference annotation projects** [6, 7]. Conventional gene annotation procedures namely focus on finding ORFs with a minimum length of 50 or 100 codons, disregarding sORFs as having occurred just by chance and therefore representing mere **translational noise**. Moreover, computational efforts to discover and annotate genes rely primarily on detecting similarities. These include sequence similarities, which suggest conservation of coding sequences to preserve functional value, and similarities to protein domains and structures with known functions, implying similar roles for novel, unannotated sequences. sORFs are commonly overlooked using these approaches, as their small sizes result in **lower quantitative conservation scores and lack of conventional protein domains**. Another standard criterium requires translation initiation from an AUG codon, but proteomic studies have shown that sORFs often initiate from near-cognate codons, which differ from AUG by a single base (e.g. CUG and GUG) [5]. Especially difficult is obtaining evidence for translation of sORF-encoded peptides. This leads to often **stochastic detection** and uncertainties regarding their experimental reproducibility, further precluding their annotation [7]. Finally, sORFs have been omitted from reference annotation databases due to their exceptionally high numbers, which present a challenge for their annotation and subsequent curation [8].

Despite all the described difficulties and challenges, sORFs are becoming increasingly studied and their translation products are now recognized as a **novel class of peptides**. These are often called sORF-encoded peptides, microproteins, micropeptides or simply peptides. A small fraction of them have been shown to form stable products with physiological functions in diverse

biological processes [9, 10, 11, 12]. Judging by recent methodological advances and growing interest, more functional microproteins are likely to follow, thus advancing our understanding of the genome and proteome compositions.

1.1.2 Emerging properties and classification

Unlike classical peptides, which arise from proteolytic processing of longer precursor proteins, sORF-encoded microproteins are produced directly from ribosomal translation of sORFs. Genome-wide studies of translation have revealed **widespread translation of sORFs** from many regions of the genome, including those considered to be noncoding [13]. This led to the classification of sORFs according to their genomic location in comparison to annotated genes. Accordingly, sORFs can be located within coding transcripts, either in their 5' and 3' untranslated regions or overlapping the annotated CDSs. Being expressed from alternative start codons means they can also be frameshifted with respect to the main CDS. sORFs have also been found in various transcripts that were previously believed to be noncoding, such as long noncoding RNAs (lncRNAs), antisense transcripts and circular RNAs. Finally, combining translation predictions with *de novo* transcriptome assemblies revealed that sORFs can be encoded by novel, previously unannotated transcripts [14].

However, this kind of classification does not provide much insight into sORF characteristics and function, and Couso and colleagues have proposed an alternative categorization based on sORF properties and possible functions [6]. They suggest the existence of three classes of sORFs (**Figure 1.1**), which represent intermediates between non-functional intergenic ORFs and canonical ORFs, which are highly translated to produce annotated proteins.

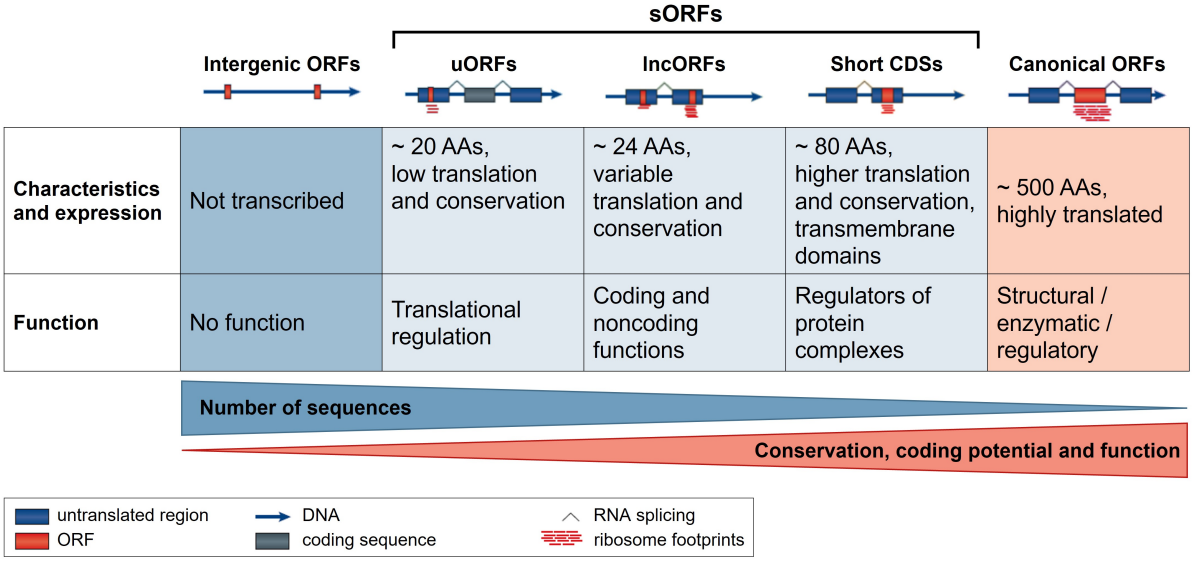

Figure 1.1: Classification of sORFs. sORFs represent intermediates between intergenic ORFs and canonical ORFs and can be divided into 3 classes according to their characteristics and function: uORFs, lncORFs and short CDSs. AAs: amino acids. Adapted from Couso and Patraquim, 2017 [6].

The first such intermediate class consists of **upstream ORFs** (**uORFs**), located in 5' untranslated regions of coding transcripts. More than half of annotated mammalian transcripts contain uORFs and their translation has been reported in several other organisms as well [15]. They have a median length of about 20 amino acids and exhibit low sequence conservation, consistent with their role in translational regulation. They mainly function as translational repressors by reducing the expression of their downstream main ORFs. They can also engage nonsense-mediated decay and direct the usage of internal ribosome entry sites, as well as alternative start codons to control the translation of protein isoforms [16]. However, many recent studies support the idea that uORFs also exhibit functional roles through their encoded peptides [3, 17].

The second class of sORFs are denoted **lncRNA-encoded ORFs** (**lncORFs**), as they localize to putative lncRNAs. These are a diverse group of polyadenylated transcripts longer than 200 nucleotides that lack conventional annotated ORFs. Many lncRNAs have functional roles related only to their RNA sequence and structure, but some of them have been shown to bind ribosomes and others even produce stable microproteins [18, 19]. This last group of translated lncRNAs mostly consists of polycistronic transcripts encoding sORFs with a median size of 24 codons. They show differential translation patterns depending on developmental stage and cell type. Translated lncRNAs are also more conserved across short evolutionary distances than those that do not engage ribosomes and have longer sORFs with preserved amino acid sequences. Their functions are generally unknown and remain to be investigated. However, it has been proposed that translated lncRNAs serve as a source of novel genes that evolve *de novo* to provide favourable adaptations to challenging environments [6, 20].

The third and final class is represented by **short coding sequences** (**short CDSs**) with a median coding size of 80 codons. These "longer sORFs" are typically found in monocistronic transcripts with shorter and more simple structures than canonical mRNAs, but can be translated just as efficiently. They display differential amino acid usage, such as enrichment of lysine and phenylalanine residues and depletion of serine residues, which is typical for transmembrane alpha-helices [2]. This suggests that many longer sORFs are uncharacterized transmembrane peptides targeted to biological membranes, which provide a favorable environment to enhance their stability. Furthermore, many of these peptides are believed to exert their functions through protein-protein interactions. While their small size limits their structural or enzymatic capabilities, it also enables them to fit into small binding clefts of larger protein complexes. This makes them especially well-suited for precise regulation of highly complex biological systems [21].

1.1.3 Discovery and identification: how to search for something small?

First sORFs and microproteins were discovered in studies looking for **regulators of different (disease) phenotypes**. The earliest examples of microproteins encoded by presumable lncRNAs were found in tarsal-less/polished rice gene in *Drosophila*, because their deletion led to developmental defects and truncated limbs in flies [22, 23]. This phenotype was attributed to the lack of 4 microproteins, with sizes ranging between 11 and 32 amino acids, that are required for proteolytic activation of a transcription factor involved in normal embryogenesis and differentiation [24]. Similarly, screening for human genes involved in β-amyloid-mediated cell death led

to the discovery of a microprotein called humanin [25]. This 24-amino acid peptide is encoded by a sORF located in mitochondrial 16S RNA and is especially effective at protecting cells from β-amyloid-mediated apoptosis. Later studies revealed that humanin exerts its functions through protein-protein interactions and thus inhibits the proapoptotic protein BAX [26].

These early studies provided first evidence that sORFs and microproteins are not just translational noise, but rather represent **functional biological molecules**. Since sORFs are a common and frequent feature of all genomes, this raised a possibility of a whole new class of small bioactive peptides. In the following years, researchers dedicated extensive effort to re-evaluating the coding potential of sORFs with the purpose of **systematically searching for biologically relevant microproteins**.

1.1.3.1 Computational approaches

Systematic discovery of sORFs has been greatly facilitated by improved computational approaches. Such *in silico* predictions are based on different strategies, with the most common one relying on comparing sequences among species to identify **evolutionary conserved sORFs**. More stringent conservation tests may also require the sequences to be of comparable length or to share syntenic regions across genomes. Evolutionary conservation implies that sORFs are less likely to be random sequences and suggests they encode functional peptides [27]. However, relying heavily on these approaches fails to predict **species-specific sORFs and evolutionary young microproteins**, that emerged *de novo* from previously noncoding sequences. Recent work has accumulated evidence that sORFs lacking sequence conservation should not be simply disregarded, as they contribute to significant phenotypic effects and engage in a variety of biological processes [28, 29]. Additionally, a fraction of sORFs may not have a coding function, which is generally mediated in a sequence-specific manner by encoded peptides. Instead, they might possess **noncoding regulatory effects**, such as engaging ribosomes and translational machinery. This has been extensively described for uORFs, where the process of their synthesis is crucial for regulation of downstream translation, but the translated peptides do not necessarily have a further biological role [30]. Similar regulatory role has also been proposed for translated lncRNAs and their lncORFs [18, 31].

Other approaches exploit **sequence features** such as the context of the start codon, because sORFs with an optimal Kozak sequence are more likely to initiate translation. Nevertheless, sORFs can also initiate from non-AUG and near-cognate codons. Equally insightful is the examination of **coding potential**, which can be determined by analyzing features such as codon usage and amino acid residue bias. This also includes statistical analysis of $\mathbf{K_a/K_s}$ (the ratio between the number of nonsynonymous substitutions per nonsynonymous site (K_a) and the number of synonymous substitutions per synonymous site (K_s)) to identify coding regions that may be under natural selection. Finally, microprotein sequences can be analyzed for the presence of **motifs responsible for protein sorting** to specific subcellular locations or to predict the existence of known, **functional protein domains** [2, 27, 32]. Although useful in initial predictions of potentially coding sORFs, these computational methods and gene prediction algorithms need to be complemented with **experimental evidence** to prove

the coding potential of sORFs. First line of evidence is usually provided by verifying their expression on transcript level, for example by RNA-seq or qRT-PCR [33]. However, confirming the transcription of sORFs is only the first step in the validation process. Next, evidence of translation is required to corroborate their protein-coding capacity.

1.1.3.2 Ribosome profiling

A rather novel and indispensable method for genome-wide discovery of sORFs is ribosome profiling (Ribo-seq). This assay is based on selective digestion of cellular RNAs by nucleases to degrade parts of transcripts that are not bound by ribosomes. The remaining **ribosome-protected fragments** or footprints (RPFs) are approximately 30 nucleotides long and can be converted into cDNA libraries, followed by deep sequencing. The obtained reads are then mapped onto an annotated genome to determine ribosomal occupancy [34]. Coverage by ribosome footprints provides a genome-wide insight into sequences undergoing translation, also termed the **translatome**. This contains exact positional information on where the translation processes take place, as well as quantitative information on how much a certain region is occupied by ribosomes. Additionally, normalizing ribosome footprints to RNA abundance of the same sample provides an estimation of translation efficiency [15].

Since its first report, Ribo-seq has been successfully adapted in thousands of studies in many different organisms to study global translation at high resolution. In addition, Ribo-seq has revealed novel insights about the process of translation, and one such insight is **pervasive translation of sORFs outside of annotated protein-coding genes** [15, 31]. However, since transcripts encoding sORFs tend to be rather short, this represents a physical barrier for ribosomal binding and consequently generates less ribosomal footprints. To improve the detection of high confidence sORFs and confirm their translation, several studies have refined the initial Ribo-seq protocol and proposed complementary approaches.

Following the standard Ribo-seq protocol, ribosomes are recovered by centrifugation, which provides little specificity and results in many ribonucleoprotein complexes (such as RNase P, telomerase and vault complex) co-sedimenting with ribosomes. To deplete these non-ribosomal contaminations, Ingolia and colleagues introduced **affinity purification of tagged ribosomes**. For this purpose, large subunit ribosomal protein L1 is first biotinylated *in vivo* and only tagged ribosomes are subsequently purified by streptavidin pulldown [13]. This enables more reliable quantification of translation by fully assembled 80S ribosomes. Later studies built upon this idea and used RiboTag mice (with HA-tagged ribosomal protein L22) crossed with a cell type-specific Cre driver, to immunoprecipitate ribosomes from desired target cell types [35, 36].

Another approach to differentiating true, productive translation from non-productive binding of ribosomes is Poly-Ribo-seq. Instead of sequencing all ribosomal-bound RNAs, **Ribo-seq is performed on polysomal fractions**, which are bound by multiple ribosomes and thus actively translated [2]. This method successfully detected hundreds of translated sORFs enriched in small polysomal fractions (containing 2 - 6 ribosomes/RNA).

Finally, computational and statistical approaches can also aid in Ribo-seq pipelines to detect actively translating sORFs. This is based on the triplet nature of genetic code, which results in ribosomes moving along RNA templates one codon at a time and exhibiting a characteristic **three-nucleotide periodicity**. Since mapping Ribo-seq reads provides subcodon resolution of ribosome position, statistical analyses allow for precise identification of the translated ORFs. First tool to implement periodicity analysis was called RiboTaper, but meanwhile this strategy has become widely adopted by several other Ribo-seq analysis pipelines [37].

1.1.3.3 Mass spectrometry

Improved computational approaches and ribosome profiling combined with transcriptomics have led to the discovery of several thousand sORFs with the potential to encode microproteins. However, these methods provide only a prediction of translation, which needs to be corroborated with experimental evidence; that is the **detection of stable protein products**. Mass spectrometry (MS) workflows optimized for peptide analysis are particularly useful in this setting, both in discovery and validation of microproteins [38].

Because microproteins are much smaller and less abundant than canonical proteins, MS-based workflows need to be adapted accordingly. Key steps involve efficient and **rapid peptide extraction** to minimize the degradation by peptidases, separation of peptides from larger proteins to maximize microprotein identification and careful **consideration of enzymatic digestion** with trypsin or other enzymes. The combination of chosen methods will greatly impact the identification of certain sets of microproteins depending on their molecular characteristics [39, 40]. Additionally, the discovery of microproteins is influenced by subsequent bioinformatical analysis of MS-generated data. Typically, MS spectra are interrogated against theoretical databases of known or predicted peptides, which are likely to lack many microprotein sequences. To solve this problem, researchers have created custom databases from RNA-seq and Ribo-seq data or even from a theoretical translation of an entire genome in all six reading frames [33, 41]. Despite significant improvements, confident discovery of microproteins using MS remains very challenging. Instead, MS-based methodology represents a valuable tool for **confirming the existence of microproteins**. By using synthetic peptides to compare their retention times and fragmentation patterns with those of endogenous ones, studies can benefit from obtaining unambiguous evidence for microprotein translation [38].

1.1.3.4 Combined approaches are particularly powerful

Initial studies of sORFs typically relied on one of the methodologies described above, but nowadays it has become widely accepted to combine these methods to **optimize systematic discovery of sORFs** (**Figure 1.2**). This chapter highlights two recent examples that applied various approaches to uncover and validate functional sORFs.

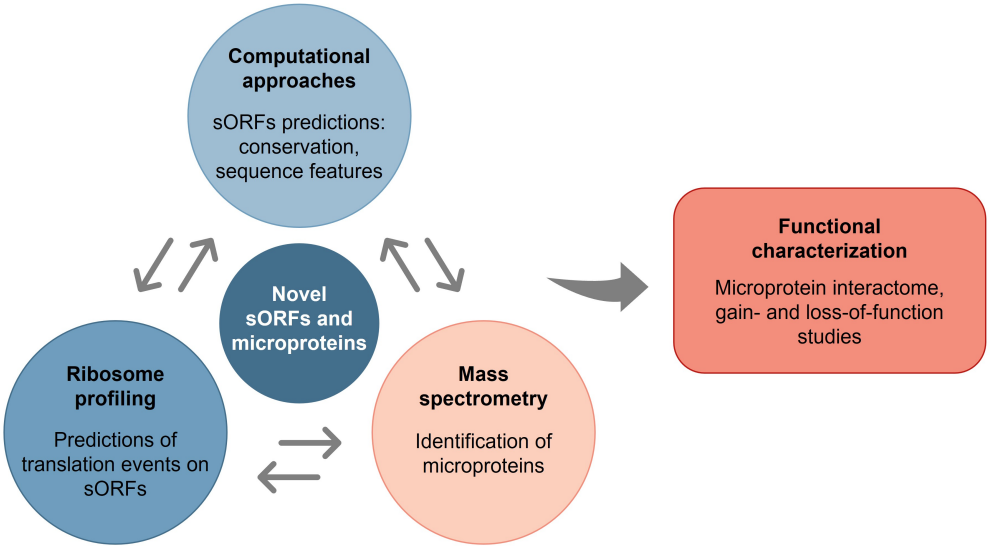

Figure 1.2: Overview of commonly used methodology to systematically search for sORFs and microproteins. Advances in computational approaches, ribosome profiling techniques and mass spectrometry have greatly facilitated the discovery of novel sORFs and microproteins. This growing field of research now needs to transition to functional characterization of known microproteins to describe their biological relevance (reviewed in chapter 1.1.4). Adapted from Kute *et al.*, 2022 [37].

Chen and colleagues successfully combined **ribosome profiling, mass spectrometry and CRISPR-based loss-of-function screen** across several human cell types to pinpoint translated and functional sORFs [3]. Starting with Ribo-seq, they identified over **3000 translated noncanonical coding sequences** that were mostly shorter than 100 amino acids and thus corresponded to sORFs. Proteomics and human leukocyte antigen peptidomics confirmed stable expression of noncanonical microproteins, but only for about 10 % of the predicted sORFs, which is likely due to their shorter length and lower abundance. For their next step, they performed CRISPR/Cas9-based knockout screen in two human cell lines to perturb the expression of over 2500 noncanonical sORFs identified by Ribo-seq. They quantified the growth defect for each targeted sORF and found **400 hits that resulted in significantly altered growth phenotype** in both cell lines. Then, they followed up on a handful of selected hits encoded by presumable lncRNAs. They determined the subcellular localizations of lncORFs by fluorescence microscopy and used coimmunoprecipitations coupled with MS to show that lncORFs form specific protein complexes consistent with their localization. They also investigated the functional effects of uORF translation and showed that growth defects upon uORF knockout are mediated by encoded peptide products, not by disrupting the translation of the main CDS. Some of the uORF peptides had localization patterns and functions independent of their main CDS, while another subset of uORF peptides formed stable complexes with their downstream encoded canonical proteins. Altogether, their large-scale study establishes widespread translation of noncanonical and unannotated sORFs in humans and provides examples of **stable and functional lncORF- and uORF-encoded peptides**.

Another study by Sandmann and colleagues utilized a recently compiled catalogue of 7000 human sORFs detected by Ribo-seq to perform **extensive conservation analysis** [29]. They observed

that 90 % of these sORFs are **not conserved in non-primate mammals** and are therefore evolutionary young. Tracing their genomic changes revealed that more than half of these **young peptides emerged *de novo*** from ancestral noncoding sequences, and over 200 were human-specific. Next, they selected 45 of recently evolved microproteins for **interactome profiling using PRISMA** (protein interaction screen on peptide matrix). This approach is based on dividing protein sequences into short peptides, which are synthesized on a membrane, incubated with a protein lysate and the resulting protein interaction complexes are analyzed by MS. Careful analysis showed that young microproteins form specific interactomes in line with their predicted subcellular localization. They also engage in interactions with canonical proteins required for cell survival, suggesting that even recently evolved microproteins are **involved in vital biological processes**. Furthermore, the authors specifically looked into extremely small peptides, ranging between 3 and 15 amino acids in length. They detected over 200 of them as translated in Ribo-seq datasets of 5 human tissues, mostly from 5' untranslated regions. Interestingly, their structures were well conserved across mammals and they showed distinct, organelle-restricted protein interaction profiles as indicated by PRISMA. Collectively, these results shed light onto commonly overlooked elements of human proteome and show that both **evolutionary novel and very short microproteins can participate in essential cellular processes**.

1.1.4 Elucidating biological functions of microproteins

Global approaches, such as Ribo-seq and proteomics, have the power to identify thousands of novel sORFs, but they do not necessarily link microproteins to their biological functions. Further experimental evidence is required to identify the roles that these peptides may play in cells.

A frequently used approach assumes that microproteins exert their **functions through interactions** with other regulatory proteins. Identifying binding partners of a microprotein can thus offer insight about its potential functions based on the known roles of interaction partners. Performing co-immunoprecipitations followed by MS has been used in many studies to identify binding partners that hint at the biological function of a microprotein [42].

Another common strategy for determining functions of microproteins is to overexpress them in cell lines or model organisms and monitor changes in selected phenotypes. The reverse experiment of **inactivation or removal of a microprotein** is more challenging. It can be difficult to determine whether observed phenotypes arise due to the disruption of sORF or changes in abundance of the transcript on which it lies (and which may perform other, noncoding functions or possess other ORFs). The advent of CRISPR-Cas9 has proven especially useful in this case, as it enables precise disruptions and replacements in coding regions [3, 10]. First insight into the function of a microprotein can then be obtained by gene expression profiling (e.g. RNA-seq, Ribo-seq). Comparing cells with overexpression or inactivation of a microprotein of interest with control cells can yield differentially regulated genes, which might be regulated by or involved in the same cellular processes as the investigated microprotein [43].

1.1.5 Membrane microproteins regulate cardiac metabolism and contractility

Another indication that can be used to infer the functions of microproteins is the presence of target peptides and functional domains. Interestingly, many microproteins are predicted to **form transmembrane α-helices** and thus represent a valuable source of membrane peptides [21]. Several examples of functionally characterized microproteins suggest that these peptides regulate various aspects of membrane biology. This has also been demonstrated for many heart-enriched membrane microproteins, which are closely related to cardiac mitochondrial metabolism, cardiac muscle contractility and protein recruitment to the membranes (**Figure 1.3**).

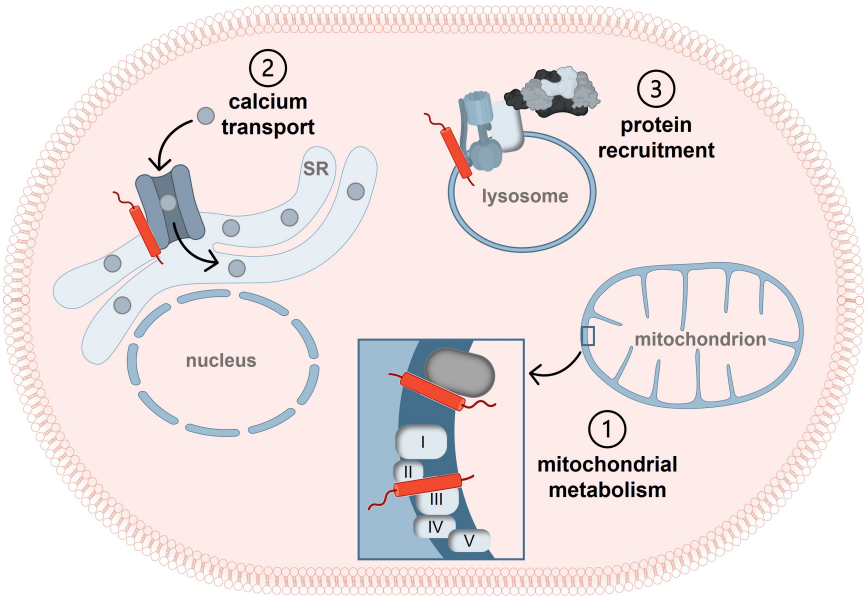

Figure 1.3: Cardiac microproteins (indicated in red) localize to different membrane domains, where they regulate vital biological processes. **1** - Many microproteins translated in the heart localize to mitochondrial membranes and interact with enzymes or subunits of ETC to precisely fine-tune cardiac mitochondrial metabolism. **2** – A family of transmembrane microproteins controls the activity of SERCA pump in the heart and thereby affects cardiac muscle contraction and relaxation. **3** – Membrane microproteins also coordinate protein recruitment to membranes to modulate signalling events, which culminate in metabolic adaptations. Examples for each of these functional groups are described in the following sections. Adapted from Makarewich, 2020 [21].

1.1.5.1 Cardiac microproteins localize to mitochondria and regulate oxidative phosphorylation

The most comprehensive analysis of cardiac microproteins to date used ribosome profiling to map and quantify translation events in 65 ventricular samples from dilated cardiomyopathy patients and 15 healthy controls [41]. Results from this study indicated that as many as 20 % of transcribed lncRNAs in the heart are actively translated and possibly produce over 300 unique microproteins. The expression of translated lncRNAs correlated particularly well with **mitochondrial genes, especially those involved in oxidative phosphorylation**. This suggests that many translated lncRNAs are involved in cardiac mitochondrial metabolism, which was corroborated by **mitochondrial localization** of these lncRNA-encoded microproteins.

Further evidence to support these conclusions is provided by functionally relevant and characterized mitochondrial microproteins in the heart. One example is evolutionary conserved **MOXI/mitoregulin**, a muscle-enriched microprotein that localizes to the inner mitochondrial membrane [44, 45]. MOXI interacts with several respiratory supercomplexes, phospholipids and protein complexes to enhance mitochondrial fatty acid oxidation. Consequently, hearts from MOXI KO mice exhibit disorganized mitochondrial network, swollen mitochondria and disrupted folding of inner mitochondrial membrane, which collectively contribute to the reduction in exercise capacity of these KO animals.

Another example is **mitolamban**, a cardiac-enriched microprotein that localizes to the inner mitochondrial membrane, where it supports the assembly and function of complex III of electron transport chain (ETC) [46]. Mitolamban KO mice show deficiencies in complex III assembly and activity, whereas cardiac-specific overexpression of mitolamban results in cardiomyopathy and premature death, attributed to mitochondrial dysfunction and increased oxidative stress.

Analogous function in regulating complex III has also been described for **small integral membrane protein 4** (**SMIM4**). This microprotein was first suggested to be involved in mitochondrial translation based on its interaction with mitochondrial ribosomes [47]. A later study found that SMIM4 localizes to the inner mitochondrial membrane, where it binds to early assembly factors of complex III of ETC and is required for its biogenesis [48]. The most recent study considering SMIM4 function provided a detailed characterization of intermediates that support early assembly of complex III [49]. SMIM4 was shown to be a part of an intermediate that is required for nascent cytochrome b (CYTB) stabilization and membrane insertion, as well as dynamic response to external metabolic cues to adapt CYTB maturation to nutrient and energy availability.

1.1.5.2 A family of SERCA-regulating microproteins control heart contractility

Sarco(endo)plasmic reticulum calcium ATPase (SERCA) is a membrane protein that pumps calcium ions from cytosol into sarcoplasmic reticulum of skeletal and cardiac muscle cells to initiate muscle relaxation. Two cardiac-enriched transmembrane microproteins function as endogenous inhibitors of SERCA by reducing its affinity for calcium ions. These are **phospholamban** (PLN, 52 amino acid residues), which is more abundant in heart ventricles, and **sarcolipin** (SLN, 31 amino acid residues), which is enriched in the atria [50].

Another regulator of SERCA in the heart is a novel microprotein named dwarf open reading frame (**DWORF**, 34 amino acids). Located within a presumable noncoding transcript, this peptide was shown to act as a stimulator of SERCA, enhancing its activity by displacing the inhibitors PLN and SLN [9]. Recently, cardiac-specific overexpression of DWORF using adeno-associated virus has been proposed as a **gene therapy for heart failure** [51, 52]. Overexpression of DWORF contributed to cardioprotection in mouse models of myocardial infarction and dilated cardiomyopathy by enhancing ventricular function and reducing cardiac dilation. Hence, this study illustrates that microproteins can also have medical implications and therapeutic potential for treatment of cardiac disease.

1.1.5.3 Microprotein SPAAR inhibits mTORC1 activation at the lysosomal membrane

Another role of membrane microproteins is the recruitment of larger proteins to membrane domains to coordinate signalling pathways. A recently identified example is small regulatory polypeptide of amino acid response (**SPAAR**), which contains a short transmembrane domain and localizes to the lysosomal membrane [10]. Mass spectrometry analysis of its protein interactome identified 4 subunits of lysosomal vacuolar ATPase complex (v-ATPase), but subsequent experiments showed that SPAAR has no effect on assembly or activity of this proton pump, neither on lysosomal morphology nor activity of lysosomal peptidases. Instead, SPAAR maintains a tightly bound supercomplex consisting of v-ATPase, Ragulator and Rags, which hampers the recruitment and complete activation of mTORC1 at the lysosome (**Figure 1.4 A**). Interestingly, the inhibitory effect of SPAAR on mTORC1 signalling is only observed in a specific setting, namely when cells are starved of amino acids and then stimulated with them again.

A recent study examined the evolutionary conservation of SPAAR and identified its **orthologs in all major mammalian groups**: placentals, marsupials (e.g. opossum) and monotremes (e.g. platypus) [53]. However, no SPAAR homologs could be found outside of mammalian lineage. Interestingly, SPAAR orthologs exhibit low amino acid identity, but have maintained the basic structure across all mammalian groups (**Figure 1.4 D**). This consists of a 23-amino-acid transmembrane domain at the N-terminus, immediately followed by a short disordered region. Accordingly, mouse homolog of *Spaar* contains a preserved sORF that shares 65 % amino acid identity with human SPAAR and highly similar peptide conformation. This enabled functional characterization of SPAAR *in vivo* by generating SPAAR KO mice. Using CRISPR/Cas9-based homology-directed repair, Matsumoto and colleagues removed the start codon of SPAAR to prevent peptide translation, while still maintaining the expression of the host transcript [10]. Since *Spaar* is highly abundant in muscle tissue, they based their experiments on toxin-induced injury of skeletal muscle. Following the administration of the toxin, SPAAR-deficient mice showed increased activation of mTORC1 signalling pathway, which promoted stem cell proliferation and maturation, finally leading to increased regenerative capacity *in vivo*.

Interestingly, mining various expression datasets revealed that *Spaar* transcript features a distinctive spatial expression pattern and is **strongly enriched in the heart**. Consistent with this observation is the differential expression of *Spaar* in cardiovascular dysfunction as identified by several studies. *SPAAR* transcript was found to be downregulated in patients with coronary artery disease and in patients with dilated cardiomyopathy [41, 54]. Furthermore, data from *in vivo* mouse models point to *Spaar* downregulation early during ischemia-reperfusion injury (**Figure 1.4 B**) and in early stages of cardiac hypertrophy (**Figure 1.4 C**). This substantiates the role of SPAAR as an mTOR inhibitor in the heart, because its downregulation agrees with high activity of mTORC1 in these adverse cardiac conditions [55].

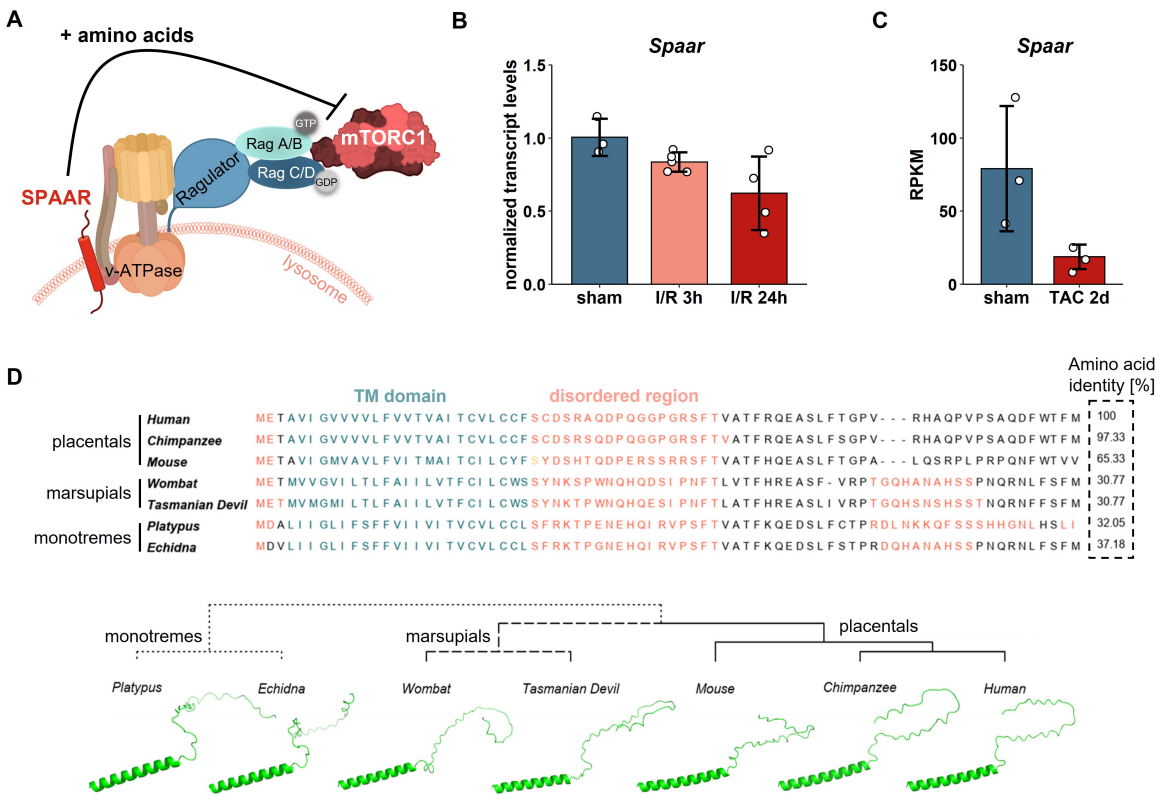

Figure 1.4: SPAAR is a recently discovered conserved microprotein that inhibits mTORC1. **A.** SPAAR interacts with subunits of v-ATPase in the lysosomal membrane and inhibits mTORC1 signalling during periods of amino acid stimulation. Adapted from Matsumoto *et al.*, 2017 [10]. **B.** *Spaar* is downregulated at early time points after ischemia-reperfusion (I/R). h: hours. **C.** *Spaar* is also downregulated in early stages of cardiac hypertrophy induced by transverse aortic constriction (TAC). d: days. **D.** Extensive evolutionary analysis revealed that SPAAR is conserved, expressed and translated in all three major mammalian groups. Despite divergent primary sequences (top panel), all SPAAR orthologs adopt highly similar protein structure (bottom panel), consisting of an N-terminal transmembrane (TM) domain, which is followed by a disordered region. Adapted from Lee *et al.*, 2021 [53].

1.2 mTOR signalling underlies cardiac biology and disease

1.2.1 Overview of canonical mTORC1 signalling

The mechanistic target of rapamycin (mTOR) is a **serine/threonine protein kinase** that is evolutionary conserved from yeast to humans. It functions as a major and essential regulator of cellular metabolism and its dysfunction is associated with various adverse conditions and diseases. mTOR represents the **catalytic subunit of two macromolecular complexes**: mTOR complex 1 (**mTORC1**), which controls cell metabolism and growth, and mTOR complex 2 (**mTORC2**), which is involved in maintaining cell polarity, proliferation and survival [56].

Especially mTORC1 has been extensively studied in the past years and has emerged as a **central signalling hub** by sensing upstream stimuli, such as nutrients, metabolic intermediates and growth factors. Once modulated by these inputs, mTORC1 transduces the signals by phosphorylating a variety of downstream substrates that are involved in fundamental cellular processes (**Figure 1.5 A**). Importantly, mTORC1 plays a dual role in modulating cellular

metabolism. It **promotes anabolic pathways**, such as protein synthesis by phosphorylating substrates like S6 kinase (S6K) and eukaryotic initiation factor 4E-binding protein (4E-BP), as well as lipid and nucleotide biosynthesis. At the same time, mTORC1 **inhibits catabolic processes**, such as lysosomal biogenesis and autophagy by phosphorylating its substrates of the MiT/TFE family and Unc-51 like autophagy activating kinase 1 (ULK1) [57].

Phosphorylation of downstream targets occurs once mTORC1 is activated in response to nutrient availability and pro-growth stimuli. This **activation occurs at the lysosome**, a major degradative organelle, that represents a source of cellular metabolites and thus a fitting platform for mTORC1 to integrate nutritional and energy inputs. The activation of mTORC1 is a two-step process, that requires coordinated action of several protein molecules (**Figure 1.5 B**).

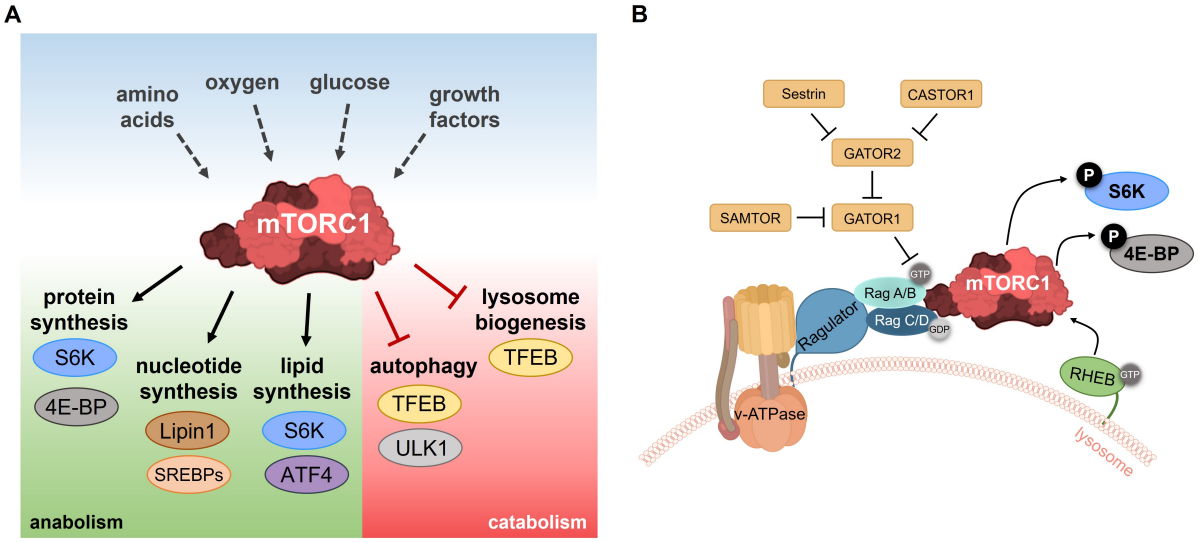

Figure 1.5: Overview of canonical mTORC1 pathway. A. mTORC1 activation incorporates environmental stimuli (top) to phosphorylate diverse downstream targets. These promote anabolic processes (bottom left) and repress catabolic processes (bottom right). SREBPs: sterol regulatory element-binding proteins, ATF4: activating transcription factor 4, TFEB: transcription factor EB. Adapted from Rabanal-Ruiz *et al.*, 2017 [57]. **B.** Canonical mTORC1 activation at the lysosomal membrane requires two inputs: activation of amino acid-stimulated Rags for mTORC1 recruitment and growth factor-activated RHEB for conformational changes in the active site of mTORC1. Adapted from Napolitano *et al.*, 2022 [58].

First, the recruitment of mTORC1 to the lysosome relies on **Ras-related GTPases** (**Rags**), which are heterodimeric proteins consisting of RagA or RagB bound to RagC or RagD. Rags localize to the lysosome due to their interaction with lysosome-anchored complex **Ragulator** [59, 60]. To recruit mTORC1, Rags must be activated and have RagA/B in GTP-bound state, which is ensured through a complex machinery that senses the availability of cytosolic and lysosomal amino acids (e.g. GATOR1/2, SAMTOR, CASTOR1, Sestrin2) [58]. Additionally, **lysosomal v-ATPase**, which maintains the proton gradient and low intra-lysosomal pH, was shown to be required for amino acid sensing. When amino acids are scarce, v-ATPase , Ragulator and Rags form a tightly bound supercomplex that cannot activate mTORC1. Upon amino acid accumulation in the lysosomal lumen, v-ATPase activates Ragulator to promote GTP-loading of RagA/B, which weakens Rag-Ragulator interactions and recruits mTORC1 to the lysosomal surface [61, 62].

Lysosomal recruitment brings mTORC1 in proximity to small GTPase **RHEB**, which also localizes to the lysosome and mediates the second step of mTORC1 activation process. Like RagA/B, RHEB is active in its GTP-bound form, which is maintained by the availability of growth factors [63]. According to structural studies, active RHEB induces conformational changes in the active site of mTORC1 to promote its kinase activity [64].

Complete activation of mTORC1 thus requires integration of multiple inputs: amino acid-stimulated Rag activity for lysosomal recruitment of mTORC1 and growth factor-induced RHEB activation for allosteric alignment of the active site in mTORC1. However, this two-step activation process is only required for phosphorylation of **canonical mTORC1 targets**, such as S6K and 4E-BP, which contain a **TOR signalling motif**. Consisting of five amino acids, this region interacts with mTORC1 via its subunit Raptor to enable substrate recruitment and phosphorylation [65].

Phosphorylation of S6K and 4E-BP represents a mechanism by which mTOR **regulates global translation rates**. Majority of mRNAs rely on cap-mediated translation, which is mostly controlled at the initiation step by ribosome recruitment to the mRNA. This initiation process is facilitated by several eukaryotic translation initiation factors (eIFs), which assemble into eIF4F complex, required for recruiting the small ribosomal subunit and eventual translation. mTORC1 promotes translation by **phosphorylation and inactivation of 4E-BP**, which would otherwise prevent interactions between eIFs and the assembly of eIF4F complex [66]. In addition to this global cap-mediated translational regulation, mTORC1 also regulates the **translation of specific subsets of mRNAs** with distinct features in their 5' untranslated regions (5' UTR). One such subset of mRNAs is characterized by **terminal oligopyrimidine motifs**, which are stretches of 4 – 15 pyrimidines in 5' UTRs of genes mostly associated with translation [67]. Other mTOR-sensitive mRNAs include those with long and complex 5' UTRs (encoding cell-cycle and survival-promoting proteins) and those with extremely short 5' UTRs (encoding mitochondrial proteins) [68]. Collectively, these findings show that mTORC1 regulates several aspects of protein synthesis and functions as a major regulator of cellular translatome, both in normal physiology and disease.

1.2.2 Non-canonical mTORC1 signalling and regulation of MiT/TFE factors

Besides the above-described targets involved in protein synthesis, mTOR also phosphorylates a family of MiT/TFE factors. This consists of four basic helix-loop-helix leucine-zipper transcription factors that share high sequence similarity: microphthalmia transcription factor (MITF), transcription factor EB (TFEB), transcription factor E3 (TFE3) and transcription factor EC (TFEC). Upon formation of homo- or heterodimers these factors bind to palindromic DNA sequences referred to as E-boxes and transcribe their target genes. Although MiT/TFE factors are ubiquitously expressed, their expression levels vary considerably across tissues, which results in tissue-specific transcriptional networks [69]. MITF regulates genes important for melanosome biogenesis and eye development [70]. TFEB and TFE3 control a largely overlapping set of genes referred to as the CLEAR (coordinated lysosomal expression and regulation) network and thus play cooperative and partially redundant roles in lysosomal biogenesis and autophagy [71, 72].

Transcriptional activity of MiT/TFE factors is regulated through their subcellular localization and **nucleo-cytoplasmic shuttling depending on nutrient availability**. When nutrients are abundant, various kinases phosphorylate selected amino acid residues in MiT/TFE factors and thereby retain them in the cytoplasm. On the contrary, nutrient deprivation causes nuclear translocation of MiT/TFE factors and enables their transcriptional activity [69]. Best studied mechanism of this process involves **phosphorylation of specific serine residues in TFEB by mTORC1** (**Figure 1.6 A**). These include phosphorylation at Ser122, Ser138, Ser142 and Ser211 under nutrient-replete conditions, which result in cytoplasmic retention of TFEB. Mechanistically, phosphorylation of Ser211 mediates the association of TFEB with cytosolic chaperone 14-3-3, which inhibits nuclear translocation of TFEB by occluding its nuclear localization signal [73, 74]. Ser138 and Ser142 are located in close proximity to nuclear export signal of TFEB and their phosphorylation is required for efficient nuclear export and cytosolic redistribution [75]. Another study has shown that Ser138 and Ser142 are also phosphorylated by GSK3β and ERK, respectively, to promote nuclear export [76]. However, the mechanisms by which phosphorylation at Ser122 promotes cytoplasmic localization of TFEB are not known and require further investigation [77].

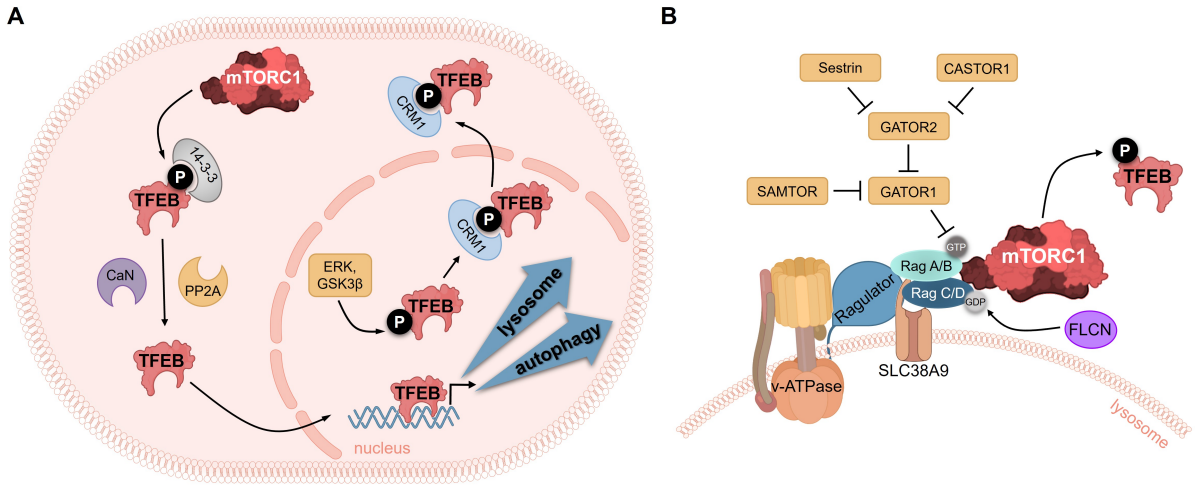

Figure 1.6: Overview of non-canonical mTORC1 signalling pathway. A. mTORC1 is a major regulator of subcellular localization and transcriptional activity of MiT/TFE factors, here exemplified by TFEB. mTORC1 phosphorylates TFEB to induce its cytosolic retention by 14-3-3 chaperones. This process is counteracted by cytosolic phosphatases calcineurin (CaN) and protein phosphatase 2 (PP2A). Unphosphorylated TFEB is translocated into the nucleus, where it transcribes genes involved in lysosomal biogenesis and autophagy. Nuclear phosphorylation of TFEB regulates its nuclear export through exportin 1 (CRM1). **B.** Non-canonical mTORC1 activation at the lysosomal membrane relies exclusively on the activation of Rags. Active RagA/B is required for mTORC1 recruitment, whereas active RagC/D is necessary for selective phosphorylation of TFEB. FLCN acts as a GTPase activating protein for RagC/D and thus plays an important role in substrate specificity of mTORC1. SLC38A9 mediates arginine-induced lysosomal efflux of essential amino acids and promotes FLCN activity. Adapted from Napolitano *et al.*, 2022 [58].

Recently, accumulating evidence has suggested that phosphorylation of MiT/TFE factors by mTORC1 differs from that of canonical targets, such as S6K and 4E-BP. MiT/TFE factors namely lack the TOR signalling motif, but instead contain an **N-terminal Rag-binding region**, which enables them to interact with Rags exclusively in their active state, that is when

RagA/B is GTP-bound and RagC/D is GDP-bound [78]. Furthermore, phosphorylation of MiT/TFE factors does not require RHEB and growth factor stimulation, but relies exclusively on amino-acid-dependent activation of Rags [79] (**Figure 1.6 B**). This highlights a double role of Rags in mTORC1 signalling, which is favoured by their dimeric composition. Active RagA/B is strictly required for lysosomal recruitment of mTORC1 and subsequent phosphorylation of all its substrates. On the other hand, **active RagC/D is strictly necessary for selective recruitment and phosphorylation of MiT/TFE factors**, but plays a minor role in mTORC1 lysosomal localization or phosphorylation of canonical targets [58].

Consequently, MiT/TFE phosphorylation is selectively modified in conditions that regulate the activity of RagC/D or folliculin (FLCN), which acts as GTPase activating protein for RagC/D [80]. This is the case in Birt-Hogg-Dubé syndrome, where mutations in FLCN cause inhibition of RagC/D and hypophosphorylation of TFEB, leading to its constitutive transcriptional activity [79]. Further examples are provided by lysosomal damage and selective autophagy following mitochondrial damage and pathogen infection, which similarly induce RagC/D inhibition and prevent TFEB phosphorylation, thereby promoting its activity [58]. These results highlight that non-canonical and substrate-specific mTORC1 signalling is equally important for cellular homeostasis, and its dysregulation can lead to the development of adverse conditions.

1.2.3 mTORC1 function in the heart and ischemic heart disease

Maintaining cardiac function requires a correct balance between anabolic and catabolic processes, as well as timely and coordinated response to nutrient availability and cardiac stress. mTOR signalling is thus indispensable for cardiac development, structure and growth, and plays crucial roles in cardiac disease.

Systemic deletions of mTOR and other subunits of mTORC1 are embryonically lethal, which points to their essential roles in early development and survival [81, 82]. Constitutive cardiomyocyte-specific deletion of mTOR has a similar effect and points to the requirement of mTOR for cardiomyocyte survival and proliferation during heart development [83]. During postnatal period, **mTOR is vital for maintaining cardiac structure and function**. This is exemplified by inducible cardiomyocyte-specific deletion of mTOR in adult mice, which results in dilated cardiomyopathy, heart failure and premature death. These mice are characterized by increased apoptosis, autophagy, mitochondrial abnormalities and accumulation of 4E-BP, which inhibits translation initiation [84]. The importance of mTOR signalling in the heart is further underscored by deletion of mTORC1 subunit Raptor or mTORC1 activator RHEB, both of which cause cardiac dilation and dysfunction, sarcomere disarray and high mortality [85, 86]. Complete deletion or inactivation of mTORC1 is therefore not favourable, but **partial inhibition of mTORC1 in aging mice has shown beneficial effects**. Reducing mTORC1 activity by administration of inhibitor rapamycin or through caloric restriction led to increased life span, improved cardiac function, reduced myocardial fibrosis and activation of autophagy during aging process [87, 88].

Additionally, increasing number of studies suggest that mTOR pathway plays an **important role in cardiac adaptation during ischemia** (**Figure 1.7**). Myocardial ischemia occurs

when blood flow to the heart muscle is reduced, for example by physical obstruction of a vessel, such as plaque build-up during coronary atherosclerosis. Reduced blood flow represents energy stress, as it reduces the supply of nutrients and oxygen, and causes irreversible injury referred to as myocardial infraction. The only way to rescue ischemic myocardium is reperfusion or restoration of blood flow, which paradoxically induces further irreversible damage, referred to as reperfusion injury [89]. These irreversible injuries contribute to the development of ischemic heart disease, which is the most prevalent type of cardiovascular disease and a major cause of death worldwide [90]. Interventions aimed at reducing myocardial injury through cardioprotective signalling pathways are therefore needed to reduce mortality.

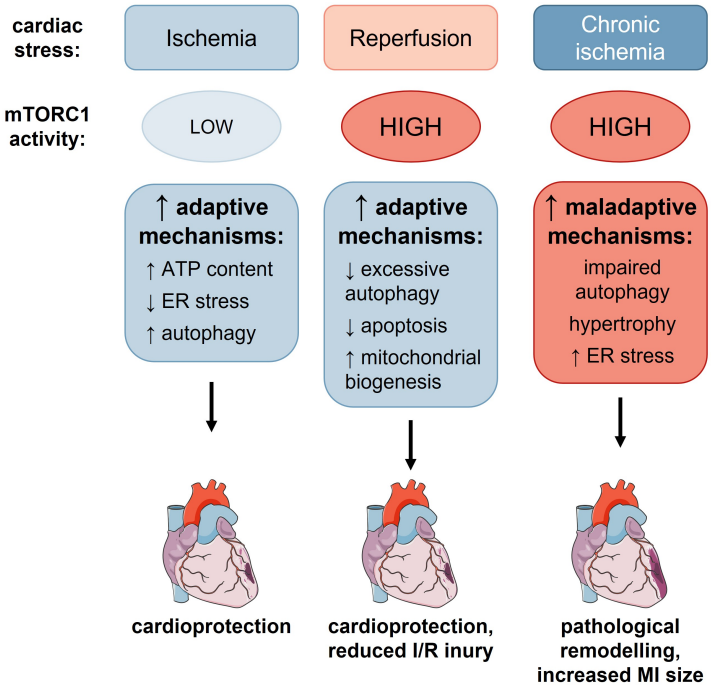

Figure 1.7: Modulation of mTORC1 pathway in cardiac ischemic disease. Ischemic conditions (left) reduce mTORC1 activity to promote adaptive mechanisms and confer cardioprotection. However, prolonged and chronic ischemia (right) leads to enhanced mTORC1 activity that promotes maladaptive mechanisms and pathological heart remodelling, resulting in larger infarct size. Reperfusion of ischemic areas (middle) invokes beneficial increase in mTORC1 activity, leading to reduced I/R injury and smaller infarct size. Adapted from Sciaretta *et al.*, 2022 [55].

Myocardial **ischemia is characterized by low mTORC1 activity**, which is an adaptive response that activates protective mechanisms, such as autophagy. Ischemic conditions lead to inactivation of RHEB, which results in mTORC1 inhibition, autophagy induction, increased ATP cellular content, decreased protein folding stress and overall improvement in cardiomyocyte survival [91]. Another mechanism for mTORC1 inhibition during ischemia is provided by activation of GSK3β, which modulates upstream regulators of mTORC1 and similarly results in improved autophagic flux and decreased myocardial injury [92].

mTORC1 inhibition is also beneficial during prolonged, chronic ischemic injury. Attenuation of mTORC1 activity via pharmacological inhibition with everolimus improved cardiac function in a rat model of myocardial infarction (MI). Cardiac dilation and infarct size were

reduced, which was attributed to increased autophagy and reduced proteasomal activity [93]. Similarly, administering rapamycin or S6K inhibitor led to improved cardiac function, decreased cardiac remodelling and reduced cardiomyocyte apoptosis after MI in mice [94]. Genetic manipulation of mTORC1 activity indicated similar outcomes, with cardiomyocyte-specific deletion of RHEB and concomitant inhibition of mTORC1 conferring protection against pathological heart remodelling following MI in mice [95]. Comparable cardioprotective results post-infraction were also obtained by cardiomyocyte-specific overexpression of mTORC1 subunit PRAS40, that led to reduced activity of the complex [96].

However, in the case of ischemia-reperfusion (I/R) injury, the role of mTORC1 activation remains ambiguous. Reperfusion phase rapidly activates mTORC1 by restoring nutrient and oxygen availability. Inhibition of mTORC1 by rapamycin in this setting only proved to be beneficial when administered prior to the onset of I/R [97], whereas administration at the time of reperfusion did not confer protective advantages [92]. Instead, mice that received branched chain amino acids prior to I/R showed a reduction in infarct size that was mediated by mTORC1 activation [98]. Similarly, genetic manipulations such as heterozygous deletion of GSK3β or cardiomyocyte-specific overexpression of CITED-4 resulted in reduced reperfusion injury through activation of mTORC1 [92, 99]. In agreement with these results, another study found that cardiomyocyte-specific overexpression of mTOR preserves heart function after I/R and reduces pathological remodelling through inhibition of necrosis and inflammatory suppression [100]. Many studies thus suggest that mTORC1 activity is beneficial during reperfusion phase, as it limits maladaptive and excess autophagy and promotes translation of mitochondrial biogenesis genes, which might support functional recovery after ischemic injury [101, 102].

1.2.4 mTOR pathway as a signature of human cardiac translational landscape

Most of our current knowledge of mTORC1 biology and its regulatory functions in the heart stems from experiments on cell lines and animals. Yet how the activity of mTORC1 is regulated in cardiovascular diseases in humans has only been addressed in few studies.

Yano and colleagues performed immunohistochemical analyses of over 50 human endomyocardial biopsies from patients with nonischaemic dilated cardiomyopathy and compared them to healthy controls [103]. Diseased samples revealed significant increases in the phosphorylation of S6K and its downstream target RPS6, which are hallmarks of mTORC1 activation. Patients with **high mTOR activity** also exhibited **higher levels of myocardial fibrosis** and were more likely to develop adverse cardiac events during follow-up. These results indicate that **persistent activity of mTORC1 is clinically relevant**, as it exacerbates human heart failure and is associated with worse prognosis.

Recently, a genome-wide study of human cardiac translation used ribosome profiling to investigate translation events in 65 ventricular samples from end-stage dilated cardiomyopathy patients and 15 controls [41]. In line with Yano and colleagues, the authors observed a transcriptional and **translational increase in the components of mTOR pathway in diseased hearts**. Consequently, mRNAs containing terminal oligopyrimidine motifs, such as ribosomal proteins, were translationally upregulated in dilated hearts, pointing to an overall increase in translational

activity. Accordingly, this work highlights the involvement of mTOR signalling in end-stage cardiac dilation and serves as an excellent resource for further studies of human cardiac translational landscape.

1.3 Diversity of cardiac cell types guides heart function

The heart is a highly complex organ composed of four distinctive chambers and several anatomical subdomains. Their activities need to be perfectly synchronized to support heart function; that is the pumping of blood along with oxygen and nutrients throughout the entire organism. The heart is also a very dynamic organ, which reacts to various stimuli that converge on heart cells during development and in adulthood. These local cues drive gene expression changes and phenotypical adaptations that eventually culminate in heart (dys)function.

Although bulk transcriptomic studies have greatly contributed to our knowledge of transcriptional signatures in heart physiology and disease, single-cell omics technologies are offering remarkable insights into gene expression and **heterogeneity at the level of individual cells**. Single-cell transcriptomics (scRNA-seq) applied to the heart has substantiated the existence of diverse cardiac cell types, which include cardiomyocytes, fibroblasts, endothelial cells, smooth muscle cells, pericytes, immune cells and rare cell types, such as neuronal cells and adipocytes. In addition, each of these cell types populates multiple cell states in different anatomical locations (**Figure 1.8**), and scRNA-seq has proven to be invaluable for identification of novel markers and specific signatures to define diversified cell states [104].

Cardiomyocytes are large rod-shaped cells that occupy the majority of heart mass and volume, although they are outnumbered by non-myocyte types. Consistent with their role in generating contractile forces, all cardiomyocytes express sarcomere and calcium-handling genes. But the most comprehensive human heart cell atlas to date identified as many as 5 ventricular and 5 atrial cardiomyocyte populations. This highlights the transcriptional diversity of myocytes in the healthy heart, which possibly arises due to differences in developmental origin, electromechanical stimulation and adaptation to specific microenvironments [105]. Another study used human samples after myocardial infarction and detected 5 distinct niches of stressed cardiomyocytes within areas of ischemic injury, repair and fibrotic remodelling [106].

As much as 10 % of the adult myocardium is comprised of **fibroblasts**, which are critical for structural and mechanical maintenance of the heart. These cells produce and regulate extracellular matrix (ECM) and collagen components and act as lineage progenitors [107]. ScRNA-seq in healthy hearts defined 7 fibroblast populations, including chamber-specific, high ECM-producing and immune response-modulating fibroblasts [105]. Following disease stimuli, fibroblasts become activated and adopt a myofibroblast state, in which they express more contractile genes and fibrillar markers. Various myofibroblast states have different pro- or anti-fibrotic gene signatures, highlighting the tight control and resolving of fibrosis [108].

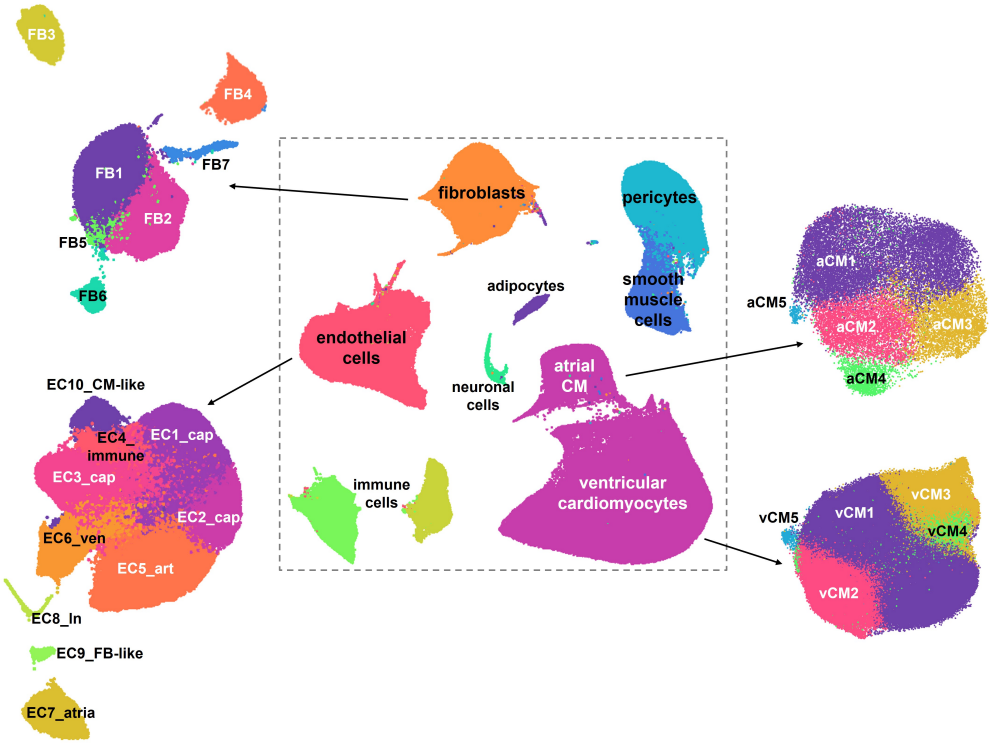

Figure 1.8: Human heart cell atlas reveals the heterogeneity of cardiac cell types and states. Single-cell transcriptomics applied to healthy human heart has identified major cardiovascular cell types (within the gray dashed outline). Each of these cell types populates various cell states. Indicated on the right are atrial cardiomyocytes (aCM1-5) and ventricular cardiomyocytes (vCM1-5), and on the left are fibroblasts (FB1-7) and endothelial cells (ECs; EC1/2/3_cap: capillary ECs, EC4_immune: immune-related ECs, EC5_art: arterial ECs, EC6_ven: venous ECs, EC7_atria: atria-enriched ECs, EC8_ln: lymphatic ECs, EC9_FB-like: ECs with fibroblast features, EC10_CM-like, ECs with cardiomyocyte features). Adapted from https://www.heartcellatlas.org [105].

Cardiac **vascular compartment** is also highly diverse and includes coronary arteries, veins, capillaries and the endocardium. Notably, cardiac microvasculature is the most abundant and in close contact with cardiomyocytes to provide optimal diffusion of oxygen and nutrients into the muscle [109]. The most in-depth scRNA-seq study of cardiac endothelial cells identified 10 distinct populations in healthy hearts, while analysis of diseased samples illustrated contrasting transcriptional signatures of these populations compared to controls [105, 110]. Especially the upregulation of transforming growth factor-β (TGF-β) signalling pathway has been noted in several studies and suggests the transition of endothelial cells towards a more invasive and migratory phenotype in cardiac disease [106, 110, 111].

Collectively, the advances in single-cell transcriptomics have reinforced the appreciation for the **entire repertoire of cardiac cells**. ScRNA-seq studies have also enabled a more detailed understanding of the cardiac cellular landscape and facilitated our understanding of cardiovascular disease mechanisms.

1.4 Preliminary data and project aims

1.4.1 Aim 1: Identifying novel cardiac microproteins

Previous studies in our research group used cardiomyocyte-specific ribosome profiling to investigate translational processes specific to the muscle tissue of the heart [36]. Ribosome footprints were therefore immunoprecipitated from heart lysates of transgenic mice expressing HA-tagged ribosomes only in cardiomyocytes (**Figure 1.9**). The cardiac myocyte translatome was then analyzed during pathological cardiac remodeling induced by transverse aortic constriction (TAC) or I/R injury in mice. Additionally, the same method was used to probe the cardiac translatome in a physiological setting induced by a swimming regime. The resulting Ribo-seq libraries were subsequently processed and analysed with *Rp-Bp*, a pipeline which features an unsupervised Bayesian approach based on the ribosome's periodic behaviour to predict translated ORFs [112].

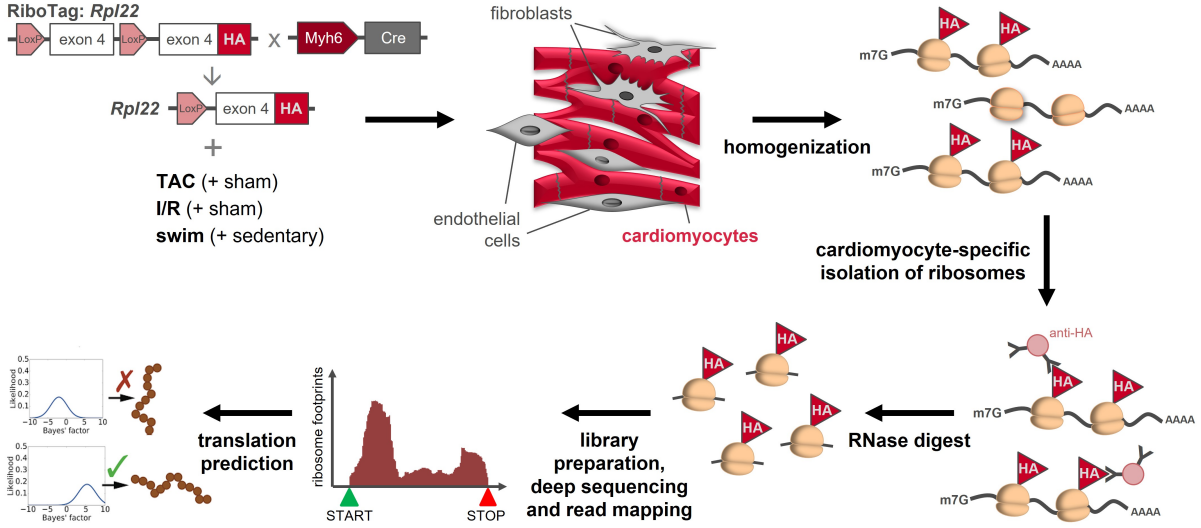

Figure 1.9: Cardiomyocyte-specific ribosome profiling provides insight into the translatome of mouse myocardium. RiboTag mice carrying an HA-tagged allele of 60S ribosomal protein L22 (RPL22-HA) were crossed with mice expressing Cre recombinase from cardiac alpha-myosin (Myh6) promoter, resulting in the expression of HA-tagged RPL22 specifically in cardiomyocytes. These mice were then exposed to either pathological (TAC or I/R) or physiological stimulus (swimming) with corresponding controls (sham or sedentary mice). Heart tissue of these animals was homogenized and cardiomyocyte HA-tagged ribosomes were immunoprecipitated using an anti-HA antibody. Following library preparation and deep sequencing, ribosome footprints were mapped onto the genome to reveal regions and ORFs undergoing active translation.

The described approach resulted in approximately 8400 sORFs (shorter than 100 codons) predicted to be translated in at least one of the experimental conditions. This extensive compilation of translated sORFs led me to hypothesize that myocardial translatome might contain previously uncharacterized cardiac microproteins involved in heart function or disease development.

Therefore, the first aim of the project was to **identify biologically functional sORFs and microproteins expressed in the cardiac muscle**. Since evolutionary conserved sequences are more likely to be functionally and structurally important, the presented dataset was expanded by including expression information from human samples. Special focus was also placed on **novel, previously unannotated microproteins**. Integrating these criteria enabled me

to assemble a list of potential candidates, from which most promising targets were selected for further, more detailed **experimental characterization in the myocardium**.

1.4.2 Aim 2: Functional characterization of cardiac-enriched microprotein SPAAR

Although microprotein SPAAR has been previously characterized by other research groups, its features and function in the heart remain highly interesting for several reasons.

1. The initial publication investigated the role of SPAAR in skeletal muscle [10], but mining various expression datasets revealed that the transcript encoding SPAAR is in fact strongly enriched in the heart, both in humans and mice (**Figure 1.10 A, B**) [105, 106, 113]. Additionally, Ribo-seq results provide evidence that SPAAR is actively translated in mouse and human hearts (**Figure 1.10 C**). Finally, ultimate proof for SPAAR translation can be found in a genome-wide study of cardiac translation, which identified a signature peptide corresponding to SPAAR in the mass spectra of human heart tissue [41].
2. SPAAR was originally described as an inhibitor of mTORC1; a protein complex, which plays a crucial role in protein synthesis, cell growth and metabolism. In the cardiac setting, mTORC1 is required for cardiovascular development, maintenance of cardiac structure and function and was also shown to be necessary for cardiac adaptation to ischemic stress and development of hypertrophy [55].
3. Furthermore, the inhibitory effect of SPAAR on mTORC1 is dependent on the stimulation with amino acids. Their fluctuating levels and defects in amino acid metabolism have been implicated in various cardiovascular diseases, including murine models of heart failure and clinical studies in patients with chronic heart failure [114, 115].

Combining these existing reports with the restricted spatial expression pattern of SPAAR in the heart, led me to hypothesize that SPAAR plays an important role in regulating mTORC1 in the myocardium, thus making it a strong candidate for further functional characterization. To comprehensively address this research problem, two specific focus areas were defined:

- **Specific focus 1** is directed at understanding **the role of SPAAR in myocardium *in vivo***. For this purpose, SPAAR KO mouse model was used to probe heart phenotype and function at baseline and following cardiac damage.
- **Specific focus 2** is intended to provide more detailed **mechanistic insights into SPAAR function**. Careful examination of cell type-specific expression pattern of *SPAAR* within the heart indicates its highest abundance in cardiac endothelial cells (**Figure 1.10 D**) [105, 106]. Therefore, isolating primary cardiac endothelial cells from SPAAR KO mice and performing various *in vitro* experiments was used to evaluate the contribution of SPAAR to endothelial cell activity.

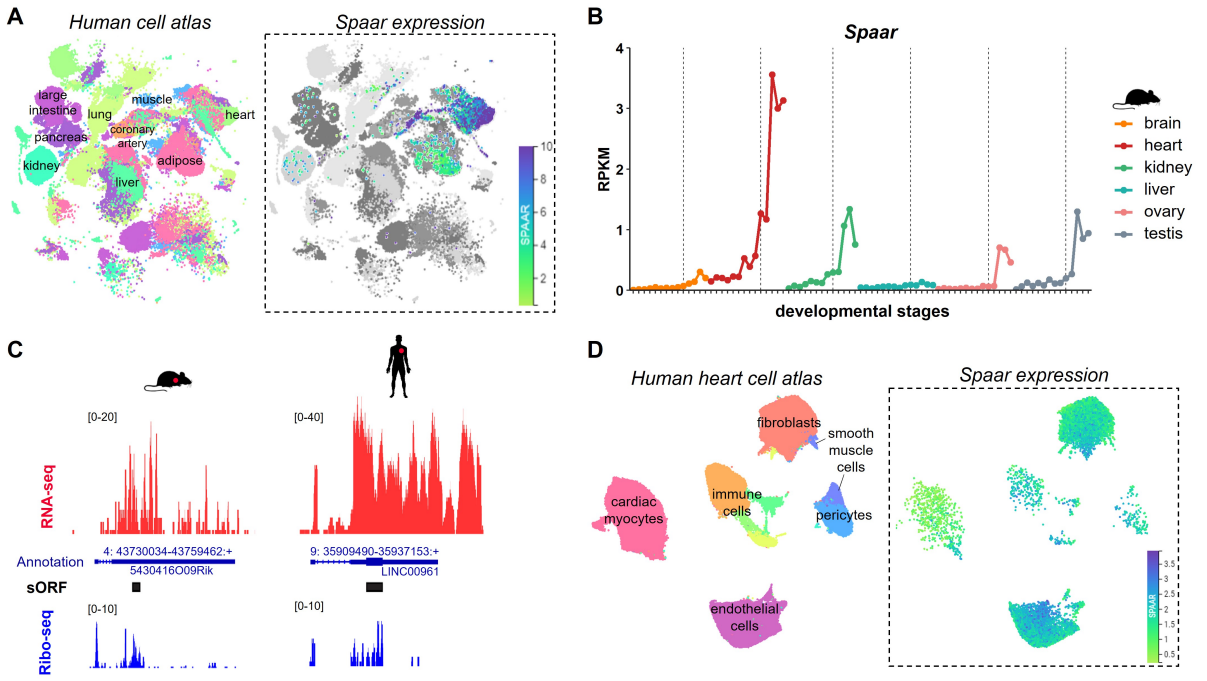

Figure 1.10: SPAAR is a cardiac endothelial-enriched microprotein. A. ScRNA-seq dataset of human organs indicates highest abundance of *SPAAR* in the heart. Adapted from Complete Tabula Sapiens Dataset available at https://tabula-sapiens-portal.ds.czbiohub.org/ [116]. **B.** Transcriptomic dataset of 6 mouse organs during development reveals the same spatial expression pattern in mouse tissues, namely enrichment of *Spaar* in the heart. Vertical dashed lines mark the beginning of postnatal development. Plotted based on the expression data published in Cardoso-Moreira *et al.*, 2019 [113]. **C.** RNA-seq and Ribo-seq libraries from mouse and human hearts show that *Spaar* is expressed and actively translated in both organisms. **D.** ScRNA-seq dataset of human heart tissue demonstrates that *SPAAR* is predominantly expressed in cardiac endothelial cells. Adapted from scRNA-seq data generated by Kuppe *et al.*, 2022 [106].

2 Materials

2.1 Chemicals and reagents

Table 2.1 List of used chemicals and reagents

Chemical or reagent	Vendor	Catalog number
2-deoxy-D-glucose	Sigma-Aldrich	D8375-5G
2-mercaptoethanol	Sigma-Aldrich	M6250-250ML
4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)	Thermo Fisher Scientific	15630056
6bK (IDE inhibitor)	Tocris / Bio-Techne	5402/500U
Acetic acid	Honeywell	64-19-7
Agarose	Carl Roth	2267.4
Ampicillin	Sigma-Aldrich	A9518-25G
Azidohomoalanine (AHA)	AnaSpec	AS-63669
Bouin's solution	Sigma-Aldrich	HT10132
Bovine serum albumin (BSA)	Sigma-Aldrich	A6003
Butanedione monoxime	Sigma-Aldrich	B0753-100G
Cadmium chloride ($CdCl_2$)	Honeywell	20899
Calcium chloride ($CaCl_2$)	Carl Roth	A119.1
Chloroform	Merck	32211
Citric acid, anhydrous	Sigma-Aldrich	C3674-500G
cOmplete Protease Inhibitor Cocktail	Sigma-Aldrich	11873580001
Cycloheximide (CHX)	Sigma-Aldrich	C1988
DAPI	Thermo Fisher Scientific	D21490
Dialyzed fetal bovine serum	Thermo Fisher Scientific	A3382001
Di-sodium hydrogen phosphate (Na_2HPO_4)	Carl Roth	P030.2
Dithiothreitol (DTT)	Sigma-Aldrich	43819
DMEM (no glucose, no glutamine, no phenol red)	Thermo Fisher Scientific	A1443001
DMEM (high glucose, with pyruvate)	Thermo Fisher Scientific	21969035
DMEM/F-12	Thermo Fisher Scientific	11330057
DMEM for SILAC	Thermo Fisher Scientific	88364
Dynabeads Sheep Anti-Rat IgG	Thermo Fisher Scientific	11035
Dynabeads Sheep Anti-Rabbit IgG	Thermo Fisher Scientific	11203D
Endothelial Cell Growth Medium 2	PromoCell	C-22111

Chemical or reagent	Vendor	Catalog number
Endothelial Cell Growth Supplement /Heparin	PromoCell	C-30120
Ethanol (EtOH)	Sigma-Aldrich	51976
Ethidium bromide	Sigma-Aldrich	E1510
Ethylenediaminetetraacetic acid (EDTA)	Sigma-Aldrich	E5134
Eukitt quick-hardening mounting medium	Sigma-Aldrich	03989-100ML
Fetal bovine serum (FBS)	Thermo Fisher Scientific	10270
Fibronectin	Sigma-Aldrich	F1141
Formaldehyde	Carl Roth	4980.4
Gelatine	Sigma-Aldrich	G9391-100G
Gene Ruler 1 kb DNA ladder	Thermo Fisher Scientific	SM0313
Glucose	Sigma-Aldrich	G7021-1KG
Glutaraldehyde solution	Sigma-Aldrich	G6257-100ML
GlycoBlue coprecipitant	Thermo Fisher Scientific	AM9516
HBSS (without Ca, Mg)	Thermo Fisher Scientific	14175053
Heparin	Ratiopharm	N68542.04
HiPerFect Transfection Reagent	Qiagen	301704
Hydrochloric acid	Honeywell	30721
Hydrogen peroxide (H_2O_2)	Merck-Millipore	107209
Isolectin IB4, Alexa Fluor 594 conjugate	Thermo Fisher Scientific	I21413
Isoflurane	Baxter	HDG9623
Isopropanol	Sigma-Aldrich	I9516
iTaq Universal SYBR Green Supermix	Bio-Rad Laboratories	1725124
L-Glutamine	Thermo Fisher Scientific	25030081
L-arginine ($^{13}C_6$)	Cambridge Isotope Laboratories	CLM-2265-H-PK
L-arginine ($^{13}C_6$,$^{15}N_4$)	Cambridge Isotope Laboratories	CNLM-539-H-PK
L-lysine (4,4,5,5-D_4)	Cambridge Isotope Laboratories	DLM-2640-PK
L-lysine ($^{13}C_6$,$^{15}N_2$)	Cambridge Isotope Laboratories	CNLM-291-H-PK
Laemmli Sample Buffer	Bio-Rad Laboratories	161-0747
LB-Agar	Carl Roth	6671
Magnesium chloride ($MgCl_2$)	Sigma-Aldrich	P9333-1KG
Magnesium sulfate ($MgSO_4$)	Sigma-Aldrich	M2643-500G
Mammalian Polysome Buffer	Illumina	ASBHMR1212
Matrigel Growth Factor Reduced Basement Membrane Matrix	Corning	356231
MEM Amino Acids Solution	Thermo Fisher Scientific	11130036

Chemical or reagent	Vendor	Catalog number
MEM Non-Essential Amino Acids Solution	Thermo Fisher Scientific	11140035
Methanol	VWR Chemicals	20847.307
Nonidet P-40 Substitute (NP-40)	Sigma-Aldrich	74385-1L
Paraffin Paraplast	Leica	39601006
Paraformaldehyde	Electron Microscopy Sciences	15713-S
Polyethylenimine hydrochloride (PEI)	Polysciences	24765-2
Penicillin-Streptomycin	Thermo Fisher Scientific	15140122
Penicillin-Streptomycin-Glutamine	Thermo Fisher Scientific	10378016
Percoll	GE Healthcare	17-0891-02
Phenol Red	Sigma-Aldrich	P3532-25G
Phenylephrine hydrochloride (PE-HCl)	Sigma-Aldrich	P6126-10G
Phosphate buffered saline (PBS)	Sigma-Aldrich	D8537-500ML
PhosSTOP phosphatase inhibitors	Sigma-Aldrich	4906837001
Ponceau S	Sigma-Aldrich	P3504
Potassium chloride (KCl)	Carl Roth	6781.1
Potassium hydrogen carbonate ($KHCO_3$)	Carl Roth	P748.2
Powdered milk	Carl Roth	T145.2
Primocin	Invivogen	ant-pm-2
primaAMP 2× PCR Master Mix Red	Steinbrenner	SL-9612
Puromycin	Sigma-Aldrich	540411
QIAzol Lysis Reagent	Qiagen	79306
RPMI 1640 medium without amino acids	US Biological	R8999-04A
Sodium chloride (NaCl)	Carl Roth	9265.2
Sodium deoxycholate	Sigma-Aldrich	D6750-100G
Sodium dodecyl sulfate (SDS)	Serva	20765.03
Sodium hydrogen carbonate ($NaHCO_3$)	Carl Roth	6885.5
Sodium hydroxide (NaOH)	Sigma-Aldrich	30620
SUPERase In RNase Inhibitor	Thermo Fisher Scientific	AM2696
SYBR Safe DNA Gel Stain	Thermo Fisher Scientific	S33102
Taurine	Sigma-Aldrich	T8691-100G
Thiazolyl Blue Tetrazolium Bromide (MTT)	Sigma-Aldrich	M5655-500MG
Tris	Carl Roth	4855.1
Tris hydrochloride (Tris-HCl)	Carl Roth	9090.2
Triton X-100	Sigma-Aldrich	X100-500ML
Trypsin	Thermo Fisher Scientific	15090046
Trypsin-EDTA	Thermo Fisher Scientific	25200056
TSA Blocking Reagent	PerkinElmer	FP1012
Tween 20	Carl Roth	25300-120

Chemical or reagent	Vendor	Catalog number
Vectashield Antifade Mounting Medium	Vector Laboratories	H-1000
Water, nuclease-free	Thermo Fisher Scientific	AM9937
Weigert's iron hematoxylin solution	Sigma-Aldrich	HT1079-1SET
WesternBright ECL	Advansta	541005
Wheat Germ Agglutinin (WGA), FITC conjugate	Sigma-Aldrich	L4895-2MG
XT MES Running Buffer	Bio-Rad Laboratories	1610789
XT MOPS Running Buffer	Bio-Rad Laboratories	1610788
Xylene	Carl Roth	9713.5

2.2 Kits

Table 2.2 List of used kits

Kit	Vendor	Catalog number
Click-iT Protein Enrichment Kit	Thermo Fisher Scientific	C10416
CyQUANT NF Cell Proliferation Assay Kit	Thermo Fisher Scientific	C35007
DC Protein Assay	Bio-Rad Laboratories	5000112
Dead Cell Apoptosis Kit for flow cytometry	Thermo Fisher Scientific	V13242
First Strand cDNA Synthesis Kit	Thermo Fisher Scientific	K1612
Hematoxylin & Eosin fast staining kit	Carl Roth	9194.1
High Sensitivity DNA Kit	Agilent Technologies	5067-4626
HiSpeed Plasmid Midi Kit	Qiagen	12643
Ide ELISA Kit Mouse	Aviva Systems Biology	OKEH03420
Masson Trichrome Staining set	Sigma-Aldrich	HT15-1KT
NEXTFLEX Small RNA-seq v3 Kit	PerkinElmer	NOVA-5132-06
peqGOLD Plasmid Miniprep Kit I	VWR Chemicals	13-6943-02
Quick-RNA MiniPrep Kit	Zymo Research	R1055
Qubit dsDNA High Sensitivity Assay Kit	Thermo Fisher Scientific	Q32855
RiboCop rRNA Depletion Kit for Human/Mouse/Rat	Lexogen	144.24
riboPOOL ribo-seq kit (mouse/rat)	siTOOLs Biotech	dp-K012-000052
Ribo-Zero Gold rRNA Removal Kit (Human/Mouse/Rat)	Illumina	NA (discontinued product)
RNA 6000 Nano Kit	Agilent Technologies	5067-1511
RNA Clean & Concentrator-5	Zymo Research	R1014
RNA Clean & Concentrator-25	Zymo Research	R1018
TruSeq Ribo Profile (Mammalian) Library Prep Kit	Illumina	NA (discontinued product)

2.3 Enzymes

Table 2.3 List of used enzymes

Enyzme	Vendor	Catalog number
Collagenase type 2	Worthington Biochemical	LS0004176
Deoxyribonuclease I	Worthington Biochemical	LS002006
DNase I (grade II)	Sigma-Aldrich	10104159001
Gateway LR Clonase II Enzyme mix	Thermo Fisher Scientific	11791020
Liberase DH	Sigma-Aldrich	5401054001
PacI	New England Biolabs	R0547L
RNase I	Invitrogen	AM2295
SacII	New England Biolabs	R0157L
T4 Polynucleotide Kinase	New England Biolabs	M0201S
TURBO DNase	Invitrogen	AM2239

2.4 Antibodies

Table 2.4 List of antibodies used for western blot

Target	Dilution	Vendor	Catalog number
FLAG tag	1:2000	Sigma-Aldrich	F1804-200UG
GAPDH	1:2000	Santa Cruz Biotechnology	sc-365062
IDE	1:3000	Proteintech	67106-1-Ig
phospho-4E-BP (Thr37/46)	1:3000	Cell Signaling Technology	2855S
phospho-AKT (Ser473)	1:1000	Cell Signaling Technology	9271S
phospho-RPS6 (Ser235/236)	1:1000	Cell Signaling Technology	2211S
phospho-S6K (Thr389)	1:1000	Cell Signaling Technology	9205L
phospho-TFEB (Ser122)	1:1000	Cell Signaling Technology	86843S
phospho-TFEB (Ser211)	1:1000	Cell Signaling Technology	37681S
phospho-ULK1 (Ser757)	1:3000	Cell Signaling Technology	14202T
peroxidase-conjugated anti-mouse IgG	1:5000	Cell Signaling Technology	715-035-151
peroxidase-conjugated anti-rabbit IgG	1:5000	Cell Signaling Technology	111-035-144
puromycin	d1:10,000	Merck-Millipore	MABE343
SMIM4	d1:1000	Biozol	BSS-BS-15183R
SPAAR	d1:500	Bethyl Laboratories	SPAR-48 (custom-made)

Table 2.5 List of antibodies used for isolation of juvenile mouse endothelial cells

Target	Vendor	Catalog number
rat anti-mouse CD102	BD Biosciences	553326
rat anti-mouse CD31	BD Biosciences	553370

Table 2.6 List of antibodies used for immunohistochemistry and immunofluorescence

Target	Dilution	Vendor	Catalog number
cardiac troponin T (cTnT)	1:200	Abcam	ab209813
Cy3-conjugated Anti-Mouse IgG	1:100	Jackson ImmunoResearch	715-165-151
Cy3-conjugated Anti-Rabbit IgG	1:100	Jackson ImmunoResearch	711-165-152
FITC-conjugated Anti-Mouse IgG	1:100	Jackson ImmunoResearch	715-095-151
FITC-conjugated Anti-Rabbit IgG	1:100	Jackson ImmunoResearch	711-095-152
FLAG tag	1:100	Sigma-Aldrich	F1804-200UG

2.5 Primers

Species specificity of each primer pair is indicated in the brackets (*Mm*: *M. musculus* transcriptome, *Rn*: *R. norvegicus* transcriptome).

Table 2.7 List of primer sequences [5' → 3']

Target	Forward primer	Reverse primer
Atp6v0d1 (Mm)	TGTGTCCTATCCTAGAGTTTGAAG	ACCTGTTCATAGTCATCAGCCC
B2m (Mm)	CTACTCGGCGCTTCAGTCG	ACTTGAATTTGAGGGGTTTTCTGG
Btub (Mm, Rn)	TCGTGGAATGGATCCCCAAC	CTCCATCTCGTCCATGCCCT
Col1a1 (Mm)	GCTCCTCTTAGGGGCCACT	CCACGTCTCACCATGGGGG
Col5a1 (Mm)	GGGGCGGGACAGTATTTGAA	AATCCATCGGAAAGGCACGT
Col18a1 (Mm)	ACCAGGACCAGGATTTGCTG	GGCTCCATCTGCTCCAACTT
Ctsf (Mm)	GTTGCCATTAACGCCTTCGG	TTAGAGCGGTTGCCATAGCC
Fkbp10 (Mm)	CCTTCGGACCCTGTGGAAAT	CCTCATAGTCGTGGGACGAGA
Gnb1 (Mm)	TGAAGGGAATGTGCGTGTGA	CAGTCTCGATGTCCCACAGG
Gaa (Mm)	CCCTGACAACAAGAGCCTGG	ACACGGTGGGAAAAGCTCAA
Hprt (Mm)	CTTCCTCCTCAGACCGCTTT	TCGCTAATCACGACGCTGG
Hprt (Rn)	TCAGACCGCTTTTCCCGCGA	TCACTAATCACGACGCTGGGACTGA
Ide (Mm)	ACCACGAGGCTATACGTCCA	GAAACTATTGCCACCCGCAC
Lamp1 (Mm)	GGCTTTCAAGGTGGACAGTG	TGAGGTAGGCAATGAGGACG
Mcoln1 (Mm)	TGCTGAGCCTCTTCATTGCA	CGACGGAACTTGCCAGATGT
Sdha (Mm)	TGTTATTGCTACTGGGGGCTAC	CTCCACGACACCCTTCTGTG
Smim4 (Mm, Rn)	AAGCAGCGATTCGGCATCTA	TCCAGCCTTCTCTGATACTGTCT
Spaar (Mm)	CAGGCTTGGGAGGGAGAAG	GCAAGGGAAGGAGGAGCAG
Spaar genotyping	CTGGCTTCCGTCTTAGCAAC	ATTGCCAACCTCACACAACA

2.6 Buffers and solutions

All buffers and solution were prepared in ddH_2O unless indicated otherwise.

2.6.1 Buffers for genotyping SPAAR KO mice

Buffer	Composition
buffer A	25 mM NaOH, 200 µM EDTA
buffer B	40 mM Tris-HCl
TAE (50×)	2 M Tris, 1 M acetic acid, 50 mM EDTA

2.6.2 Buffers for cell and tissue lysate preparation

Buffer	Composition
Mammalian Lysis Buffer	1× Mammalian Polysome Buffer (Illumina), 1 % (v/v) Triton X-100, 1 mM DTT, 0.01 U/µL TURBO DNase, 0.1 mg/mL cycloheximide
RIPA buffer	150 mM NaCl, 25 mM Tris-HCl (pH=7.4), 1 % (v/v) NP-40, 0.5 % (w/v) sodium deoxycholate, 1 % (w/v) SDS, cOmplete Protease Inhibitor Cocktail, PhosSTOP phosphatase inhibitors

2.6.3 Buffers for immunoprecipitation of SPAAR from mouse heart lysates

Buffer	Composition
lysis buffer	150 mM NaCl, 25 mM Tris-HCl (pH=7.4), 1 % (v/v) NP-40, 0.5 % (w/v) sodium deoxycholate, 0.1 % (w/v) SDS, cOmplete Protease Inhibitor Cocktail
wash buffer	lysis buffer with 500 mM NaCl

2.6.4 Buffers for immunohistochemistry and immunofluorescence

Buffer	Composition
Citrate buffer	10 mM citric acid (anhydrous), pH=6.0
PBS	137 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH=7.4
TN buffer	0.1 M Tris, 0.15 M NaCl, pH=7.5-8.0
TNB blocking solution	0.5 TSA blocking reagent in TN buffer

2.6.5 Buffers for western blotting

Buffer	Composition
Ponceau S solution	0.1 % (w/v) Ponceau S, 5 % (v/v) acetic acid
TBST	150 mM NaCl, 20 mM Tris, 0.1 % (v/v) Tween 20, pH=7.4
Transfer buffer	190 mM glycine, 25 mM Tris, 20 % (v/v) methanol

2.6.6 Cell culture solutions and media for neonatal rat cardiomyocytes

Solution	Composition
Digestion Solution (in HBSS)	0.25 % (v/v) trypsin, 300 U/mL DNase I (grade II), 0.4 M HEPES, 0.0016 M NaOH, 2% (v/v) Penicillin-Streptomycin
Stop Solution (in HBSS)	4 % (v/v) FBS, 300 U/mL DNase I (grade II), 2% (v/v) Penicillin-Streptomycin
ADS (10×)	1.16 M NaCl, 180 mM HEPES, 8.45 mM Na_2HPO_4, 55.5 mM glucose, 53.7 mM KCl, 8.31 mM $MgSO_4$, pH=7.35±0.5
red ADS buffer	1× ADS, 0.002 % Phenol Red
Percoll stock	90 % (v/v) Percoll, 10 % ADS (10×)
top Percoll (density: 1.059 g/mL)	83% (v/v) Percoll stock, 17% ADS (1×)
bottom Percoll (density: 1.082 g/mL)	65% (v/v) Percoll stock, 35% (v/v) red ADS
Percoll gradient	4 mL top Percoll (pipetted first), 3 mL bottom Percoll (pipetted underneath)
NRCM Culture Medium	DMEM/F-12, 10% (v/v) FBS, 1% (v/v) Penicillin-Streptomycin-Glutamine
NRCM Treatment Medium	DMEM/F-12, 0.5% (v/v) FBS, 1% (v/v) Penicillin-Streptomycin-Glutamine

2.6.7 Cell culture solutions and media for adult mouse cardiomyocytes

Solution	Composition
Perfusion Buffer	113 mM NaCl, 4.7 mM KCl, 0.6 mM KH_2PO_4, 0.6 mM Na_2HPO_4, 1.2 mM $MgSO_4$, 12 mM $NaHCO_3$, 10 mM $KHCO_3$, 10 mM HEPES, 30 mM taurine, 5.5 mM glucose, 10 mM butanedione monoxime
Digestion Buffer	0.042 mg/mL Liberase DH, 0.025 % (v/v) trypsin, 12.5 µM $CaCl_2$ in Perfusion Buffer
Stop Solution	10 % (v/v) FBS in Perfusion Buffer

2.6.8 Cell culture solutions and media for juvenile mouse endothelial cells

Solution	Composition
enzyme solution (in HBSS without Ca, Mg)	1 mg/mL Collagenase type 2, 0.1 mg/ml Deoxyribonuclease I
gelatine solution	0.5% gelatine in WFI
basal endothelial medium	DMEM (high glucose, with pyruvate), 20% (v/v) FBS, 25 mM HEPES, 1% (v/v) Penicillin-Streptomycin-Glutamine
full endothelial medium	DMEM (high glucose, with pyruvate), 20% (v/v) FBS, 25 mM HEPES, 1% (v/v) Penicillin-Streptomycin-Glutamine, 1% (v/v) MEM Non-Essential Amino Acids Solution, 0.8% (v/v) Endothelial Cell Growth Supplement/Heparin
ischemic medium	DMEM (no glucose, no glutamine, no phenol red), 0.5 % dialyzed serum, 20 mM 2-deoxy-D-glucose, 1 % (v/v) Penicillin-Streptomycin
RPMI medium without amino acids	RPMI 1640 medium without amino acids, 0.5 % dialyzed serum, 25 mM HEPES, 1 % (v/v) Penicillin-Streptomycin, pH=7.4
scratch starvation medium	DMEM (high glucose, with pyruvate), 2 % (v/v) FBS, 25 mM HEPES, 1 % (v/v) Penicillin-Streptomycin-Glutamine
serum-free medium	DMEM (high glucose, with pyruvate), 25 mM HEPES, 1 % (v/v) Penicillin-Streptomycin-Glutamine, 1 % (v/v) MEM Non-Essential Amino Acids Solution

3 Methods

3.1 Animal models

Animal procedures for this project were approved by the regional council of Baden-Württemberg and executed in accordance with the Animal Welfare Regulation Governing Experimental Animals.
All *in vivo* experiments were performed in 10- to 12-week-old male mice. SPAAR KO mice were a gift from P. P. Pandolfi and were generated using C57BL/6J mouse strain as previously described [10]. C57BL/6JRj were used as wild-type controls and were purchased from Janvier Labs.
Animals were kept and bred in the central animal facility of University of Heidelberg, in a temperature- and humidity-controlled environment with a 12:12-hour light-dark cycle and with *ad libitum* access to standard chow.

3.1.1 Genotyping SPAAR KO mouse

To isolate genomic DNA, ear clips from SPAAR KO mice were boiled in 50 µL Buffer A for 35 min at 95 °C. 50 µL Buffer B was then added and samples were thoroughly mixed. Compositions of these buffers are given in section 2.6.1.
PCR amplification was performed using primaAMP 2× PCR Master Mix Red, 800 nM *Spaar* genotyping forward primer, 800 nM *Spaar* genotyping reverse primer and 4 µL of ear clip extract. Cycling conditions were as follows: 2 min at 94 °C, followed by 35 cycles of 10 sec at 96 °C, 15 sec at 60 °C, 45 sec at 72 °C and a final 10 min extension at 72 °C.
PCR products were then digested with restriction enzyme SacII for 1 hour at 37 °C. Finally, digested PCR products were loaded on 1.5 % agarose gel with ethidium bromide, ran in 1× TAE buffer and visualized under a UV light. Expected amplicons are shown in **Figure 3.1**.

A

Spaar	genotype	fragment sizes [nt]
+/+	WT	362 + 227
+/-	HET	589 + 362 + 227
-/-	KO	589

B

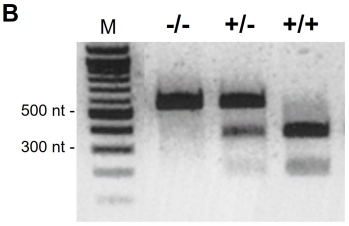

Figure 3.1: Digested PCR products after genotyping SPAAR KO mice. A. Expected fragment sizes after PCR amplification and SacII digestion vary depending on *Spaar* genotype. **B**. Digested PCR products are separated on an agarose gel. Indicated on top are *Spaar* genotypes as given in A. M: DNA marker, nt: nucleotides.

3.1.2 Echocardiography

Ultrasound echocardiography was performed to non-invasively monitor heart phenotype and function by using Vevo 2100 Imaging Platform (VisualSonics) coupled with MS550D transducer. Prior to measurements, mouse chest hair was removed and animals were anesthetized with isoflurane. Mice were then transferred onto a heating plate (37 °C) and fixed in supine position. Anesthesia was maintained throughout the measurement via a facial mask and a minimum dose of isoflurane (2.0 %). Recordings were taken in parasternal long axis in brightness mode (B-mode), while heart rate was kept between 400 and 500 beats per minute for consistent characterization of heart function. Recordings were analyzed in Vevo LAB software (VisualSonics).

3.1.3 Myocardial ischemia-reperfusion injury *in vivo*

Mice were randomly assigned to experimental groups (sham or I/R). Surgical procedures to induce I/R injury were carried out by Ole M. Schwerdt and Adrian Serafin. Prior to surgery, each mouse was injected with buprenorphine (0.1 mg/kg) and anesthetized in an induction chamber with 5 % isoflurane and oxygen flow rate of 1.0 L/min. The mouse was then placed in a supine position on a platform preheated to 37 °C and intubated. Artificial ventilation was provided with a rodent respirator (MiniVent model 845, Harvard Apparatus) with respiratory volume set to 260 µL/stroke and respiratory rate to 180 strokes/min. Anesthesia was maintained with 1.5 – 2 % isoflurane and oxygen flow rate of 1.0 L/min. Mouse chest hair was removed and 250 µL sterile 0.9 % saline solution was injected subcutaneously.
The surgical procedure was performed as described previously [117]. Briefly, approximately 1 cm long skin incision was made left of the sternal border in the direction of the left axilla. The pectoralis muscles were pierced with a 6-0 Prolene suture (Ethicon, 8695G) and fixed next to the mouse on the surgical platform to expose the ribs and intercostal muscles. An incision was made in the third intercostal space, another 6-0 Prolene suture was inserted around the third rib and attached on the surgical platform to spread the ribs and expose the heart. The pericardium was carefully dissected apart and 8-0 Prolene suture (Ethicon, EPM8739) was passed underneath the left anterior descending artery approximately 1 mm caudal to the left atrium and perpendicular to the vessel for a length of 2 - 3 mm. A piece of PE-10 tubing (Thermo Fisher Scientific, 22-204008), was placed over the LAD and ischemia was induced by knotting a suture over the tube and maintained for 60 min. Isoflurane content was reduced to 0.75 % - 1 % and the open thorax was covered with a moist gauze. Before the onset of reperfusion, isoflurane was again increased to 1.5 – 2 %. The PE-10 tube was removed and 8-0 Prolene knot was cut and removed as well. The chest cavity, muscle layers and skin were closed with continuous sutures. Isoflurane flow was ceased, but the mouse was ventilated with oxygen till it resumed with normal breathing. The intubation tube was then removed and the mouse was monitored until it was moving again. Post-operative analgesia was provided by subcutaneous injections of buprenorphine (0.1 mg/kg) every 6 hours for the next 24 hours. Sham surgeries were performed the same way to expose the heart, but the LAD was subsequently not ligated.

3.1.4 Measurement of cardiac troponin T

24 hours after the myocardial I/R surgery, animals were anesthetized with isoflurane and blood samples were taken from retrobulbar sinus using heparinized hematocrit capillaries (BRAND, 749311). Blood samples were then centrifuged at 1000 ×g for 10 min at 4 °C. Clear blood serum supernatant was collected and a small aliquot was diluted with PBS for a routine measurement performed in the central laboratory of Heidelberg University Hospital.

3.1.5 Preparation of mouse heart lysates

Mice were sacrificed by cervical dislocation, the chest was opened and the heart was carefully removed. Following a brief washing step in cold PBS, hearts were weighed and frozen in liquid nitrogen.
Frozen heart tissue was then homogenized in 700 µL RIPA buffer, in screw-cap tubes with ceramic beads, using Precellys tissue homogenizer (Bertin Technologies). To ensure complete lysis, tissue homogenates were passed several times through a syringe with a 25-gauge needle (Braun, 10162148) and incubated on ice for 5 min before centrifugation at 20,000 ×g for 15 min at 4 °C. Only the clear supernatant was retained for further analyses.

3.1.5.1 Immunoprecipitation of SPAAR from mouse heart lysates

To validate the translation of SPAAR on protein level, an immunoprecipitation procedure was optimized to detect it in WT mouse hearts and prove its depletion in SPAAR KO hearts. For each heart sample, 125 µL Dynabeads Sheep Anti-Rabbit IgG were washed twice with Lysis Buffer on a magnetic stand and coupled with 10 µg custom-made anti-SPAAR antibody for 1 hour at room temperature with gentle agitation. Afterwards, beads were washed twice with Lysis Buffer and resuspended again in a small volume of the same buffer.
Heart samples were homogenized as described above (section 3.1.5) and protein concentrations were measured using DC Protein Assay to adjust total protein content to 8.5 mg in 500 µL Lysis Buffer. These lysates were subsequently mixed with antibody-coupled beads in low retention tubes and rotated overnight at 4 °C. Next day, beads were separated on a magnetic stand and washed three times with Wash Buffer. Each wash step was performed for 5 min with rotation at room temperature. Precipitated proteins were eluted in 30 µL Lysis Buffer by boiling for 7 min at 95 °C. Following magnetic separation, only clear eluates were retained. Their protein concentration was measured and western blotting was performed as described in section 3.4.1, using equal amounts of total protein in inputs and eluates. Finally, membranes were probed with an anti-SPAAR antibody. Compositions of buffers used for immunoprecipitation are given in section 2.6.3.

3.1.6 Mouse heart fixation for histology

Mice were sacrificed by cervical dislocation and the chest was promptly opened to expose the beating heart. 1 mL of 100 mM $CdCl_2$ was carefully injected into the left ventricle to stop beating in diastole. A small incision was made in the right atrium to enable drainage and 5

mL of heparin (diluted in PBS to 100 U/mL) were slowly perfused through the right ventricle into the left ventricle to remove excess blood. 10 mL of 10 % formaldehyde (in PBS) were then perfused the same way to fix the heart tissue. Finally, the heart was excised, cleaned and kept in 4 % formaldehyde (in PBS) for 24 hours at 4 °C, before being transferred to fresh PBS for short-term storage or immediately embedded in paraffin.

3.2 Histological methods

3.2.1 Mouse heart embedding and sectioning

After fixation, mouse hearts were dehydrated and infiltrated with paraffin using HistoCore PEARL tissue processor (Leica). Next, hearts were embedded in paraffin blocks using HistoCore Arcadia H embedding station (Leica). Solid paraffin blocks were kept at -20 °C overnight before being cut into 10 µm histological sections using a microtome (RM2145, Leica). Heart sections were transferred onto glass slides and left to dry for at least 2 hours on a flattening table (HI1220, Leica) preheated to 37 °C. Tissue slides were then kept at room temperature until further use.
Before histological staining was performed, heart sections were deparaffinized in xylene (2× 5 min) and rehydrated in a descending ethanol series (2× absolute EtOH, 1× 95 % EtOH, 1× 70 % EtOH, 5 min each) before 3 final washes in ddH_2O (3 min each).

3.2.2 Hematoxylin and eosin staining

The staining was performed using Hematoxylin & Eosin fast staining kit as per manufacturer's instructions. Briefly, tissue slides were stained for 6 min in Hematoxylin, washed for 10 sec in tap water and differentiated for 10 sec in 0.1 % HCl. Next, slides were rinsed for 6 min in running tap water before they were stained for 30 sec in Eosin and finally washed for 30 sec in running tap water. Slides were then dehydrated (1 min in 95 % EtOH, 2× 1 min in absolute EtOH and 2× 3 min in xylene), mounted with Eukitt quick-hardening mounting medium and covered with a coverslip. Images of sections were taken on Axio Vert.A1 microscope (Zeiss) and quantification of hematoxylin-positive regions was done in ImageJ.

3.2.3 Masson Trichrome staining

The staining was performed using Masson Trichrome Stain Kit as per manufacturer's instructions. Briefly, tissue slides were incubated in preheated Bouin's Solution at 60 °C for 15 min, then washed with running tap water for 5 min and stained in Weigert's Iron Hematoxylin Solution for 5 min. Following another washing step in running tap water, slides were quickly rinsed with ddH_2O and incubated in Bieberich Scarlet-Acid Fuchsin solution for 5 min. After another rinse with ddH_2O, slides were placed in Working Phosphotungstic/Phosphomolybdic Acid Solution for 5 min and then directly transferred to Aniline Blue Solution for 5 min. Slides were incubated in 1 % acetic acid for 2 min and rinsed with ddH_2O. After dehydration series (1 min in 95 % EtOH, 2× 1 min in absolute EtOH and 2× 3 min in xylene), heart sections were

mounted with Eukitt quick-hardening mounting medium and covered with a coverslip. Images of sections were taken on Axio Vert.A1 microscope (Zeiss) and quantification of collagen deposits displaying purple-blue coloration was done in ImageJ.

3.2.4 Immunohistochemistry

Antigen retrieval was performed by boiling heart tissue slides in citrate buffer in the microwave (700 W for 3 min, followed by 450 W for 12 min). Slides were cooled at 4 °C for 20 - 30 min, washed with ddH_2O (2× 3 min) and equilibrated in TN buffer (2× 3 min). Blocking was achieved with TNB blocking solution for 60 min at room temperature. The following antibodies and reagents were diluted in TNB blocking solution and used for staining: anti-cardiac troponin T antibody (1:200), WGA–FITC (1:50) and isolectin IB4–Alexa Fluor 594 (1:100). They were pipetted onto heart sections and incubated in a humid chamber overnight at 4 °C. Next day, slides were washed with TN buffer (3× 3 min) and if required, incubated with a secondary antibody (diluted in TNB blocking solution) for 2 hours at room temperature. Used antibodies and their dilutions are listed in section 2.4. After washing with TN buffer (3× 3 min), sections were mounted with Vectashield Antifade Mounting Medium containing 5 µg/mL DAPI. Specimens were imaged on Axio Observer fluorescence microscope (Zeiss) and image analyses were carried out using ImageJ.

3.3 Cell culture methods

3.3.1 Isolation of neonatal rat cardiomyocytes

Neonatal rat cardiomyocytes (NRCMs) were isolated from hearts of 1- to 3-day-old Wistar rats (RjHan:WI, Janvier Labs). Following decapitation, hearts were excised and washed in cold HBSS. Atria and any remaining vessels were removed, and ventricles were minced into small pieces. These were then resuspended in Digestion Solution and incubated for 10 min at 37 °C with gentle agitation. The digest was stopped by the addition of Stop Solution, followed by gentle resuspension by pipetting to ensure proper tissue homogenization. Undigested material was allowed to settle and supernatant containing isolated cells was collected and kept on ice. Digestion was repeated as described for 6 - 8 times with the remaining undigested tissue. Collected supernatants were pooled together and centrifuged at 500 ×g for 15 min at 4 °C. The cell pellet was resuspended in NRCM Culture Medium and cells were counted using a hemocytometer. Next, cardiomyocytes were purified on a Percoll density gradient as described previously [118]. Cells were therefore collected by brief centrifugation at 250 ×g for 5 min at 4 °C and resuspended in 2 mL red ADS buffer (resulting in 20 - 30 million cells per mL). This cell suspension was carefully transferred on the top of Percoll gradient and centrifuged at 15,000 ×g for 30 min at room temperature in Eppendorf 5810R centrifuge using a swinging bucket rotor and no deceleration brake. After centrifugation, pure NRCM cell population was concentrated in the layer between the lower red ADS layer and the middle clear ADS layer. NRCMs were carefully collected and washed twice with ADS buffer, and finally resuspended in NRCM Culture Medium. Compositions of used solutions and media are given in section 2.6.6.

After determining the cell count, NRCMs were plated in NRCM Culture Medium at a cell density of approximately 50,000 cells/cm^2. Plastic culture plates were precoated with 5 µg/mL fibronectin for one hour before plating. Isolated cultures were kept in a humidified incubator at 37 °C and 5 % CO_2. The following day, cells were washed with PBS and cultured in NRCM Treatment Medium for subsequent experiments.

3.3.1.1 Transfection of NRCMs with small interfering RNA

24 hours after isolation and plating of NRCMs, cells were transfected with small interfering RNA (siRNA) to achieve a transient knockdown of SMIM4. For that purpose, control scrambled siRNA (siScr; Thermo Fisher Scientific, 4390843) or siRNA targeting SMIM4 transcript (siSMIM4; Thermo Fisher Scientific, s166843) were delivered into NRCMs using HiPerfect transfection reagent according to the manufacturer's protocol. Briefly, 5 µL of HiPerfect transfection reagent was used per well of a 12-well plate and final siRNA concentration was kept at 100 nM. NRCMs were incubated with transfection mixes overnight, then the medium was replaced by fresh NRCM Treatment Medium and cells were cultured for another 48 hours to ensure efficient knockdown.

3.3.1.2 Treatment of NRCMs with phenylephrine

48 hours after transfection with siRNA, NRCMs were treated with alpha-1 adrenergic receptor agonist phenylephrine (PE), which simulates pathological hypertrophy by inducing NRCM growth. A fresh stock solution of PE (1 mg/mL in PBS) was prepared before each experiment and diluted in NRCM Treatment Medium to a final concentration of 50 µM. NRCMs were kept in PE-containing medium for 24 hours.

3.3.2 Isolation of adult mouse cardiomyocytes

Adult mouse cardiomyocytes were isolated from 10-week-old mice as previously described [119]. Briefly, mice were sacrificed by cervical dislocation and the heart was carefully cut out before being placed onto a cannula and mounted on a Langendorff perfusion system preheated to 37 °C. The heart was first washed with Perfusion Buffer for 3 min before being digested with Digestion Buffer for 10 min. The heart was subsequently removed from the Langendorff apparatus and transferred onto a dish containing the digestion solution with cells. The tissue was minced into small pieces and the digest was stopped by the addition of Stop Solution, followed by gentle resuspension by pipetting to ensure proper tissue homogenization. The resulting suspension was passed through a 100 µm cell strainer and cardiomyocytes were allowed to settle for 8 min. The supernatant (containing dead cells and non-myocytes) was discarded and the remaining cardiomyocytes were washed twice with cold PBS. Compositions of used solutions and media are given in section 2.6.7.
One portion of isolated cardiomyocytes was fixed with 3 % glutaraldehyde in PBS, plated in 12-well culture plates and used for brightfield microscopy to measure their cell sizes. Pictures were taken with Axio Vert.A1 microscope (Zeiss) and cell length and width of individual cells were measured in ImageJ.

The other half of isolated cardiomyocytes was briefly centrifuged, the remaining supernatant was removed and the cell pellet was resuspended in QIAzol Lysis Reagent. RNA was subsequently isolated as described in section 3.4.2.

3.3.3 Isolation of juvenile mouse endothelial cells

Primary juvenile endothelial cells were isolated from hearts of 8- to 10-day-old mice as previously described [120]. Briefly, following decapitation and opening of the chest cavity, heparin (300 U/mL in PBS) was injected into the left ventricle. The heart was excised, washed in DMEM and cleaned from atria and large vessels. Hearts were then transferred into enzyme solution, minced into small pieces and transferred into tubes, which were incubated for 45 min at 37 °C with rotation. Following resuspension by pipetting to ensure proper tissue homogenization, the suspension was passed through a 70 µm cell strainer and the digest was stopped with basal endothelial medium. Tubes were centrifuged at 400 ×g for 8 min at 4 °C, the pellet was resuspended in basal endothelial medium and CD31-coated Dynabeads Sheep Anti-Rat IgG were added. Tubes were then rotated for 30 min at room temperature, before the beads with bound cells were washed five times with basal endothelial medium on a magnetic rack. Finally, beads were resuspended in full endothelial medium and plated on gelatine-precoated T75 flasks and kept in a humidified incubator at 37 °C and 5 % CO_2. When confluent (approximately three days later), isolated cells were trypsinized, resuspended in basal endothelial medium and CD102-coated Dynabeads Sheep Anti-Rat IgG were added. Tubes were rotated for 30 min at room temperature and then the beads with bound endothelial cells were washed five times with basal endothelial medium on a magnetic rack. Finally, beads were resuspended in full endothelial medium and plated on gelatine-precoated T75 flasks. Compositions of used solutions and media are given in section 2.6.8 and used antibodies are listed in section 2.4.
When confluent (approximately three days later), purified endothelial cells were trypsinized and counted. They were plated in full endothelial medium at a cell density of approximately 20,000 cells/cm^2 on gelatine-coated culture plates for subsequent experiments.

3.3.3.1 *In vitro* treatments of endothelial cells

For amino acid starvation, full endothelial medium was removed from cells, followed by two washes with PBS and subsequent incubation in RPMI medium without amino acids for 3 hours. Next, cells were stimulated by the addition of amino acid mixture (1× MEM Amino Acids Solution, 1× MEM Non-essential Amino Acids Solution, 1× L-glutamine and 150 mM NaOH to neutralize pH) for 1 hour.
For insulin starvation, full endothelial medium was removed from cells, followed by two washes with PBS and subsequent incubation in serum-free medium for 3 hours. Next, cells were stimulated with insulin (Insuman Rapid 40 I.E./mL, 1.4 µM) for 1 hour.
6bK, pharmacological inhibitor of IDE, was dissolved in ddH_2O and then diluted in the required culture medium to a final concentration of 500 nM.
For simulated ischemia, full endothelial medium was removed from cells, followed by two washes with PBS. Next, cells were kept in ischemic medium in a hypoxia incubator (37 °C, 5 % CO_2,

1.0 % O2) for 6 hours. Following the treatment, the conditioned medium was collected and kept at -20 °C for further experiments.

3.3.4 Adenovirus production and viral transduction

Entry clones were ordered from BioCat in pDONR221 backbone with the following coding DNA sequences (CDS): SPAAR (*M. musculus* CDS) with FLAG tag, TMEM192 (*H. sapiens* CDS) with 3× HA tag and TFEB (*M. musculus* CDS) with V5 tag.
These were subcloned into pAd/CMV/V5-DEST expression vector (Invitrogen, V49320) by performing the LR reaction using Gateway LR Clonase II Enzyme mix per manufacturer's instructions. The obtained construct was transformed into One Shot TOP10 Chemically Competent E. coli cells (Thermo Fisher Scientific, C404010) as indicated in the product information sheet, spread onto plates with LB-agar and ampicillin (100 µg/mL) and incubated at 37 °C overnight. Next day, colonies were picked and plasmids were isolated using peqGOLD Plasmid Miniprep Kit I. Correct insert sequence was confirmed by Sanger sequencing (Mix2Seq Service, Eurofins). Glycerol stock of initial bacterial transformation was then used to prepare a liquid culture and expression plasmid was isolated using HiSpeed Plasmid Midi Kit.
6 µg of each expression plasmid was linearized with restriction enzyme PacI for 3 hours at 37 °C, before the enzyme was inactivated for 20 min at 65 °C. The linearized plasmid was transfected into HEK293A cells (10 cm plate) using PEI transfection reagent. Two days later, the cells were split into three 10 cm plates and cultured in antibiotic-free medium. Plates were inspected every day under a microscope and were harvested once 30 % of cells detached. Following 3 freeze-thaw cycles, the lysate was used to reinfect HEK293A cells (T175 flask), which were harvested once 70 % of cells detached. After 3 freeze-thaw cycles, benzonase was added to the lysate and tubes were incubated for 90 min in a heated water bath (37 °C). Finally, tubes were centrifuged at 3700 rpm for 40 min at 4 °C, the supernatant was collected and aliquoted for further use.
To determine the number of adenoviral particles, a small aliquot of each virus was diluted in 0.1 % SDS and absorbance at 260 nm was measured on NanoDrop Lite spectrophotometer (Thermo Fisher Scientific). The number of viral particles per mL was calculated as follows: A260 × dilution factor × 1.1 × 10^{12} particles.
Before viral transduction, cell culture medium was removed and then replaced with fresh one. Adenoviruses were added directly into the medium and plates were incubated for 24 hours to achieve overexpression. Control samples were transduced with a control adenovirus (without a CDS), using the same number of viral particles.

3.3.5 Cell-based functional assays

3.3.5.1 Proliferation assay

The assay was performed using CyQUANT NF Cell Proliferation Assay Kit as per manufacturer's instructions. Primary endothelial cells were seeded in black 96-well culture plates with clear bottom (5000 cells per well). Shortly before the measurement, 1× dye binding solution was prepared by diluting HBSS buffer and CyQUANT NF dye reagent in an appropriate volume of ddH_2O. Culture medium was then removed from cells, 100 µL of 1× dye binding solution was

pipetted in each microplate well and plates were returned to the incubator for 1 hour. Finally, fluorescence intensity (excitation: 485 nm, emission: 530 nm) was measured at 25 different points in each well using a microplate reader (Victor Nivo Microplate Reader, PerkinElmer). These values were averaged and corrected for background absorbance values (blank) created by the dye solution.

3.3.5.2 MTT assay

The assay was performed in transparent 96-well culture plates with 15,000 NRCMs or 5000 primary endothelial cells seeded per well. After the desired experimental treatments were performed, 30 µL of MTT stock solution (5 mg/mL in PBS) was added to each well and plates were returned to the incubator for 4 hours. Afterwards, 100 µL of solubilization solution (10 % SDS with 0.01 M HCl) was added per well and the solution was carefully mixed by pipetting. Plates were kept in the incubator overnight and the following day, the absorbance at 570 nm was measured on a spectrophotometer (Victor Nivo Microplate Reader, PerkinElmer). Measured absorbances were corrected for background absorbance values (blank) created by the medium and solutions.

3.3.5.3 Puromycin incorporation assay

After the desired experimental treatments were performed, Puromycin was added to the cell culturing medium to a final concentration of 0.5 µg/mL. Cultured cells were returned to the incubator for 30 min before they were harvested in RIPA buffer. Puromycin-labelled peptides were then detected on a western blot with an anti-puromycin antibody to provide information about global protein synthesis.

3.3.5.4 Scratch assay

Primary endothelial cells were seeded on 24-well plates (50,000 cells per well), cultured until confluent and then starved in scratch starvation medium overnight to stop cell proliferation. Next day, a 1000 µL pipette tip was used to make a scratch across each well. After washing away dead cells with PBS, fresh scratch starvation medium was added again. Plates were returned to the incubator and images of the scratch were taken on Axio Vert.A1 microscope (Zeiss) at the following time points: 0, 5, 9 and 24 hours. Scratch size was quantified in ImageJ.

3.3.5.5 Tube formation assay

Transparent 96-well culture plates were kept on ice while 50 µL of Matrigel Growth Factor Reduced Basement Membrane Matrix was pipetted in each well. Plates were then incubated at 37 °C for 1 hour to enable polymerization of Matrigel. Meanwhile, primary endothelial cells were trypsinized, counted and 20,000 cells were seeded per well in freshly prepared and fully supplemented Endothelial Cell Growth Medium 2. Plates were returned to the incubator and images of the tubular network were taken on Axio Vert.A1 microscope (Zeiss) after 7 hours. Images were quantified using Angiogenesis Analyzer in ImageJ [121].

3.3.6 Cell lysate preparation

Cultured cells were washed with PBS and then harvested in RIPA buffer with added protease and phosphatase inhibitors. Cell lysates were collected by carefully scraping cell culture dishes and subsequently centrifuged at 20,000 xg for 15 min at 4 °C. Only soluble fractions (supernatants) were retained for further use.

3.4 Molecular and biochemical methods

3.4.1 Western blot

Protein concentration of lysates was measured in 96-well microplates using DC Protein Assay according to the manufacturer's protocol. 15 - 30 µg of total protein per sample were mixed with Laemmli Sample Buffer (with added 10 % 2-mercaptoethanol) and boiled for 5 min at 95 °C. Denatured samples were then subjected to SDS-PAGE on 4-12 % Bis-Tris polyacrylamide precast gels (Bio-Rad Laboratories, 3450124/3450125) using XT MOPS or XT MES running buffer and 200 V. Proteins were transferred onto PVDF membrane (Immobilon P with pore size 0.45 µm, Merck-Millipore, IPVH00010) with a wet transfer (100 V, 35 min). Membranes were subsequently blocked with 5 % milk solution (prepared in TBST) for 1 hour at room temperature and then incubated with primary antibodies in 3 % BSA solution (prepared in TBST) overnight at 4 °C. The following day, membranes were washed three times with TBST (10 min each wash) and incubated with the corresponding peroxidase-conjugated secondary antibodies in 5 % milk for 1 hour at room temperature. Finally, membranes were washed three more times with TBST (10 min for each wash) before being visualized using WesternBright ECL and Fusion FX Imaging System (Vilber). Used antibodies and their dilutions are listed in section 2.4 and buffer compositions are given in section 2.6.5.
Following chemiluminescent detection of proteins, membranes were also stained with Ponceau S for 10 min at room temperature with gentle shaking. After several washes with ddH_2O to remove unwanted background, membranes were left to air-dry before scanning.
Acquired images and scans were then used for semi-quantitative densitometry analysis in ImageJ. Densitometry values for proteins of interest were normalized to the values of loading control (Ponceau S-stained membranes).

3.4.2 RNA isolation

Cultured cells were harvested in RNA Lysis Buffer and isolation of total RNA was performed by using Quick-RNA MiniPrep Kit per manufacturer's instructions (including on-column DNase I digest).
RNA isolation from mouse heart tissue was done by first homogenizing the organs as described above in section 3.1.5. 100 µL of the heart lysate was then thoroughly mixed with 1 mL of QIAzol Lysis Reagent and 200 µL of chloroform and centrifuged at 12,000 ×g for 15 min at 4 °C. Upper aqueous phase was carefully collected and mixed with 1 volume of isopropanol and isolation of total RNA was completed by using Quick-RNA MiniPrep Kit according to the manufacturer's protocol (including on-column DNase I digest).

RNA concentration and purity were measured on NanoDrop Lite spectrophotometer (Thermo Fisher Scientific).

3.4.3 Reverse transcription and quantitative real-time PCR

700 – 1000 ng of total RNA were subjected to reverse transcription with First Strand cDNA Synthesis Kit and oligo-dT primers were used to specifically detect mature polyadenylated mRNAs and produce a more consistent qRT-PCR signal. The resulting cDNA samples were diluted ten- to twenty-fold with nuclease-free H_2O and used for quantitative real-time PCR (qRT-PCR) analysis with iTaq universal SYBR Green Supermix. Each reaction consisted of 5 µL iTaq universal SYBR Green Supermix, 300 nM forward primer, 300 nM reverse primer and 3 µL diluted cDNA. Samples were pipetted into 96- or 384-well microplates and analyzed with QuantStudio5 real-time PCR System (Thermo Fisher Scientific).
The sequences of primers used are given in section 2.5. Each primer pair was tested prior to use to ensure an amplification efficiency between 90 % and 110 %. Additionally, melting curves were carefully inspected and agarose gel electrophoresis was performed to verify that a single gene-specific amplicon of the correct size was produced. C_t values of genes of interest were normalized to the geometric mean of at least two reference genes (*B2m*, *Btub*, *Hprt*, *Sdha*). Relative abundances and changes in expression were calculated using the $2^{-\Delta\Delta CT}$ method.

3.4.4 Immunofluorescence

Cells were cultured in 12-well plates that contained glass coverslips. After the desired experimental treatments were performed, cells were washed with PBS and fixed with 4 % paraformaldehyde for 10 min at room temperature. Following another washing step with PBS, cells were permeabilized with 0.3 % Triton X-100 (in PBS) for 10 min at room temperature. Blocking was performed for 1 hour in 10 % horse serum at room temperature. Coverslips were then incubated overnight at 4 °C with primary antibodies in 10 % horse serum. Next day, coverslips were washed three times with PBS (5 min each wash) and incubated with secondary antibodies in 10 % horse serum for 1 hour at room temperature in the dark. Used antibodies and their dilutions are listed in section 2.4. After 3 final washing steps with PBS (5 min each wash), coverslips were mounted on glass slides using Vectashield Antifade Mounting Medium containing 5 µg/mL DAPI. Specimens were then imaged on Axio Observer fluorescence microscope (Zeiss) or SP8 Lightning confocal microscope (Leica). Microscopy image analyses such as colocalization analysis and cell size measurements were carried out using ImageJ.

3.4.5 Flow cytometry for assessing apoptosis and necrosis *in vitro*

The assay was performed using Dead Cell Apoptosis Kit with annexin V-FITC and PI for flow cytometry per manufacturer's instructions. Shortly, NRCMs were kept in endothelial conditioned medium (full or ischemic) while apoptosis was induced with 200 µM H_2O_2 for 4 hours. Following the treatment, medium was retained to collect detached cells, while the rest were trypsinized, pooled with the medium from the respective sample and collected by centrifugation

at 500 ×g for 5 min at 4 °C. After a washing step with PBS, cell pellets were resuspended in staining solution (annexin-binding buffer with annexin V-FITC and propidium iodide) and incubated in the dark for 15 min at room temperature. Subsequently, stained cells were analyzed by flow cytometry using BD FACSVerse Cell Analyzer (BD Biosciences) and data analysis was done in BD FACSuite software.

3.4.6 ELISA for secreted IDE levels

To determine IDE levels in mouse blood serum samples, Ide ELISA Kit Mouse was used as per manufacturer's instructions. Briefly, 100 µL of serially titrated IDE standard, samples (diluted 1:10) and blank were pipetted into wells of IDE microplate. Covered plate was incubated at 37 °C for 2 hours, then the liquid in the wells was discarded and 100 µL of 1× Biotinylated Ide Detector Antibody was added to each well. Next, covered plate was incubated at 37 °C for 1 hr, the liquid in the wells was discarded and the wells were washed 3 times with 1× Wash Buffer (200 µL per well, 1 min incubation at room temperature for each wash). 100 µL of 1× Avidin-HRP Conjugate was pipetted into each well next, covered plate was incubated at 37 °C for 1 hour and then washed again 5 times as described. After washing, 90 µL of TMB Substrate was added per well and the plate was incubated at room temperature in the dark for 5 min before 50 µL of Stop Solution was added to each well. The absorbance at 450 nm was measured on a spectrophotometer (Victor Nivo Microplate Reader, PerkinElmer) with background correction set to 540 nm. Measured absorbances at 450 nm were corrected for microplate background and background created by the solutions (blank). IDE standard was used to plot a standard curve and linear regression was used to calculate IDE concentration in the tested samples.

3.4.7 Ribosome profiling of human heart samples

3.4.7.1 Human primary material

The use of myocardial tissue samples was approved by the local ethics committee after obtaining written informed consent from patients and the Medical Faculty of Heidelberg. Biopsy specimens were obtained from the apical part of the free left ventricular wall from patients diagnosed with ischemic cardiomyopathy (sample 'hHeart1') or dilated cardiomyopathy (samples 'hHeart5' and 'hHeart8'). The samples were briefly rinsed in 0.9 % NaCl, snap-frozen in liquid nitrogen and stored at –80 °C until use.

3.4.7.2 Library preparation

Initial RNA quality check was performed using Agilent RNA 6000 Nano Kit and 2100 Bioanalyzer Instrument (Agilent) to ensure the starting material had a RIN value of at least 8.0.
Library preparation was done using TruSeq Ribo Profile (Mammalian) Library Prep Kit with selected modifications reported in McGlincy and Ingolia, 2017 [122]. Each piece of frozen heart tissue was homogenised in 1 mL precooled Mammalian Lysis Buffer using Precellys 24 Tissue Homogenizer (Bertin Technologies). Remaining tissue clumps were dissociated by multiple passes through a syringe with a 25-gauge needle (Braun, 10162148) and samples were incubated

on ice for 5 min to ensure complete lysis before centrifugation at 20,000 ×g for 10 min at 4 °C. Per sample, 400 – 500 µL of lysate (supernatant) was collected and used to generate ribosome footprints. A small aliquot of the lysate was diluted 1:10 to determine optical density at 260 nm (OD260) using a NanoDrop Lite spectrophotometer (Thermo Fisher Scientific). 7.5 U RNase I was then added per 1 unit of OD260 to each lysate and samples were rotated for 45 min at room temperature. Digestion was stopped by adding 15 µL SUPERase In RNase Inhibitor per sample and lysates were loaded onto pre-washed MicroSpin S-400 HR columns (Cytiva, 27-5140-01). Following centrifugation at 600 ×g for 4 min at room temperature, flow-through was collected and ribosome footprints were isolated using RNA Clean & Concentrator-25 kit. The obtained footprints were subsequently cleaned up on a denaturing 15 % TBE-urea gel (Thermo Fisher Scientific, EC6885BOX). The gel was stained with SYBR Safe DNA Stain and slices corresponding to the region between 28 nt and 30 nt were excised, transferred to gel breaker tubes (IST Engineering, 3388-100) and fragmented by centrifugation at 12,000 ×g for 2 min at room temperature. RNA was eluted using 450 µL of extraction buffer (0.45 M ammonium acetate and 0.045 % SDS) with gentle agitation at 4 °C overnight. Next day, gel pieces were removed by centrifugation at 2000 ×g for 3 min at room temperature. Supernatant was retained and RNA was precipitated in 700 µL 100 % isopropanol and 2 µL GlycoBlue for 2 hours at –20 °C. Following centrifugation, washing with 80 % ethanol and air-drying, the pellet was resuspended in 20 µL nuclease-free water. End repair was performed exactly as indicated in the TruSeq Ribo Profile protocol. For 3' adapter ligation, 0.5 µL TruSeq Ribo Profile 3' Adapter was added to each sample along with Ligation Master Mix and incubated for 2 hours at room temperature. 2 µL TruSeq Ribo Profile Adapter Removal Enzyme was then added to each reaction and incubated for 2 hours at 30 °C. RNA was purified using RNA Clean & Concentrator-5 kit and ribosomal RNA depletion was achieved using Ribo-Zero Gold rRNA Removal Kit for sample 'hHeart1' and RiboCop rRNA Depletion Kit for samples 'hHeart5' and 'hHeart8', as described in the respective protocols. rRNA-depleted RNA was then purified using RNA Clean & Concentrator-5 kit. Reverse transcription was performed exactly as indicated in the TruSeq Ribo Profile protocol and cDNA was concentrated using RNA Clean & Concentrator-5 kit before cleaned up on a denaturing 10 % TBE-urea gel (Thermo Fisher Scientific, EC6875BOX). The gel was stained with SYBR Safe DNA Stain and slices corresponding to the region between 70 nt and 80 nt were excised. cDNA was extracted from the gel and precipitated as already described above. Next, cDNA was circularized as indicated in the TruSeq Ribo Profile protocol and amplified in 12 PCR cycles to obtain a 150-nt band, corresponding to Ribo-seq libraries. Quality check of these libraries was done using Agilent High Sensitivity DNA Kit and exact concentration was measured by Qubit dsDNA High Sensitivity Assay Kit.

3.4.7.3 Sequencing and bioinformatical data processing

The obtained Ribo-seq libraries were multiplexed to a final concentration of 4 nM and sequenced on HiSeq 2000 instrument (Illumina) to an average depth of about 100 million raw reads. Our collaborator Etienne Boileau subsequently processed these data with *Rp-Bp* pipeline [112]. Briefly, adapters and low-quality reads were removed using flexbar package and reads aligned to a custom ribosomal index using bowtie2 were eliminated as well. Remaining reads were aligned to the

human genome (GRCh38) with a splice-aware aligner STAR and multimapping reads were discarded. Periodic fragment lengths and ribosome P-site offsets were determined for each sample using automatic Bayesian selection. Only periodic reads were used to construct ORF profiles (based on Ensembl v94), which were used for calculations to infer whether they undergo active translation or not.

3.4.8 Ribosome profiling of mouse heart samples

3.4.8.1 Collecting mouse heart samples for ribosome profiling

Three WT and three SPAAR KO mice were sacrificed by cervical dislocation, the chest was opened and the heart was carefully removed. Following a brief washing step in cold PBS with 0.1 mg/mL cycloheximide, atria and right ventricle were discarded and the remaining left ventricle was frozen in liquid nitrogen.

3.4.8.2 Library preparation

Each piece of frozen left ventricular tissue was homogenised in 0.5 mL precooled Mammalian Lysis Buffer using PowerLyzer 24 Homogenizer (Qiagen). Remaining tissue clumps were dissociated by multiple passes through a syringe with a 25-gauge needle (Braun, 10162148) and samples were incubated on ice for 5 min to ensure complete lysis before centrifugation at 20,000 ×g for 10 min at 4 °C. Per sample, 200 – 300 µL of lysate (supernatant) was collected and used to generate ribosome footprints as already described in section 3.4.7.2. Following the purification of footprints using RNA Clean & Concentrator-25 kit, ribosomal RNA depletion was achieved using riboPOOL ribo-seq kit. rRNA-depleted RNA was purified using RNA Clean & Concentrator-5 kit and loaded on a denaturing 15 % TBE-urea gel (Thermo Fisher Scientific, EC6885BOX). The gel was stained with SYBR Safe DNA Stain and slices corresponding to the region between 28 nt and 30 nt were excised and processed as described in section 3.4.7.2 to extract and precipitate RNA. End repair was performed using T4 Polynucleotide Kinase with 1× supplied T4 Polynucleotide Kinase Reaction Buffer for 1 hour at 37 °C. Following purification with RNA Clean & Concentrator-5 kit, library preparation was done using NEXTFLEX Small RNA-seq v3 Kit per manufacturer's instructions. Following reverse transcription and bead cleanup, a small aliquot of cDNA was diluted and used for a qRT-PCR reaction to pick the number of PCR amplification cycles based on the obtained C_t value. After PCR amplification and bead cleanup, the size distribution of libraries was determined using Agilent High Sensitivity DNA Kit and exact concentration was measured by Qubit dsDNA High Sensitivity Assay Kit.

3.4.8.3 Sequencing and bioinformatical data processing

The obtained Ribo-seq libraries were multiplexed to a final concentration of 4 nM and sequenced on HiSeq 2000 instrument (Illumina) to an average depth of about 50 million raw reads. Our collaborator Etienne Boileau subsequently processed these data with *Rp-Bp* pipeline, as already described in section 3.4.7.3. Reads were aligned to the mouse genome (GRCm39) and ORF profiles were constructed based on Ensembl v107.

3.4.9 Whole transcriptome sequencing of primary cardiac endothelial cells

Cardiac endothelial cells were isolated and purified from WT and SPAAR KO juvenile mice as described in section 3.3.3. After two days of culturing, cells were harvested and total RNA was isolated as described in section 3.4.2. RNA quality check was performed on 5300 Fragment Analyzer System (Agilent), which resulted in RQN value of 10.0 for all samples (4 technical replicates for WT cells and 4 technical replicates for SPAAR KO cells). 1 µg of total RNA per sample was then sent for transcriptome sequencing at BGI Genomics. Following a proprietary library construction protocol and DNBSEQ Sequencing System, approximately 22 million clean paired-end reads were obtained per library.

Sequencing data processing and analysis was done using a local instance of Galaxy [123]. Reads aligned to a custom ribosomal index using bowtie2 [124] (v2.3.4.1) were eliminated. Remaining reads were aligned to the mouse genome (GRCm38) with a splice-aware aligner STAR [125] (v2.7.8.a). Only uniquely mapping reads were used for generating count matrices using featureCounts [126] (subread package v2.0.1 and Ensembl v97). Differential gene expression analysis was done using R (v4.0.2) and DESeq2 [127] (v1.28.1) with a code adapted from Chothani *et al.* [128] and modified to analyze only mRNA count matrices. Gene ontology enrichment analysis was done using the Database for Annotation, Visualization and Integrated Discovery [129] (DAVID v2022q4, https://david.ncifcrf.gov/).

3.4.10 Secretome analysis of primary cardiac endothelial cells

Isolated WT and SPAAR KO endothelial cells were grown on 10 cm culture dishes to reach 90 % confluency, with 3 technical replicates used per genotype. Full endothelial medium was then removed, cells were washed with PBS and incubated in depletion medium (DMEM for SILAC, 10 % dialyzed fetal bovine serum, 4mM L-glutamine, 100 mg/L primocin) for 30 min. Next, the depletion medium was removed and replaced with labelling medium (depletion medium with 0.1 mM azidohomoalanine (AHA) and stable isotope labelled amino acids; 146 µg/mL [4,4,5,5-D_4] L-lysine and 84 µg/mL [$^{13}C_6$] L-arginine for intermediate form or 146 µg/mL [$^{13}C_6$,$^{15}N_2$] L-lysine and 84 µg/mL [$^{13}C_6$,$^{15}N_4$] L-arginine for heavy form). The cells were returned to the incubator for 24 hours to allow stable isotope labelling of newly synthesized proteins. After the 24-hour incubation, media were collected by pooling intermediate and heavy samples of each replicate and centrifuged at 1000 ×g for 5 min. Supernatants were retained and cOmplete Protease Inhibitor was added before freezing them at -80 °C.

Our collaborator Tim C. Kuhn then processed the samples and used them for proteomics data acqusition and analysis as previously described [130]. Briefly, newly synthesized proteins with incorporated AHA were enriched using Click-iT Protein Enrichment Kit, and trypsin was used for on-bead digestion of proteins at 37 °C overnight, followed by desalting. Single-shot mass spectrometry was performed on Q Exactive Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher Scientific). Raw data were analyzed by MaxQuant, and Andromeda search engine was used to search the obtained spectra in *M. musculus* database in UniProt. Only proteins detected in more than 50 % of replicates were used for further analysis. Statistical analysis was performed using R and limma package to determine differentially secreted proteins. I performed all subsequent

analyses, such as functional annotation enrichment by using the Database for Annotation, Visualization and Integrated Discovery [129] (DAVID v2022q3, https://david.ncifcrf.gov/). Protein interaction network was obtained by performing a search with differentially secreted proteins in STRING database [131] (https://string-db.org/).

3.5 Figures

Parts of figures were created with BioRender (https://www.biorender.com/) and illustrations provided by Servier Medical Art (https://smart.servier.com/).

3.6 Data analyses

3.6.1 Identification of annotated small proteins enriched in the heart

Data filtering for identification of small cardiac-enriched proteins was done in R (v3.6.1) and RStudio (v1.2.5019-6). Table with expression data for seven organs throughout their development in human was downloaded from https://apps.kaessmannlab.org/evodevoapp/. A gene was classified as heart-enriched if its postnatal mean expression in the heart was at least 2-fold higher than in other tissues. biomaRt package (v2.42.0) was then used to retrieve amino acid sequences of all known protein isoforms (annotated in Ensembl v98) that are encoded by heart-enriched genes. From this list, small proteins were identified by applying a threshold of maximum 100 amino acids. Finally, only those small proteins were retained that were previously detected as translated in human hearts (listed in Table S1 from [41]).

3.6.2 Statistics

All statistical analyses were performed using R (v4.0.2) and RStudio (v2022.02.3, build 492). For each experiment, normal distribution of data was first tested by Shapiro-Wilk's test, followed by F-test to check for equality of variances. To compare means of two independent groups, unpaired T-test was used as a parametric test and Wilcoxon test was used as a non-parametric alternative. For comparing means of more than two groups, ANOVA was used as a parametric test followed by Tukey's test. Kruskal-Wallis test was used as a non-parametric alternative to ANOVA, followed by multiple pairwise comparison with Benjamini-Hochberg correction for multiple testing. Calculated P-values were symbolized in figures as follows: $0 < *** < 0.001 < ** < 0.01 < * < 0.05 < \# < 0.1 < \text{ns} < 1$. Whenever possible, results were plotted as individual observations with indicated mean and error bars representing standard deviation.

4 Results

4.1 Identifying novel cardiac microproteins

The aim of the first part of the project was to pinpoint functional, cardiac-enriched microproteins with the focus on novel, previously unannotated peptides. The most promising targets were then selected for further characterization in the heart.

4.1.1 Novel cardiac sORFs in mouse are poorly conserved in the human genome

Preliminary dataset on mouse heart translatome (**Figure 1.9**) indicated that there are thousands of sORFs present in (diseased) myocardium with the potential to be translated. To validate their protein-coding capacity, further evidence was sought in published proteomics data. Sequences of sORFs were therefore compared with unique peptides originating from healthy and hypertrophic mouse hearts that were analysed by liquid chromatography-tandem MS [132]. Notably, only about 20 % of these sORFs (1600 out of 8400 predicted) were successfully matched with a peptide detected by MS. One half of these peptides were derived from canonical, already annotated sORFs, and less than 200 originated from sORFs localized in presumable noncoding RNAs and novel transcripts.

These novel mouse sORFs with peptide evidence were then analysed for conservation in *H. sapiens*. Alignments showed that for approximately 40 candidates there is no comparable region in the human genome. Most sORFs were aligned to multiple shorter regions in *H. sapiens*, which did not correspond to an ORF defined by a start codon and in-frame stop codon. Some sORFs were aligned to annotated human protein-coding genes, which suggests these cases might be mouse pseudogenes or imperfect gene annotations in mouse.

Altogether these findings indicate that the majority of predicted sORFs appear to be mouse-specific and do not have a conserved translation initiation site in the human genome, making them less relevant for further characterization.

4.1.2 Analysis of human cardiac translatomes reveals rare cases of conserved translated microproteins

To circumvent the problem of finding mouse-specific novel sORFs and to advance clinical relevance, I prepared Ribo-seq libraries from human heart samples to search for translated and functional sORFs. For that purpose, I have successfully adjusted some of the critical steps in the Ribo-seq protocol, such as prolonged RNase I digest and improved rRNA depletion protocol. This enabled me to prepare 3 libraries from human diseased heart samples that contained up to 24 million usable reads each (**Figure 4.1 A**). These footprints showed a clear trinucleotide periodicity (**Figure 4.1 B**) and were analyzed by Evgeny Leushkin to identify translated sORFs, followed by conservation analysis. This revealed a microprotein of 27 amino acids, whose sequence is conserved across placental mammals (**Figure 4.1 C, D**). It is located on a novel transcript that is expressed and translated in human hearts, but unfortunately not in the mouse heart, which critically limits available experimental models for further characterization.

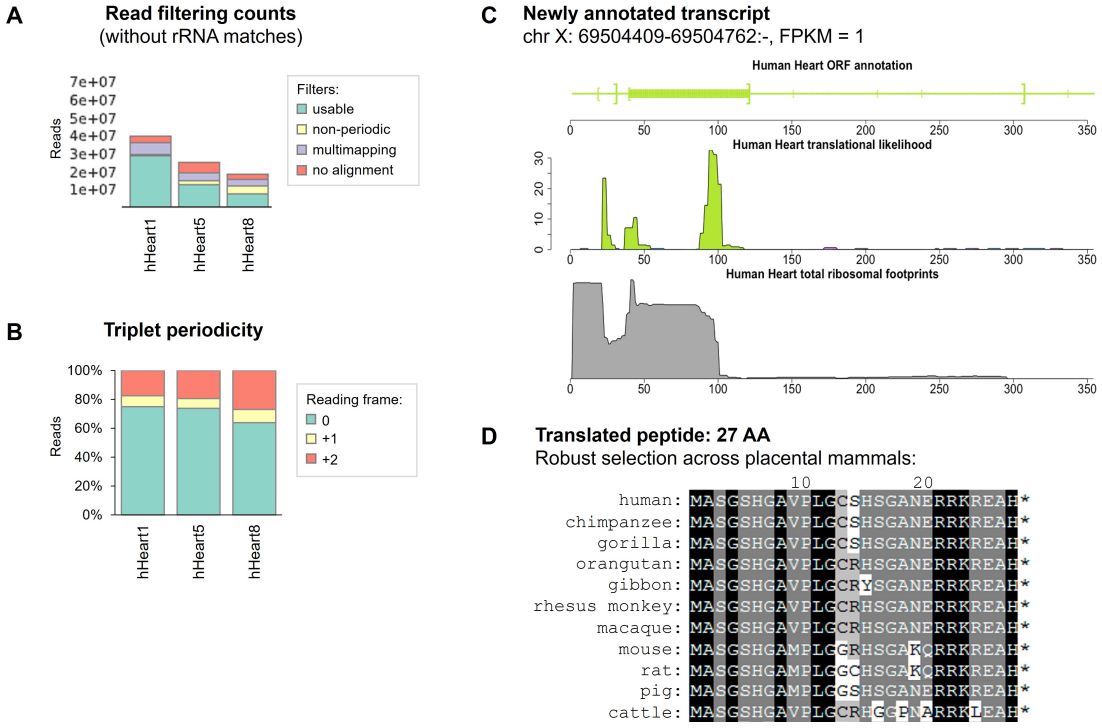

Figure 4.1: Conservation analysis of Ribo-seq libraries from human hearts reveals a novel microprotein. A. Pre-processing of 3 libraries from diseased human heart samples shows the number of periodic reads (ranging between 10 and 24 million per library, here marked in light green as "usable"). **B**. The percentage of reads matching three possible reading frames shows a distinct three-nucleotide periodicity in each library. **C**. Analysis of human cardiac translatomes predicts the translation of a novel microprotein encoded by a previously unannotated transcript. **D**. The predicted microprotein has 27 amino acids and is conserved across placental mammals. Bioinformatical analyses in A and B were performed by Etienne Boileau and conservation analysis in C and D was conducted by Evgeny Leushkin.

Meanwhile, a study has been published describing the translational landscape of the human heart. The authors prepared Ribo-seq libraries from 65 diseased human hearts and 15 control samples and identified 169 lncRNAs that encode novel microproteins. Their conservation analysis revealed that about a third of these lncRNAs exhibit positional conservation in mouse or rat, but only 5 of these lncRNAs encode sORFs that have a conserved translation initiation site in mouse or rat.

Building upon this study, my supervisor and I approached the authors to discuss potential conserved candidates that could be used for further experiments. They were willing to share a list of 5 candidate sORFs (**Table 4.1**) that were conserved and found to be translated in human as well as in mouse hearts. After careful examination and consultation with Evgeny Leushkin, I decided not to pursue any of them. All candidate microproteins were ubiquitously expressed and thus less likely to exhibit a cardiac phenotype when generating a cardiac-specific knockout mouse model. Some of the proposed microproteins were encoded by uORFs, which make it difficult to ascertain whether these microproteins are independent translation events or merely parts of larger alternative protein isoforms. Furthermore, sORFs located on antisense transcripts were not considered to be promising candidates, as their conservation signal could be entirely driven by the protein-coding genes encoded on the sense strand.

Table 4.1 Characteristics of conserved candidate sORFs (provided by Sebastiaan van Heesch)

Candidate	IFRD1 uORF	LCOR uORF	TUG1	KLF9-AS1	ZBTB11-AS1
sORF type/region	uORF	uORF	novel lncRNA isoform	novel antisense transcript	antisense transcript
phyloCSF score	+81	+137	+350	+81	-1332
GTEx expression	ubiquitous	ubiquitous	ubiquitous	NA	ubiquitous
Mouse cardiac translatome	translated	translated	translated	translated	translated
Human cardiac translatome	translated	translated	translated	translated	translated
blastp positives (human to mouse)	90%	100%	92%	89%	82%
Immuno-fluorescence with overexpression	cytoskeleton	not expressed	mitochondrion and nucleus	not tested	cytoplasm
Detected by MS in human heart [133]	yes	yes	no	no	no

Collectively, these findings demonstrate that novel cardiac microproteins are rare occurrences, which are unlikely to be conserved between humans and most common animal models, such as mouse and rat. While describing human microproteins would be much more relevant in terms of clinical translation, it would restrict experimental characterization to performing all experiments *in vitro*, for example in human iPSC-derived cardiomyocytes.

4.1.3 Exploring annotated small cardiac-enriched proteins identifies more promising candidates

The attempts so far suggested that the criteria applied for identifying novel cardiac microproteins were too stringent to select potential targets. Focusing on novel microproteins that would be conserved in small rodents and humans, as well as enriched and translated in the heart tissue of these organisms did not yield any candidates for further studies.

Therefore, I decided to simplify the search for microproteins by focusing on known protein-coding transcripts which encode small proteins. For this purpose, I used a publicly available transcriptome dataset which contains gene expression data for seven organs throughout their development in 7 animal species [113]. This information allowed me to first filter for heart-enriched transcripts. Since the described RNA-seq dataset contains information on protein-coding transcripts, which are mostly well conserved among selected species, the filtered cardiac-enriched transcripts were conserved and expressed in humans as well as in small rodents. I then used Ensembl annotations to keep only those transcripts that encode small proteins (with a maximum of 100 amino acids). Moreover, to validate their translation in the heart, I used published Ribo-seq data from human heart samples to retain only sORFs that are predicted to be translated [41].

The final list contained approximately 250 small cardiac-enriched microproteins, which I continued inspecting manually to select most promising candidates. I confirmed their spatial expression pattern in other available RNA-seq projects and I checked for annotated and alternative transcript isoforms, predicted microprotein characteristics and existing publications. I also considered the availability of experimental tools, such as antibodies and siRNAs, to study the function of these microproteins in more detail (**Figure 4.2 A**).

Finally, I decided for two candidates that are heart-enriched, annotated as protein-coding genes and expressed in both murine and human hearts (**Figure 4.2 B, C**). However, at the time of this analysis their relatively small protein products have not been characterized in the cardiac setting. The first candidate is SMIM4, which has multiple annotated transcript isoforms with protein products up to 102 amino acids. The second chosen candidate is SMCO1, which is strictly speaking not a microprotein, as the canonical isoform encodes a protein of 214 amino acids. It was retained in my pipeline because of a shorter isoform that encodes a peptide with premature termination codon, resulting in nonsense mediated decay. Interestingly, our collaborator Daniel Oehler identified yet another SMCO1 transcript isoform, which is cardiac-specific and contains an ORF with 202 codons.

The subsequent experimental characterization of SMIM4 and SMCO1 was performed by Nicholas A. Rüdinger, a medical student under my supervision. After validating the expression of both candidates in mouse heart tissue by qRT-PCR, I realized that SMCO1 was below detection limit in the most common *in vitro* model, primary NRCMs. For this reason, I decided to first focus our efforts on characterizing the function of SMIM4 in NRCMs.

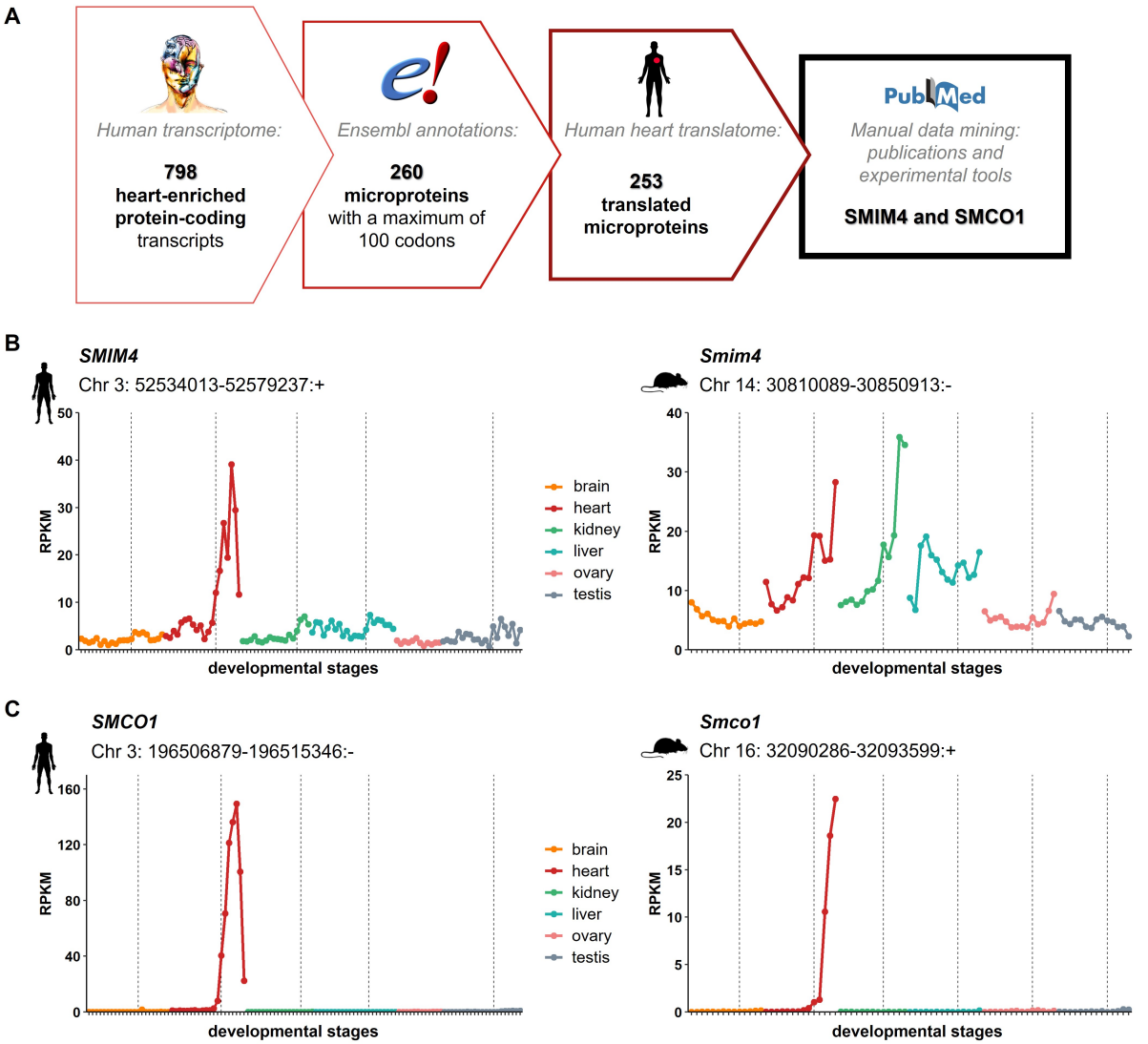

Figure 4.2: Focusing on known small proteins enriched in the heart identifies two new candidates. **A**. Overview of the pipeline I used to search for annotated small proteins that are enriched and translated in the human heart. Human transcriptome data was taken from Cardoso-Moreira *et al.*, 2019 [113] and human heart translatomes from van Heesch *et al.*, 2019 [41]. **B**. Spatial expression pattern of *SMIM4* during development indicates enrichment in the human heart (left panel) and expression in the mouse heart (right panel). **C**. Spatial expression pattern of *SMCO1* during development indicates highly specific expression in the human heart (left panel) as well as in the mouse heart (right panel). B and C were plotted based on the expression data published in Cardoso-Moreira *et al.*, 2019 [113]. Vertical dashed lines mark the beginning of postnatal development.

4.1.4 SMIM4 is a cardiac microprotein required for mitochondrial metabolism and cell growth in NRCMs

To address the role of SMIM4, a loss-of-function approach was adopted in NRCMs by transfecting them with siRNA targeting the transcript encoding SMIM4 (siSMIM4) (**Figure 4.3 A**). A successful knockdown of approximately 50 % was confirmed on both RNA and protein level (**Figure 4.3 B, C**). Next, various cell-based assays were performed, in which NRCMs were additionally stressed by treatment with hypertrophic stimulus phenylephrine (PE).

The results of MTT assay showed a decrease in mitochondrial metabolism and NRCM viability upon SMIM4 knockdown, and no further changes after PE treatment (**Figure 4.3 D**). On the contrary, puromycin incorporation to assess protein synthesis rates was not affected by depletion of SMIM4 (**Figure 4.3 E**). Measuring NRCM size after immunofluorescence staining revealed a blunted growth response in cells treated with SMIM4 siRNA alone or in combination with PE as compared to respective controls (**Figure 4.3 F**).

Collectively, these results point to the involvement of SMIM4 in cardiomyocyte mitochondrial metabolism, viability and cellular growth. These initial data will be expanded by performing further experiments *in vivo*, in SMIM4-deficient mice generated with CRISPR/Cas9-based approach (mouse model in preparation by collaborators from the group of Prof. Dr. Freichel).

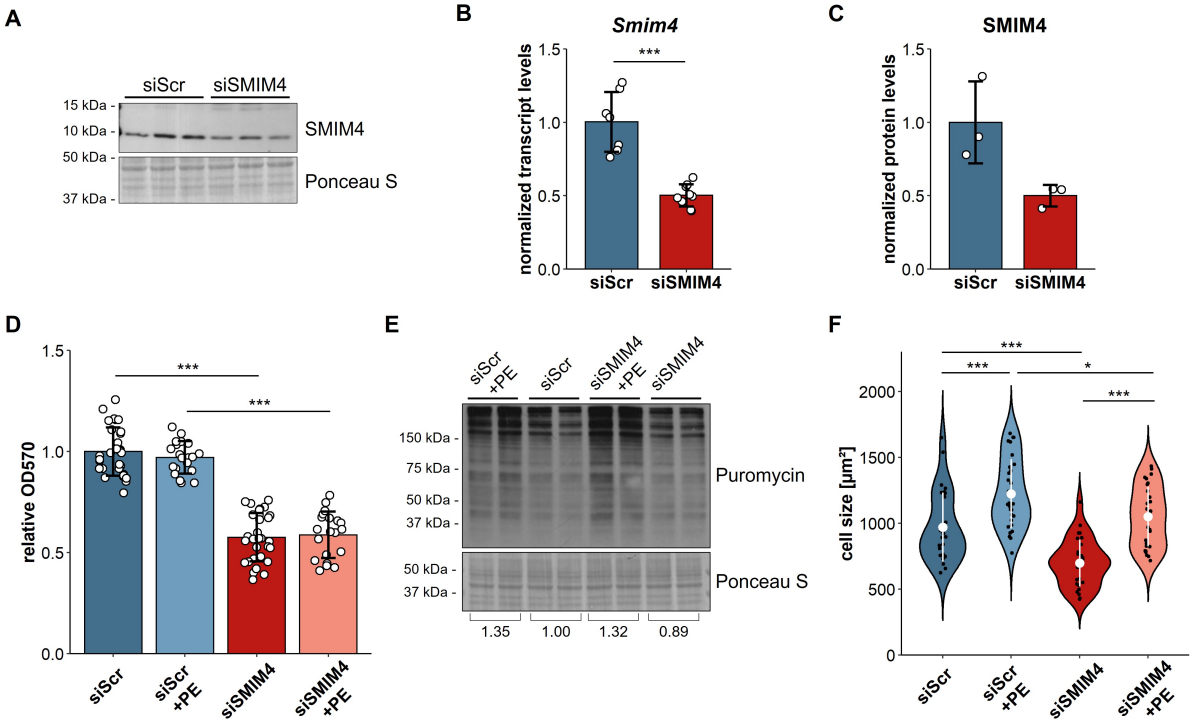

Figure 4.3: SMIM4 is required for mitochondrial metabolism and cell growth in NRCMs. A. NRCMs were transfected with control siRNA (siScr) or siRNA targeting the transcript encoding SMIM4 (siSMIM4) to achieve transient knockdown. **B**. Knockdown efficiency was measured by qRT-PCR and amounted to approximately 50 % in NRCMs treated with siSMIM4. *** $P < 0.001$ by Wilcoxon test. **C**. Quantification of A: 50 % knockdown efficiency was confirmed on protein level by western blot. **D**. Results of MTT assay show a significant decrease in mitochondrial metabolism and NRCM viability upon SMIM4 knockdown. *** $P < 0.001$ by one-way ANOVA and Tukey's test. **E**. Puromycin incorporation is not affected by SMIM4 depletion. Quantification of puromycin signal is shown at the bottom of the blot. Values are normalized to Ponceau S and represent averages of two biological replicates per condition. **F**. NRCM size is significantly reduced upon SMIM4 depletion. A smaller but significant decrease in cell size is also observed after SMIM4 knockdown and treatment with PE. * $P < 0.05$, *** $P < 0.001$ by pairwise Wilcoxon test with p-value adjustment. Experimental data for this figure were produced jointly with Nicholas A. Rüdinger.

4.2 Functional characterization of cardiac-enriched microprotein SPAAR

The aim of the second part of the project was to provide more insight into how SPAAR regulates mTORC1 in the heart and how that shapes the cardiac phenotype.

Therefore, two specific focus areas were defined. The first one was aimed at describing the overall heart phenotype and function following a loss-of-function approach *in vivo*. The second one focused on *in vitro* experiments using isolated primary cardiac cells to better understand the mechanisms of SPAAR function.

4.2.1 Specific focus 1: the role of SPAAR in myocardium *in vivo*

4.2.1.1 SPAAR KO mice lack the start codon to prevent SPAAR translation

To study how SPAAR contributes to the overall heart phenotype I concentrated on a loss-of-function approach. Our research group acquired a previously described global SPAAR KO mouse model from P. P. Pandolfi [10]. In this model, only the start codon of SPAAR was removed by using CRISPR/Cas9 homology-directed repair (**Figure 4.4 A**). This trinucleotide (ATG) deletion prevents SPAAR peptide expression, while transcript levels of *Spaar* remain unchanged. Utilizing this KO model thus offers the advantage of studying specifically the role of SPAAR peptide, and not the entire transcript.

To check whether the deletion of ATG codon indeed prevents all translation events involving *Spaar* transcript, I performed Ribo-seq on wild-type (WT) and SPAAR KO mouse hearts. Sequencing data analysis showed that ribosome footprints map onto *Spaar* transcript only in WT mouse hearts, whereas SPAAR KO hearts do not show any signs of active translation at this locus (**Figure 4.4 C**).

Next, I also wanted to prove the translation of SPAAR on protein level. Since the endogenous levels of SPAAR are very low and could not be detected in whole cell or tissue lysates by western blotting, I optimized an immunoprecipitation protocol for its detection. A custom-made anti-SPAAR antibody was first used to precipitate the peptide from heart lysates and then to detect SPAAR by western blotting. This enabled me to validate SPAAR translation in WT mouse hearts and prove its depletion in SPAAR KO hearts (**Figure 4.4 B**).

Finally, I also confirmed by qRT-PCR that *Spaar* transcript levels do not differ between WT and SPAAR KO hearts (**Figure 4.4 D**). In addition, RNA-seq of isolated cardiac endothelial cells (which exhibit the highest expression levels of *Spaar*) corroborated that SPAAR KO cells lack the ATG codon (**Figure 4.4 C**), but express comparable levels of *Spaar* transcript as WT cells (**Figure 4.4 E**).

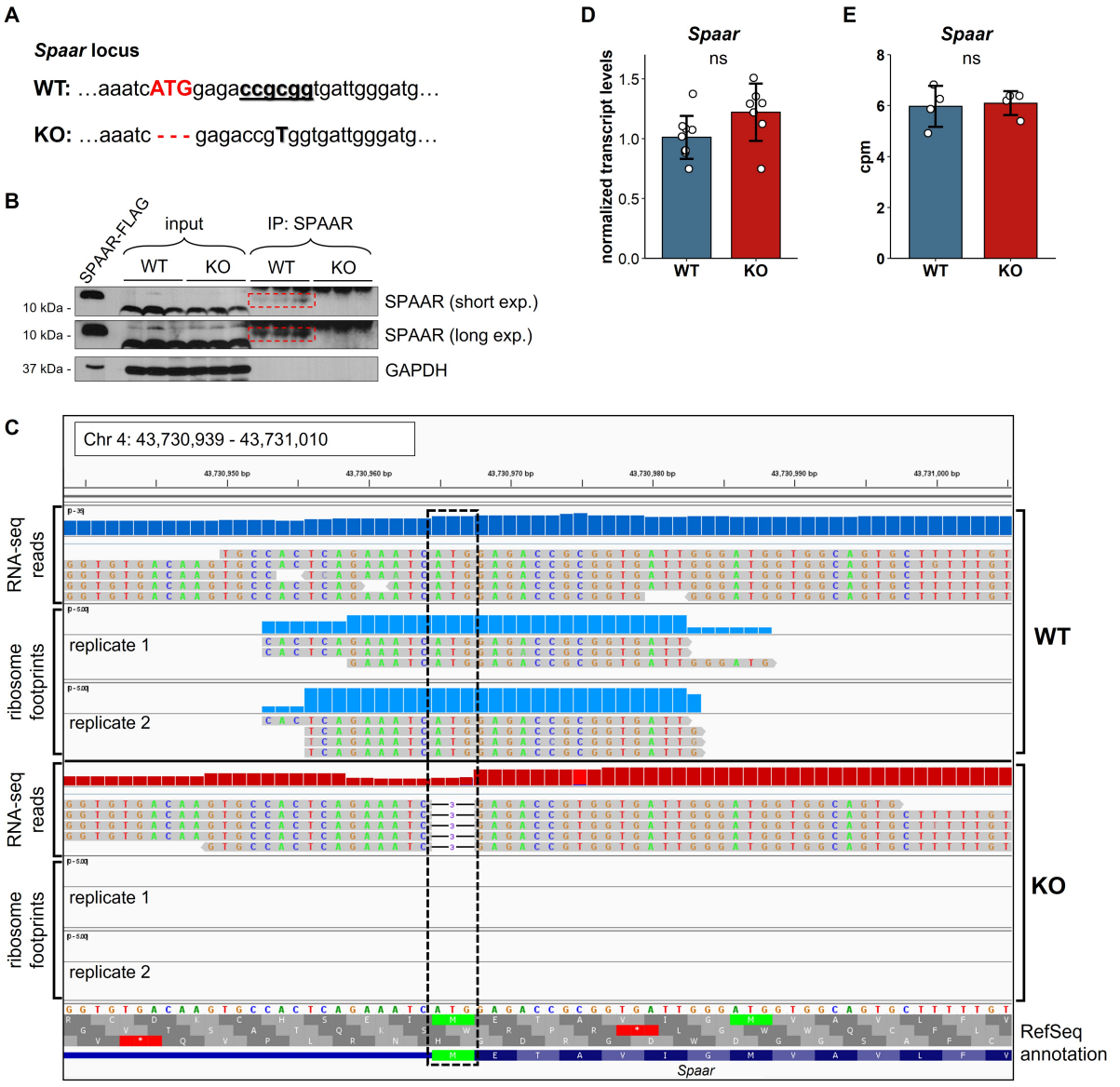

Figure 4.4: Deletion of ATG codon of SPAAR prevents its translation while maintaining the integrity of the host transcript. A. *Spaar* locus with start (ATG) codon in WT mice indicated in red. Highlighted in bold and underlined text is SacII restriction enzyme site. In SPAAR KO mice, the ATG codon is deleted and SacII site is mutated to distinguish WT and SPAAR KO alleles by genotyping. **B**. Immunoprecipitation (IP) of SPAAR from mouse heart lysates followed by western blotting. SPAAR peptide is marked with red dashed outline in WT mice and is depleted in SPAAR KO mice. Cell lysate with overexpression of FLAG-tagged SPAAR construct (SPAAR-FLAG) on the left side of the blot serves as a positive control (note that it migrates slightly higher due to a linker and FLAG tag). Exp.: exposure time. **C.** A snapshot from Integrative Genomics Viewer showing RNA-seq reads and ribosome footprints aligned to the start codon of *Spaar* in WT (top) and SPAAR KO mouse hearts (bottom). RNA-seq reads from SPAAR KO cells carry a three-nucleotide ATG deletion in *Spaar* (marked with black dashed outline). Ribosome footprints indicate active translation of SPAAR in WT hearts, but not in SPAAR KO. Bioinformatical analysis of ribosome profiling data was performed by Etienne Boileau. **D.** qRT-PCR analysis of WT and SPAAR KO mouse heart lysates shows that *Spaar* transcript levels are not significantly different between the genotypes. $P > 0.1$ by T-test. **E.** Differential expression analysis after RNA-seq confirms that *Spaar* transcript levels do not differ between WT and SPAAR KO cardiac endothelial cells.

4.2.1.2 SPAAR KO mice do not display a distinct cardiac phenotype at baseline

I performed basal characterization of SPAAR KO mice in 12-week old adult male mice using ultrasound echocardiography in parasternal long axis view. Basal cardiac dimensions, such as left ventricular area in systole and diastole, systolic and diastolic volume and stroke volume were not different in SPAAR KO mice as compared to WT (**Figure 4.5 A-E**). Similarly, cardiac function as calculated by cardiac output and left ventricular ejection fraction (EF), did not differ between SPAAR KO and WT mice (**Figure 4.5 F, G**). Furthermore, heart weight/body weight ratio did not show any significant changes between the genotypes (**Figure 4.5 H**). However, isolating cardiomyocytes from adult mice revealed small, but significant increase in their width in SPAAR KO cells, whereas no differences were observed in their length (**Figure 4.6**).

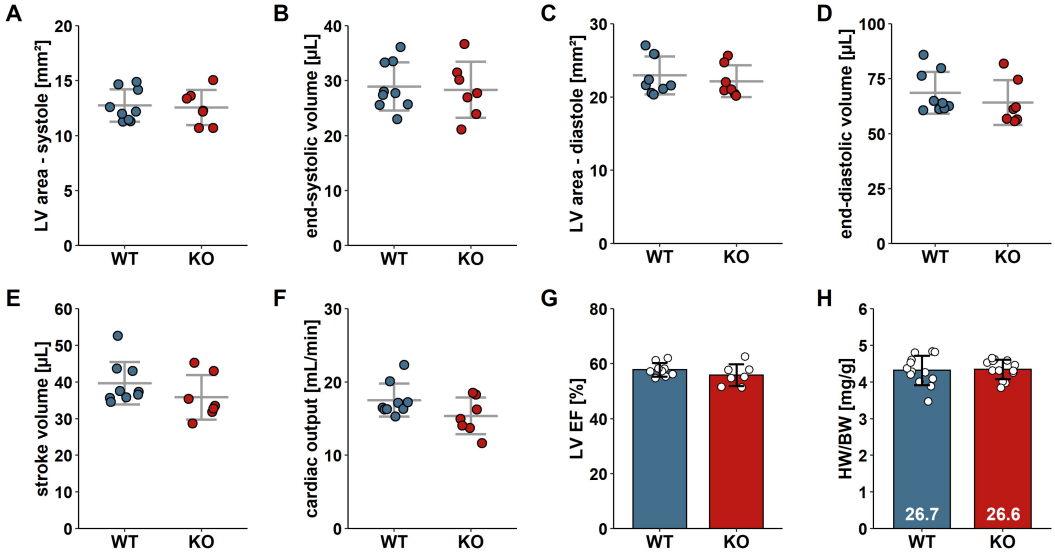

Figure 4.5: Basal cardiac dimensions and function evaluated by echocardiography in 12-week old male mice. **A**. Left ventricular systolic area. **B**. End-systolic volume. **C**. Left ventricular diastolic area. **D**. End-diastolic volume. **E**. Stroke volume. **F**. Cardiac output. **G**. Left ventricular ejection fraction. **H**. Heart weight/body weight ratio with average body weight indicated at the bottom of the bars. None of the comparisons between WT and KO are significant.

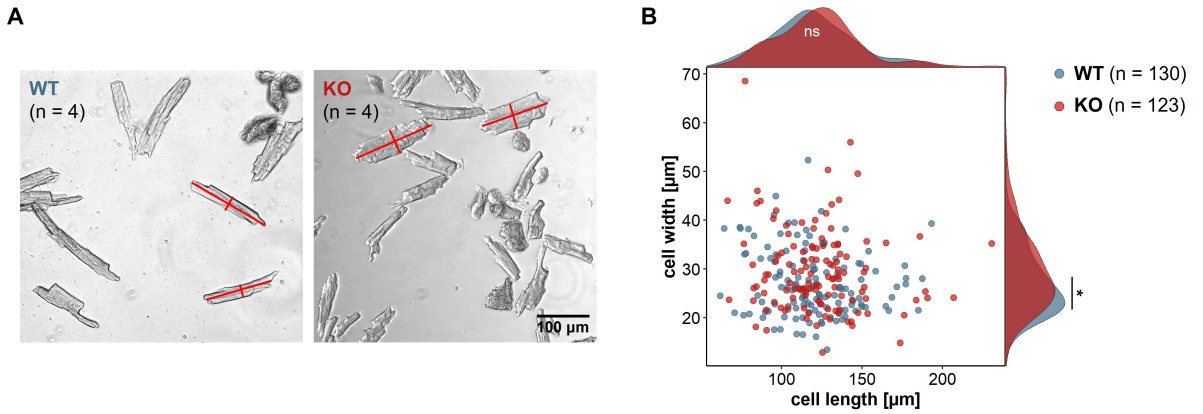

Figure 4.6: Cardiomyocytes isolated from adult SPAAR KO mice exhibit larger cell width. **A**. Representative brightfield images of adult cardiomyocytes isolated from 4 WT mice and 4 SPAAR KO mice. **B**. Quantification of cell width and length reveals a significant increase in the width of SPAAR KO cells compared to WT. * $P < 0.05$ by Wilcoxon test.

4.2.1.3 SPAAR KO mice are protected from cardiac damage early after I/R injury

Since SPAAR KO mice were indistinguishable from WT mice at baseline conditions, the next step involved exposing them to cardiac injury induced by ischemia-reperfusion (I/R). Following a period of ischemia, when reduced blood supply leads to deficiencies in nutrients and oxygen, reperfusion is intended to re-establish blood flow to salvage ischemic cardiac tissue. At the same time, reperfusion also produces adverse cellular and tissue damage through rupture of mitochondria, increased generation of reactive oxygen species and induction of various proteases [89]. I/R thus partially simulates the conditions in which SPAAR was originally described, namely after periods of amino acid deprivation and stimulation. Therefore, I reasoned that I/R injury would represent a relevant *in vivo* model to investigate the role of SPAAR in cardiac damage.

Surgical procedures to induce I/R injury in mice were carried out by medical students Ole M. Schwerdt and Adrian Serafin. Left anterior descending artery (LAD) was ligated for 60 minutes to induce an infarction, followed by release of the ligature to restore blood flow to the affected areas of the heart [117]. Cardiac phenotype and function were then monitored over a course of 2 weeks, before mice were sacrificed for further histological and biochemical analyses (**Figure 4.7 A**).

Approximately 24 hours after the induction of I/R, I collected blood samples for determination of cardiac troponin T (cTnT). cTnT is a subunit of troponin complex, essential for cardiac muscle contraction, and is released from damaged cardiomyocytes upon disruption of sarcolemma. Its levels in peripheral blood peak at about 24 hours after the cardiac insult, making cTnT a standard biochemical marker for myocardial infarction and damage [134]. As expected, WT mice had significantly increased levels of cTnT in blood serum following I/R, confirming that the chosen experimental model leads to acute cardiac damage. Surprisingly, SPAAR KO mice exhibited a protective response already at this early time point and displayed decreased levels of cTnT compared to WT mice (**Figure 4.7 B**).

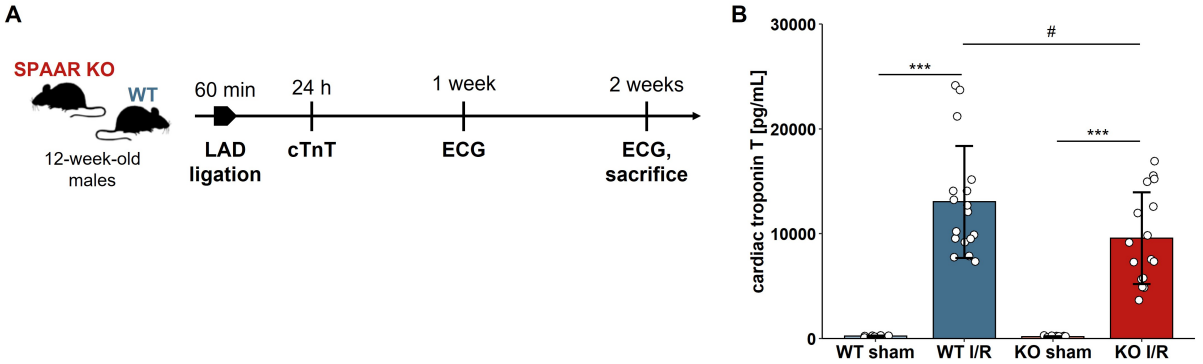

Figure 4.7: Myocardial ischemia-reperfusion injury leads to less short-term cardiac damage in SPAAR KO mice. **A**. Experimental timeline used to study the influence of SPAAR depletion following myocardial I/R injury. LAD ligation was performed by Ole M. Schwerdt and Adrian Serafin. **B**. 24 h after LAD ligation cTnT levels in blood serum were increased in WT and SPAAR KO mice subjected to I/R, but less so in SPAAR KO animals. # $0.05 < P < 0.1$, *** $P < 0.001$ after outlier removal and pairwise Wilcoxon test with p-value adjustment.

One week post-surgery, I used ultrasound echocardiography to non-invasively monitor heart function. Although there was a trend towards increased end-systolic volume and consequently decreased stroke volume and cardiac output in mice subjected to I/R, there were no significant changes in left ventricular EF among experimental groups (**Figure 4.8**).

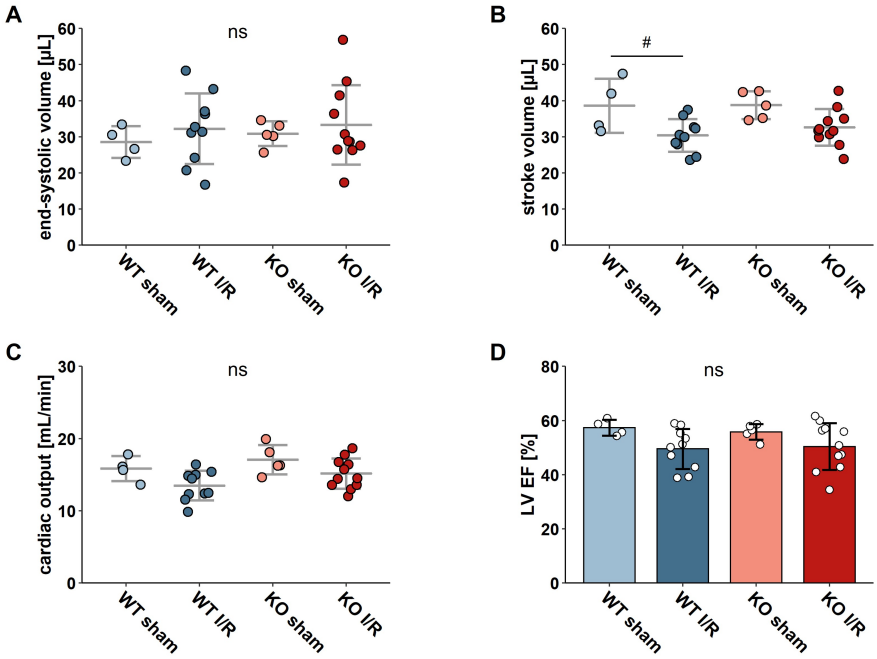

Figure 4.8: Cardiac dimensions and function evaluated by echocardiography 1 week after the induction of I/R. A. End-systolic volume. **B**. Stroke volume. **C**. Cardiac output. **D**. Left ventricular ejection fraction. # $0.05 < P < 0.1$ by one-way ANOVA and Tukey's test.

4.2.1.4 SPAAR KO mice maintain a cardioprotective phenotype at later stages of reperfusion

I repeated the echocardiography analysis two weeks post-surgery to assess heart function at a later stage of reperfusion. WT mice that underwent I/R surgery had a significantly increased end-systolic volume and decreased stroke volume as well as cardiac output (**Figure 4.9 A-C**). This led to a decreased heart function as indicated by reduced EF in comparison to WT sham mice. On the contrary, SPAAR KO mice subjected to I/R had a slightly decreased cardiac output, but no differences in end-systolic volume or stroke volume, thereby resulting in sustained EF compared to SPAAR KO sham mice. Overall, this contributed to SPAAR KO I/R mice having significantly improved EF when compared to WT I/R mice at two weeks of reperfusion (**Figure 4.9 D**).

The I/R model thus induced cardiac injury and systolic heart impairment with reduced EF in WT mice, but not in SPAAR KO mice. Complex changes that guide such heart dysfunction are referred to as pathological cardiac remodelling and include molecular, cellular and structural changes that manifest as changes in mass, shape and function of the heart [135]. To evaluate the extent of cardiac remodelling after I/R, all mice were sacrificed two weeks post-surgery. Their gross heart phenotype was compared among experimental groups, but no differences were observed in heart weight normalized to the respective body weight (**Figure 4.9 E**).

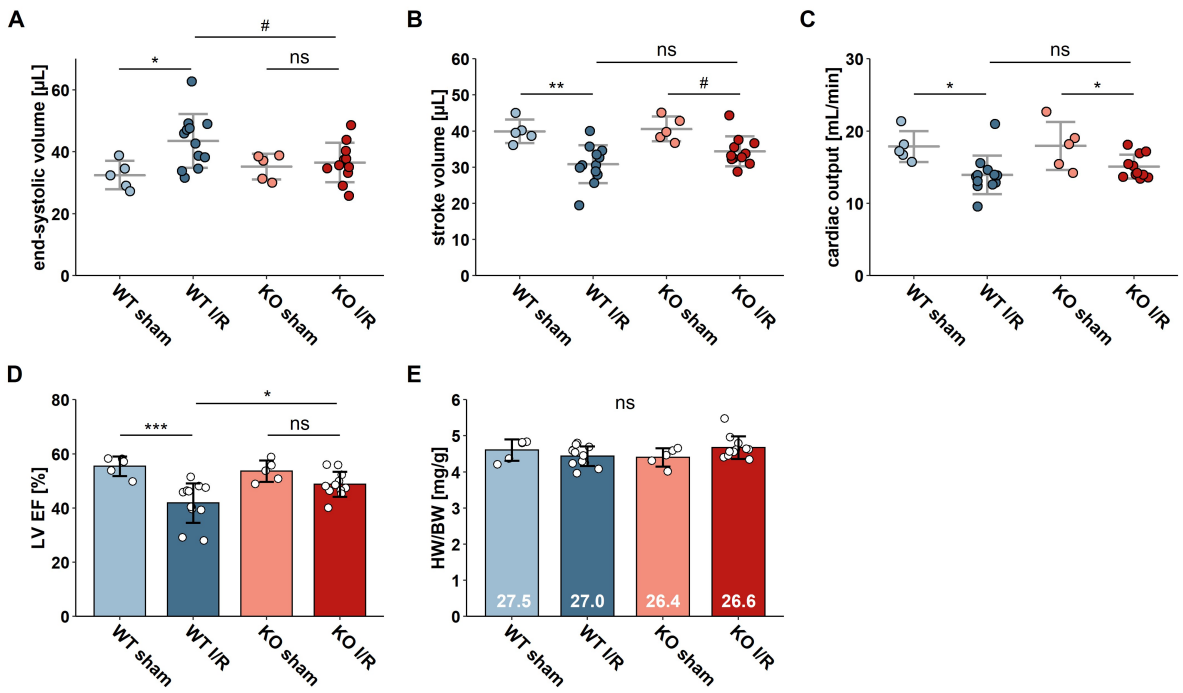

Figure 4.9: Cardiac dimensions and function evaluated by echocardiography 2 weeks after the induction of I/R. A. End-systolic volume. **B**. Stroke volume. **C**. Cardiac output. **D**. Left ventricular ejection fraction. **E**. Heart weight/body weight ratio with average body weight indicated at the bottom of the bars. # $0.05 < P < 0.1$, * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$ by one-way ANOVA and Tukey's test (in A, B, D, E) or by pairwise Wilcoxon test with p-value adjustment (in C).

Next, I wanted to verify these findings on cellular level and therefore stained heart sections using wheat germ agglutinin (WGA), which binds to glycoproteins on cell membranes, and anti-cTnT antibody, which distinguishes cardiomyocytes from other cell types (**Figure 4.10 A**). Quantification of cross-sectional area of cardiomyocytes revealed a small increase in SPAAR KO sham hearts compared to WT sham animals (**Figure 4.10 B**), which is in agreement with previously described increase in cell width of isolated SPAAR KO cardiomyocytes (**Figure 4.6**). There was also a small, but significant increase in cell size of WT cardiomyocytes following I/R, whereas this increase was not apparent when comparing SPAAR KO sham hearts with KO I/R hearts. Overall, this meant that cardiomyocyte area in mice subjected to I/R does not differ between WT and SPAAR KO animals. However, these results also indicate that cross-sectional area of cardiomyocytes alone does not entirely reflect the gross heart phenotype described above (**Figure 4.9 E**).

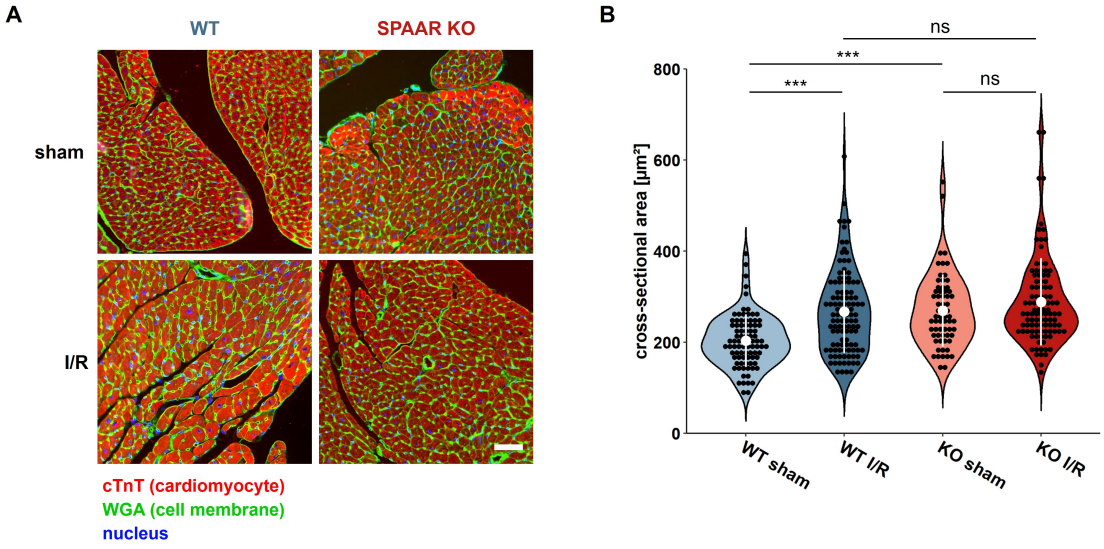

Figure 4.10: WGA staining of heart sections reveals differences in cardiomyocyte cell size. **A**. Representative fluorescent images of WGA and cTnT staining 2 weeks after the induction of I/R. Scale bar: 50 µm. **B**. Quantification of cardiomyocyte cross-sectional area shows an increase in cell size after I/R in WT mice, but not in SPAAR KO mice. However, KO mice have larger cardiomyocytes than WT already in control sham condition. *** $P < 0.001$ by pairwise Wilcoxon test with p-value adjustment.

To investigate major histopathological changes in the cardiac tissue of these mice, I then performed hematoxylin and eosin staining of heart sections (**Figure 4.11 A**). Regions without myocytes (eosin-free) and rich with atypical nuclei (hematoxylin-positive) were quantified and the analysis revealed a higher hematoxylin content after I/R, indicating expansion and infiltration of non-cardiomyocyte cell types. Furthermore, SPAAR KO I/R mice exhibited a trend towards decreased hematoxylin-positive areas (**Figure 4.11 B**).

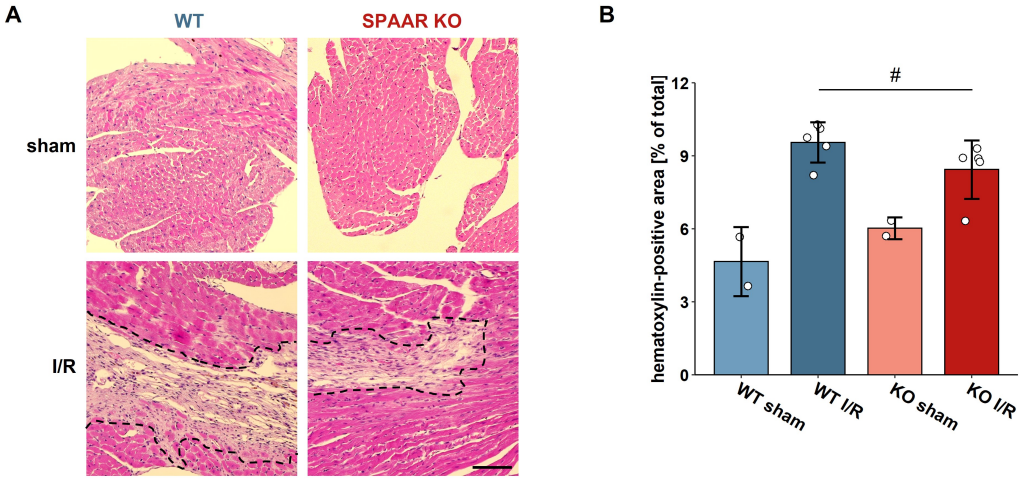

Figure 4.11: Hematoxylin and eosin staining of heart sections implies cell expansion and infiltration after I/R. A. Representative brightfield images of heart tissue sections after 2 weeks of I/R. Eosin-free and hematoxylin-rich areas are marked with a dashed outline. Scale bar: 100 µm. **B**. Quantification of hematoxylin-positive areas shows an increase after I/R, which is less pronounced in SPAAR KO mouse hearts. # $0.05 < P < 0.1$ by Wilcoxon test.

These results already point to another aspect of pathological cardiac remodelling, namely to changes in cellular composition of the heart. Since cardiomyocytes have a very limited regenerative capacity, such damaged cells are replaced by a fibrotic scar. This process is termed fibrosis and is characterized by extensive proliferation and activation of fibroblasts and deposition of extracellular matrix proteins, such as collagens [136]. To compare the extent of collagen deposition between WT and SPAAR KO I/R mice, I performed Masson Trichrome staining on heart sections, followed by quantification of scar size (**Figure 4.12 A**). This revealed that SPAAR KO I/R mice had significantly less scar tissue, which correlated with reduced cTnT levels measured 24 hours post-surgery (**Figure 4.12 B**). Furthermore, qRT-PCR analysis of heart lysates confirmed that SPAAR KO I/R mice express less collagen type 1 (*Col1a1*), a major constituent of fibrotic tissue, compared to WT I/R mice (**Figure 4.12 C**).

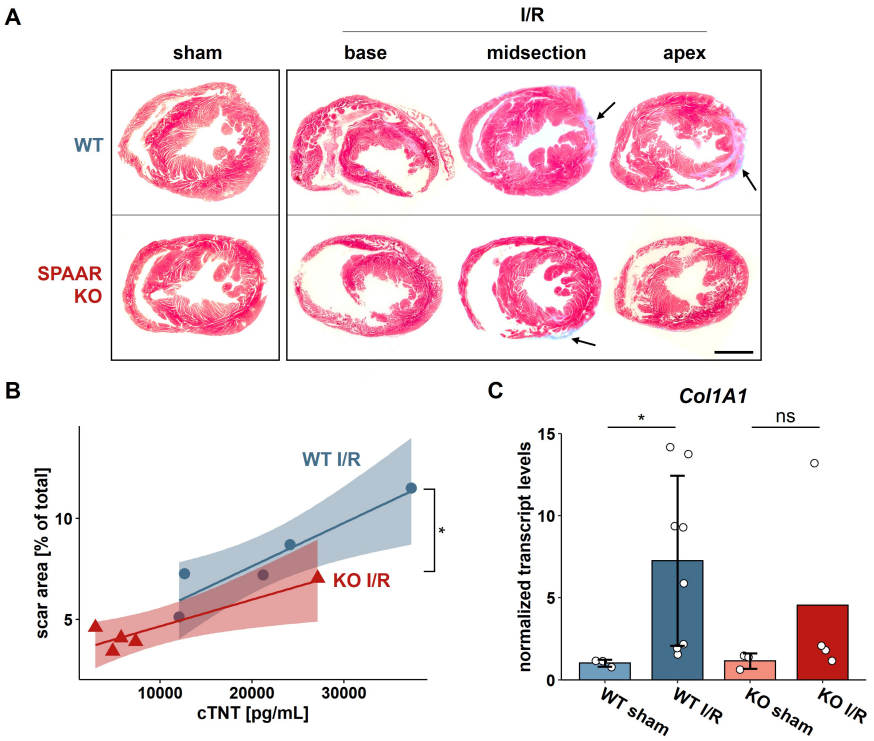

Figure 4.12: SPAAR KO mice have smaller fibrotic scars and less collagen deposits following I/R. A. Representative images of Masson Trichrome staining on heart sections after 2 weeks of I/R. Collagen deposits displaying purple-blue coloration are indicated with black arrows. Scale bar: 2 mm. **B**. Quantification of collagen deposits revealed a significant decrease in scar area in SPAAR KO mice subjected to I/R, as compared to WT I/R mice. * $P < 0.05$ by T-test. **C.** qRT-PCR analysis for *Col1A1* indicates a strong increase after I/R in WT mice, whereas SPAAR KO I/R mice express it to a smaller degree. * $P < 0.05$ by Wilcoxon test.

Besides cardiomyocytes, endothelial cells are also sensitive to I/R damage leading to vasoconstriction and disruption of the capillary microcirculation, further exacerbating cardiac injury [137]. However, endothelial cells are capable of proliferation and thus formation of new vessels by sprouting angiogenesis [138]. To check whether capillary density differs between WT and SPAAR KO mice following I/R, I stained heart sections using isolectin IB4, which binds to galactose residues on endothelial cell membranes, and WGA (**Figure 4.13 A**). However, determining the ratio of capillaries (IB4- and WGA-positive) per cardiomyocyte (only WGA-positive) did not reveal any differences between WT and SPAAR KO mice (**Figure 4.13 B**).

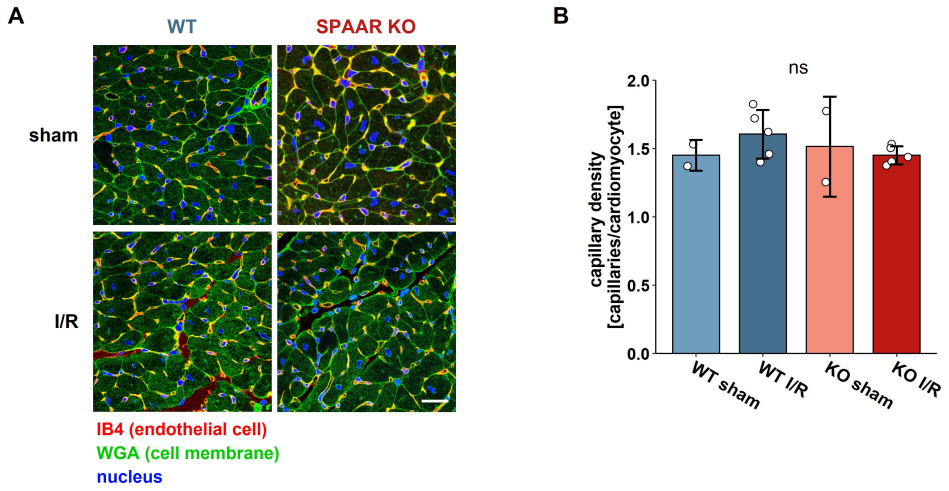

Figure 4.13: Isolectin IB4 staining of heart sections does not reveal differences in capillary density after I/R. **A**. Representative fluorescent images of isolectin IB4 and WGA staining on heart sections after 2 weeks of I/R. Membranes of endothelial cells are stained with isolectin IB4 and WGA and thus appear yellow, whereas other cell membranes are stained with WGA only and are therefore green. Scale bar: 25 µm. **B**. Quantification of capillary density expressed as a ratio of capillaries per cardiomyocyte does not reveal any significant differences between experimental groups.

4.2.1.5 Summary of Specific focus 1: SPAAR depletion is cardioprotective after I/R *in vivo*

The aim of Specific focus 1 was to describe how SPAAR shapes the overall heart phenotype and function by adopting a loss-of-function approach *in vivo*. I used SPAAR KO mice, which carry a global deletion of ATG codon to hinder SPAAR translation, but still preserve the integrity of *Spaar* transcript. At baseline, these mice did not exhibit a particular heart phenotype and were basically indistinguishable from WT mice. However, when subjected to myocardial I/R injury, SPAAR KO mice showed a distinctive cardioprotective phenotype. Already early during reperfusion they demonstrated decreased levels of biomarker for cardiac damage, cTnT. Later on, this protective effect was maintained as evidenced by sustained left ventricular EF and reduced fibrotic depositions.

4.2.2 Specific focus 2: mechanistic insights into SPAAR function

Results from Specific focus 1 indicate that SPAAR KO mice are protected from cardiac damage induced by I/R injury *in vivo* and show improved heart function compared to WT animals. But ensuring proper heart function requires simultaneous activity of diverse cardiac cell populations, such as cardiomyocytes, fibroblasts, endothelial cells and many others. Inspecting *Spaar* transcript levels in these various cell types within the heart surprisingly demonstrates its highest expression in endothelial cells and only very low expression in cardiomyocytes. Therefore, all experiments aimed at understanding the role of SPAAR in cardioprotection were performed in primary cardiac endothelial cells isolated from WT and SPAAR KO mice.

4.2.2.1 SPAAR is localized in the lysosomal membrane of cardiac endothelial cells

To confirm the subcellular localization of SPAAR, cardiac endothelial cells were infected with adenoviruses to overexpress SPAAR-FLAG and lysosomal marker TMEM192-3xHA [139]. Following immunofluorescence staining and subsequent imaging, intensity-based colocalization analysis revealed extensive overlap between the two fluorescent channels, confirming that SPAAR localizes to the lysosomes in cardiac endothelial cells (**Figure 4.14**).

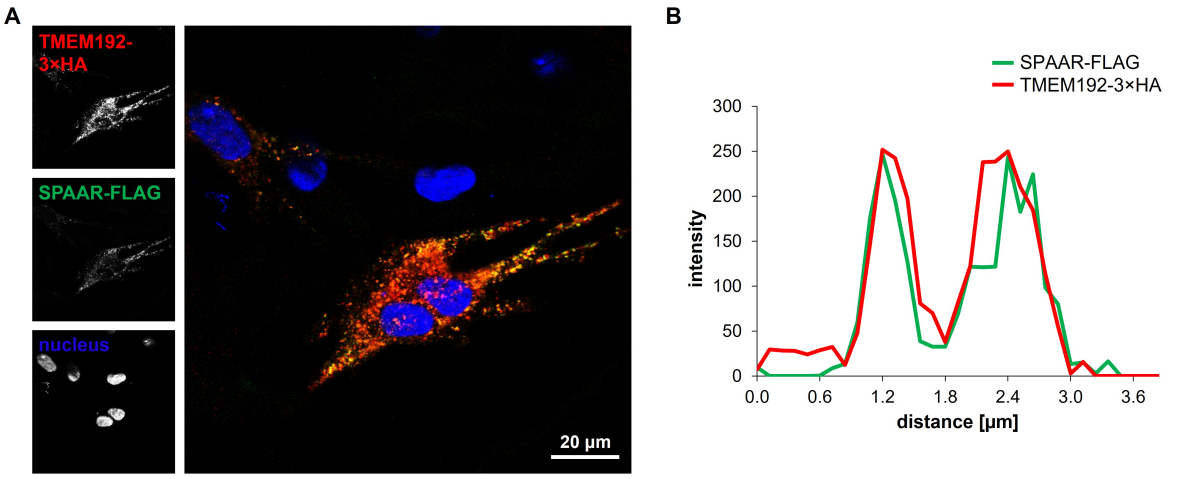

Figure 4.14: SPAAR is a lysosomal microprotein. A. Representative fluorescent image of primary endothelial cells overexpressing SPAAR-FLAG (green) and lysosomal marker TMEM192-3×HA (red). **B**. Colocalization analysis reveals considerable overlap between the overexpression constructs, pointing to colocalization of SPAAR with lysosomes.

4.2.2.2 SPAAR depletion selectively activates mTORC1 signalling in cardiac endothelial cells

The first report of SPAAR function described this peptide as a regulator of mTORC1 [10]. When cells were kept in culture medium without amino acids and then stimulated with a mixture of all 20 proteinogenic amino acids for a short period, *Spaar* depletion led to increased activation of mTORC1 signalling pathway. This was demonstrated by western blotting as increased phosphorylation levels of canonical mTORC1 targets S6K, RPS6 and 4E-BP.

To reproduce these findings in cardiac endothelial cells, I performed amino acid depletion and stimulation experiments in WT and SPAAR KO endothelial cells. However, I could not observe any significant increase in the phosphorylation levels of previously reported mTORC1 targets in SPAAR KO cells. I also examined the levels of phosphorylated ULK1, which is an mTORC1 target required for starvation-induced autophagy, but there were no changes in its phosphorylation levels in SPAAR KO cells (**Figure 4.15 A, B**).

Furthermore, I investigated phosphorylation of transcription factor EB (TFEB), an mTORC1 target involved in lysosomal biogenesis. I assessed phosphorylation levels of two different serine residues of TFEB, that were previously reported to be phosphorylated by mTOR: Ser122 and Ser211 [77]. Both at baseline and following amino acid starvation and stimulation, SPAAR KO cells showed significantly increased levels of phospho-Ser122 compared to WT cells. This

increase could be reversed by the addition of adenovirus encoding SPAAR-FLAG. Interestingly, no phospho-Ser211 signal was detected after amino acid re-stimulation neither in WT nor in SPAAR KO cells (**Figure 4.15 A, B**).

Phosphorylation of TFEB is a post-translational modification that regulates its subcellular localization: phosphorylated TFEB is retained in the cytoplasm by binding to 14-3-3 cytosolic chaperones, whereas unphosphorylated TFEB can translocate into the nucleus, where it transcribes genes important for lysosomal biogenesis and autophagy [73]. To indirectly assess whether increased phosphorylation of TFEB in SPAAR KO cells affects its localization and thus activity, I performed qRT-PCR analysis to determine transcript levels of selected lysosomal genes that are known TFEB targets. However, I could not observe any differences in their transcript levels between WT and SPAAR KO cells, neither at baseline nor after amino acid starvation and stimulation (**Figure 4.15 C**). These results imply that differential phosphorylation of TFEB at Ser122 does not affect its transcriptional activity in SPAAR KO cells.

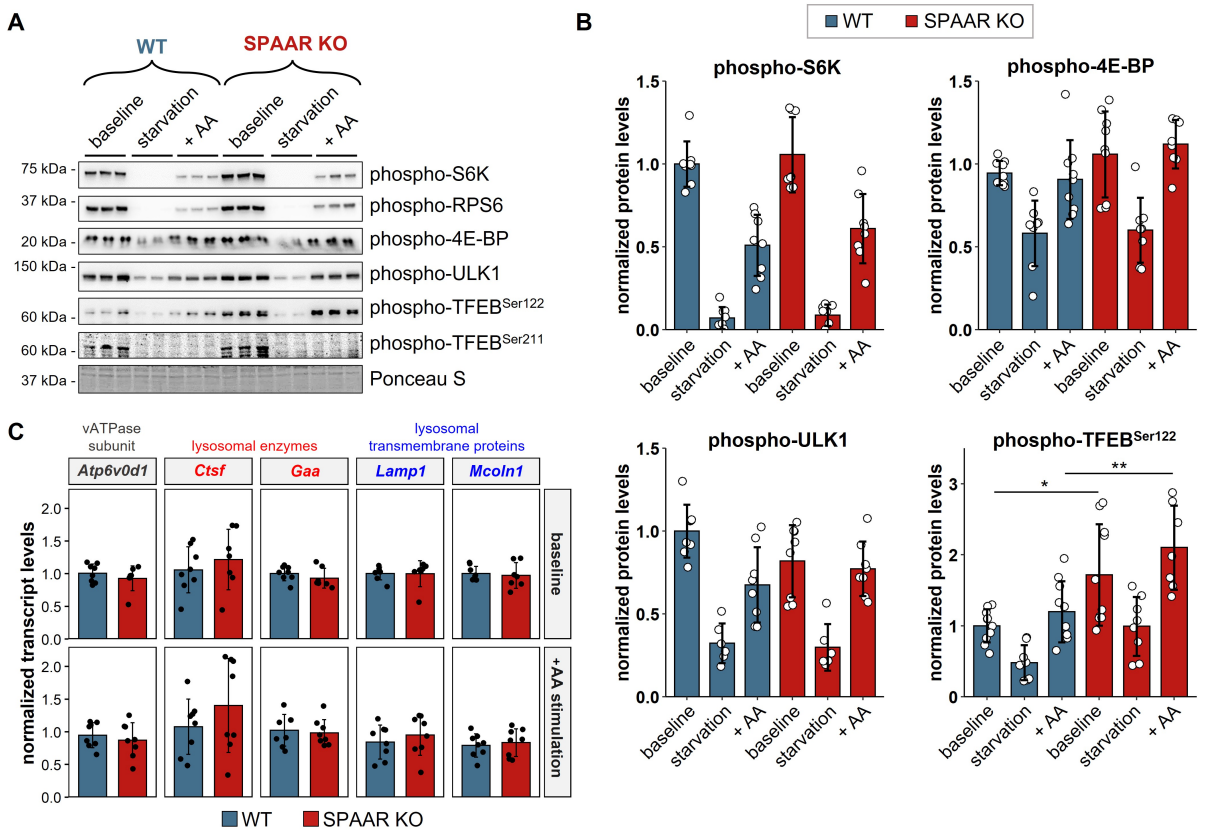

Figure 4.15: SPAAR deletion leads to selective upregulation of mTORC1 signalling upon amino acid stimulation. **A**. Representative western blots for phosphorylated canonical mTORC1 targets in WT and SPAAR KO endothelial cells. Starvation was carried out for 3 h in culture medium without amino acids. +AA denotes starvation followed by 1 h of stimulation with a mixture of all 20 proteinogenic amino acids. **B**. Quantification of selected targets from A shows significant increase in the phosphorylation of TFEB at Ser122 in SPAAR KO cells at baseline and after amino acid stimulation. N = 6 - 9 per condition from 4 independent cell isolations. * $P < 0.05$, ** $P < 0.01$ by pairwise Wilcoxon test with p-value adjustment. **C.** qRT-PCR analysis of selected TFEB targets does not show any differences in their transcript levels at baseline nor after amino acid stimulation, pointing to comparable transcriptional activity of TFEB in WT and SPAAR KO cells. N = 7 - 8 per condition from 2 independent cell isolations. None of the comparisons between WT and KO are significant.

Additionally, I also checked the activity of mTORC1 and mTORC2 signalling pathways following insulin starvation and stimulation. For this purpose, WT and SPAAR KO endothelial cells were kept in serum-free medium and then stimulated with insulin for a short period. There were no significant changes in the phosphorylation of AKT at Ser473, which is a target of mTORC2, in SPAAR KO cells compared to WT. Furthermore, downstream targets S6K, RPS6 and 4E-BP did not show any differential phosphorylation in SPAAR KO cells either (**Figure 4.16**). These results indicate that SPAAR depletion does not influence the activity of mTOR towards canonical targets following insulin starvation and stimulation.

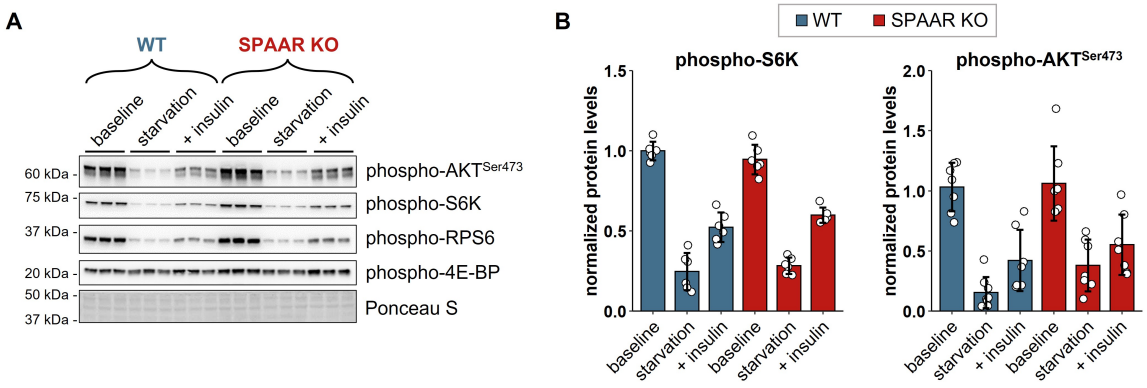

Figure 4.16: SPAAR deletion does not affect canonical mTOR signalling upon insulin stimulation. A. Representative western blots for phosphorylated canonical mTORC1 and mTORC2 targets in WT and SPAAR KO endothelial cells. Starvation was carried out for 3 h in culture medium without serum. +insulin denotes starvation followed by 1 h of stimulation with insulin. **B**. Quantification of selected targets from A reveals there are no differences in the activity of mTORC1 (represented by phospho-S6K status) and mTORC2 (represented by phospho-AKT^{Ser473}) in SPAAR KO cells compared to WT. N = 6 - 7 per condition from 3 independent cell isolations.

4.2.2.3 SPAAR depletion does not affect proliferation, metabolic activity nor protein synthesis of cardiac endothelial cells

To perform more extensive phenotypic evaluation of SPAAR KO endothelial cells, I carried out several cell-based assays. First, CyQUANT Cell Proliferation assay was used, which is based on determination of cellular DNA content after binding of a fluorescent dye. WT and SPAAR KO cells were cultured in basal endothelial medium (containing serum) and full endothelial medium (containing serum, as well as additional growth supplements and non-essential amino acids) and their proliferation rates were compared over a course of three days. Results of this assay did not indicate any significant differences in the proliferation of SPAAR KO and WT endothelial cells in either of the tested culture media (**Figure 4.17 A**).

Next, MTT assay was performed, which is based on the activity of mitochondrial reductases to detect metabolically active cells. Similar to the proliferation assay, WT and SPAAR KO endothelial cells were cultured in basal and full endothelial medium and their metabolic rates were compared across three days. The measurements showed highly comparable rates of mitochondrial metabolism between WT and KO cells in both culture media (**Figure 4.17 B**).

Finally, puromycin incorporation assay was completed to estimate protein translation rates. Since SPAAR function was initially characterized in cells subjected to amino acid starvation and stimulation, I included this treatment in the puromycin assay. Furthermore, adenovirus encoding SPAAR-FLAG (Ad-SPAAR-FLAG) was used to infect endothelial cells and compare SPAAR overexpression (WT + Ad-SPAAR-FLAG) with KO rescue (SPAAR KO + Ad-SPAAR-FLAG) (**Figure 4.17 C**). However, western blot analysis of puromycylated proteins did not show any significant differences in protein synthesis between WT and KO cells at baseline, nor after amino acid starvation and stimulation (**Figure 4.17 D**). Addition of Ad-SPAAR-FLAG led to a small decrease in global translation, which was not significant across independent cell isolations.

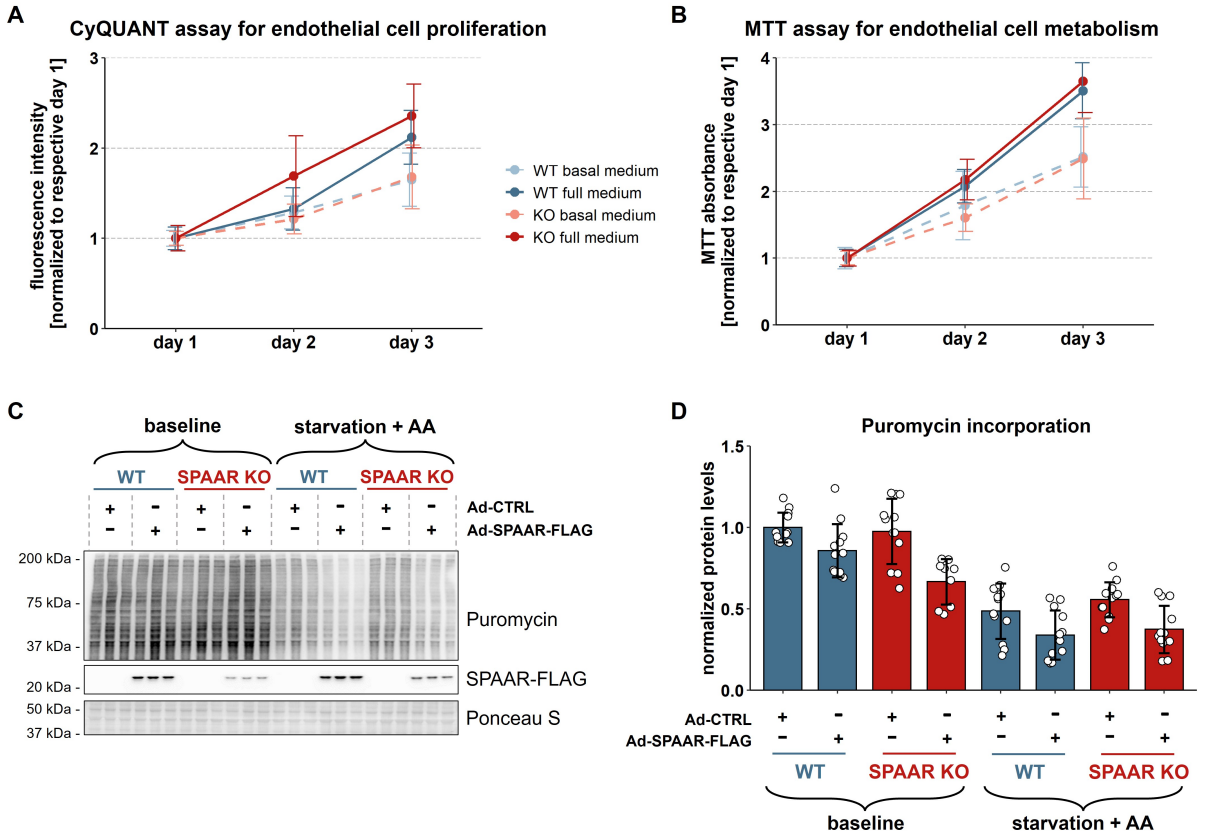

Figure 4.17: SPAAR deletion does not affect endothelial cell proliferation, mitochondrial metabolism nor global protein synthesis. **A**. Results of proliferation assay in basal and full endothelial medium indicate comparable proliferation rates of WT and SPAAR KO endothelial cells. N = 22 per condition from 3 independent cell isolations. **B**. Results of MTT assay show equivalent rates of mitochondrial metabolism in WT and SPAAR KO endothelial cells. N = 24 per condition from 3 independent cell isolations. **C**. Representative western blot of puromycin incorporation in WT and SPAAR KO endothelial cells. Starvation + AA denotes 3 h incubation in culture medium without amino acids followed by 1 h stimulation with a mixture of all 20 proteinogenic amino acids. **D**. Quantification of puromycin signal from C reveals no significant changes in protein synthesis between WT and SPAAR KO endothelial cells regardless of the treatment. N = 12 per condition from 4 independent cell isolations.

4.2.2.4 Transcriptome analysis of cardiac endothelial cells reveals differentially expressed genes in SPAAR KO cells

To study transcriptional changes in cardiac endothelial cells upon SPAAR depletion, whole transcriptome RNA-seq was performed with isolated WT and SPAAR KO endothelial cells. Subsequent differential gene expression analysis revealed 608 genes that were differentially expressed (adjusted p-value < 0.05) in SPAAR KO compared to WT cells. Of those, 25 were upregulated and 4 were downregulated in SPAAR KO cells compared to WT ($|\log_2$FC(KO vs WT)$|> 0.7$, **Figure 4.18 A**).

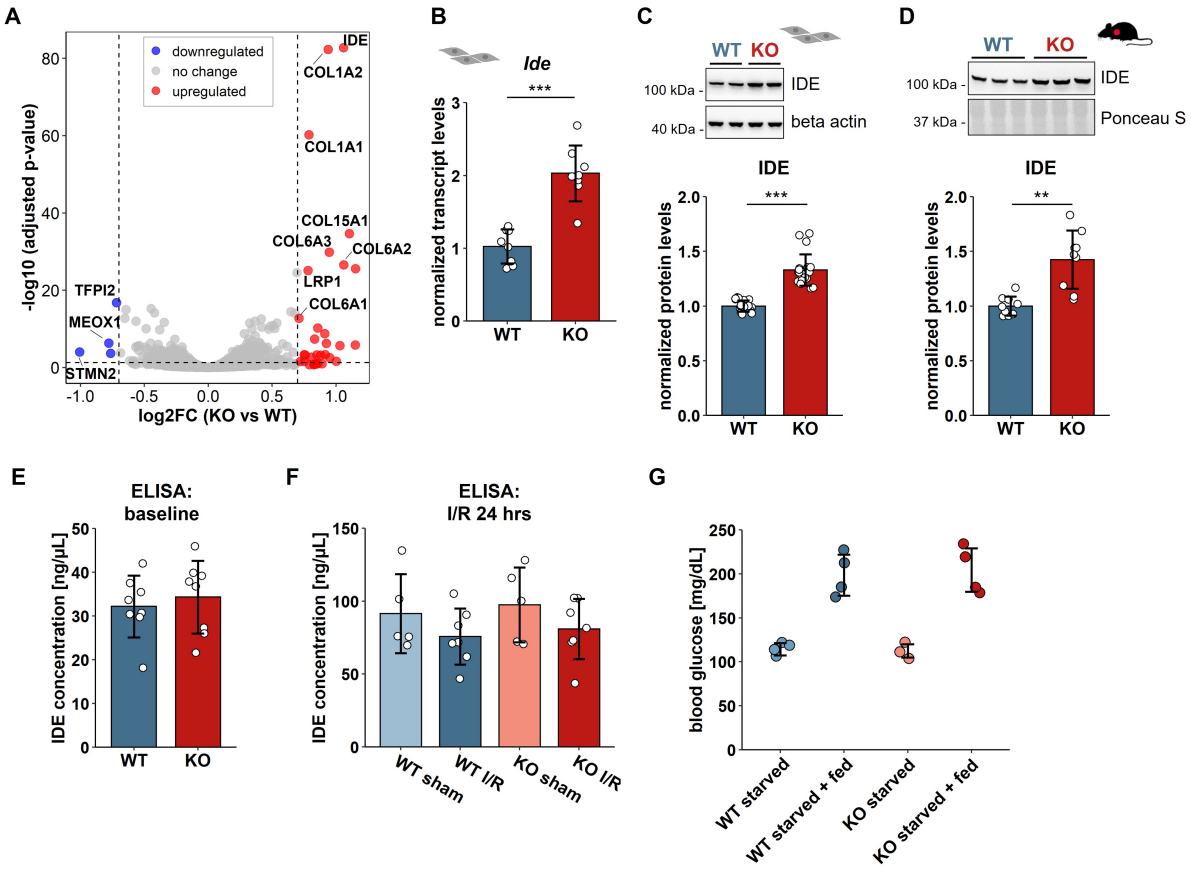

Figure 4.18: Transcriptome analysis identifies differentially expressed genes in SPAAR KO endothelial cells. **A**. Volcano plot showing the results of RNA-seq and differential gene expression analysis in WT and SPAAR KO endothelial cells. Applying a threshold of $|\log_2$FC(KO vs WT)$|> 0.7$ defines 25 upregulated genes (red) and 4 downregulated genes (blue) in SPAAR KO cells. **B**. Two-fold upregulation of *Ide* in SPAAR KO endothelial cells was confirmed by qRT-PCR. N = 8 per genotype from 2 independent cell isolations. *** $P < 0.001$ by T-test. **C and D**. IDE is also upregulated on protein level in SPAAR KO endothelial cells (N = 20 per genotype from 3 independent cell isolations) and mouse hearts (N = 9 per genotype). ** $P < 0.01$, *** $P < 0.001$ by Wilcoxon test. **E and F**. Results of ELISA indicate that IDE is not differentially secreted in SPAAR KO mice at baseline (N = 8 per genotype) nor after I/R injury (N = 5 for sham, 7 for I/R). **G**. Blood glucose levels do not differ between WT and SPAAR KO mice starved overnight, nor between WT and KO animals starved overnight and then fed *ad libitum* for 4 h, indicating comparable insulin activity in both genotypes. N = 3 – 4 per condition.

One of the most significantly upregulated genes was insulin degrading enzyme (IDE), which showed a two-fold increase in expression in SPAAR KO cells. IDE is a ubiquitously expressed peptidase that cleaves insulin, natriuretic peptides and amyloid beta peptides, and has been reported to be both secreted and localized intracellularly [140]. I first confirmed the upregulation of *Ide* on transcript level by qRT-PCR and further validated it on protein level in isolated SPAAR KO endothelial cells and whole hearts (**Figure 4.18 B-D**). However, assessing secreted IDE levels by performing ELISA on mouse blood serum samples did not show any significant differences between SPAAR KO and WT mice at baseline nor after I/R injury (**Figure 4.18 E, F**). I also checked for insulin clearance by measuring glucose levels in mouse blood serum, but observed no differences between SPAAR KO and WT animals after fasting or feeding (**Figure 4.18 G**). Collectively, these results indicate that increase in IDE in SPAAR KO mice is rather maintained intracellularly.

4.2.2.5 Angiogenic activity of cardiac endothelial cells is not influenced by SPAAR depletion

To better understand functional profiles of genes that were detected as differentially expressed in RNA-seq of cardiac endothelial cells, I performed gene ontology (GO) enrichment analysis. Annotation of biological processes showed an enrichment of terms related to development and formation of blood vessels (angiogenesis, **Figure 4.19 A**). Since endothelial cells crucially contribute to these processes, I performed two functional assays aimed at evaluating angiogenic activity of SPAAR KO endothelial cells.

First, scratch assay tests the ability of endothelial cells to migrate into a cell-free zone, created by a scratch in a confluent cell monolayer. Analysis of images taken over a course of 24 hours following scratch formation enables quantification of the scratch size. However, comparing these scratch sizes did not reveal any significant differences in cell migration between WT and SPAAR KO cells (**Figure 4.19 B, C**).

Second, tube formation assay measures the ability of endothelial cells to form capillary-like tubes on a basement membrane. Image analysis of these tubular structures with Angiogenesis Analyzer [40] provides various readouts, but none of them were significantly different when comparing WT and SPAAR KO cardiac endothelial cells (**Figure 4.19 D, E**).

Finally, to study whether angiogenic properties of SPAAR KO cells depend on the upregulation of IDE described in the section above, I repeated these assays with the addition of pharmacological IDE inhibitor 6bK [141]. Results of scratch assay were very comparable to the earlier ones and did not indicate any differences in cell migration between WT and SPAAR KO cells, not even in the presence of 6bK (**Figure 4.19 F**). Results of tube formation assay were rather preliminary due to low number of replicates, but suggest a role for IDE in tube formation. The addition of 6bK namely resulted in the formation of irregular tubular network with unusually large meshes (**Figure 4.19 G**). However, more replicates are required to ensure reproducibility and assess differences between WT and SPAAR KO cells treated with 6bK.

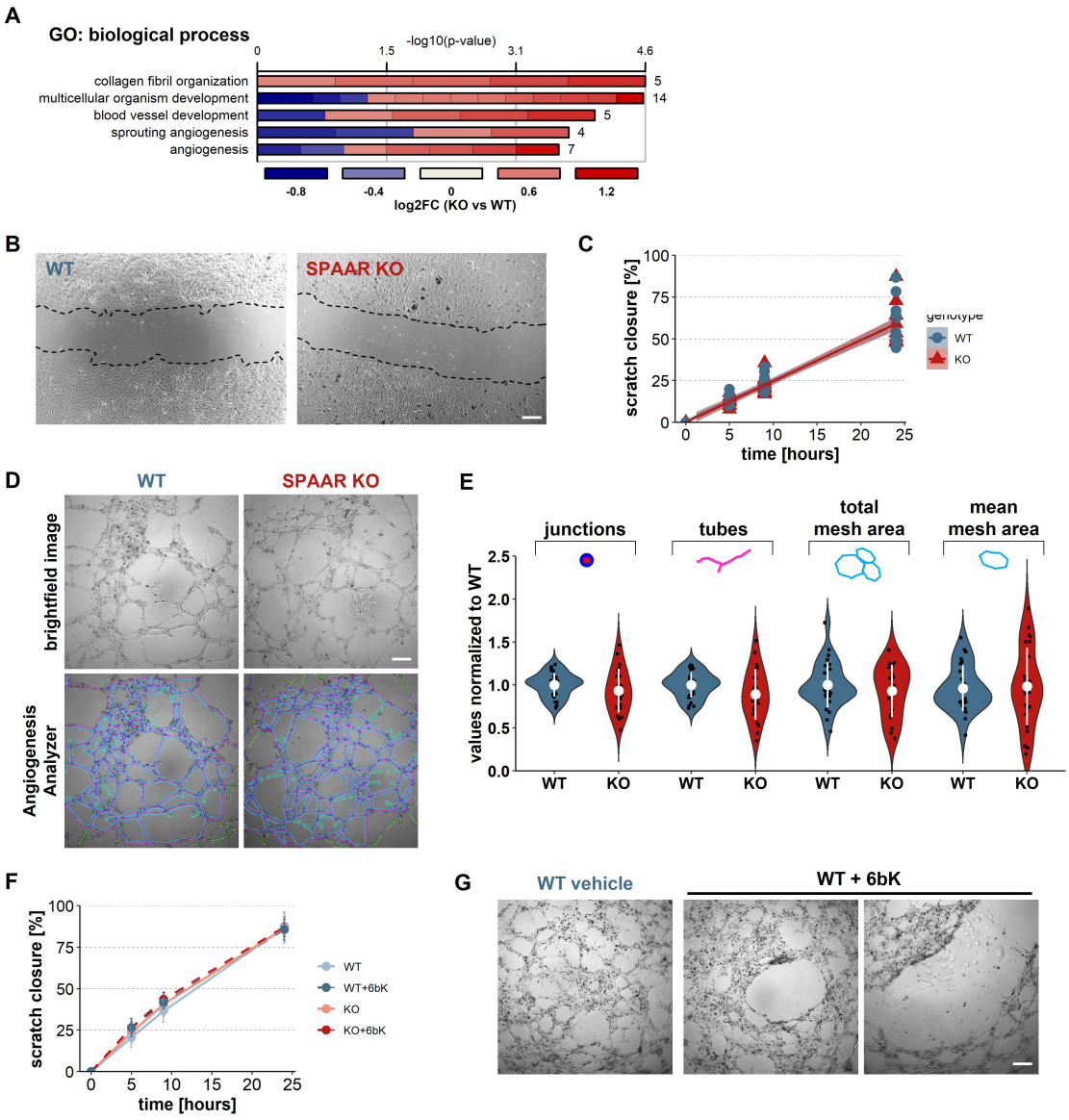

Figure 4.19: SPAAR depletion does not affect angiogenic activity of cardiac endothelial cells. **A**. GO enrichment analysis of differentially expressed genes in SPAAR KO endothelial cells (adjusted p-value < 0.05 and $|\log_2 FC(KO \text{ vs } WT)| > 0.5$) points to biological processes associated with angiogenesis. **B**. Representative brightfield images of scratch assay in WT and SPAAR KO endothelial cells. Border of the scratch is marked with a dashed outline. Scale bar: 200 µm. **C**. Quantification of scratch size over a course of 24 h shows equivalent migration rates of WT and SPAAR KO endothelial cells. N = 12 per genotype from 2 independent cell isolations. **D**. Representative brightfield images of tube formation assay in WT and SPAAR KO endothelial cells (top) and image analysis of tubular structures with Angiogenesis Analyzer (bottom). Scale bar: 200 µm. **E**. Angiogenesis Analyzer quantifies typical characteristics of a tubular network, such as junctions, tubes and mesh area. Comparison of these values between WT and SPAAR KO cells indicates highly similar tube formation. N = 25 per genotype from 3 independent cell isolations. **F**. Quantification of scratch size over a course of 24 h in the presence of IDE inhibitor 6bK does not reveal any differences in cell migration rates. N = 6 per condition. **G**. Tube formation after the addition of 6bK is impaired and exhibits unusually large meshes. Scale bar: 200 µm.

4.2.2.6 Exploring secretory functions of SPAAR KO endothelial cells

Interestingly, GO analysis of differentially expressed genes in RNA-seq of cardiac endothelial cells further showed a consistent enrichment of terms related to extracellular matrix (ECM) and secreted proteins (**Figure 4.20 A**). This led me to hypothesize that perhaps it is not cell-autonomous functions of SPAAR KO endothelial cells, but rather paracrine activity of these cells that lead to cardiac protection. Since SPAAR KO endothelial cells express more ECM constituents and collagens, they might provide local protective cues for cardiomyocytes and improve their survival or function.

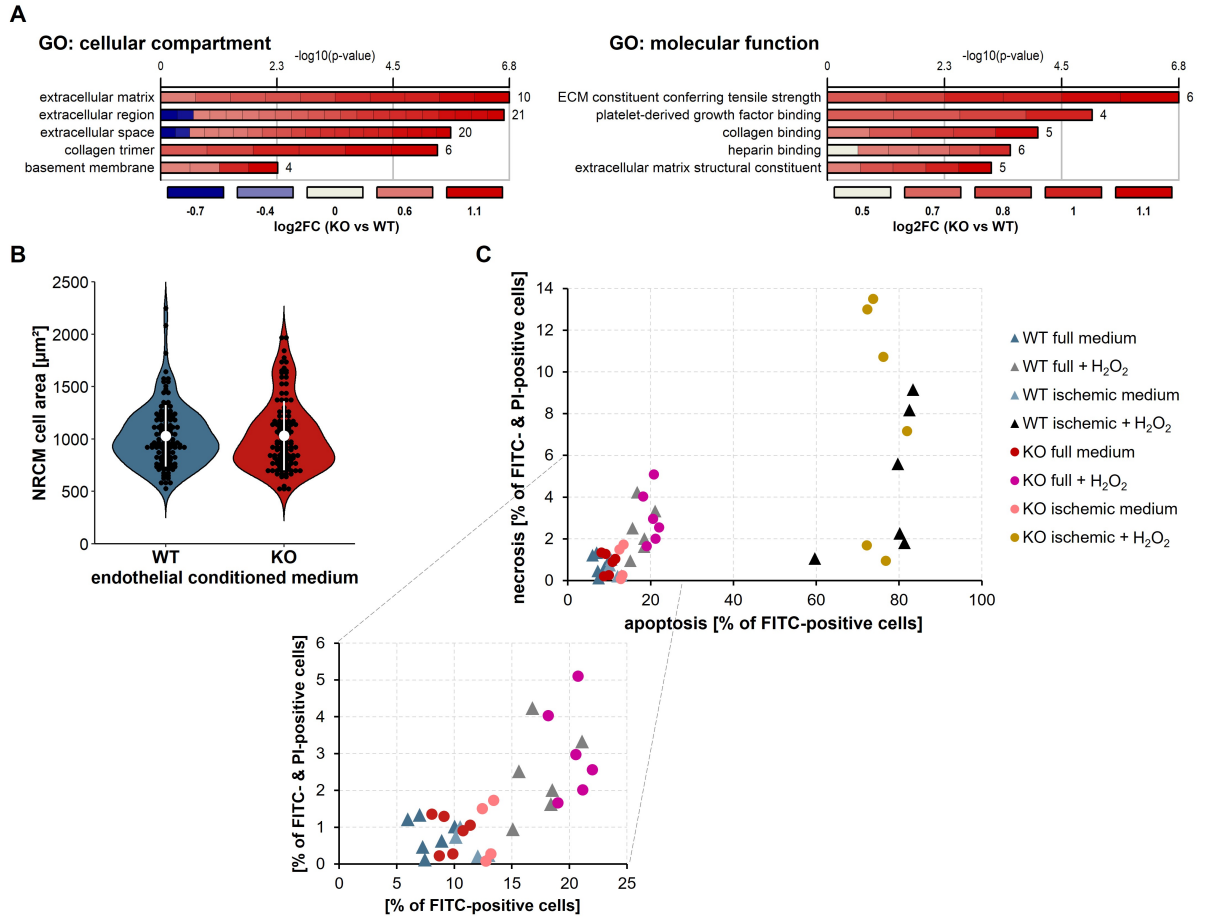

Figure 4.20: Conditioned media from WT and SPAAR KO endothelial cells do not influence NRCM size nor survival. A. GO analysis of differentially expressed genes in SPAAR KO endothelial cells (adjusted p-value < 0.05 and $|\log_2 FC(\text{KO vs WT})| > 0.5$) shows enrichment of terms associated with ECM and secreted components, implying the importance of paracrine functions of cardiac endothelial cells. **B**. Cell size analysis of NRCMs grown in conditioned full medium from WT and SPAAR KO endothelial cells for 24 hrs does not reveal any differences in their cell size. **C.** Flow cytometry analysis of apoptosis and necrosis in NRCMs treated with H_2O_2 for 4 h in conditioned full or ischemic endothelial medium. Quantification reveals the same percentage of apoptotic and necrotic NRCMs in WT and SPAAR KO endothelial conditioned medium. N = 4 – 6 per condition from 3 independent cell isolations.

To test this hypothesis, I used endothelial conditioned medium collected from endothelial cells cultured in full medium for 2 days or cultured in minimal ischemic medium and exposed to hypoxic conditions for 6 hours. These conditioned media were transferred onto isolated NRCMs that were additionally stressed by the addition of H_2O_2, which mimics cellular damage induced

by reactive oxygen species. Analysis of NRCMs kept in conditioned full medium from WT and SPAAR KO endothelial cells did not show any differences in their cell size (**Figure 4.20 B**). Next, I completed flow cytometry analysis of NRCMs to assess apoptosis and necrosis following H_2O_2 treatment in full or ischemic endothelial conditioned medium. Similarly, no differences could be observed in NRCM apoptosis or necrosis kept in WT or SPAAR KO conditioned medium in any of the experimental conditions (**Figure 4.20 C**).

Meanwhile, we also performed secretome analysis of cardiac endothelial cells (**Figure 4.21 A**). For this purpose, I cultured WT and SPAAR KO endothelial cells in growth medium containing amino acids labelled with intermediate or heavy stable isotopes to differentiate between the two genotypes. Additionally, the medium contained methionine homolog AHA, which was used in the downstream processing to enrich for proteins synthesized only during the labelling period of 24 h. Sample processing along with MS data acqusition and analysis to detect differentially secreted proteins was performed by our collaborator Tim C. Kuhn.

A single-shot MS analysis was sufficient to detect about 500 most abundant secreted proteins. Relative quantification to compare their secretion between WT and SPAAR KO endothelial cells pointed to a few differentially secreted proteins (**Figure 4.21 B**). GO analysis of these proteins confirmed their annotation and enrichment for various components of ECM, thereby validating the chosen approach to detect secreted proteins (**Figure 4.21 C**).

Initially, I focused on a candidate-based approach and selected COL5A1 for further validation. This collagen was one of the top upregulated hits in SPAAR KO cells and previously reported as beneficial for mechanical properties of the scar following myocardial infarction [142]. Inspecting my earlier results from RNA-seq of endothelial cells revealed that *Col5a1* is also significantly upregulated on transcript level in SPAAR KO cells, which is consistent with its increased secretion. However, I could not confirm these observations by performing qRT-PCR analysis on different batches of isolated endothelial cells, neither on mouse heart lysates (**Figure 4.21 D**). Furthermore, detecting COL5A1 with antibody-based approaches has proven to be very difficult, as I could not detect it by ELISA in endothelial conditioned medium and mouse blood serum, nor by western blotting of endothelial cell and mouse heart lysates (data not shown).

I then turned to other differentially secreted proteins to verify whether their transcript levels are consistent with differences in their secretion (**Figure 4.21 E**). qRT-PCR analysis revealed a significant decrease in *Col18a1* levels in SPAAR KO cells, concordant with a significant decrease in RNA-seq results and decreased secretion in these cells. Similarly, *Fkbp10* was significantly decreased in SPAAR KO cells analyzed by qRT-PCR and showed decreased secretion in SPAAR KO cells. On the other hand, GNB1 was downregulated exclusively by secretion, as there were no significant changes in its transcript levels in RNA-seq nor in qRT-PCR analysis.

Finally, I also looked at protein-protein interactions among differentially secreted proteins using STRING database [131]. The analysis predicted several functional associations and revealed two interacting clusters of differentially secreted proteins: collagen molecules and collagen chaperones (**Figure 4.21 E**). The latter include SERPINH1, an essential chaperone required for the assembly of typical triple-helical collagen structure [143]. This process is also supported by FKBP1A and FKBP10, which catalyze the isomerization of proline peptide bonds to assist triple-

helix formation in collagen [144]. Yet another differentially secreted protein is PLOD3, which catalyzes a critical step in collagen biosynthesis – the hydroxylation of lysine residues [145]. Considering these functional interactions and only partial experimental validation of selected targets, further research will benefit from addressing a broader collagen-interacting network.

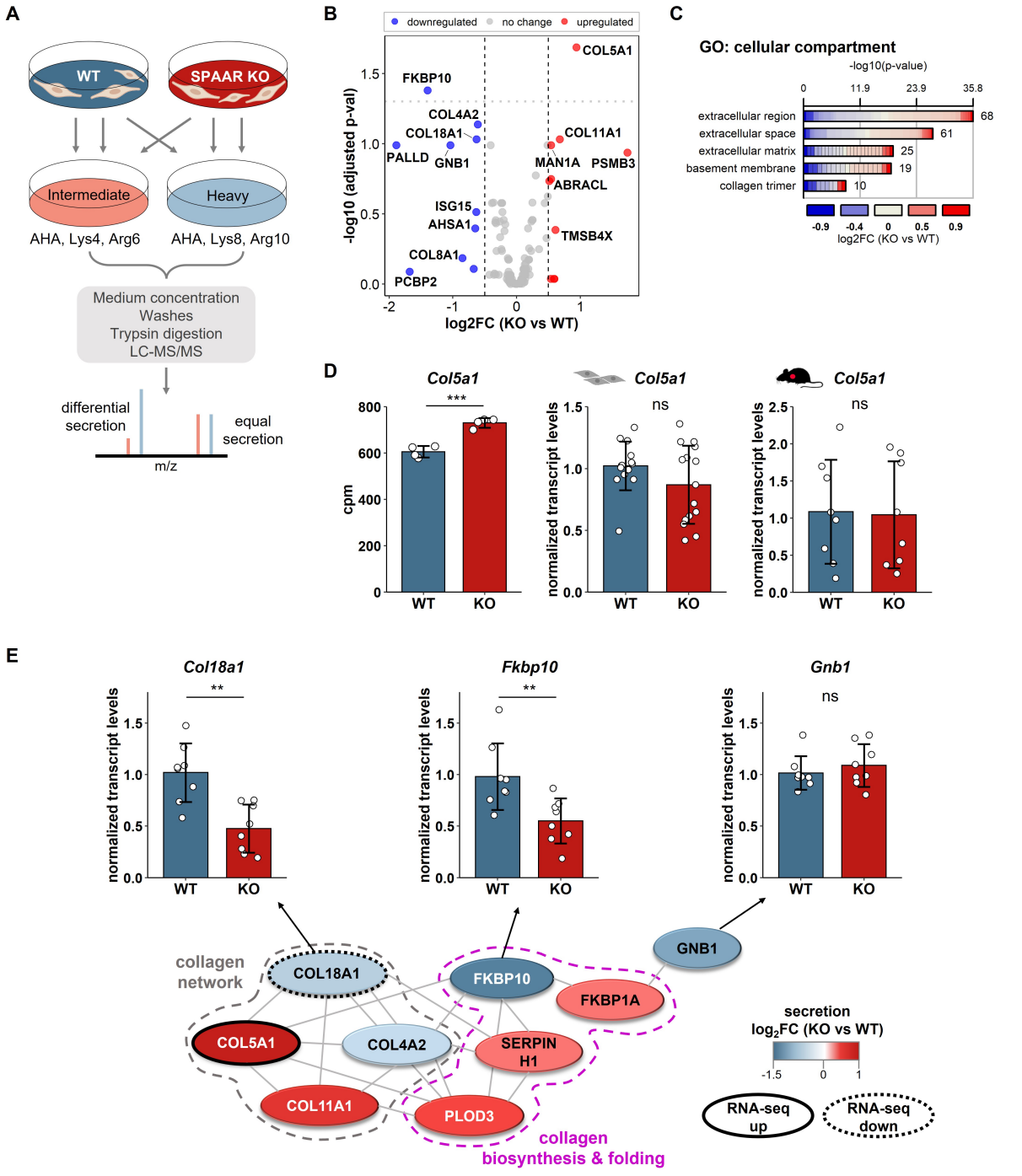

Figure 4.21: Secretome analysis of cardiac endothelial cells points to differential secretion of collagens in SPAAR KO cells. A. Workflow for secretome analysis of cardiac endothelial cells. WT and SPAAR KO cells were labelled with intermediate or heavy stable isotopes to differentiate between the two genotypes and AHA was added to enrich for newly synthesized proteins. Sample processing and MS analysis (light gray rectangle) were accomplished by Tim C. Kuhn, who also determined differentially secreted proteins. [*Figure caption continues on the next page.*]

B. Volcano plot showing differentially secreted proteins in WT and SPAAR KO endothelial cells. Applying a threshold of $|\log_2 FC(KO \text{ vs } WT)| > 0.5$ defines 10 proteins with increased secretion (red) and 10 proteins with decreased secretion (blue) in SPAAR KO cells. **C**. GO analysis of proteins secreted from cardiac endothelial cells shows enrichment of terms associated with ECM and collagens, which supports the utility of secretome analysis for detection of secreted proteins. **D**. Validation of COL5A1 as a selected target indicates coordinated upregulation of *Col5a1* transcript in SPAAR KO endothelial cells as shown by RNA-seq, *** adjusted p-value < 0.001 (left). qRT-PCR analyses of several batches of isolated endothelial cells (middle, N = 15 per genotype from 4 independent cell isolations) and mouse heart lysates (right, N = 10 per genotype) do not support these results and do not show any difference in *Col5a1* transcript levels between WT and KO mice. **E**. Network analysis of differentially secreted proteins reveals complex interactions between several collagen molecules (gray dashed outline) and collagen chaperones (purple dashed outline). Interaction network was drawn based on the results from STRING database [131]. Black solid outline indicates targets detected as differentially upregulated in RNA-seq of endothelial cells, whereas black dotted outline indicates transcriptionally downregulated targets. Arrows point to qRT-PCR analyses of isolated endothelial cells for 3 secretome targets: *Col18a*, *Fkbp10* and *Gnb1*. N = 8 per genotype from 4 independent cell isolations. ** $P < 0.01$ by T-test.

4.2.2.7 Summary of Specific focus 2: characterizing SPAAR depletion in cardiac endothelial cells identifies several potential targets that mediate its function

Since results of Specific focus 1 indicated a consistent cardioprotective phenotype in SPAAR KO mice following I/R injury *in vivo*, the aim of Specific focus 2 was to follow up on these findings and decipher mechanistic insights by which SPAAR depletion offers protective advantage. As SPAAR is predominantly expressed in cardiac endothelial cells, I focused on evaluating its contribution to endothelial cell phenotype and activity. Noteworthy are the following 3 observations:

1. increased phosphorylation of TFEB at Ser122 in SPAAR KO cells at baseline and following amino acid starvation and stimulation. This modification does not influence the transcriptional activity of TFEB and further studies are required to describe its functional consequences.

2. Increased expression of IDE in SPAAR KO endothelial cells and mouse hearts. Since IDE has not been studied in the cardiac setting, I cannot speculate how its upregulation would contribute to cardioprotection. Preliminary results of tube formation assay with IDE inhibitor 6bK indicate that this enzyme might be required for proper structure of endothelial tubular network.

3. Differential expression of collagen molecules and collagen-regulating chaperones in two omics datasets highlights the importance of secretory functions of cardiac endothelial cells. SPAAR KO cells show distinct transcriptional and secretory regulation of selected collagens and other extracellular constituents, which might provide local cues for endothelial and other cardiac cell types.

5 Discussion and future directions

The main focus of this thesis were sORFs - continuous DNA stretches defined with an initiation codon and in-frame stop codon, which encode a maximum of 100 amino acids. Due to their small size, sORFs were frequently overlooked by gene prediction pipelines and reference annotation projects. However, recent years have witnessed substantial advances in computational approaches and mass spectrometry-based methods, as well as the development of ribosome profiling. These methods have facilitated systematic discovery of millions of sORFs in different species and increasing number of studies are reporting the presence and functional roles of novel sORF-encoded microproteins.

5.1 Identifying novel cardiac microproteins

Motivated by the existence and high prevalence of sORFs, I aimed to identify them in the data previously generated by using cardiomyocyte-specific ribosome profiling in mouse hearts [36, 112]. This refined protocol enabled the investigation of translational processes specific to cardiac muscle tissue and predicted the translation of over 8000 sORFs under physiological and pathological conditions. I hypothesized that this extensive collection of translated sORFs could contain novel, yet uncharacterized cardiac microproteins that are involved in heart function or disease. The first part of this thesis therefore focused on the **identification of biologically functional microproteins expressed in the cardiac muscle**. Initially, special focus was placed on novel, previously unannotated microproteins with the goal of **experimentally characterizing them in the myocardium**.

To assemble a list of potential candidates for further characterization, I needed to downsize the number of mouse sORFs predicted to be translated. For this purpose, the existence of sORF-encoded microproteins was validated using published proteomics data from healthy and hypertrophic mouse hearts [132]. This approach confirmed the **existence of stable peptides for only about 20 % of predicted sORFs**, and just 2.5 % of all predictions were matched with peptides corresponding to novel, previously uncharacaterized microproteins. These low numbers are not too surprising, as other studies also report between 10 and 40 % of investigated sORFs to have MS-based experimental support [3, 41]. Notably, a recent comparative study of seven datasets with Ribo-seq data from human cell lines and tissues found only 8 % of predicted Ribo-seq sORFs to have unique peptide evidence [7]. This suggests that **Ribo-seq has increased sensitivity for discovery of sORFs** and can be used to identify rare translation events, which are not readily detectable by MS-based methods. There are several further reasons why microproteins are infrequently detected by MS. Besides low abundance, microproteins are more susceptible to protease digestion and rapid degradation. Their small sizes make them less likely

to be detected by standard MS approaches, as they may not generate observable tryptic peptides [38]. Consequently, this also leads to their exclusion from proteome annotation projects, as most microproteins do not meet the strict guidelines requiring at least two unique non-overlapping peptides detected by MS [146]. Overall, these findings and considerations illustrate the **need for combining multiple independent methods to support reliable sORF discovery**.

For my next step I therefore retained only sORFs with peptide evidence and focused on those encoded by presumable lncRNAs and novel transcripts. Evaluation of their evolutionary conservation showed that none of these candidates corresponds to novel transcripts with preserved sORFs in *H. sapiens*. Therefore, to overcome the problem of finding mouse-specific novel sORFs, I prepared Ribo-seq libraries from diseased human heart biopsies. Subsequent conservation analysis revealed a conserved microprotein of 27 amino acids, which is located on a novel transcript expressed and translated in human hearts. Unfortunately, this sORF is not translated in the mouse heart, which restricts the experimental characterization to performing all experiments *in vitro*, for example in human iPSC-derived cardiomyocytes. Collectively, these results substantiate that **novel cardiac microproteins are singular occurrences, which are unlikely to be conserved** across vertebrates or even mammals. Again, these seemingly disappointing results are not unanticipated in the light of recent findings. The above-mentioned comparative study of seven Ribo-seq datasets detected over 7000 translated sORFs, of which mere 2.4 % showed potential protein-coding capacity (as identified by PhyloCSF) [7]. Additionally, more studies are coming forward with the finding that **sORFs are often species-specific and encode evolutionary young microproteins**, that emerged *de novo* from noncoding sequences. Although such cases would be missed by traditional methods based on coding potential assessment, they contribute to observable phenotypes and engage in various cellular processes [28, 29, 147].

In hindsight, I consider the initial criteria set to search for novel cardiac microproteins far too stringent. Consequently, requiring *novel* microproteins to be conserved (at least) across placental mammals, as well as enriched and translated in the heart tissue of these organisms did not yield any candidates for further studies. Therefore, I focused on *annotated* protein-coding transcripts, which encode **small peptides with previously uncharacterized functions in the heart**. This led me to the identification of **SMIM4**, a conserved and heart-enriched transmembrane microprotein, which I considered a promising candidate for further characterization. Results of MTT assay showed that SMIM4 is **required for NRCM viability and mitochondrial metabolism** and cell size measurements pointed to the role of SMIM4 in **regulating cardiomyocyte growth**. These results are almost identical to those of Dennerlein and colleagues, who observed a similar, almost 50 % reduction in cell viability upon SMIM4 knockdown *in vitro* [48]. Their findings corroborate and help explain the results described here, as they identified SMIM4 as an early assembly factor of complex III of ETC. This was confirmed by a further mechanistic study detailing SMIM4 involvement in the stabilization of cytochrome b in complex III [49]. Although these studies clarified the mechanism of SMIM4 function, collaborators from the group of Prof. Dr. Freichel are developing a SMIM4 KO mouse model using CRISPR/Cas9 technology. Possible next step would thus involve characterizing heart phenotype and function of these KO mice to provide insight into SMIM4 contribution to cardiac (patho)physiology.

5.2 Functional characterization of cardiac-enriched microprotein SPAAR

The second part of this thesis was focused on SPAAR – a recently discovered microprotein, which is **evolutionary conserved across mammals** [10, 53]. Most species possess a short isoform of SPAAR with 75 codons, only primates express an additional longer isoform with 90 codons, which emerged *de novo* from nongenic sequences. SPAAR contains an N-terminal transmembrane domain and localizes to the lysosomal membrane, where it interacts with vacuolar ATPase and **hinders mTORC1 activation**. This inhibitory effect is very specific and only apparent after amino acid starvation and restimulation. Initially, SPAAR was described as a muscle-enriched microprotein, whose function was characterized in skeletal muscle tissue [10]. However, having thoroughly examined different multiple-organ expression datasets I learnt that *Spaar* is in fact **strongly enriched in the cardiac tissue**. Its function in the heart is further underscored by downregulation of *Spaar* transcript in pathological cardiac conditions, both in human patients and mouse models of heart disease [41, 54].

Importantly, these results report on the dysregulation of *Spaar* transcript, rather than the SPAAR peptide. This distinction is especially noteworthy because transcripts encoding sORFs have been reported to have dual functions: noncoding ones mediated by the transcript and coding ones mediated by the translated microprotein [148]. Indeed, a study by Spencer and colleagues suggests that *Spaar* locus produces two endothelial-enriched molecules, ***Spaar* transcript and SPAAR peptide**, which have **opposing effects on angiogenesis** [149]. The transcript inhibits vessel formation, whereas the peptide exhibits pro-angiogenic properties as measured by tube formation assay. Furthermore, pull-downs coupled with mass spectrometry revealed that both the transcript and the peptide interact with actin binding proteins. *Spaar* transcript was shown to interact with thymosin beta 4, cofilin-1 and profilin-1, which are all associated with reorganization of actin cytoskeleton. SPAAR peptide was bound to nesprin-1, which functions as a linker between cell nucleus and actin cytoskeleton. The authors therefore suggested that opposing roles of *Spaar* transcript and peptide in vessel formation could be attributed to differential regulation of actin cytoskeleton dynamics.

Since my main interest and the focus of this thesis lie in the characterization of microproteins, the aim of the second part of the project was to study specifically the role of SPAAR peptide. For this purpose, I used a previously established global SPAAR KO mouse model, in which the first ATG codon of *Spaar* was deleted using CRISPR/Cas9 [10]. Consequently, these KO mice do not show any evidence of SPAAR translation as demonstrated by SPAAR immunoprecipitation and western blotting, nor by analyzing their translatomes with Ribo-seq. Since the chosen KO model preserves the integrity of the host transcript, *Spaar* transcript levels remain unchanged in SPAAR KO mice as confirmed by qRT-PCR and RNA-seq. Therefore, this KO mouse model is particularly well-suited for studying the specific effects of SPAAR peptide.

Considering the enrichment of *Spaar* in heart tissue, I focused on characterizing **SPAAR function in shaping the cardiac phenotype**. To comprehensively address this research question, two focus areas were defined. The first one was aimed at describing the overall heart phenotype by performing *in vivo* experiments with SPAAR KO mice, while the second one aimed to provide

a more detailed understanding of the mechanisms underlying SPAAR function *in vitro*.

5.2.1 Specific focus 1: SPAAR depletion is cardioprotective after I/R *in vivo*

I began cardiac phenotyping of SPAAR KO mice by performing ultrasound echocardiography in male adult mice at baseline. The analysis did not reveal any differences in cardiac dimensions or heart function between WT and SPAAR KO mice, which contradicts some of the previous results obtained by Spiroski and colleagues. They performed a similar cardiac assessment in *Spaar* KO mouse model, in which the entire *Spaar* locus is deleted, thereby preventing the expression of transcript and peptide alike [149]. They found smaller left ventricular systolic and diastolic areas and decreased end-systolic and diastolic volumes in *Spaar* KO female mice compared to WT [150]. Despite being significantly decreased, these changes were so small that they did not affect the overall heart function at baseline, leading to unchanged cardiac output and ejection fraction in *Spaar* KO mice. It is possible that observing these minimal differences in cardiac dimensions stems from using a different KO mouse model and including only female mice into the analysis, whereas my experiments were conducted exclusively on male animals.

The **absence of a distinct (cardiac) phenotype in SPAAR KO mice at steady state** is not surprising. Other studies using loss-of-function approaches to address the biological relevance of microproteins have noted similar results and failed to observe dramatic phenotypes at baseline [45, 46]. Many microproteins, including SPAAR, exert their functions through interactions with other, often larger proteins and thus make precise adjustments to complex processes. These can be complemented through alternative pathways and stressing the model system might be necessary to reveal the functional role of the microprotein [42].

Therefore, SPAAR KO mice were exposed to **cardiac injury induced by ischemia - reperfusion**. Ischemia is characterized by reduced blood flow, shortage of nutrients and oxygen, and low mTORC1 activity. The ensuing reperfusion phase is intended to re-establish blood supply in ischemic cardiac tissue and features high mTORC1 activity. Given the adaptive and protective mechanisms of mTOR in this process [55], along with fluctuating amino acid levels (that elicit SPAAR function *in vitro*), I reasoned that I/R injury would represent a relevant stress model *in vivo*.

Surprisingly, already 24 hours after induction of I/R **SPAAR KO mice exhibited a protective response**. They displayed **decreased serum levels of cardiac troponin T**, a standard marker for myocardial damage, as compared to WT mice. This protective effect continued at later time points, so that 2 weeks post-surgery SPAAR KO mice showed **sustained left ventricular ejection fractions**, whereas WT mice developed systolic impairment of heart function with reduced ejection fraction. SPAAR KO mice also exhibited reduced collagen deposits and **smaller fibrotic scars**. My results thus point to cardioprotection in male SPAAR KO mice following I/R *in vivo*.

This conclusion differs from observations of Spiroski and colleagues, who reported increased area at risk in male *Spaar* KO mice following an acute, short period of ischemia [150]. Yet after prolonged ischemia (1 week of permanent occlusion of LAD artery), they did not observe any differences in cardiac dimensions and function between WT and *Spaar* KO female mice. It is worth noting that their model of cardiac injury is more severe, as it features a permanent ligation of LAD artery. Instead, the model used in my project is based on a temporary ligation of LAD artery (for 60 min), hence the results are not directly comparable. Strangely, the authors also do not comment on why they chose to separate genders between different experiments and readouts by using male mice for acute damage and female mice for chronic injury. This might prove important for the cardiac setting addressed here, because sex differences have been well established in models of myocardial injury. Female mice are frequently protected from cardiac damage induced by temporary or permanent coronary ligation, while male mice show higher myocardial rupture, increased hypertrophy and dilation and worsened heart dysfunction [151, 152]. The discrepancies between cardiac phenotypes of KO mice after injury could thus be partially attributed to sex differences, as my experiments were carried out exclusively on male mice. The *Spaar* KO model used by Spiroski and colleagues further complicates data interpretation. Since it features a global deletion of the entire *Spaar* locus, it is impossible to dissect whether the increase in acute damage in *Spaar* KO mice is a consequence of deleted transcript or peptide. Also, previously reported antagonizing effects of the two molecules might simply cancel each other out and lead to the absence of overt phenotype in *Spaar* KO mice after 1 week of ischemia.

Therefore, it is more reasonable to compare my results to those of Matsumoto and colleagues, as they generated SPAAR KO mice with peptide depletion and I used this KO model throughout my project. Although their *in vivo* experiments focused on a different tissue and were based on **toxin-induced injury in skeletal muscle**, they also revealed **protective effect in SPAAR KO mice** [10]. Following acute muscle injury, SPAAR KO mice showed increased mTORC1 activation, higher proliferation capacity and faster maturation of myofibers and overall improvement in muscle regeneration.

Matsumoto and colleagues also noted an increase in size of regenerating SPAAR KO myofibers, which is in line with the role of mTORC1 in supporting cell growth. Their observation was specific to injured and regenerating muscle, as no such increase was observed at steady state. In my experiments, cell size measurements were carried out on cardiomyocytes with two independent methods: one enabled measurements of cell width and length of isolated cardiomyocytes, whereas the second method was based on determining their cross-sectional area in heart tissue sections. Both measurements revealed a consistent result, namely that **SPAAR KO mice have slightly increased cardiomyocyte width and area already at baseline**. However, this small increase does not contribute to overall gain in heart weight or increased heart weight/body weight ratio. This discrepancy challenges the biological significance of increased cardiomyocyte size in SPAAR KO mice, as opposed to mere mathematical significance of cell size measurements. Since SPAAR depletion was previously shown to activate mTORC1 only under stress conditions, the observed increase in cell size at baseline cannot be easily explained through the effect of SPAAR on mTOR pathway. Interestingly, adding the stress element in the form of I/R

injury *in vivo* only elicits a response in WT mice. These show a modest, but significant increase in cardiomyocyte area consistent with previous reports using the same model of cardiac injury [153], whereas SPAAR KO mice do not show any further increase in cardiomyocyte size after I/R.

Although my results after cardiac I/R injury and those previously published by Matsumoto and colleagues concur on the overall protective effect of SPAAR KO, they remain incomplete. Combining these conclusions namely suggests that SPAAR deficiency is beneficial and confers protective advantage following cardiac and skeletal muscle damage, but SPAAR function at steady state remains elusive. Since SPAAR has existed at least since the emergence of mammals and is evolving under purifying selection at amino acid sequence and protein structure levels, it most likely **represents a functional coding gene** [53]. It thus seems improbable that only its depletion can be favourable under stress conditions. Consistent with the role of SPAAR in mTORC1 signalling, the choice of cardiac stress might be crucial here. Depending on the type of cardiac insult, mTORC1 is differentially activated to elicit distinct (mal)adaptive mechanisms, culminating in divergent heart phenotypes [55]. Further investigation is thus required to dissect the complete function of SPAAR in other disease conditions, as well as at baseline.

I addressed this topic in the second focus area, by studying the underlying mechanisms of SPAAR function *in vitro*. A crucial indication for this part of experimental work was provided by advances in single-cell transcriptomics, which uncovered the intricate spatial expression pattern of SPAAR. As previously reported, the transcript encoding **SPAAR is enriched in the cardiac and skeletal muscle**, but its expression is also detected in other organs, such as lungs, placenta, pancreas and adipose tissue. Although phenotypically and functionally diverse, these tissues are all highly vascularized and metabolically active. Using two recently generated single-cell transcriptomic atlases of human and mouse organs enabled me to compare cell types shared across tissues, such as endothelial, epithelial, stromal and immune cells from various anatomical locations [116, 154]. This comparison revealed a striking cell type-specific expression pattern, whereby ***SPAAR* is strongly expressed in capillary endothelial cells of highly vascular tissues**, such as heart, muscle, adipose and pancreatic endothelial cells. Since the results from the first focus area *in vivo* indicated a cardiac protective phenotype in SPAAR KO mice, I followed up on them by performing more detailed experiments in primary cardiac endothelial cells isolated from WT and SPAAR KO mice.

5.2.2 Specific focus 2: studying SPAAR in cardiac endothelial cells identifies several potential targets that mediate its function

Although I performed a variety of *in vitro* experiments to understand the functional role of SPAAR, I would like to highlight the following three most consistent observations in SPAAR KO endothelial cells: increased phosphorylation of TFEB, increased expression of IDE and differential production and secretion of ECM components.

5.2.2.1 Increased phosphorylation of TFEB in SPAAR KO endothelial cells

Initial report classified SPAAR as an mTORC1 inhibitor based on the observation that SPAAR overexpression leads to a decrease in the phosphorylation of canonical mTORC1 targets, such as S6K, RPS6 and 4E-BP, after amino acid stimulation. Conversely, SPAAR KO led to increased phosphorylation of these targets [10]. However, canonical mTORC1 signalling was not affected by SPAAR in my experiments conducted with cardiac endothelial cells. Instead, I observed an **increase in the phosphorylation of TFEB at Ser122 in SPAAR KO endothelial cells**. This finding suggests that **SPAAR is involved in the substrate-specific activation of mTORC1** and is supported by several recent studies [79, 155, 156]. Emerging evidence is challenging the longstanding view of mTORC1 as an "on or off" kinase that in response to favourable stimuli phosphorylates all its substrates without any bias. Recent work rather suggests that certain conditions can induce selective phosphorylation of specific mTOR targets. This is referred to as non-canonical mTORC1 signalling and impinges on the activation of RagC/D to selectively phosphorylate factors of MiT/TFE family, such as TFEB and TFE3.

Molecular details by which SPAAR affects selective phosphorylation of TFEB were not experimentally addressed in this thesis, but previous work established that SPAAR localizes to the lysosomal membrane, where it interacts with several subunits of lysosomal v-ATPase [10]. During amino acid starvation, v-ATPase maintains a tight complex with Ragulator and Rags. When amino acids accumulate, v-ATPase activates Ragulator, which in turn promotes the activity of RagA/B through weakened interactions and results in the recruitment of mTORC1 [61, 62]. Immunoprecipitations in HEK293T cells showed that **SPAAR interferes with interactions between v-ATPase, Ragulator and Rags at the lysosomal membrane** and impedes favorable conformational changes, leading to reduced mTORC1 recruitment and activity [10]. This result explains the differences in phosphorylation of canonical mTORC1 targets observed in the initial report on SPAAR, but it does not explain the specificity for TFEB phosphorylation I detected in cardiac endothelial cells. Recent structural studies namely suggest that **TFEB phosphorylation strictly requires two active RagC molecules** and stresses the importance for RagC GTPase activating protein folliculin (FLCN) to ensure the active, GDP-bound conformation of RagC [157]. Careful revision of my previous experiments did not reveal any differences in FLCN expression on transcript nor protein level between WT and SPAAR KO endothelial cells (data not shown). However, three-dimensional structures of various lysosomal complexes show that FLCN activity is mediated by elaborate conformational changes. These are brought about by nutrient-activated SLC38A9, a lysosomal amino acid transporter, which competes with FLCN and displaces it from its binding cleft between Rag dimers, thus promoting FLCN activity and active RagC conformation [158, 159]. Since SPAAR is a lysosomal transmembrane protein itself, it might form a part of these large protein complexes at the lysosomal membrane and cooperate in regulating the activity of RagC/D. The small size of SPAAR seems particularly beneficial in this case, as it enables it to fit into very small gaps and thus fine-tune the intricate FLCN-RagC/D-mTORC1-TFEB signalling axis. This hypothesis could be tested by performing immunoprecipitation experiments in cardiac endothelial cells. During amino acid starvation, Ragulator, Rags and FLCN exist in a tight lysosomal folliculin complex (LFC) [158]. Upon amino acid stimulation, FLCN dissociates to promote active RagC conformation. Ac-

cordingly, immunoprecipitating Rags from endothelial cells stimulated with amino acids and subsequently detecting FLCN would be expected to show decreased interaction between FLCN and Rags in SPAAR KO cells compared to WT. This would hint at higher FLCN and RagC activity in SPAAR KO cells, which would explain the specific increase in TFEB phosphorylation. In addition, amino acid starvation in endothelial cells could be followed by specific stimulation with arginine, which activates SLC38A9 and promotes its conformational changes [159]. If SPAAR interacts with and regulates structural rearrangements in SLC38A9, arginine-specific stimulation would result in differential phosphorylation of TFEB in SPAAR KO cells.

Next question that emerges is the **functional role of increased TFEB phosphorylation specifically at Ser122** in SPAAR KO cells. Most studies investigating the phosphorylation and concomitant localization of TFEB focus on the assessment of Ser211, a conserved residue that binds cytosolic chaperones 14-3-3 once it is phosphorylated. This interaction retains TFEB in the cytoplasm and prevents its translocation to the nucleus, where it would otherwise transcribe genes involved in lysosomal biogenesis and autophagy [73, 74]. While Ser122 also represents a conserved residue regulated in an mTORC1-dependent manner, the mechanisms by which it affects TFEB localization are not known [77]. Since Ser122 is distant to 14-3-3 binding motif of TFEB, it is unlikely to be regulated the same way as Ser211. One explanation could be that phosphorylation at Ser122 results in the binding of other proteins, which regulate the shuttling of TFEB between cytoplasm and nucleus. In addition, these interaction partners could impact the stability of TFEB. This is exemplified by phosphorylation at another serine residue of TFEB (Ser142), which mediates binding to E3 ubiquitin ligase STUB1, resulting in TFEB ubiquitination and targeting for proteasomal degradation [160].

Finally, another layer of complexity is added by **tissue- and compartment-specific functions of TFEB**. Initially described as an oncogene, TFEB was found to be upregulated in a set of renal carcinomas and later found to be overexpressed in many cancer types. It is also known as a pro-angiogenic factor, required for proper vascularization both in embryonic development and in adult life [161]. Moreover, TFEB positively regulates angiogenesis after ischemic insult. Endothelial-specific overexpression of TFEB was shown to promote angiogenesis and improve blood flow in a mouse model of hindlimb ischemia, whereas endothelial-specific TFEB KO resulted in attenuated vessel formation and decreased blood flow. The described pro-angiogenic effect of TFEB was dependent on the activation of AMPK signalling and increased expression of lysosomal and autophagy genes [162]. Other studies highlighted an anti-inflammatory effect of TFEB and implicated its role in regulation of endothelial cell cycle, both of which function through lysosomal- and autophagy-independent mechanisms [161, 163]. The emerging results from the vascular compartment thus illustrate that **TFEB represents an unexplored hub of transcriptional networks**, which are likely regulated in a cell- and context-specific manner. Combined with highly restricted expression of SPAAR in cardiac endothelial cells and its role in TFEB phosphorylation, the function of this transcription factor necessitates further investigation in heart tissue and specifically in cardiac endothelial cells.

5.2.2.2 Increased expression of IDE in SPAAR KO endothelial cells

To identify additional targets and mechanisms that would help me decipher SPAAR function, whole transcriptome RNA-seq was performed with isolated WT and SPAAR KO endothelial cells. Subsequent differential expression analysis revealed that transcriptional changes in SPAAR-depleted cells are rare and pointed to a few significantly upregulated genes. One of them was IDE, with a **two-fold increase in transcript levels in SPAAR KO cells and concomitant increase on protein level**, both in isolated endothelial cells and heart lysates from SPAAR KO mice.

Despite the substrate specificity suggested by its name, IDE is a ubiquitously expressed zinc metallopeptidase that cleaves not only insulin, but many other hormones (glucagon, amylin), growth factors (insulin-like growth factor 2), neuropeptides, cytokines and amyloid beta peptides [164]. It has been highly investigated in the context of metabolic and neurodegenerative diseases, where its modulation could represent a potential therapeutic option, for example by treating diabetes with pharmacological inhibitors of IDE to slow insulin clearance [165]. In my experiments, comparing SPAAR KO and WT mice (after fasting or feeding) revealed no difference in insulin clearance as indicated by comparable blood glucose levels in both genotypes.

The subcellular localization of IDE is just as broad as its list of substrates, as it has been found in the cytoplasm, mitochondria, peroxisomes and also in extracellular space, despite the lack of conventional signal peptide [166]. However, the results of ELISA did not indicate any significant differences in secreted IDE levels between SPAAR KO and WT mice at baseline nor after I/R injury. In conclusion, all my results indicate that **increase in IDE in SPAAR KO mouse hearts is maintained rather intracellularly**.

The immediate next question is **the functional role of IDE in SPAAR KO cardiac endothelial cells** and whether it contributes to the cardioprotection I observed in SPAAR KO mice *in vivo*. Unfortunately, no study to date has addressed the relevance of IDE in the heart or in cardiovascular diseases. My preliminary experiments with IDE inhibitor 6bK in cardiac endothelial cells did not show any difference in cell migration between WT and SPAAR KO cells. However, the addition of 6bK caused the formation of irregular tubular network with unusually large meshes, which suggests a role for IDE in tube formation. In this case, more replicates with appropriate dosage are required to compare the differences and to establish whether SPAAR KO cells are partially protected from the anti-angiogenic effect of IDE inhibition. A more time-consuming approach would involve using 6bK in mice *in vivo*. Administering 6bK prior to I/R surgery, followed by the assessment of heart phenotype (e.g. ejection fraction and fibrosis) could clarify whether the catalytic activity of IDE mediates the cardioprotective effects I noted previously. Importantly, even inhibitor-based approaches could prove to be ineffective, as previous studies identified catalytically independent functions of IDE mediated through its chaperone-like activity or its protein binding partners (e.g. proteasomal subunits) [140, 167].

5.2.2.3 Differential production and secretion of ECM components in SPAAR KO endothelial cells

Besides upregulation of IDE, RNA-seq analysis of isolated WT and SPAAR KO cardiac endothelial cells revealed differential expression of several other genes. These were used to perform GO enrichment analysis to identify broader gene signatures. All aspects of GO (biological process, molecular function and cellular component) were enriched for terms associated with extracellular constituents, secreted proteins (such as collagens) and extracellular region. RNA-seq results were then validated by qRT-PCR on a panel of selected genes, which confirmed **differential regulation of extracellular components in SPAAR KO endothelial cells**. Furthermore, secretome analysis of cardiac endothelial cells corroborated these results and showed **differential secretion of several collagen molecules**. Inspecting their protein interactions *in silico* suggested that collagens form functional associations with other differentially secreted proteins, many of which function as collagen chaperones to support collagen biosynthesis and maturation.

Based on these results I hypothesized that SPAAR KO endothelial cells might exhibit distinct paracrine functions and secrete local protective molecules. These could potentially improve the survival and function of cardiomyocytes and contribute to cardioprotection *in vivo*. Indeed, the top upregulated hit in the secretome of SPAAR KO cells was COL5A1, which was found to limit adverse cardiac remodelling following myocardial infarction [142]. However, I could not validate its upregulation by qRT-PCR in endothelial cells, nor could I confirm its secretion by ELISA. Instead, I adopted a more general approach for investigating protective paracrine functions of SPAAR KO endothelial cells. I performed flow cytometry analysis for assessing the apoptosis of cardiomyocytes treated with endothelial-conditioned medium from WT or SPAAR KO cells. But even this experiment did not show any differences; cardiomyocyte apoptosis was highly comparable between conditioned media from WT and SPAAR KO cells, even when NRCMs were additionally stressed by H_2O_2. Also, cardiomyocyte size did not differ when they were grown in endothelial-conditioned medium from WT or SPAAR KO cells. Altogether, these findings **do not support a beneficial paracrine effect of SPAAR KO endothelial cells** on cardiomyocyte survival and growth.

Interestingly, **collagen production by cardiac endothelial cells has been noted in previous studies**. These cells were shown to produce collagen type 1 and type 3, and further enhance their synthesis upon cardiac pressure overload. This effect was specific to cardiac endothelial cells, as umbilical vein endothelial cells produced and secreted significantly less collagen type 1 [119]. Additionally, other studies found an upregulation of ECM components and collagen fibrils in cardiac endothelial cells following myocardial infarction. Performing scRNA-seq at different time points after ischemic injury identified a cluster of endothelial cells, which showed upregulation of genes associated with ECM within the first week after infarction [111]. Increased expression of these genes is associated with **endothelial-mesenchymal transition (EndMT)**, a process during which endothelial cells undergo phenotypical changes towards mesenchymal cells (e.g. myofibroblasts or smooth muscle cells). EndMT has a vital role in organismal development, but recent studies indicate that pathological conditions can induce EndMT even in adult

endothelial cells [168]. This is also the case in myocardial infarction, which induces a transient mesenchymal activation in a cluster of endothelial cells within the first week post-infarction. Notably, these cells never completely lose their endothelial identity and thus exhibit a partial EndMT, which is reversed to homeostatic levels at later time points after infarction [111]. I compared the expression of mesenchymal markers from this study with differentially expressed genes detected in RNA-seq analysis of isolated WT and SPAAR KO cardiac endothelial cells. I discovered that many of the most significant RNA-seq hits were differentially expressed according to reported mesenchymal activation, implying that **SPAAR KO cells exhibit some aspects of partial EndMT** (**Supplemental figure S1**).

Admittedly, these gene expression changes are rather modest in baseline conditions, when RNA-seq analysis was performed. This fact probably explains why I did not detect changes in functional characteristics between WT and SPAAR KO cells at baseline. According to EndMT activation, endothelial cells are expected to lose their ability to form tubules in culture, while at the same time they show increased proliferation, invasion and migration characteristics [168]. Since most of my *in vitro* experiments with cardiac endothelial cells were performed at baseline, I did not observe any changes in proliferation, tube formation, nor migration between WT and SPAAR KO cells. Importantly, previous studies noted the induction of partial EndMT specifically after ischemia and associated it with the formation of new, fully functional capillaries [169]. However, my own immunohistochemical assessment of capillary density in heart sections after I/R *in vivo* did not reveal any changes between WT and SPAAR KO mice. I performed this evaluation on heart sections collected two weeks after the ischemic insult, which is likely too late to observe the transient nature of EndMT [111]. To determine the contribution of mesenchymal activation to I/R injury and cardioprotection *in vivo*, I would need to assess capillary density at earlier stages of reperfusion. Finally, it would be most worthwhile to investigate endothelial cell heterogeneity and distinct adaptive responses by implementing scRNA-seq analysis in WT and SPAAR KO hearts following I/R injury *in vivo*.

6 Conclusions and outlook

The study described here investigated translated sORFs in the heart and in cardiac remodelling. By using a combination of ribosome profiling, mass spectrometry and evolutionary analysis I set out to identify novel, yet uncharacterized cardiac microproteins. Consolidation of these methodologies revealed that novel cardiac microproteins are unlikely to show signs of conservation between mice and humans. While there is no theoretical reason to believe that microproteins with reduced evolutionary constraints cannot have meaningful functions, working with species-specific microproteins hampers either their experimental characterization *in vivo* (in case of human- or primate-specific candidates) or their translational and clinical perspective (in case of murine-specific microproteins). Although I then focused on annotated but poorly characterized small proteins in the heart, the work presented here reinforces the idea of pervasive translation of sORFs and thus supports a broader view of protein-coding potential of the genome. This view is challenging many traditional beliefs by embracing very small peptides that are difficult to uncover, but still participate in vital biological processes. Emerging studies in the field are expanding this view even further, to include also functional microproteins with little to no signs of conservation.

A recently discovered microprotein SPAAR was chosen for more detailed functional characterization, which I describe in the second part of the thesis. Intrigued by the heart-enriched spatial expression pattern of SPAAR, I studied its contribution to cardiac phenotype following I/R injury *in vivo*. Surprisingly, SPAAR KO mice were protected from myocardial damage as evidenced by several independent methods. Next, the availability of scRNA-seq data from mammalian hearts presented another unexpected finding, namely the enrichment of *Spaar* in capillary endothelial cells. As a result, I turned to primary cardiac endothelial cells to investigate the molecular mechanisms of SPAAR activity. I pinpointed three aspects by which this microprotein modulates endothelial cell function: increased phosphorylation of TFEB, upregulated expression of IDE and differential production and secretion of ECM components upon SPAAR depletion. These observations advance the existing knowledge of the intricate functional role of SPAAR, but also raise additional exciting questions.

What is the purpose of highly restricted spatial expression pattern of SPAAR? The advances in single-cell transcriptomics have provided information about its enrichment in capillary endothelial cells in a specific subset of highly vascularized organs, such as the heart, skeletal muscle, lung, pancreas and adipose tissue. On the other hand, liver and brain also represent highly vascular organs, but virtually do not express SPAAR. Future single-cell studies of endothelial cells are needed to decipher their heterogenous expression signatures, which will contribute to a better understanding of SPAAR function.

What is the contribution of SPAAR to mTORC1 activation during I/R and how does this lead to cardioprotection *in vivo*? Most of my observations in mice come from relatively late time points, when signalling pathways and transient molecular changes already return to homeostasis. Assessing TFEB phosphorylation, capillary density and transcriptional changes underlying EndMT in SPAAR KO mice would be more insightful at earlier stages of reperfusion. In addition, my results suggest that endothelial-specific activation of mTORC1 through SPAAR depletion is beneficial after I/R injury. This finding contrasts with some previous studies [97] and unpublished data from our research group, which highlight promising outcomes following mTORC1 inhibition during I/R, albeit in cardiomyocytes. Perhaps SPAAR regulates conformational changes at the lysosomal membrane specifically in endothelial cells to fine-tune mTORC1 activity within a desirable, physiological range.

Also, what are the roles of TFEB phosphorylation at Ser122 and IDE in cardiac endothelial cells? Both proteins have been intensively studied, but their molecular changes described here and their respective functions cannot be explained by existing knowledge. Further research will thus be required to clarify the importance of these highly reproducible observations in SPAAR KO endothelial cells.

Finally, if SPAAR KO cardiac endothelial cells show characteristics of partial EndMT, what is the (patho)physiological role of such mesenchymal activation in the heart? Findings from this field of research are divergent, showing that pro-invasive and pro-migratory phenotypes of endothelial cells can have vastly different effects. While some report the formation of new blood vessels, other groups propose the role of EndMT in fibrosis and atherosclerosis [168]. Future investigations may provide more insights into this process, especially as the research focus is shifting from cardiomyocyte-centric mechanisms of cardiac disease towards exploring the unprecedented roles of non-myocyte cell types.

References

[1] Ariel A Bazzini et al. 'Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation.' In: *The EMBO journal* (2014), pp. 981–993. DOI: 10.1002/embj.201488411.

[2] Julie L. Aspden et al. 'Extensive translation of small Open Reading Frames revealed by Poly-Ribo-Seq'. In: *eLife* (2014), e03528. DOI: 10.7554/eLife.03528.

[3] Jin Chen et al. 'Pervasive functional translation of noncanonical human open reading frames.' In: *Science (New York, N.Y.)* (2020), pp. 1140–1146. DOI: 10.1126/science.aay0262.

[4] Xiaohan Yang et al. 'Discovery and annotation of small proteins using genomics, proteomics, and computational approaches.' In: *Genome research* (2011), pp. 634–641. DOI: 10.1101/gr.109280.110.

[5] Jiao Ma et al. 'Discovery of human sORF-encoded polypeptides (SEPs) in cell lines and tissue.' In: *Journal of proteome research* (2014), pp. 1757–1765. DOI: 10.1021/pr401280w.

[6] Juan-Pablo Couso and Pedro Patraquim. 'Classification and function of small open reading frames.' In: *Nature reviews. Molecular cell biology* (2017), pp. 575–589. DOI: 10.1038/nrm.2017.58.

[7] Jonathan M Mudge et al. *Standardized annotation of translated open reading frames.* 2022. DOI: 10.1038/s41587-022-01369-0.

[8] M A Basrai, P Hieter and J D Boeke. 'Small open reading frames: beautiful needles in the haystack.' In: *Genome research* (1997), pp. 768–771. DOI: 10.1101/gr.7.8.768.

[9] Benjamin R Nelson et al. 'A peptide encoded by a transcript annotated as long noncoding RNA enhances SERCA activity in muscle.' In: *Science (New York, N.Y.)* (2016), pp. 271–275. DOI: 10.1126/science.aad4076.

[10] Akinobu Matsumoto et al. 'mTORC1 and muscle regeneration are regulated by the LINC00961-encoded SPAR polypeptide'. In: *Nature* (2016), pp. 1–25. DOI: 10.1038/nature21034.

[11] Douglas M. Anderson et al. 'A micropeptide encoded by a putative long noncoding RNA regulates muscle performance'. In: *Cell* (2015), pp. 595–606. DOI: 10.1016/j.cell.2015.01.009.

[12] Nadia G D'Lima et al. 'A human microprotein that interacts with the mRNA decapping complex'. In: *Nature Chemical Biology* (2016), pp. 174–180. DOI: 10.1038/nchembio.2249.

[13] Nicholas T Ingolia et al. 'Ribosome profiling reveals pervasive translation outside of annotated protein-coding genes.' In: *Cell reports* (2014), pp. 1365–1379. DOI: 10.1016/j.celrep.2014.07.045.

[14] Mona Wu Orr et al. 'Alternative ORFs and small ORFs: shedding light on the dark proteome.' In: *Nucleic acids research* (2020), pp. 1029–1042. DOI: 10.1093/nar/gkz734.

[15] Nicholas T Ingolia, Liana F Lareau and Jonathan S Weissman. 'Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes.' In: *Cell* (2011), pp. 789–802. DOI: 10.1016/j.cell.2011.10.002.

[16] Anna M McGeachy and Nicholas T Ingolia. 'Starting too soon: upstream reading frames repress downstream translation'. In: *The EMBO Journal* (2016), pp. 699–700. DOI: 10.15252/embj.201693946.

[17] Divya Ram Jayaram et al. 'Unraveling the hidden role of a uORF-encoded peptide as a kinase inhibitor of PKCs.' In: *Proceedings of the National Academy of Sciences of the United States of America* (2021). DOI: 10.1073/pnas.2018899118.

[18] Pedro Patraquim et al. 'Translation and natural selection of micropeptides from long non-canonical RNAs.' In: *Nature communications* (2022), p. 6515. DOI: 10.1038/s41467-022-34094-y.

[19] Hongwei Wang et al. 'Global and cell-type specific properties of lincRNAs with ribosome occupancy'. In: *Nucleic Acids Research* (2017), pp. 2786–2796. DOI: 10.1093/nar/gkw909.

[20] Jorge Ruiz-Orera et al. 'Long non-coding RNAs as a source of new peptides.' In: *eLife* (2014), e03523. DOI: 10.7554/eLife.03523.

[21] Catherine A Makarewich. 'The hidden world of membrane microproteins.' In: *Experimental cell research* (2020), p. 111853. DOI: 10.1016/j.yexcr.2020.111853.

[22] Máximo Ibo Galindo et al. *Peptides Encoded by Short ORFs Control Development and Define a New Eukaryotic Gene Family*. Ed. by Alfonso Martinez Arias. 2007. DOI: 10.1371/journal.pbio.0050106.

[23] Takefumi Kondo et al. 'Small peptide regulators of actin-based cell morphogenesis encoded by a polycistronic mRNA'. In: *Nature Cell Biology* (2007), pp. 660–665. DOI: 10.1038/ncb1595.

[24] T Kondo et al. 'Small peptides switch the transcriptional activity of Shavenbaby during Drosophila embryogenesis.' In: *Science (New York, N.Y.)* (2010), pp. 336–339. DOI: 10.1126/science.1188158.

[25] Y Hashimoto et al. 'A rescue factor abolishing neuronal cell death by a wide spectrum of familial Alzheimer's disease genes and Abeta.' In: *Proceedings of the National Academy of Sciences of the United States of America* (2001), pp. 6336–6341. DOI: 10.1073/pnas.101133498.

[26] Bin Guo et al. 'Humanin peptide suppresses apoptosis by interfering with Bax activation.' In: *Nature* (2003), pp. 456–461. DOI: 10.1038/nature01627.

[27] Shea J. Andrews and Joseph A. Rothnagel. 'Emerging evidence for functional peptides encoded by short open reading frames'. In: *Nature Reviews Genetics* (2014), pp. 193–204. DOI: 10.1038/nrg3520.

[28] Nikolaos Vakirlis et al. 'De novo birth of functional microproteins in the human lineage.' In: *Cell reports* (2022), p. 111808. DOI: 10.1016/j.celrep.2022.111808.

[29] Clara-L Sandmann et al. 'Evolutionary origins and interactomes of human, young microproteins and small peptides translated from short open reading frames.' In: *Molecular cell* (2023). DOI: 10.1016/j.molcel.2023.01.023.

[30] Timothy G Johnstone, Ariel A Bazzini and Antonio J Giraldez. 'Upstream ORFs are prevalent translational repressors in vertebrates'. In: *The EMBO Journal* (2016), pp. 706–723. DOI: 10.15252/embj.201592759.

[31] Guo-Liang Chew et al. *Ribosome profiling reveals resemblance between long non-coding RNAs and 5 leaders of coding RNAs*. 2013. DOI: 10.1242/dev.098343.

[32] Sebastian D Mackowiak et al. 'Extensive identification and analysis of conserved small ORFs in animals.' In: *Genome biology* (2015), p. 179. DOI: 10.1186/s13059-015-0742-x.

[33] Sarah A Slavoff et al. 'Peptidomic discovery of short open reading frame–encoded peptides in human cells'. In: *Nature Chemical Biology* (2012), pp. 59–64. DOI: 10.1038/nchembio.1120.

[34] Nicholas T Ingolia et al. 'Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling.' In: *Science (New York, N.Y.)* (2009), pp. 218–223. DOI: 10.1126/science.1168978.

[35] Elisenda Sanz et al. *Cell-type-specific isolation of ribosome-associated mRNA from complex tissues*. 2009. DOI: 10.1073/pnas.0907143106.

[36] S. Doroudgar et al. 'Monitoring cell-Type-specific gene expression using ribosome profiling in vivo during cardiac hemodynamic stress'. In: *Circulation Research* (2019). DOI: 10.1161/CIRCRESAHA.119.314817.

[37] Preeti Madhav Kute et al. 'Small Open Reading Frames, How to Find Them and Determine Their Function.' In: *Frontiers in genetics* (2021), p. 796060. DOI: 10.3389/fgene.2021.796060.

[38] Bertrand Fabre, Jean-Philippe Combier and Serge Plaza. 'Recent advances in mass spectrometry-based peptidomics workflows to identify short-open-reading-frame-encoded peptides and explore their functions.' In: *Current opinion in chemical biology* (2021), pp. 122–130. DOI: 10.1016/j.cbpa.2020.12.002.

[39] Jiao Ma et al. 'Improved Identification and Analysis of Small Open Reading Frame Encoded Polypeptides.' In: *Analytical chemistry* (2016), pp. 3967–3975. DOI: 10.1021/acs.analchem.6b00191.

[40] Tristan Cardon et al. 'Optimized Sample Preparation Workflow for Improved Identification of Ghost Proteins.' In: *Analytical chemistry* (2020), pp. 1122–1129. DOI: 10.1021/acs.analchem.9b04188.

[41] Sebastiaan van Heesch et al. 'The Translational Landscape of the Human Heart.' In: *Cell* (2019), 242–260.e29. DOI: 10.1016/j.cell.2019.05.010.

[42] Catherine A Makarewich and Eric N Olson. 'Mining for Micropeptides.' In: *Trends in cell biology* (2017), pp. 685–696. DOI: 10.1016/j.tcb.2017.04.006.

[43] Qian Chu, Jiao Ma and Alan Saghatelian. 'Identification and characterization of sORF-encoded polypeptides'. In: *Critical Reviews in Biochemistry and Molecular Biology* (2015), pp. 134–141. DOI: 10.3109/10409238.2015.1016215. arXiv: 15334406.

[44] Colleen S Stein et al. 'Mitoregulin: A lncRNA-Encoded Microprotein that Supports Mitochondrial Supercomplexes and Respiratory Efficiency.' In: *Cell reports* (2018), 3710–3720.e8. DOI: 10.1016/j.celrep.2018.06.002.

[45] Catherine A Makarewich et al. 'MOXI Is a Mitochondrial Micropeptide That Enhances Fatty Acid β-Oxidation.' In: *Cell reports* (2018), pp. 3701–3709. DOI: 10.1016/j.celrep.2018.05.058.

[46] Catherine A Makarewich et al. 'The cardiac-enriched microprotein mitolamban regulates mitochondrial respiratory complex assembly and function in mice.' In: *Proceedings of the National Academy of Sciences of the United States of America* (2022). DOI: 10.1073/pnas.2120476119.

[47] Jakob D Busch et al. 'MitoRibo-Tag Mice Provide a Tool for In Vivo Studies of Mitoribosome Composition.' In: *Cell reports* (2019), 1728–1738.e9. DOI: 10.1016/j.celrep.2019.09.080.

[48] Sven Dennerlein et al. 'Defining the interactome of the human mitochondrial ribosome identifies SMIM4 and TMEM223 as respiratory chain assembly factors.' In: *eLife* (2021). DOI: 10.7554/eLife.68213.

[49] Chao Liang et al. 'Mitochondrial microproteins link metabolic cues to respiratory chain biogenesis.' In: *Cell reports* (2022), p. 111204. DOI: 10.1016/j.celrep.2022.111204.

[50] Sana A. Shaikh, Sanjaya K. Sahoo and Muthu Periasamy. 'Phospholamban and sarcolipin: Are they functionally redundant or distinct regulators of the Sarco(Endo)Plasmic Reticulum Calcium ATPase?' In: *Journal of Molecular and Cellular Cardiology* (2016), pp. 81–91. DOI: 10.1016/j.yjmcc.2015.12.030.

[51] Catherine A Makarewich et al. 'The DWORF micropeptide enhances contractility and prevents heart failure in a mouse model of dilated cardiomyopathy.' In: *eLife* (2018). DOI: 10.7554/eLife.38319.

[52] Catherine A Makarewich et al. *Gene Therapy With the DWORF Micropeptide Attenuates Cardiomyopathy in Mice.* 2020. DOI: 10.1161/CIRCRESAHA.120.317156.

[53] Jiwon Lee, Aaron Wacholder and Anne-Ruxandra Carvunis. *Evolutionary Characterization of the Short Protein SPAAR.* 2021. DOI: 10.3390/genes12121864.

[54] C-T Wu, S Liu and M Tang. 'Downregulation of linc00961 contributes to promote proliferation and inhibit apoptosis of vascular smooth muscle cell by sponging miR-367 in patients with coronary heart disease.' In: *European review for medical and pharmacological sciences* (2019), pp. 8540–8550. DOI: 10.26355/eurrev_201910_19168.

[55] Sebastiano Sciarretta et al. 'The complex network of mTOR signalling in the heart.' In: *Cardiovascular research* (2022), pp. 424–439. DOI: 10.1093/cvr/cvab033.

[56] Joungmok Kim and Kun-Liang Guan. 'mTOR as a central hub of nutrient signalling and cell growth'. In: *Nature Cell Biology* (2019), pp. 63–71. DOI: 10.1038/s41556-018-0205-1.

[57] Yoana Rabanal-Ruiz and Viktor I Korolchuk. 'mTORC1 and Nutrient Homeostasis: The Central Role of the Lysosome.' In: *International journal of molecular sciences* (2018). DOI: 10.3390/ijms19030818.

[58] Gennaro Napolitano, Chiara Di Malta and Andrea Ballabio. 'Non-canonical mTORC1 signaling at the lysosome.' In: *Trends in cell biology* (2022), pp. 920–931. DOI: 10.1016/j.tcb.2022.04.012.

[59] Yasemin Sancak et al. 'The Rag GTPases bind raptor and mediate amino acid signaling to mTORC1.' In: *Science (New York, N.Y.)* (2008), pp. 1496–1501. DOI: 10.1126/science.1157535.

[60] Yasemin Sancak et al. 'Ragulator-Rag complex targets mTORC1 to the lysosomal surface and is necessary for its activation by amino acids.' In: *Cell* (2010), pp. 290–303. DOI: 10.1016/j.cell.2010.02.024.

[61] Roberto Zoncu et al. 'mTORC1 senses lysosomal amino acids through an inside-out mechanism that requires the vacuolar H(+)-ATPase.' In: *Science (New York, N.Y.)* (2011), pp. 678–683. DOI: 10.1126/science.1207056.

[62] Liron Bar-Peled et al. 'Ragulator is a GEF for the rag GTPases that signal amino acid levels to mTORC1.' In: *Cell* (2012), pp. 1196–1208. DOI: 10.1016/j.cell.2012.07.032.

[63] Ken Inoki et al. 'Rheb GTPase is a direct target of TSC2 GAP activity and regulates mTOR signaling.' In: *Genes & development* (2003), pp. 1829–1834. DOI: 10.1101/gad.1110003.

[64] Haijuan Yang et al. 'Mechanisms of mTORC1 activation by RHEB and inhibition by PRAS40'. In: *Nature* (2017), pp. 368–373. DOI: 10.1038/nature25023.

[65] Stefanie S Schalm and John Blenis. 'Identification of a Conserved Motif Required for mTOR Signaling'. In: *Current Biology* (2002), pp. 632–639. DOI: 10.1016/S0960-9822(02)00762-5.

[66] Neethi Nandagopal and Philippe P Roux. 'Regulation of global and specific mRNA translation by the mTOR signaling pathway.' In: *Translation (Austin, Tex.)* (2015), e983402. DOI: 10.4161/21690731.2014.983402.

[67] Carson C Thoreen et al. 'A unifying model for mTORC1-mediated regulation of mRNA translation.' In: *Nature* (2012), pp. 109–113. DOI: 10.1038/nature11083.

[68] Valentina Gandin et al. 'nanoCAGE reveals 5' UTR features that define specific modes of translation of functionally related MTOR-sensitive mRNAs.' In: *Genome research* (2016), pp. 636–648. DOI: 10.1101/gr.197566.115.

[69] Rosa Puertollano et al. 'The complex relationship between TFEB transcription factor phosphorylation and subcellular localization.' In: *The EMBO journal* (2018). DOI: 10.15252/embj.201798804.

[70] Eiríkur Steingrímsson, Neal G Copeland and Nancy A Jenkins. 'Melanocytes and the microphthalmia transcription factor network.' In: *Annual review of genetics* (2004), pp. 365–411. DOI: 10.1146/annurev.genet.38.072902.092717.

[71] Michela Palmieri et al. 'Characterization of the CLEAR network reveals an integrated control of cellular clearance pathways.' In: *Human molecular genetics* (2011), pp. 3852–3866. DOI: 10.1093/hmg/ddr306.

[72] José A Martina et al. 'The nutrient-responsive transcription factor TFE3 promotes autophagy, lysosomal biogenesis, and clearance of cellular debris.' In: *Science signaling* (2014), ra9. DOI: 10.1126/scisignal.2004754.

[73] Agnes Roczniak-Ferguson et al. 'The transcription factor TFEB links mTORC1 signaling to transcriptional control of lysosome homeostasis.' In: *Science signaling* (2012), ra42. DOI: 10.1126/scisignal.2002790.

[74] Jose A Martina et al. 'MTORC1 functions as a transcriptional regulator of autophagy by preventing nuclear transport of TFEB.' In: *Autophagy* (2012), pp. 903–914. DOI: 10.4161/auto.19653.

[75] Gennaro Napolitano et al. 'mTOR-dependent phosphorylation controls TFEB nuclear export.' In: *Nature communications* (2018), p. 3312. DOI: 10.1038/s41467-018-05862-6.

[76] Linxin Li et al. 'A TFEB nuclear export signal integrates amino acid supply and glucose availability.' In: *Nature communications* (2018), p. 2685. DOI: 10.1038/s41467-018-04849-7.

[77] Silvia Vega-Rubin-de-Celis et al. 'Multistep regulation of TFEB by MTORC1.' In: *Autophagy* (2017), pp. 464–472. DOI: 10.1080/15548627.2016.1271514.

[78] Jose A Martina and Rosa Puertollano. 'Rag GTPases mediate amino acid-dependent recruitment of TFEB and MITF to lysosomes.' In: *The Journal of cell biology* (2013), pp. 475–491. DOI: 10.1083/jcb.201209135.

[79] Gennaro Napolitano et al. 'A substrate-specific mTORC1 pathway underlies Birt-Hogg-Dubé syndrome.' In: *Nature* (2020), pp. 597–602. DOI: 10.1038/s41586-020-2444-0.

[80] Zhi-Yang Tsun et al. 'The folliculin tumor suppressor is a GAP for the RagC/D GTPases that signal amino acid levels to mTORC1.' In: *Molecular cell* (2013), pp. 495–505. DOI: 10.1016/j.molcel.2013.09.016.

[81] David A Guertin et al. 'Ablation in mice of the mTORC components raptor, rictor, or mLST8 reveals that mTORC2 is required for signaling to Akt-FOXO and PKCalpha, but not S6K1.' In: *Developmental cell* (2006), pp. 859–871. DOI: 10.1016/j.devcel.2006.10.007.

[82] Yann-Gaël Gangloff et al. 'Disruption of the mouse mTOR gene leads to early postimplantation lethality and prohibits embryonic stem cell development.' In: *Molecular and cellular biology* (2004), pp. 9508–9516. DOI: 10.1128/MCB.24.21.9508-9516.2004.

[83] Yi Zhu et al. 'Mechanistic target of rapamycin (Mtor) is essential for murine embryonic heart development and growth.' In: *PloS one* (2013), e54221. DOI: 10.1371/journal.pone.0054221.

[84] Denghong Zhang et al. 'MTORC1 regulates cardiac function and myocyte survival through 4E-BP1 inhibition in mice.' In: *The Journal of clinical investigation* (2010), pp. 2805–2816. DOI: 10.1172/JCI43008.

[85] Pankaj Shende et al. 'Cardiac raptor ablation impairs adaptive hypertrophy, alters metabolic gene expression, and causes heart failure in mice.' In: *Circulation* (2011), pp. 1073–1082. DOI: 10.1161/CIRCULATIONAHA.110.977066.

[86] Takahito Tamai et al. 'Rheb (Ras homologue enriched in brain)-dependent mammalian target of rapamycin complex 1 (mTORC1) activation becomes indispensable for cardiac hypertrophic growth after early postnatal period.' In: *The Journal of biological chemistry* (2013), pp. 10176–10187. DOI: 10.1074/jbc.M112.423640.

[87] James M Flynn et al. 'Late-life rapamycin treatment reverses age-related heart dysfunction.' In: *Aging cell* (2013), pp. 851–862. DOI: 10.1111/acel.12109.

[88] Dao-Fu Dai et al. 'Altered proteome turnover and remodeling by short-term caloric restriction or rapamycin rejuvenate the aging heart.' In: *Aging cell* (2014), pp. 529–539. DOI: 10.1111/acel.12203.

[89] Gerd Heusch. 'Myocardial ischaemia-reperfusion injury and cardioprotection in perspective.' In: *Nature reviews. Cardiology* (2020), pp. 773–789. DOI: 10.1038/s41569-020-0403-y.

[90] Moien Ab Khan et al. 'Global Epidemiology of Ischemic Heart Disease: Results from the Global Burden of Disease Study.' In: *Cureus* (2020), e9349. DOI: 10.7759/cureus.9349.

[91] Sebastiano Sciarretta et al. 'Rheb is a critical regulator of autophagy during myocardial ischemia: pathophysiological implications in obesity and metabolic syndrome.' In: *Circulation* (2012), pp. 1134–1146. DOI: 10.1161/CIRCULATIONAHA.111.078212.

[92] Peiyong Zhai et al. 'Differential roles of GSK-3β during myocardial ischemia and ischemia/reperfusion.' In: *Circulation research* (2011), pp. 502–511. DOI: 10.1161/CIRCRESAHA.111.249532.

[93] Sebastian J Buss et al. 'Beneficial effects of Mammalian target of rapamycin inhibition on left ventricular remodeling after myocardial infarction.' In: *Journal of the American College of Cardiology* (2009), pp. 2435–2446. DOI: 10.1016/j.jacc.2009.08.031.

[94] Ruomin Di et al. 'S6K inhibition renders cardiac protection against myocardial infarction through PDK1 phosphorylation of Akt.' In: *The Biochemical journal* (2012), pp. 199–207. DOI: 10.1042/BJ20110033.

[95] Xiangqi Wu et al. 'Genetic and pharmacological inhibition of Rheb1-mTORC1 signaling exerts cardioprotection against adverse cardiac remodeling in mice.' In: *The American journal of pathology* (2013), pp. 2005–2014. DOI: 10.1016/j.ajpath.2013.02.012.

[96] Mirko Völkers et al. 'Mechanistic target of rapamycin complex 2 protects the heart from ischemic damage.' In: *Circulation* (2013), pp. 2132–2144. DOI: 10.1161/CIRCULATIONAHA.113.003638.

[97] Anindita Das et al. 'Rapamycin protects against myocardial ischemia-reperfusion injury through JAK2-STAT3 signaling pathway.' In: *Journal of molecular and cellular cardiology* (2012), pp. 858–869. DOI: 10.1016/j.yjmcc.2012.09.007.

[98] Shiho Satomi et al. 'Branched-chain amino acids-induced cardiac protection against ischemia/reperfusion injury.' In: *Life sciences* (2020), p. 117368. DOI: 10.1016/j.lfs.2020.117368.

[99] Vassilios J Bezzerides et al. 'CITED4 induces physiologic hypertrophy and promotes functional recovery after ischemic injury.' In: *JCI insight* (2016). DOI: 10.1172/jci.insight.85904.

[100] Toshinori Aoyagi et al. 'Cardiac mTOR protects the heart against ischemia-reperfusion injury.' In: *American journal of physiology. Heart and circulatory physiology* (2012), H75–85. DOI: 10.1152/ajpheart.00241.2012.

[101] Yutaka Matsui et al. 'Distinct roles of autophagy in the heart during ischemia and reperfusion: roles of AMP-activated protein kinase and Beclin 1 in mediating autophagy.' In: *Circulation research* (2007), pp. 914–922. DOI: 10.1161/01.RES.0000261924.76669.36.

[102] John T Cunningham et al. 'mTOR controls mitochondrial oxidative function through a YY1-PGC-1alpha transcriptional complex.' In: *Nature* (2007), pp. 736–740. DOI: 10.1038/nature06322.

[103] Toshiyuki Yano et al. 'Clinical impact of myocardial mTORC1 activation in nonischemic dilated cardiomyopathy.' In: *Journal of molecular and cellular cardiology* (2016), pp. 6–9. DOI: 10.1016/j.yjmcc.2015.12.022.

[104] Antonio M A Miranda et al. 'Single-cell transcriptomics for the assessment of cardiac disease.' In: *Nature reviews. Cardiology* (2022). DOI: 10.1038/s41569-022-00805-7.

[105] Monika Litviňuková et al. 'Cells of the adult human heart'. In: *Nature* (2020), pp. 466–472. DOI: 10.1038/s41586-020-2797-4.

[106] Christoph Kuppe et al. 'Spatial multi-omic map of human myocardial infarction.' In: *Nature* (2022), pp. 766–777. DOI: 10.1038/s41586-022-05060-x.

[107] Michelle D Tallquist and Jeffery D Molkentin. 'Redefining the identity of cardiac fibroblasts.' In: *Nature reviews. Cardiology* (2017), pp. 484–491. DOI: 10.1038/nrcardio.2017.57.

[108] Nona Farbehi et al. 'Single-cell expression profiling reveals dynamic flux of cardiac stromal, vascular and immune cells in health and injury'. In: *eLife* (2019). Ed. by Edward Morrisey and Harry C Dietz, e43882. DOI: 10.7554/eLife.43882.

[109] William C Aird. 'Phenotypic Heterogeneity of the Endothelium'. In: *Circulation Research* (2007), pp. 174–190. DOI: 10.1161/01.RES.0000255690.03436.ae.

[110] Andrew L Koenig et al. 'Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure'. In: *Nature Cardiovascular Research* (2022), pp. 263–280. DOI: 10.1038/s44161-022-00028-6.

[111] Lukas S Tombor et al. 'Single cell sequencing reveals endothelial plasticity with transient mesenchymal activation after myocardial infarction.' In: *Nature communications* (2021), p. 681. DOI: 10.1038/s41467-021-20905-1.

[112] Brandon Malone et al. 'Bayesian prediction of RNA translation from ribosome profiling.' In: *Nucleic acids research* (2017), pp. 2960–2972. DOI: 10.1093/nar/gkw1350.

[113] Margarida Cardoso-Moreira et al. 'Gene expression across mammalian organ development.' In: *Nature* (2019), pp. 505–509. DOI: 10.1038/s41586-019-1338-5.

[114] Roberto Aquilani et al. 'Plasma Amino Acid Abnormalities in Chronic Heart Failure. Mechanisms, Potential Risks and Targets in Human Myocardium Metabolism.' In: *Nutrients* (2017). DOI: 10.3390/nu9111251.

[115] Haipeng Sun et al. 'Catabolic Defect of Branched-Chain Amino Acids Promotes Heart Failure.' In: *Circulation* (2016), pp. 2038–2049. DOI: 10.1161/CIRCULATIONAHA.115.020226.

[116] Tabula Sapiens Consortium et al. 'The Tabula Sapiens: A multiple-organ, single-cell transcriptomic atlas of humans'. In: *Science* (2023), eabl4896. DOI: 10.1126/science.abl4896.

[117] Zhaobin Xu et al. 'A murine model of myocardial ischemia-reperfusion injury through ligation of the left anterior descending artery.' In: *Journal of visualized experiments : JoVE* (2014). DOI: 10.3791/51329.

[118] Shirin Doroudgar et al. 'Hrd1 and ER-Associated Protein Degradation, ERAD, are Critical Elements of the Adaptive ER Stress Response in Cardiac Myocytes.' In: *Circulation research* (2015), pp. 536–546. DOI: 10.1161/CIRCRESAHA.115.306993.

[119] Natali Froese et al. 'Analysis of myocardial cellular gene expression during pressure overload reveals matrix based functional intercellular communication.' In: *iScience* (2022), p. 103965. DOI: 10.1016/j.isci.2022.103965.

[120] Natali Froese et al. 'Endothelial Cell GATA2 Modulates the Cardiomyocyte Stress Response through the Regulation of Two Long Non-Coding RNAs.' In: *Biology* (2022). DOI: 10.3390/biology11121736.

[121] Gilles Carpentier et al. 'Angiogenesis Analyzer for ImageJ — A comparative morphometric analysis of "Endothelial Tube Formation Assay" and "Fibrin Bead Assay"'. In: *Scientific Reports* (2020), p. 11568. DOI: 10.1038/s41598-020-67289-8.

[122] Nicholas J McGlincy and Nicholas T Ingolia. 'Transcriptome-wide measurement of translation by ribosome profiling.' In: *Methods (San Diego, Calif.)* (2017), pp. 112–129. DOI: 10.1016/j.ymeth.2017.05.028.

[123] Galaxy Community. 'The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update'. In: *Nucleic Acids Research* (2022), W345–W351. DOI: 10.1093/nar/gkac247.

[124] Ben Langmead and Steven L Salzberg. 'Fast gapped-read alignment with Bowtie 2.' In: *Nature methods* (2012), pp. 357–359. DOI: 10.1038/nmeth.1923.

[125] Alexander Dobin et al. 'STAR: ultrafast universal RNA-seq aligner.' In: *Bioinformatics (Oxford, England)* (2013), pp. 15–21. DOI: 10.1093/bioinformatics/bts635.

[126] Yang Liao, Gordon K Smyth and Wei Shi. 'featureCounts: an efficient general purpose program for assigning sequence reads to genomic features.' In: *Bioinformatics (Oxford, England)* (2014), pp. 923–930. DOI: 10.1093/bioinformatics/btt656.

[127] Michael I Love, Wolfgang Huber and Simon Anders. 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2'. In: *Genome Biology* (2014), p. 550. DOI: 10.1186/s13059-014-0550-8.

[128] Sonia Chothani et al. 'deltaTE: Detection of Translationally Regulated Genes by Integrative Analysis of Ribo-seq and RNA-seq Data.' In: *Current protocols in molecular biology* (2019), e108. DOI: 10.1002/cpmb.108.

[129] Glynn Dennis et al. 'DAVID: Database for Annotation, Visualization, and Integrated Discovery'. In: *Genome Biology* (2003), R60. DOI: 10.1186/gb-2003-4-9-r60.

[130] Tim Christian Kuhn et al. 'Secretome Analysis of Cardiomyocytes Identifies PCSK6 (Proprotein Convertase Subtilisin/Kexin Type 6) as a Novel Player in Cardiac Remodeling After Myocardial Infarction.' In: *Circulation* (2020), pp. 1628–1644. DOI: 10.1161/CIRCULATIONAHA.119.044914.

[131] Damian Szklarczyk et al. 'STRING v10: protein-protein interaction networks, integrated over the tree of life.' In: *Nucleic acids research* (2015), pp. D447–52. DOI: 10.1093/nar/gku1003.

[132] Edward Lau et al. 'A large dataset of protein dynamics in the mammalian heart proteome.' In: *Scientific data* (2016), p. 160015. DOI: 10.1038/sdata.2016.15.

[133] Sophia Doll et al. 'Region and cell-type resolved quantitative proteomic map of the human heart'. In: *Nature Communications* (2017), p. 1469. DOI: 10.1038/s41467-017-01747-2.

[134] Ru-Yi Xu et al. 'High-sensitive cardiac troponin T.' In: *Journal of geriatric cardiology : JGC* (2013), pp. 102–109. DOI: 10.3969/j.issn.1671-5411.2013.01.015.

[135] Paula S Azevedo et al. 'Cardiac Remodeling: Concepts, Clinical Impact, Pathophysiological Mechanisms and Pharmacologic Treatment.' In: *Arquivos brasileiros de cardiologia* (2016), pp. 62–69. DOI: 10.5935/abc.20160005.

[136] Svenja Hinderer and Katja Schenke-Layland. 'Cardiac fibrosis - A short review of causes and therapeutic strategies.' In: *Advanced drug delivery reviews* (2019), pp. 77–82. DOI: 10.1016/j.addr.2019.05.011.

[137] Arun K Singhal et al. 'Role of Endothelial Cells in Myocardial Ischemia-Reperfusion Injury.' In: *Vascular disease prevention* (2010), pp. 1–14. DOI: 10.2174/1874120701007010001.

[138] Xuekun Wu et al. 'Angiogenesis after acute myocardial infarction'. In: *Cardiovascular Research* (2021), pp. 1257–1273. DOI: 10.1093/cvr/cvaa287.

[139] Monther Abu-Remaileh et al. 'Lysosomal metabolomics reveals V-ATPase- and mTOR-dependent regulation of amino acid efflux from lysosomes.' In: *Science (New York, N.Y.)* (2017), pp. 807–813. DOI: 10.1126/science.aan6298.

[140] Grazia R Tundo et al. 'Multiple functions of insulin-degrading enzyme: a metabolic crosslight?' In: *Critical reviews in biochemistry and molecular biology* (2017), pp. 554–582. DOI: 10.1080/10409238.2017.1337707.

[141] Juan Pablo Maianti et al. 'Anti-diabetic activity of insulin-degrading enzyme inhibitors mediated by multiple hormones.' In: *Nature* (2014), pp. 94–98. DOI: 10.1038/nature13297.

[142] Tomohiro Yokota et al. 'Type V Collagen in Scar Tissue Regulates the Size of Scar after Heart Injury.' In: *Cell* (2020), 545–562.e23. DOI: 10.1016/j.cell.2020.06.030.

[143] Christine Widmer et al. 'Molecular basis for the action of the collagen-specific chaperone Hsp47/SERPINH1 and its structure-specific client recognition'. In: *Proceedings of the National Academy of Sciences* (2012), pp. 13243–13247. DOI: 10.1073/pnas.1208072109.

[144] Yoshihiro Ishikawa, Kazunori Mizuno and Hans Peter Bächinger. 'Ziploc-ing the structure 2.0: Endoplasmic reticulum-resident peptidyl prolyl isomerases show different activities toward hydroxyproline'. In: *Journal of Biological Chemistry* (2017), pp. 9273–9282. DOI: 10.1074/jbc.M116.772657.

[145] Yifei Qi and Ren Xu. *Roles of PLODs in Collagen Synthesis and Cancer Progression.* 2018.

[146] Gilbert S Omenn et al. 'Progress on the HUPO Draft Human Proteome: 2017 Metrics of the Human Proteome Project.' In: *Journal of proteome research* (2017), pp. 4281–4287. DOI: 10.1021/acs.jproteome.7b00375.

[147] Jorge Ruiz-Orera et al. 'Translation of neutrally evolving peptides provides a basis for de novo gene evolution'. In: *Nature ecology & evolution* (2018), pp. 890–896. DOI: 10.1038/s41559-018-0506-6.

[148] Jin-Wu Nam, Seo-Won Choi and Bo-Hyun You. 'Incredible RNA: Dual Functions of Coding and Noncoding.' In: *Molecules and cells* (2016), pp. 367–374. DOI: 10.14348/molcells.2016.0039.

[149] Helen L Spencer et al. 'The LINC00961 transcript and its encoded micropeptide SPAAR regulate endothelial cell function.' In: *Cardiovascular research* (2020). DOI: 10.1093/cvr/cvaa008.

[150] Ana-Mishel Spiroski et al. 'The Influence of the LINC00961/SPAAR Locus Loss on Murine Development, Myocardial Dynamics, and Cardiac Response to Myocardial Infarction.' In: *International journal of molecular sciences* (2021). DOI: 10.3390/ijms22020969.

[151] Xiao-Jun Du. 'Gender modulates cardiac phenotype development in genetically modified mice'. In: *Cardiovascular Research* (2004), pp. 510–519. DOI: 10.1016/j.cardiores.2004.03.027.

[152] Richard D Patten. 'Models of Gender Differences in Cardiovascular Disease.' In: *Drug discovery today. Disease models* (2007), pp. 227–232. DOI: 10.1016/j.ddmod.2007.11.002.

[153] Abhijit Takawale et al. 'Myocardial recovery from ischemia-reperfusion is compromised in the absence of tissue inhibitor of metalloproteinase 4.' In: *Circulation. Heart failure* (2014), pp. 652–662. DOI: 10.1161/CIRCHEARTFAILURE.114.001113.

[154] Tabula Muris Consortium. 'Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris.' In: *Nature* (2018), pp. 367–372. DOI: 10.1038/s41586-018-0590-4.

[155] Shogo Wada et al. 'The tumor suppressor FLCN mediates an alternate mTOR pathway to regulate browning of adipose tissue.' In: *Genes & development* (2016), pp. 2551–2564. DOI: 10.1101/gad.287953.116.

[156] Bridget S Gosis et al. 'Inhibition of nonalcoholic fatty liver disease in mice by selective inhibition of mTORC1.' In: *Science (New York, N.Y.)* (2022), eabf8271. DOI: 10.1126/science.abf8271.

[157] Zhicheng Cui et al. 'Structure of the lysosomal mTORC1-TFEB-Rag-Ragulator megacomplex.' In: *Nature* (2023), pp. 572–579. DOI: 10.1038/s41586-022-05652-7.

[158] Rosalie E Lawrence et al. 'Structural mechanism of a Rag GTPase activation checkpoint by the lysosomal folliculin complex.' In: *Science (New York, N.Y.)* (2019), pp. 971–977. DOI: 10.1126/science.aax0364.

[159] Simon A Fromm, Rosalie E Lawrence and James H Hurley. 'Structural mechanism for amino acid-dependent Rag GTPase nucleotide state switching by SLC38A9.' In: *Nature structural & molecular biology* (2020), pp. 1017–1023. DOI: 10.1038/s41594-020-0490-9.

[160] Youbao Sha et al. 'STUB1 regulates TFEB-induced autophagy-lysosome pathway.' In: *The EMBO journal* (2017), pp. 2544–2552. DOI: 10.15252/embj.201796699.

[161] Gabriella Doronzo et al. 'TFEB controls vascular development by regulating the proliferation of endothelial cells.' In: *The EMBO journal* (2019). DOI: 10.15252/embj.201798250.

[162] Yanbo Fan et al. 'Endothelial TFEB (Transcription Factor EB) Positively Regulates Postischemic Angiogenesis.' In: *Circulation research* (2018), pp. 945–957. DOI: 10.1161/CIRCRESAHA.118.312672.

[163] Haocheng Lu et al. 'TFEB inhibits endothelial cell inflammation and reduces atherosclerosis.' In: *Science signaling* (2017). DOI: 10.1126/scisignal.aah4214.

[164] Laetitia Lesire et al. 'Insulin-Degrading Enzyme, an Under-Estimated Potential Target to Treat Cancer?' In: *Cells* (2022). DOI: 10.3390/cells11071228.

[165] Malcolm A Leissring et al. 'Targeting Insulin-Degrading Enzyme in Insulin Clearance.' In: *International journal of molecular sciences* (2021). DOI: 10.3390/ijms22052235.

[166] Malcolm A Leissring. 'Insulin-Degrading Enzyme: Paradoxes and Possibilities.' In: *Cells* (2021). DOI: 10.3390/cells10092445.

[167] Diego Sbardella et al. 'Proteasome Activity Is Affected by Fluctuations in Insulin-Degrading Enzyme Distribution.' In: *PloS one* (2015), e0132455. DOI: 10.1371/journal.pone.0132455.

[168] Jason C Kovacic et al. 'Endothelial to Mesenchymal Transition in Cardiovascular Disease: JACC State-of-the-Art Review.' In: *Journal of the American College of Cardiology* (2019), pp. 190–209. DOI: 10.1016/j.jacc.2018.09.089.

[169] Yosif Manavski et al. 'Clonal Expansion of Endothelial Cells Contributes to Ischemia-Induced Neovascularization.' In: *Circulation research* (2018), pp. 670–677. DOI: 10.1161/CIRCRESAHA.117.312310.

Supplemental figures

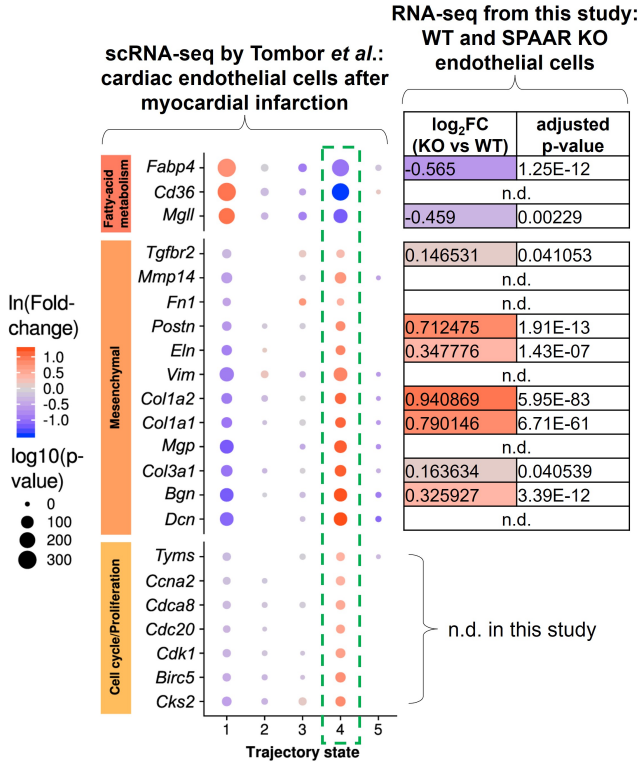

Figure S1: RNA-seq results from cardiac endothelial cells suggest that SPAAR KO cells are undergoing partial EndMT. Tombor and colleagues performed scRNA-seq analysis of non-myocyte cells following myocardial infarction [111]. They identified different endothelial cells states, with most cells found in state 4 (dashed green outline). This state exhibits expression changes indicative of EndMT: upregulation of mesenchymal genes and proliferation markers, as well as downregulation of genes involved in fatty acid metabolism. Comparing these results with the RNA-seq data from this study (table right) revealed expression changes consistent with partial mesenchymal activation in SPAAR KO endothelial cells. n.d.: not differentially expressed.

Acknowledgements

Back in 2017 when I started my PhD, I had no idea I was in for such a bumpy ride... But nevertheless, I have very much enjoyed the way, and that is mostly because of the great support I have received.

First, an immense thank you goes to my PhD supervisor, Dr. Mirko Völkers, whose guidance and recommendations throughout the project have been invaluable. I am deeply grateful for your patience and unfaltering understanding all these years. I will never forget your endless optimism and big picture thinking, which provided a much-needed balance to my detail-oriented mind.

I would also like to thank Prof. Dr. Stoecklin and Prof. Dr. Kaessmann, for taking their time to review my progress every year and for many helpful and interesting discussions.

I would like to extend my thanks to Prof. Dr. Heineke for being part of my examination committee and for hosting me in his lab on several occasions. I would like to mention Merve Keles, Steve Grein, Shruthi Hemanna and Nina Weinzierl – all members of Heineke lab, who shared their skills with me and taught me valuable methods.

I owe my thanks to many more talented researchers and students, who helped me with the project along the road: Tim C. Kuhn, Evgeny Leushkin, Etienne Boileau, Ole M. Schwerdt, Adrian Serafin and Nicholas A. Rüdinger.

In addition, I would like to acknowledge Heidelberg Biosciences International Graduate School for their financial support during the first year of my PhD studies.

Special thanks goes to the wonderful members of the Völkers lab for making the past years a much more delightful and fun experience. Especially to Vivi and Aga – our lunch sessions and jokes kept me sane on this long journey and I am thankful for all your advice and help.

I am truly grateful to Michi, for being always curious and supportive of my work. Thank you for all the memorable adventures outside the lab and many more exciting plans ahead. Many thanks also to my fantastic group of climbing friends, who cheered me up whenever I felt troubled.

Finally, I wish to thank my family for their unwavering love and support. You encouraged me all along and taught me to trust and believe in myself. Hvala, mami in oči, rada vaju imam tja gor do lune in nazaj.