Technische Universität München

Fakultät für Medizin

Contribution of the novel tumor suppressor SASH1 and the CRK protein family to tumor progression and metastasis formation

Fabian Christoph Franke

Vollständiger Abdruck der von der Fakultät für Medizin der Technischen Universität München zur Erlangung des akademischen Grades eines Doktors der Naturwissenschaften (rer. nat.) genehmigten Dissertation.

Vorsitzender:	Prof. Dr. Dieter Saur
Prüfende der Dissertation:	1. apl. Prof. Dr. Klaus-Peter Janssen
	2. Prof. Dr. Bernhard Küster
	3. Prof. Dr. Andreas Jung

Die Dissertation wurde am 02.07.2019 bei der Technischen Universität München eingereicht und durch die Fakultät für Medizin am 11.02.2020 angenommen.

Summary

The candidate tumor suppressor SASH1 has been implicated in metastasis formation of colorectal cancer and other solid tumors. However, causal evidence was lacking, and the underlying mechanisms remained largely elusive. To investigate the role of SASH1 in metastasis formation, SASH1-deficient colorectal cancer cells were generated using the CRISPR/Cas9 system. Loss of SASH1 was sufficient to induce epithelial-mesenchymal transition (EMT), generating highly invasive cells with increased cell survival and resistance to chemotherapy. Forced expression of *SASH1* in turn counteracted cytokine-induced EMT. Mechanistically, SASH1 physically associated with the signal adaptor CRKL, counteracting CRKL-mediated SRC kinase signaling, and thus EMT, invasiveness and chemoresistance. Strikingly, SASH1-deficient cells spawned significantly more metastases *in vivo*, entirely dependent on its novel interaction partner CRKL. Next, deficiency of either CRKL or its closely related homolog CRK induced a pronounced epithelial phenotype in colorectal cancer cell lines, suggesting functional redundancy of the CRK family in EMT. Of note, loss of the entire CRK family generated highly epithelial cells that completely lacked invasiveness and the ability to undergo EMT. As underlying mechanism, CRK proteins were identified to act as central amplifiers of SRC/FAK signaling by mediating a novel small GTPase-dependent positive feedback loop. Genetic activation of SRC was sufficient to induce sustained SRC/FAK signaling, invasiveness and EMT in the parental line, but not in CRK family-deficient cells. Therefore, CRK family proteins facilitate the at least partially independent activation of the kinases SRC and FAK to promote EMT. This was further verified in pancreatic cancer cells, indicating a role across tumor entities. Additionally, CRK family-deficient cells exhibited reduced ERK1/2 phosphorylation, MYC protein levels and transcriptome-wide MYC target gene expression, as well as decreased cell survival, chemoresistance and proliferation. Clinically, expression of *CRK* family genes was directly correlated with the EMT marker *ZEB1* in locally advanced colorectal cancer. High intratumoral expression of *CRK* and/or *CRKL*, as well as of *ZEB1*, were associated with poor survival. Finally, a high affinity peptide was demonstrated to bind both CRKL and CRK, inhibiting their N-terminal SH3 domains and counteracting CRK family-mediated SRC/FAK signaling. Therefore, in addition to verifying SASH1 as tumor and metastasis suppressor and identifying the underlying molecular mechanism, this thesis highlights CRK family proteins as promising therapeutic targets to counteract EMT, invasiveness, proliferation, chemoresistance, metastasis formation, and minimal residual disease.

Zusammenfassung

SASH1 ist ein potentieller Tumorsuppressor, der mit der Metastasierung kolorektaler Karzinome und anderer Tumorentitäten in Verbindung gebracht wurde. Bislang konnte jedoch kein kausaler Zusammenhang nachgewiesen werden und die zugrundeliegenden Mechanismen waren größtenteils unbekannt. Um die Rolle von SASH1 bei der Metastasierung zu untersuchen, wurden SASH1-defiziente Darmkrebszellen mit Hilfe des CRISPR/Cas9-Systems generiert. Das Fehlen von SASH1 war ausreichend, um eine epithelial-mesenchymale Transition (EMT), sowie hoch invasive Zellen mit erhöhtem Zellüberleben und Resistenz gegen Chemotherapie zu induzieren. Eine verstärkte Expression von *SASH1* konnte dagegen eine Zytokin-induzierte EMT hemmen. Als Mechanismus konnte gezeigt werden, dass SASH1 auf Proteinebene mit dem Signaladapter CRKL interagiert, wodurch der CRKL-vermittelte Signalweg der Kinase SRC und somit mittelbar auch EMT, Invasivität und Chemoresistenz inhibiert werden. Bemerkenswerter Weise bildeten SASH1-defiziente Zellen signifikant mehr Metastasen im Mausmodell *in vivo*, was vollständig von dessen neuartigem Interaktionspartner CRKL abhängig war. Weiterhin wurde gezeigt, dass die Defizienz von CRKL oder seinem Homolog CRK einen ausgeprägt epithelialen Phänotyp hervorruft. Ein vollständiges Fehlen der gesamten CRK-Familie führte zu einem noch stärker ausgeprägten epithelialen Phänotyp und blockierte Invasivität und EMT vollständig. Die CRK-Proteine konnten als zentrale Verstärker des SRC/FAK-Signalweges identifiziert werden, indem sie eine von kleinen GTPasen-abhängige positive Rückkopplungsschleife vermitteln. Eine genetische Aktivierung von SRC war zudem ausreichend, um den SRC/FAK-Signalweg, die Invasivität und EMT in parentalen Zellen zu verstärken. In CRK-Familie-defizienten Zellen war dies jedoch nicht möglich. Somit vermitteln CRK-Proteine eine zumindest partiell unabhängige Aktivierung der Kinasen SRC und FAK, weshalb sie für die EMT wichtige Faktoren darstellen. Dies wurde ebenfalls in Bauchspeicheldrüsenkrebszellen gezeigt, wodurch eine Rolle bei verschiedenen Tumorentitäten vermutet werden kann. Zusätzlich wiesen CRK-Familie-defiziente Zellen eine reduzierte Phosphorylierung von ERK1/2, einen verringerten MYC Proteinspiegel und eine verminderte Transkriptom-weite Expression von MYC Zielgenen auf, wobei Zellüberleben, Chemoresistenz und Proliferation ebenfalls reduziert waren. Die Expression von Genen der *CRK*-Familie korrelierte direkt mit dem EMT-Marker *ZEB1* in Patienten mit lokal fortgeschrittenem Darmkrebs. Zudem waren eine hohe intratumorale Expression von *CRK* und/oder *CRKL* sowie von *ZEB1* mit einer schlechten Prognose assoziiert. Zuletzt konnte

gezeigt werden, dass ein hochaffines Peptid dem CRK-Familie-vermittelten SRC/FAK-Signalweg durch Inhibition der N-terminalen SH3 Domänen von CRKL und CRK entgegenwirkt. Zusätzlich zu der Verifizierung von SASH1 als Tumor- und Metastasierungssuppressor und der Aufklärung der zugrundeliegenden molekularen Mechanismen bestätigt diese Arbeit Proteine der CRK-Familie als vielversprechende therapeutische Zielstrukturen, um EMT, Invasivität, Proliferation, Chemoresistenz, Metastasierung und einer minimalen Resterkrankung entgegenzuwirken.

Table of contents

Abbreviations

A.U.	Arbitrary Unit
aa	Amino Acid
ABL	V-Abl Abelson Murine Leukemia Viral Oncogene Homolog 1
AKT	V-Akt Murine Thymoma Viral Oncogene Homolog 1
APC	Adenomatous Polyposis Coli
ATP	Adenosine Triphosphate
BRAF	V-Raf Murine Sarcoma Viral Oncogene Homolog B
BSA	Bovine Serum Albumin
C3G	Rap Guanine Nucleotide Exchange Factor 1
Cas9	CRISPR-associated 9
CDC42	Cell Division Cycle 42
CDH1	*Cadherin 1*
cDNA	Complementary Deoxyribonucleic Acid
CH domain	Calponin Homology Domain
c-MYC	V-Myc Avian Myelocytomatosis Viral Oncogene Homolog
CPP	Cell Penetrating Peptide
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRK	V-Crk Avian Sarcoma Virus CT10 Oncogene Homolog
CRKL	V-Crk Avian Sarcoma Virus CT10 Oncogene Homolog-Like
CSC	Cancer Stem Cell
CSK	C-Terminal Src Kinase
DAPI	4’,6-Diamidino-2-Phenylindole
DMEM	Dulbecco’s Modified Eagle Medium
DMSO	Dimethyl Sulfoxide
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
DOCK180	Dedicator of Cytokinesis 1
DTT	Dithiothreitol
DUSP	Dual Specificity Phosphatase
E2F	E2F Transcription Factor
ECM	Extracellular Matrix

EDTA	Ethylenediaminetetraacetic Acid
EGF	Epidermal Growth Factor
EGFR	Epidermal Growth Factor Receptor
EGTA	Ethylene Glycol-bis(β-aminoethyl ether)-N,N,N′,N′-Tetraacetic Acid
EMT	Epithelial-Mesenchymal Transition
EphB	Ephrin Receptor B
ERBB2	Erb-B2 Receptor Tyrosine Kinase 2
ERK	Extracellular Signal-Regulated Kinase
FAK	Focal Adhesion Kinase
FCS	Fetal Calf Serum
FITC	Fluorescein Isothiocyanate
GAB1/2	GRB2 Associated Binding Protein 1/2
Gag	Group Antigens
GEF	Guanine Nucleotide Exchange Factor
GFP	Green Fluorescent Protein
GSEA	Gene Set Enrichment Analysis
GTP	Guanosine Triphosphate
H&E	Hematoxylin and Eosin
HAP	High Affinity Peptide
HGF	Hepatocyte Growth Factor
HIV1	Human Immunodeficiency Virus-1
HRP	Horseradish Peroxidase
IC50	Half Maximal Inhibitory Concentration
IGF1	Insulin-Like Growth Factor 1
(Co-)IP	(Co-)Immunoprecipitation
IκB kinase α/β	Inhibitor of Kappa B Kinase Alpha/Beta
JNK	JUN N-Terminal Kinase
KRAS	Kirsten Rat Sarcoma Viral Oncogene Homolog
LB	Lysogeny Broth
LPS	Lipopolysaccharide
MAPK	Mitogen-Activated Protein Kinase
MAPKK	MAP Kinase Kinase

MAPKKK	MAP Kinase Kinase Kinase
MEK1/2	Mitogen-Activated Protein Kinase Kinase 1/2
MET	Mesenchymal-Epithelial Transition
MKP	MAPK phosphatase
MMP	Matrix Metalloproteinase
mRNA	Messenger RNA
MS	Mass Spectrometry
MWCO	Molecular Weight Cut-Off
NF-κB	Nuclear Factor Kappa B
NGS	Next Generation Sequencing
NLS	Nuclear Localization Signal
NOTCH	Neurogenic Locus Notch Homolog Protein
p130CAS	P130 Crk-Associated Substrate
p90RSK	90 KDa Ribosomal Protein S6 Kinase 1
PBS(T)	Phosphate-Buffered Saline (with Tween 20)
PCR	Polymerase Chain Reaction
PDB	Protein Data Bank
PDGF	Platelet Derived Growth Factor
PH domain	Pleckstrin Homology Domain
PI3K	Phosphatidylinositol-4,5-Bisphosphate 3-Kinase
PMSF	Phenylmethylsulfonyl Fluoride
qRT-PCR	Quantitative Real-Time Polymerase Chain Reaction
RAC1	Ras-Related C3 Botulinum Toxin Substrate 1
RAF	Rapidly Accelerated Fibrosarcoma Proto-Oncogene Serine/Threonine-Protein Kinase
RAP1	Ras-Related Protein 1
RAS	Rat Sarcoma Viral Oncogene Homolog
RFP	Red Fluorescent Protein
RHO	Ras Homolog Family Member
RIPA	Radioimmunoprecipitation Assay
RMSD	Root Mean Square Deviation
RNA	Ribonucleic Acid

RoI	Region of Interest
RSK	Ribosomal Protein S6 Kinase
RTK	Receptor Tyrosine Kinase
SAM	Sterile Alpha Motif
SASH1	SAM and SH3 Domain Containing 1
SCID	Severe Combined Immunodeficiency
SDS	Sodium dodecyl sulfate
SH2 domain	Src Homology 2 Domain
SH3 domain	Src Homology 3 Domain
SH3N domain	N-Terminal Src Homology 3 Domain
SLUG	Snail Family Transcriptional Repressor 2
SLY	SH3 Domain Containing Protein Expressed in Lymphocytes
SMAD	Mothers Against Decapentaplegic Homolog
SNAIL	Snail Family Transcriptional Repressor 1
SOC	Super Optimal Broth with Catabolite Repression
SOS1	Son of Sevenless Homolog 1
STRING	Search Tool for the Retrieval of Interacting Genes/Proteins
TAK1	Transforming Growth Factor Beta-Activated Kinase 1
TAT	Trans-Activator of Transcription
TCGA	The Cancer Genome Atlas
TGFβ	Transforming Growth Factor Beta
TLR4	Toll Like Receptor 4
TNF	Tumor Necrosis Factor
TP53	*Tumor Suppressor P53*
TRAF6	TNF Receptor Associated Factor 6
TWIST	Twist Family BHLH Transcription Factor 1
UICC	Union International Contre le Cancer
UPL	Universal Probe Library
VIM	*Vimentin*
WNT	Wingless-Type MMTV Integration Site Family
WT	Wild Type
ZEB1/2	Zinc Finger E-Box Binding Homeobox 1/2

1. Introduction

1.1 Tumor progression and metastasis formation

1.1.1 Cancer evolution as fundamental principle of tumor progression

Locally restricted tumors are rarely lethal. However, benign tumor cells can progress into malignant, highly invasive cells, which are able to spawn distant metastases. Given that metastases account for the majority of cancer related-deaths, it is crucial to understand the mechanisms underlying tumor progression and metastatic spread (Gupta and Massague 2006). During the progression of epithelial-derived cancers, which account for the large majority of all cancer cases worldwide, tumor cells sequentially undergo several genetic alterations that increase their aggressiveness, ultimately generating invasively growing, pro-metastatic carcinoma cells (Vogelstein et al. 1988; Gupta and Massague 2006; Siegel, Miller, and Jemal 2018). Genes that are frequently altered during this process can be generally classified into (proto-)oncogenes and tumor suppressors. Oncogenes foster tumorigenesis and tumor progression, arising from proto-oncogenes by elevated expression or activating mutations. *KRAS*, for instance, is the most commonly mutated and thereby constitutively activated proto-oncogene in human cancer (Bos 1989; Haigis 2017). Tumor suppressors in turn counteract carcinogenesis and are therefore frequently lost at the genomic level or downregulated at the transcriptional level during tumor progression. The most prominent example is p53, which induces cell-cycle arrest, senescence and apoptosis upon diverse triggers like oncogenic signaling or DNA damage (Brady and Attardi 2010). Accordingly, mutational inactivation of the *TP53* gene is frequently observed in human cancers (Hollstein et al. 1991). In the case of colorectal cancer, the most frequently altered tumor suppressor gene is *APC* (Adenomatous polyposis coli), its loss constituting the earliest observed genetic aberration in this disease (Powell et al. 1992).

Tumor progression is believed to be mediated by clonal evolution (Nowell 1976). Certain genetic or epigenetic alterations occurring in a cancer cell facilitate an evolutionary advantage over other cells, such as increased cell proliferation and survival by acquisition of constitutively active $KRAS^{G12D}$ or $BRAF^{V600E}$ variants (Nowell 1976; Bos et al. 1987; Guo, Gong, et al. 2018). Cancer cells that have acquired such alterations are more likely to prevail. Therefore, malignant traits accumulate in the cancer cell population, while traits that counteract tumorigenesis or tumor progression are reduced or completely lost. The process of clonal

evolution is fostered by genome instability, loss of DNA repair mechanisms and changes in epigenetic regulators, leading to a rapid accumulation of chromosomal rearrangements, mutations, and changes in gene expression, respectively (Negrini, Gorgoulis, and Halazonetis 2010; Heng et al. 2013; Schnekenburger et al. 2016). This way, a genetically diverse pool of cancer cells is generated, from which traits are selected that constitute an evolutionary benefit. Due to the stochastic nature of cancer evolution, high levels of heterogeneity are not only found between individual patients, but also within an individual tumor (Nowell 1976; Gerlinger et al. 2012). This complexity is even further boosted by the emergence of cancer cell internal (e.g. DNA damage) and external, microenvironment-dependent stresses (e.g. hypoxia). Stresses foster cancer evolution by increasing the system dynamics (Heng et al. 2010; Carnero and Lleonart 2016). In general, the tumor microenvironment plays an important role during tumor progression, as it facilitates remodeling of the extracellular matrix (ECM), immune escape, secretion of growth-promoting and pro-angiogenic signals, as well as promoting the invasive phenotype of cancer cells (Yuan et al. 2016).

Genetic alterations involved in tumor formation and progression mostly affect fundamental biochemical and cellular processes, which are disrupted, altered or increased. Cancer cells need to overcome growth suppressors, senescence and apoptosis, induce angiogenesis to get access to sufficient nutrients and oxygen, counteract destruction by the immune system, and promote proliferative signaling (Hanahan and Weinberg 2011). Genetic and epigenetic alterations (in cooperation with microenvironmental influences) also allow the acquisition of a motile and invasive phenotype, leading to cancer cell detachment, invasion of adjacent healthy tissue, and potentially resulting in the formation of distant metastases (Hanahan and Weinberg 2011).

1.1.2 The metastatic cascade: life-threatening dissemination of cancer cells

The outgrowth and systemic dissemination of carcinoma cells, generating metastases at distant organs, accounts for most cancer-related deaths (Mehlen and Puisieux 2006; Gupta and Massague 2006). In colorectal cancer for instance, which accounts for the third most frequent cancer-related deaths in the United States, the presence of distant lesions is associated with drastically reduced survival (Siegel, Miller, and Jemal 2018). Patients with localized colorectal cancer (UICC stage I) have a 5-year survival of 96%, as opposed to patients presenting distant metastasis at the time of diagnosis, who have a dismal prognosis with a

5-year survival of only 13% (Nitsche et al. 2016). Due to the inherent difficulty of detecting and surgically removing metastases, it is of huge clinical importance to understand the mechanisms that are involved in their formation. The multi-step process during which cancer cells detach from the primary tumor and form a distant lesion is termed the metastatic cascade, as reviewed by (Pantel and Brakenhoff 2004). Carcinoma cells need to penetrate the basement membrane, infiltrate adjacent healthy tissue, enter the lymph or blood system, survive the harsh conditions during the circulation, leave the lymph or blood system, and colonize tissues with properties distinct from the primary tumor to finally spawn metastases (Figure 1) (Pantel and Brakenhoff 2004; Valastyan and Weinberg 2011).

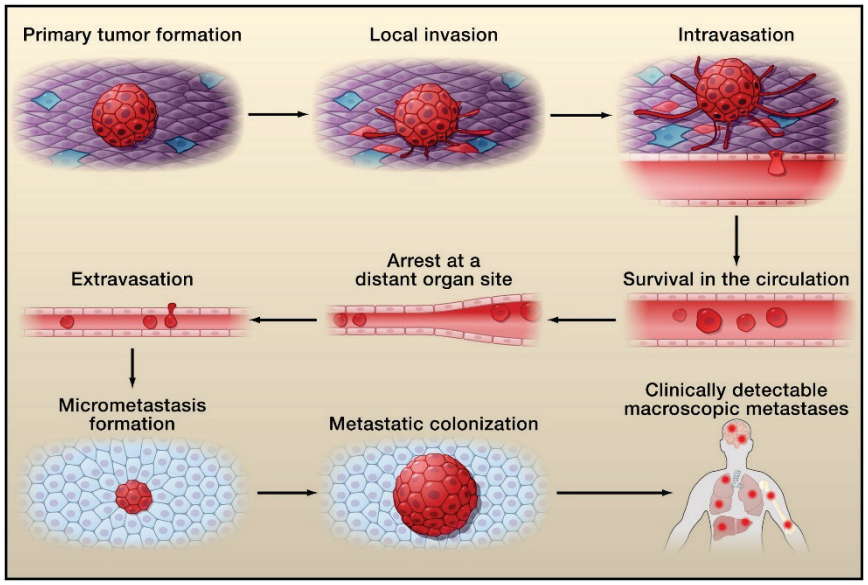

Figure 1: The metastatic cascade. To generate metastases tumor cells need to detach from the primary tumor, invade adjacent healthy tissue, enter the blood or lymph circulation (intravasation), survive the harsh conditions during the circulation, arrest within small capillaries at a distant organ, exit from the vasculature, survive in a "foreign" microenvironment at the distant site, and start again to proliferate extensively (figure from (Valastyan and Weinberg 2011)).

At the first step of the metastatic cascade, the infiltration of adjacent healthy tissue, cancer cells utilize distinct modes of migration. They can undergo collective cell migration as epithelial strands or clusters, as well as individual, single cell migration by acquisition of a mesenchymal phenotype or via amoeboid migration (Friedl and Wolf 2003). In addition to a motile phenotype, invasiveness also requires active remodeling of the ECM. For instance, ECM remodeling involves degradation of the ECM by matrix metalloproteinases to overcome steric obstructions like the basement membrane, or matrix crosslinking to increase stiffness and thereby invasiveness via integrin/focal adhesion signaling (Deryugina and Quigley 2006; Peng et al. 2016; Levental et al. 2009; Kessenbrock, Plaks, and Werb 2010). Endothelial cells, (myo-)fibroblasts, pericytes, lymphocytes, macrophages, and various other cell types are

present in the reactive tumor stroma. These cells are in continuous crosstalk with tumor cells, resulting in the acquisition of characteristics typical for wound healing and inflammation. This also contributes to the invasive phenotype of cancer cells (Joyce and Pollard 2009; Grivennikov, Greten, and Karin 2010). By means of remodeling the ECM and active migration, cancer cells can detach from the primary tumor and invade healthy tissue. As the second step of the metastatic cascade, cancer cells enter the blood or lymph system, a process called intravasation. Given that tumor cells need to enter the circulation in order to disseminate systemically and generate distant metastases, intravasation is an important step during the metastatic cascade (Pantel and Brakenhoff 2004; Valastyan and Weinberg 2011). Both active invasion of highly metastatic cancer cells that move towards blood vessels, as well as passive shedding of tumor cells represent mechanisms for intravasation (Condeelis, Wyckoff, and Segall 2000; Bockhorn, Jain, and Munn 2007). (Lymph-)angiogenesis in the primary tumor fosters metastasis formation by increasing the likelihood of cancer cells entering the circulation (Yamamura et al. 2001; Tien et al. 2001). In contrast to healthy tissue, tumor angiogenesis also induces the formation of an abnormal neo-vasculature that is continuously remodeled and leaky, thus being permissive for tumor cell penetration and intravasation (Carmeliet and Jain 2011). Once tumor cells have entered the lymphatic or blood circulation, they may potentially disseminate systemically to various distant organs. However, circulating tumor cells need to survive shear forces within the circulation, exposition to cells of the innate and adaptive immune system, as well as detachment from the ECM, which normally induces anoikis (Strilic and Offermanns 2017). Cancer cells can protect themselves from immune detection and shear forces by covering themselves with platelet aggregates (Jurasz, Alonso-Escolano, and Radomski 2004). Further, due to their large size, especially in association with platelet aggregates, circulating tumor cells quickly get stuck in small capillaries downstream of the intravasation event, and residence time of circulating tumor cells is thus very brief (Valastyan and Weinberg 2011). In colorectal cancer, most circulating tumor cells metastasize to the liver due to the direct anatomic passage through the portal vein (Gupta and Massague 2006). Once trapped in small capillaries, cancer cells can directly undergo extravasation by actively penetrating the endothelial and pericyte layer to reach the tissue parenchyma (Gassmann et al. 2009; Martin et al. 2010). Alternatively, cancer cells also proliferate within the capillary, and might thus mechanically rupture the vessel walls (Al-Mehdi et al. 2000). Disseminated tumor cells that have reached the site of future metastasis need to survive in a

foreign environment, as the ECM, stromal cells and tissue architecture are different to the primary tumor. Since the survival of cancer cells is highly dependent on microenvironmental clues, cancer cells utilize genetic and epigenetic programs to adapt to the foreign microenvironment. These programs depend on both the entity of the disseminated tumor cell (e.g. breast or prostate cancer) and the host tissue to which the cancer cells disseminated (e.g. lung or liver) (Valastyan and Weinberg 2011). The foreign tissue can also be primed by signals from the primary tumor to form a premetastatic niche that is permissive for survival and outgrowth of extravasated cancer cells (Psaila and Lyden 2009; Seubert et al. 2015). In addition to cell-intrinsic determinants, vessel anatomy, and physical accessibility of the future metastatic site, a permissive microenvironment also contributes to the tissue tropism of metastasis (Chambers, Groom, and MacDonald 2002; Fidler 2003).

Cancer cells that have been able to survive and adapt to the foreign microenvironment, in which they have settled, must undergo extensive proliferation to spawn clinically detectable, macroscopic metastases. A prerequisite for the outgrowth of metastases is therefore a high self-renewal capacity of disseminated tumor cells. According to the cancer stem cell (CSC) concept, tumors are hierarchically organized with a low number of CSCs producing a heterogenous population of rapidly proliferating, more differentiated cells (Clevers 2011). Due to their high self-renewal capacity, CSCs (or tumor-initiating cells) are discussed to be responsible for the formation of metastases (Gupta, Chaffer, and Weinberg 2009; Clevers 2011; Valastyan and Weinberg 2011). Therefore, in addition to the numerous barriers that cancer cells have to overcome during the metastatic cascade, the low abundance of disseminated cancer cells with self-renewal capacity contributes to the high inefficiency of the metastatic process (Chambers, Groom, and MacDonald 2002). Instead of activating proliferative programs, disseminated cancer cells can also form dormant micrometastases (e.g. if they fail to induce angiogenesis), which are resistant to most cancer therapies and might lead to metastatic relapse after years or even decades (Chambers, Groom, and MacDonald 2002).

Taken together, the metastatic cascade represents a complex sequential process, during which cell biological events (e.g. acquisition of an invasive phenotype) and non-autonomous interactions (e.g. platelet aggregates or tumor-associated stroma) are required. Cancer cells need to fulfill various requirements to accomplish the metastatic process. This suggests that the formation of metastases occurs as a late event during tumor progression, when highly

aggressive cells have been selected. However, this is still controversially discussed, as cancer cells could also potentially disseminate early and progress independently from the primary tumor (Klein 2009). Independent of the question of early or late initiation, the process of clonal evolution drives the metastatic cascade (Weigelt, Peterse, and van 't Veer 2005; McGranahan and Swanton 2017). Clonal evolution towards an aggressive phenotype, which is required to successfully pass the metastatic cascade, is also dependent on the microenvironment. For instance, hypoxia has been shown to promote an invasive phenotype and immune evasion (Rankin and Giaccia 2016). Metastasis can thus be seen as an interplay between a permissive microenvironment and a cancer cell subpopulation that acquired a pro-metastatic phenotype by clonal evolution (Fidler 2003). In support of this theory, highly metastatic cancer cells exhibit a higher genetic instability than non-metastatic clones (Fidler 2003; Gupta and Massague 2006). Since metastases account for the majority of cancer-related deaths, identification of tumor cell autonomous and non-autonomous key players of the metastatic cascade is of central clinical importance (Gupta and Massague 2006).

1.2 Epithelial-mesenchymal transition as a key process in the metastatic cascade

The transdifferentiation process of epithelial-mesenchymal transition (EMT) confers cancer cells many traits that are required to undergo the metastatic cascade (Drasin, Robin, and Ford 2011). During EMT, epithelial cells lose epithelial characteristics and acquire a mesenchymal phenotype (Kalluri and Weinberg 2009). This way, epithelial cells shed their apical-basal polarity and intercellular adhesions, while cell motility is increased. EMT is a fundamental physiological process required during embryonic development (type I EMT), e.g. to form the mesoderm and neural crest cells, as well as during wound healing and tissue regeneration (type II EMT), e.g. to form highly migratory cells that close the wound (Kim et al. 2017). However, EMT can also be hijacked by carcinoma cells during tumor progression (type III EMT) to acquire a highly invasive, pro-metastatic cellular phenotype (Kim et al. 2017). Due to the transient nature of EMT, with epithelial cells being able to shift between EMT and the reverse process, termed mesenchymal-epithelial transition (MET), EMT is considered to be a program that mediates cell plasticity, thereby allowing cancer cells to adapt to the various microenvironments that they encounter during the metastatic cascade (Kalluri and Weinberg 2009).

In general, EMT is governed by transcription factors like the ZEB family, TWIST, SNAIL and SLUG, which are primarily induced by signals from the tumor-associated stroma and repress the expression of epithelial genes, while inducing the expression of mesenchymal factors (Kim et al. 2017; Dongre and Weinberg 2019). Downregulation of epithelial and induction of mesenchymal genes by EMT promoting transcription factors further involves the action of chromatin modifiers and other epigenetic regulators (Sanchez-Tillo et al. 2010; Lin et al. 2010; Kiesslich, Pichler, and Neureiter 2013). EMT promoting transcription factors are under the control of microRNAs like the miR200 family, miR205 and miR30a (Park et al. 2008; Gregory et al. 2008; Kumarswamy et al. 2012). Expression of these microRNAs in turn is counteracted by EMT promoting transcription factors, leading to reciprocal negative regulatory feedbacks and thus "bistable switches", e.g. between the ZEB and miR200 families (Tian, Zhang, and Xing 2013). Despite of this, carcinoma cells are thought to primarily undergo a partial EMT and rarely enter a fully mesenchymal state (Dongre and Weinberg 2019).

EMT has been reported to be induced by a large variety of intracellular signaling pathways, with TGFβ signaling being the most prominent (Lamouille, Xu, and Derynck 2014). Besides canonical activation of SMADs, which induce the expression of EMT promoting transcription factors, TGFβ-induced EMT also involves activation of mitogen-activated protein kinases (MAPKs) and the PI3K/AKT pathway, as well as the spatiotemporal regulation of the small GTPases RHO, RAC and CDC42 that act on the actin cytoskeleton and cell-cell adhesions (Derynck, Muthusamy, and Saeteurn 2014; Lamouille, Xu, and Derynck 2014). WNT signaling has also been reported to induce EMT in various carcinomas, partially by repression of SNAIL and SLUG degradation (Yook et al. 2005; Wu et al. 2012; Duan et al. 2017). Of note, the epithelial hallmark protein E-cadherin acts as negative regulator of WNT signaling by sequestering β-catenin to adherens junctions at the cell membrane. Dissolution of adherens junctions during EMT induces increased β-catenin transcriptional activity (Nelson and Nusse 2004). As a result, nuclear β-catenin is increased in carcinoma cells with a more pronounced mesenchymal phenotype at the invasive front (Brabletz et al. 2001). Activation of NOTCH signaling has also been shown to induce EMT by transcriptional regulation of EMT inducing transcription factors (Timmerman et al. 2004; Sahlgren et al. 2008; Fre et al. 2009; Saad et al. 2010). Receptor tyrosine kinase signaling has been demonstrated to induce EMT as well, e.g. downstream of the ligands HGF, EGF, IGF1 and PDGF (Lamouille, Xu, and Derynck 2014). Activation of the RAS/RAF/ERK pathway and PI3K/AKT signaling, which are induced by

receptor tyrosine kinases, have been shown to promote EMT, e.g. by interfering with the interaction between E-cadherin and β-catenin (Xie et al. 2004; Makrodouli et al. 2011; Doehn et al. 2009; Irie et al. 2005; Wang et al. 2013; Janssen et al. 2006). Pro-inflammatory cytokines like TNF are also capable EMT-inducing stimuli (Bates and Mercurio 2003; Wang et al. 2013; Ricciardi et al. 2015). Further, the non-receptor tyrosine kinase SRC is an important key player during EMT, linking ECM-mediated activation of integrins to EMT (Zhao et al. 2012; Woodcock et al. 2009; Nagaharu et al. 2011). Importantly, there is extensive crosstalk between these EMT inducing signaling pathways (Maschler et al. 2005; Kim et al. 2009). *In vivo*, it is likely a combination of various extracellular signals released by the tumor-associated stroma and cancer cell-intrinsic properties that drive EMT (Gao et al. 2012).

An important hallmark of EMT is the downregulation of *CDH1* (encodes for the epithelial cell-cell adhesion molecule E-cadherin) expression by EMT inducing transcription factors to destabilize adherens junctions (Batlle et al. 2000; Comijn et al. 2001; Aigner et al. 2007). Upon downregulation of expression of *CDH1* and other genes involved in epithelial cell-cell adhesion and cell polarity, epithelial cells detach from each other, individualize and acquire a typical fibroblast-like morphology (Dongre and Weinberg 2019). Loss or reduction of E-cadherin may be compensated by increased N-cadherin levels, which promotes motility and invasiveness (Nieman et al. 1999; Hazan et al. 2000). Besides these changes in cell-cell adhesion molecules, EMT involves alterations in the integrin repertoire (Maschler et al. 2005; Yilmaz and Christofori 2009). Intermediate filaments are also subject of change, as expression of genes encoding for cytokeratins is downregulated, while vimentin levels are increased (Kalluri and Weinberg 2009). Vimentin in turn contributes to the mesenchymal, fibroblast-like morphology, as well as to increased cell motility and altered biophysical properties of metastatic cells (Mendez, Kojima, and Goldman 2010). Remodeling of the actin cytoskeleton is also crucial for EMT and thus the acquisition of a motile phenotype (Haynes et al. 2011; Peng et al. 2018). Additionally, the extracellular matrix is modified, e.g. by expression and secretion of matrix components or of proteinases like MMPs by carcinoma cells with a mesenchymal phenotype (Miyoshi et al. 2004; Zeisberg and Neilson 2009).

In summary, changes in gene expression by EMT promoting transcription factors generates mesenchymal properties in cells that are thus endowed with increased cell motility and invasiveness. Based on this, EMT is recognized as a crucial process for the initiation of the metastatic cascade (Kalluri and Weinberg 2009). However, it becomes more and more

apparent that EMT contributes to almost every step of the metastatic cascade. Both intra- and extravasation are promoted by a mesenchymal phenotype (Tsuji et al. 2008; Frose et al. 2018). Within the circulation, cells that underwent EMT are less prone to anoikis and promote coagulation (Bao et al. 2013; Smit et al. 2009; Howe, Cochrane, and Richer 2011; Bourcy et al. 2016). Further, EMT was described to counteract destruction by the immune system (Akalay et al. 2013; Chen et al. 2014; Tripathi et al. 2016). As presented before, these traits increase the likelihood of carcinoma cells that underwent EMT to survive in the circulation. EMT has also been associated with elevated resistance against radio- and chemotherapy due to various mechanisms, including downregulation of intrinsic pro-apoptotic genes and upregulation of ABC transporters and genes involved in DNA damage response (Fischer et al. 2015; Zheng et al. 2015; Saxena et al. 2011; Hsu et al. 2010; Theys et al. 2011; Kurrey et al. 2009). The anti-apoptotic phenotype of cells that underwent EMT also promotes survival of disseminated tumor cells that are confronted with the unfamiliar microenvironment of the distant tissue parenchyma. As discussed earlier, another natural barrier that needs to be overcome for the formation of metastases is the normally limited self-renewal capacity of disseminated tumor cells, as well as the low number of CSCs that are able to initiate the growth of a secondary tumor (Clevers 2011). EMT has been associated with a CSC-like phenotype, enabling highly efficient formation of mammospheres in the case of primary breast cancer, colony formation in soft agar and generation of xenograft tumors, which represent the current gold standards to assess cancer stemness (Mani et al. 2008; Morel et al. 2008). Finally, the reversible nature of EMT allows for the generation of highly proliferative, more epithelial-type cells through MET, once disseminated cancer cells have settled down at a distant site. EMT has been associated with reduced proliferation, and reversion of EMT thus restarts tumor cell proliferation and macrometastasis formation (Tsai and Yang 2013). Therefore, MET explains the epithelial phenotype of macrometastases, which frequently resemble the epithelial state of the primary tumor from which they originate, depending on the tumor grading (Yao, Dai, and Peng 2011).

Taken together, EMT plays a pro-metastatic role during most (if not all) steps of the metastatic cascade by endowing cancer cells with various malignant traits like migration, invasion, immune evasion, chemo- and radioresistance, reduced anoikis and apoptosis, as well as a CSC-like phenotype with multipotency and self-renewal capacity (Drasin, Robin, and Ford 2011).

1.3 Novel players modulating signaling pathways related to tumor progression and metastasis formation

1.3.1 SASH1: a candidate multi-tissue tumor suppressor and putative scaffold protein

The candidate tumor suppressor SAM and SH3 Domain Containing 1 (SASH1) belongs to the SLY family of predicted scaffold and signal adaptor proteins (Zeller et al. 2003). In contrast to SLY1 and SLY2, which are primarily expressed in lymphocytes, SASH1 shows a broad tissue expression, but is not expressed in T-cells (Beer et al. 2001; Uchida et al. 2001; Claudio et al. 2001; Zeller et al. 2003; Rimkus et al. 2006; von Holleben et al. 2011). The human *SASH1* locus consists of 20 exons and is located on chromosome 6q24.3, translating into a large protein of 1247 residues (Zeller et al. 2003). While the central region of SASH1 is similar to the other members of the SLY family, SASH1 additionally contains N- and C-terminal extensions (Martini et al. 2011). SASH1 features a central SH3 and SAM domain, which is shared among the SLY family, a bipartite nuclear localization signal, as well as a proline-rich region and a second SAM domain in the C-terminal part (Figure 2) (Martini et al. 2011). While SH3 domains are in general known to specifically bind polyproline motifs, SAM domains rather exhibit diverse modes of interactions, associating with other SAM domains, but also with non-SAM domain containing proteins and RNA (Kim and Bowie 2003; Kurochkina and Guha 2013).

Figure 2: Domain architecture of SASH1. SASH1 is a large scaffold protein of 1247 residues with an N-terminal bipartite nuclear localization signal, a central SH3 and SAM domain, as well as a proline-rich region and a second SAM domain in the C-terminal part (N/C = N- and C-terminus; NLS = nuclear localization signal; SH3 = Src homology 3 domain; SAM = sterile alpha motif; Pro-rich = proline-rich region).

SASH1 has been initially identified as a gene that is frequently lost in breast cancer, which was associated with poor prognosis and metastasis formation, suggesting a role as tumor suppressor (Zeller et al. 2003). Since its discovery, *SASH1* has been implicated as candidate tumor suppressor in a wide variety of cancer entities. The chromosomal region in which *SASH1* is located has been reported to be lost also in ovarian cancer, uterine cervical adenocarcinoma, prostate cancer, B-cell non-Hodgkin's lymphoma, endocrine pancreatic tumors, as well as in colorectal, thyroid and lung cancer (Shridhar et al. 1999; Acevedo et al. 2002; Srikantan et al. 1999; Zhang et al. 1997; Barghorn et al. 2001; Alcock et al. 2003; Zeller et al. 2003). Further, the clinical relevance of *SASH1* as putative tumor suppressor has been

demonstrated in several solid tumor entities, including by our own group for several independent patient cohorts with colorectal cancer, but also in breast cancer, gastric cancer, glioma and osteosarcoma (Rimkus et al. 2006; Nitsche et al. 2012; Alcock et al. 2003; Zeller et al. 2003; Burgess, Bolderson, Saunus, et al. 2016; Zhou et al. 2018; Yang et al. 2015; Meng et al. 2013). Work from our clinic showed that *SASH1* expression is downregulated especially in later stages of colorectal cancer, indicating a role in tumor progression (Rimkus et al. 2006). Indeed, downregulation of *SASH1* has been associated with the formation of metachronous distant metastases, as well as poor post-operative survival due to metastatic disease recurrence in stage II colorectal cancer (Rimkus et al. 2006; Nitsche et al. 2012). *SASH1* is thus a candidate multi-tissue tumor suppressor that potentially counteracts tumor progression and metastasis formation.

In addition to cancer, SASH1 has been implicated in atherosclerosis. *SASH1* expression is increased in circulating monocytes from smokers, which positively correlates with the number of carotid plaques (Verdugo et al. 2013). *SASH1* expression is also increased in atherosclerotic carotids of smokers (Weidmann et al. 2015). Downregulation of *SASH1* in turn results in increased endothelial cell proliferation, migration and angiogenesis (Weidmann et al. 2015). Further, point mutations within the central region of *SASH1* have been associated with the formation of genetic skin pigmentation defects, characterized by hyper- and hypopigmented macules (Xing et al. 2003; Zhou et al. 2013; Shellman et al. 2015; Courcet et al. 2015). Importantly, a mutation within *SASH1* leading to an E617K exchange has been associated with a predisposition to recurrent spinocellular carcinoma (Courcet et al. 2015).

In the context of cancer, several studies assign SASH1 a negative regulatory role in a wide variety of signaling pathways, including Hedgehog and PI3K/AKT signaling (He et al. 2016; Zong et al. 2016; Sun et al. 2015; Pan and Liu 2016). In accordance, SASH1 has been independently identified as negative regulator of insulin-induced AKT phosphorylation (Huang et al. 2009). Functionally, SASH1 has been described to promote apoptosis, while counteracting malignant traits like migration, invasiveness, cell viability and proliferation in several cancer cell lines (Chen et al. 2012; Yang et al. 2012; Lin et al. 2012; Meng et al. 2013; Martini et al. 2011). Nevertheless, the data was primarily obtained from cell lines recombinantly expressing *SASH1*, which does not reflect its downregulation or loss during tumor progression and should thus be regarded with caution. Burgess and colleagues have suggested a caspase-3-dependent mechanism for SASH1-mediated apoptosis (Burgess, Bolderson, Adams, et al. 2016). Cleavage

of SASH1 by caspase-3 leads to nuclear translocation of the C-terminal part of SASH1 (amino acids 231-1247) and NF-κB-dependent apoptosis (Burgess, Bolderson, Adams, et al. 2016). SASH1 has also been linked to NF-κB activation by acting as scaffolding factor in endothelial TLR4 signaling, independently binding to TRAF6, TAK1, IκB kinase α, and IκB kinase β (Dauphinee et al. 2013). This way, SASH1 induces increased activation of NF-κB, JNK, and p38 in response to LPS, promoting endothelial cell migration (Dauphinee et al. 2013). Nevertheless, the increase in endothelial cell migration upon downregulation of *SASH1* seems to be LPS-specific, as an opposite, inhibitory effect was observed in human aortic endothelial cells (Weidmann et al. 2015). Further, SASH1 has been shown to associate with the F-actin cytoskeleton, being enriched at F-actin rich membrane ruffles and lamellipodia (Martini et al. 2011). Accordingly, SASH1 co-precipitates with the actin regulatory protein cortactin, and both proteins exhibit a pronounced co-localization at membrane ruffles (Martini et al. 2011). Of note, a similar phenotype was observed for the homolog SLY2 (von Holleben et al. 2011).

Several hints point towards a regulation of SASH1 by phosphorylation. Phosphoproteomic analyses have identified serine phosphorylation of SASH1 at S407, S837 and S839 in response to ionizing radiation (Yang et al. 2010). Another prominent phosphorylation occurs at S90, which is likely mediated downstream of PI3K and p90RSK, providing a binding site for 14-3-3 proteins (Dubois et al. 2009). Within screenings for EphB signaling effectors, SASH1 was found to be tyrosine phosphorylated at Y759 in response to ephrinB stimulation (Zhang et al. 2006; Zhang, Fenyo, and Neubert 2008). In addition to posttranslational modifications, *SASH1* is likely subjected to epigenetic regulation by DNA methylation, as well as negative regulation by microRNAs, which represent potential mechanisms of *SASH1* downregulation during tumor progression (Sheyu et al. 2013; Peng, Wei, and Liren 2014; Zhu et al. 2019).

In summary, many studies point towards a variety of functions of SASH1, which play a role during tumor progression, e.g. negative regulation of proliferation and invasiveness. Nevertheless, the tumor suppressive role of SASH1 remains to be elucidated *in vivo*. Furthermore, the underlying molecular mechanisms of how SASH1 potentially counteracts tumor progression and metastasis formation are unknown. Of note, experimental data generated within a cooperation with Dr. Ewa Ninio (Sorbonne University, Paris, France) prior to the present thesis project allowed the identification of potential SASH1 interaction partners by the yeast two-hybrid system, proposing the CRK family of signal adaptor proteins as promising candidates.

1.3.2 Role of the CRK family of signal adaptors in tumor progression

The CRK family of signal adaptors plays an important role during tumor progression by mediating the activation of several oncogenic signaling pathways. CRK (CT10 regulator of kinase) was initially discovered in 1988 as viral oncogene v-*Crk* that induced massive tyrosine phosphorylation and was sufficient to transform chicken embryonic fibroblasts, despite the lack of any catalytical activity (Mayer, Hamaguchi, and Hanafusa 1988). The human *CRK* family consists of *CRK* on chromosome 17p13 and its closely related homolog *CRKL* (CRK-like) on chromosome 22q11 (ten Hoeve et al. 1993; Feller 2001; Bell and Park 2012). CRKL and CRK share a high sequence and structural homology, especially regarding the SH2 and SH3 domains (ten Hoeve et al. 1993). Due to alternative splicing, *CRK* is expressed as three different transcripts encoding for CRK-I, CRK-II and CRK-III (Figure 3) (Matsuda et al. 1992; Prosser et al. 2003).

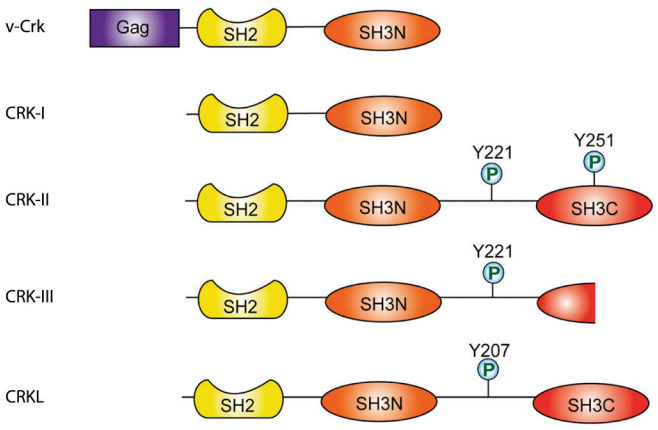

Figure 3: Domain architecture of the CRK family of signal adaptor proteins. The CRK family was initially discovered by identification of an oncovirus encoding the v-Crk fusion protein. The human genome in turn encodes *CRK*, which is transcribed as three distinct splice variants (CRK-I, CRK-II and CRK-III), and the homolog *CRKL*. The SH domains of CRK and CRKL have a high sequence similarity and the same ligand binding preferences (Gag = viral capsid protein; SH2/3 = Src homology 2/3 domain; P = phosphorylation site; figure adapted from (Bell and Park 2012)).

The SH2 domains of CRK and CRKL preferentially bind phospho-YXXP motifs of a large variety of target proteins (Feller 2001). Interaction partners of the CRK family SH2 domains include integrin-associated scaffolds like paxillin and p130CAS, as well as receptor tyrosine kinase-associated proteins such as GAB1 and GAB2 (Birge et al. 1993; Sakai et al. 1994; Garcia-Guzman et al. 1999; Sakkab et al. 2000; Salgia et al. 1995; Salgia et al. 1996; Crouin et al. 2001). SH3 domains typically bind to polyproline motifs (Kurochkina and Guha 2013). However, the C-terminal SH3 domains of CRK adaptors are believed to be inactive in that regard (Muralidharan et al. 2006). In contrast, the N-terminal SH3 domains specifically recognize the

polyproline type II motif with the PXXPX[K/R] consensus sequence, but preferentially to PXXPXK (Posern et al. 1998; Bell and Park 2012). This way, CRK adaptors associate with PXXPXK motif-containing guanine nucleotide exchange factors (GEFs) like C3G, DOCK180 and SOS1 (Knudsen, Feller, and Hanafusa 1994; Tanaka et al. 1994; Matsuda et al. 1994; Hasegawa et al. 1996). CRK family proteins act as signal adaptors by binding to membrane-associated, tyrosine phosphorylated proteins via the SH2 domains, and thereby recruiting effector proteins via the N-terminal SH3 domains into close proximity to the membrane (Feller 2001; Bell and Park 2012). CRK family mediated recruitment of guanine nucleotide exchange factors to the membrane in turn leads to local GTP loading and activation of membrane-associated small GTPases like RAS (via SOS1), RAC1 (via DOCK180) and RAP1 (via C3G) (Feller 2001). In summary, CRK family adaptor proteins spatiotemporally couple upstream signals from various receptor tyrosine kinases and integrins to downstream activation of small GTPases. This way, CRK proteins promote a diverse set of cellular processes, including proliferation, adhesion, migration and invasion, as well as alterations of the actin cytoskeleton (Figure 4) (Bell and Park 2012). Activation of RAP1 downstream of CRK proteins and C3G induces increased integrin-mediated adhesion, indicating a role in integrin inside-out signaling (Arai et al. 1999; Bos et al. 2003). Additionally, CRK adaptors are implicated in upstream regulation of kinases, in particular ABL and SRC (Sriram et al. 2011; Sabe, Shoelson, and Hanafusa 1995; Watanabe et al. 2009). Focal adhesion kinase (FAK) was also described to be activated in a manner dependent on the increased activation of SRC (Akagi et al. 2002; Iwahara et al. 2004). It has been proposed that CRK proteins mediate SRC activation by binding to and inhibiting C-terminal Src kinase (CSK), a negative regulatory kinase of SRC (Sabe, Shoelson, and Hanafusa 1995; Watanabe et al. 2009). However, the molecular mechanism of how CRK proteins mediate activation of SRC and other components of integrin signaling is still not fully understood (Bell and Park 2012).

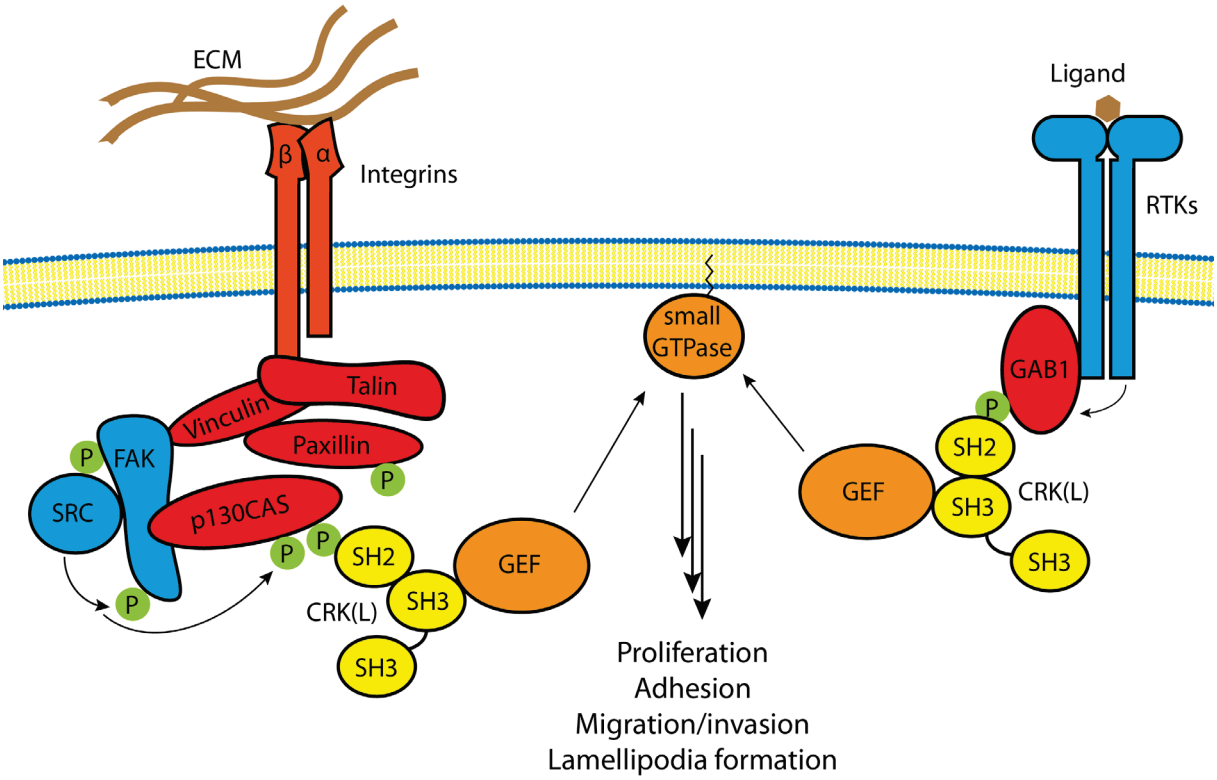

Figure 4: CRK adaptor proteins mediate activation of small GTPases downstream of integrin and receptor tyrosine kinase signaling. Upon engagement of integrins by the extracellular matrix (ECM) or of receptor tyrosine kinases (RTKs) by their ligands, scaffold proteins are phosphorylated (indicated by P). The SH2 domains of CRK and CRKL associate with tyrosine phosphorylated scaffolds, leading to membrane recruitment of guanine nucleotide exchange factors (GEFs) via the N-terminal SH3 domains of CRK proteins. Following membrane recruitment, GEFs activate small GTPases to induce downstream signaling, which in turn activates diverse cellular processes, depending on the activated small GTPase and the cellular context.

Despite the high sequence homology of the SH2 and SH3 domains of CRK family proteins, the tertiary structure of CRK-II and CRKL is different, especially regarding domain organization of the SH2 and N-terminal SH3 domains (Kobashigawa et al. 2007; Jankowski et al. 2012). While CRK-I and CRKL have an overall open conformation, CRK-II forms a compact structure (Bell and Park 2012). This is due to a hydrophobic core within the linker between the SH3 domains of CRK-II, leading to partial occlusion of the N-terminal SH3 domain (Kobashigawa et al. 2007; Bell and Park 2012). Further, CRK proteins feature an autoinhibitory mechanism, as the SH2 domains form an intramolecular interaction with phosphorylated Y221 (for CRK-II) and Y207 (for CRKL), thereby abolishing interactions involving the SH2 domains (Rosen et al. 1995; Senechal et al. 1998). However, upon CRK-II phosphorylation at residue Y221, the N-terminal SH3 domain is also inhibited, while CRKL Y207 phosphorylation has only minor effects on its N-terminal SH3 domain (Jankowski et al. 2012). Due to these discrepancies in domain organization, and notably despite of the overall high structural and functional similarity of the SH2 and N-terminal SH3 domains, CRK and CRKL have been shown to differentially modulate

the signaling pathways in which they are involved (Antoku and Mayer 2009; Yanagi et al. 2012).

The CRK family adaptor proteins CRK-I, v-Crk and CRKL have been shown to be sufficient to transform fibroblasts in culture, highlighting their role as (proto-)oncogenes (Mayer, Hamaguchi, and Hanafusa 1988; Matsuda et al. 1992; Senechal, Halpern, and Sawyers 1996). Since this discovery, many studies have demonstrated a pro-tumorigenic role of the *CRK* family in various human cancer entities, including leukemia, lung cancer, glioblastoma, breast cancer, gastric cancer, hepatocellular carcinoma, ovarian carcinoma, bladder cancer and colon cancer (Bell and Park 2012; Tsuda and Tanaka 2012; Guo, Liu, and Sun 2014). In general, high expression of *CRK* family genes in carcinomas has been associated with advanced stages, high-grade diseases, metastasis formation and poor patient survival (Bell and Park 2012; Guo, Liu, and Sun 2014). However, transgenic mouse models carrying *Crkl* amplifications or recombinantly expressed *Crk-II* developed tumors only with low penetrance and long latency, suggesting that additional oncogenic events are necessary for tumor initiation (Hemmeryckx et al. 2001; Fathers et al. 2010). Thus, CRK family proteins are rather important for tumor progression and metastasis formation, mediated by increased cell survival, proliferation, motility and invasiveness (Bell and Park 2012). Recently, CRK proteins were described to induce EMT in several cancer cell lines, which potentially represents a mechanistic link between the CRK family and tumor progression (Cheng, Guo, Yang, and Yang 2015; Han et al. 2015; Matsumoto et al. 2015; Elmansuri et al. 2016). It remains to be addressed, on which pathways CRK proteins primarily act to induce EMT. As described above, unpublished data obtained by our working group proposed the CRK family, especially CRKL, as new interaction partners of SASH1. The functional analysis of this putative interaction was a central aim of the present thesis.

1.3.3 The phosphatase DUSP5 as a negative regulator of ERK1/2 signaling

In addition to the factors discussed above, mitogen-activated protein kinases (MAPKs) are highly conserved kinase modules crucially involved in tumor progression and metastasis formation (Widmann et al. 1999; Reddy, Nabha, and Atanaskova 2003). A total of six groups of MAPKs have been described in mammals: extracellular signal-regulated kinase (ERK) 1/2, ERK3/4, ERK5, ERK7/8, Jun N-terminal kinase (JNK) 1/2/3 and four p38 isoforms (Dhillon et al. 2007). The kinases ERK1/2 are the best characterized MAPKs, which are primarily activated

downstream of receptor tyrosine kinases that mediate GTP loading of the small GTPase RAS. Activation of RAS in turn induces membrane recruitment and activation of the mitogen-activated protein kinase kinase kinase (MAPKKK) RAF, which then phosphorylates the mitogen-activated protein kinase kinases (MAPKK) MEK1/2. The activated kinases MEK1/2 directly activate ERK1/2 by dual phosphorylation at T202 and Y204, whereby both phosphorylation events are required for activation (Payne et al. 1991; Lai and Pelech 2016; Shaul and Seger 2007; Keyse 2008). Once activated, ERK1/2 act as potent protein kinases that phosphorylate a large number of cytosolic and nuclear substrates, including transcription factors, ribosomal s6 kinases (RSKs) and other protein kinases, as well as phosphatases, cytoskeleton- and apoptosis-related proteins and various other signaling factors (Yoon and Seger 2006; Shaul and Seger 2007). This way, the ERK1/2 pathway plays important roles during tumor initiation and progression by controlling cell proliferation, apoptosis, differentiation, migration and invasiveness (Dhillon et al. 2007; Wada and Penninger 2004; Kohno and Pouyssegur 2006). ERK activation is frequently upregulated in a wide variety of cancer entities (Kohno and Pouyssegur 2006). Upregulation of this pathway frequently occurs by mutational activation of RAS GTPases and BRAF, as well as by amplification of EGFR, ERBB2 and other receptor tyrosine kinases (Janssen et al. 2002; Downward 2003; Dhillon et al. 2007). ERK signaling is estimated to be deregulated in approximately 30% of human tumors, highlighting its tumor promoting functions (Keyse 2008). Similar to the ERK pathway, JNK and p38 activation involves upstream acting MAPKKK and MAPKK modules that induce activation loop phosphorylation of the TXY motif (Dhillon et al. 2007).

To allow regulation of duration and magnitude of MAPK signaling, serine/threonine phosphatases, tyrosine-specific phosphatases and dual-specificity phosphatases (DUSPs) are employed, which specifically dephosphorylate MAPKs, rendering them inactive (Keyse 2000, 2008). MAP kinase phosphatases (MKPs) belong to the DUSP family and are able to specifically dephosphorylate both threonine and tyrosine residues within MAPKs (Kondoh and Nishida 2007). MKPs contain an N-terminal MAPK-binding domain with a kinase interaction motif, which mediates substrate recognition and thus specific binding, as well as a C-terminal catalytic dual-specificity phosphatase domain (Muda et al. 1998; Keyse 2008). The MKP family is subdivided into three distinct groups. Type I MKPs comprise DUSP1, DUSP2, DUSP4 and DUSP5, which contain an N-terminal nuclear localization signal and thereby exhibit a nuclear localization. Type I MKPs act as feedback regulators of nuclear MAPK signaling, and are thus

highly inducible (Kondoh and Nishida 2007). Type II MKPs feature a cytoplasmic localization, while type III MKPs localize to both cytoplasm and nucleus (Kondoh and Nishida 2007).

Due to their inducible nature and varying substrate specificities, the role of MKPs in cancer is rather complex (Keyse 2008). DUSP1 for example exerts anti-apoptotic functions likely by inactivation of JNK, and increased expression is associated with poor progression-free survival in ovarian carcinomas (Magi-Galluzzi et al. 1997; Denkert et al. 2002). DUSP1 has also been described to promote chemoresistance, likely by counteracting JNK activation, indicating a tumor promoting role (Small et al. 2007; Sanchez-Perez et al. 2000). An opposing role was shown for the ERK-specific type II MKP DUSP6 in pancreatic cancer, as its expression counteracted ERK activation and cell growth (Furukawa et al. 1998; Furukawa et al. 2003). Thus, depending on substrate specificity, regulation, localization and cellular context, MKPs can act as tumor promoting or suppressing factors. In contrast to the other type I MKP members, which exhibit a broad substrate specificity, DUSP5 is highly specific for ERK1/2 (Keyse 2008; Mandl, Slack, and Keyse 2005). In addition to dephosphorylation of ERK1/2, DUSP5 has been shown to anchor ERK2 in the nucleus (Mandl, Slack, and Keyse 2005). *DUSP5* has also been described as a transcriptional target of p53, as well as a downstream of ERK1/2 activation (Ueda, Arakawa, and Nakamura 2003; Kucharska et al. 2009). Elevated expression of *DUSP5* in turn is associated with better disease-free survival in advanced human colorectal cancer, highlighting *DUSP5* as independent prognostic factor (Yan et al. 2016). Nevertheless, the role of DUSP5 in cancer is not yet fully understood. While DUSP5 likely counteracts elevated cell proliferation, potentially counteracting tumorigenesis and tumor progression, functional evidence *in vivo* is essentially required for a better understanding (Ueda, Arakawa, and Nakamura 2003; Wang et al. 2010). Unpublished data from our group, based on the analysis of genetic and orthotopic xenograft mouse models, indicated a specific metastasis-suppressive role of DUSP5 in colorectal cancer.

1.4 Motivation

Given that the systemic dissemination of tumor cells with the subsequent formation of metastases is the major cause for cancer-related deaths, it is crucial to identify new players in the metastatic cascade, and to understand their function and mechanism of action. *SASH1* is a candidate tumor suppressor in various cancer entities and its downregulation has been clinically implicated in metastasis formation of colorectal cancer. However, a causal relationship was lacking, and the underlying mechanism was still largely unknown.

This thesis aims to investigate the role of the candidate tumor suppressor SASH1 in tumor progression in the context of colorectal cancer. The focus was to study the underlying molecular and cellular mechanisms of how SASH1 acts as tumor suppressor, as well as to identify if and how SASH1 counteracts metastasis formation. Analysis of the putative contribution of the adaptor protein CRKL as novel interaction partner of SASH1 was also of major interest. Furthermore, the candidate tumor suppressor DUSP5 was analyzed in a complementary approach for its regulatory role in ERK activation and cell proliferation in the context of colorectal cancer.

Based on the molecular mechanisms unraveled within this thesis, another aim was finally to identify potential therapeutic target structures, which could be inhibited to counteract metastasis formation and minimal residual disease.

2. Material and methods

Important buffer compositions are listed in Table 2, important chemicals in Table 5, important consumables in Table 6, commercial kits in Table 7, bacteria, cell lines and mouse models in Table 8, and technical devices in Table 9.

2.1 Molecular biological methods

2.1.1 Isolation of genomic DNA

Cells or tissue were lysed in 100 µl lysis buffer (0.1 M NaCl, 50 mM Tris/HCl (pH 7.5), 0.5% (w/v) SDS, 5 mM EDTA). Tissue lysates were additionally incubated with 1 mg ml^{-1} Proteinase K (Sigma-Aldrich) overnight at 56 °C. Afterwards, 200 µl isopropanol was added to the lysates, mixed and incubated for 10 min at room temperature. Centrifugation was performed at 18400 g for 10 min at 4 °C. The supernatant was discarded, and the sediment was washed once with 70% ethanol. The sediment was dried, resuspended in 100 µl ddH_2O, and DNA concentration was spectrophotometrically estimated with a NanoDrop 1000 (Thermo Fisher).

2.1.2 Generation of plasmids

To clone human *CRKL* into the plasmid pmRFP-N2, RNA was purified from HEK293 cells and transcribed into cDNA as described in section 2.1.7. The *CRKL* coding sequence was isolated by PCR using primers #1/2 (Table 3) with the Phusion Hot Start II DNA polymerase system (Thermo Fisher) according to the manufacturer's protocol. All oligonucleotides were synthesized by Metabion (Planegg, Germany). After purification using the QIAquick PCR Purification Kit (Qiagen), the PCR product, as well as the empty vector, were digested with *KpnI* and *XmaI* (Thermo Fisher) for 1 h at 37 °C. Empty vector and PCR product were subjected to agarose gel electrophoresis and purified using the QIAquick Gel Extraction Kit (Qiagen), before ligation was performed with a 5:1 molar ratio (PCR product : plasmid), using T4 DNA ligase (New England Biolabs) for 10 min at 4 °C, and subsequently at room temperature for further 10 min. Next, the ligation preparation was subjected to exonuclease digestion using Plasmid-Safe ATP-Dependent DNase (Lucigen) according to the manufacturer's protocol. Chemically competent *E. coli* TOP10 (Thermo Fisher) were thawed on ice for 15 min, before 4 µL of the exonuclease reaction were added. After 30 min at 4 °C, the solution was incubated for 90 s at 42 °C, and then allowed to cool down on ice for 2 min. Finally, 400 µL SOC medium was added, and bacteria were incubated for 1 h at 37 °C under continuous shaking at 300 rpm,

before they were plated onto agar plates containing 40 µg ml^{-1} Kanamycin (Sigma-Aldrich). Successful cloning was verified by sequencing (Eurofins Genomics, Ebersberg, Germany) of plasmids, which were obtained from bacteria grown in LB medium and purified using the QIAGEN Plasmid Mini Kit (Qiagen) according to the manufacturer's protocol. The N-terminal SH3 domain of CRKL was cloned in the plasmid pET-21a with a similar protocol, using primers #3/4 (Table 3) for PCR-based isolation of the domain from the template pmRFP-N2-CRKL, *NdeI* and *XhoI* (Thermo Fisher) for restriction digest and 100 µg ml^{-1} Ampicillin (Sigma-Aldrich) for selection. After amplification in *E. coli* TOP10 (Thermo Fisher), the vector was transformed into *E. coli* One Shot BL21 Star (DE3) (Thermo Fisher) for protein expression. The vector pEGFP-HAP was generated with a similar approach, using synthetic oligos #55/56 (Table 3) as insert, and *EcoRI* and *BamHI* (Thermo Fisher) for restriction digest of pEGFP-C2. Before ligation, oligos were phosphorylated and annealed as described in section 2.1.3. The vector pIRES2-AcGFP1-3XFLAG-DUSP5 was obtained from Dr. Bernhard Holzmann (Technical University of Munich, Germany) and was used for recombinant expression of *DUSP5*.

2.1.3 Cloning of guide RNA target sequences into PX459

The vector pSpCas9(BB)-2A-Puro (PX459) V2.0 was a gift from Feng Zhang (Addgene plasmid #62988; http://n2t.net/addgene:62988; RRID:Addgene_62988) (Ran et al. 2013). To avoid off-target effects, two different guide RNA sequences per gene were designed using the MIT CRISPR design tool (crispr.mit.edu). Cloning of the guide RNA sequences (oligonucleotides #13-28, Table 3) into PX459 was performed according to (Ran et al. 2013). Briefly, oligonucleotides were phosphorylated and annealed with the following reaction mixture:

Component	Volume [µl]
guide RNA (top strand; 100 µM)	1
guide RNA (bottom strand; 100 µM)	1
T4 ligation buffer (10x)	1
T4 Polynucleotide Kinase	1
ddH_2O	to 10

The reaction was performed with a thermocycler at 37 °C for 30 min. After incubation for 5 min at 95 °C, oligonucleotides were annealed by decreasing the temperature to 25 °C at -5 °C min^{-1}.

The oligonucleotides were then cloned into PX459 using a combined restriction/ligation approach:

Component	Volume [µl]
PX459 (100 ng/µl)	1
Annealed oligonucleotides (1:200 dilution)	2
Tango buffer (10x)	2
DTT (10 mM)	1
ATP (10 mM)	1
FastDigest *BbsI*	1
T7 ligase	0.5
ddH_2O	to 20

The reaction was performed by cycling between 37 °C for 5 min and 21 °C for 5 min for 6 times.

To digest non-ligated linear DNA, the following reaction was performed:

Component	Volume [µl]
Ligation reaction from previous step	11
PlasmidSafe buffer (10x)	1.5
ATP (10 mM)	1.5
PlasmidSafe Exonuclease	1

The reaction was performed for 30 min at 37 °C, followed by incubation for 30 min at 70 °C. Finally, 4 µl of the mixture was used to transform chemically competent bacteria as described in section 2.1.2.

2.1.4 Site-directed mutagenesis

Site-directed mutagenesis was performed by the protocol of Zheng and colleagues (Zheng, Baumann, and Reymond 2004). One pair of primers was used for each PXXPXK motif within pEGFP-C2-SASH1 (vector obtained from (Martini et al. 2011)), as well as for the R39K and W160R exchanges within pmRFP-N2-CRKL. Primers #5/6 were used for mutagenesis of PXXP1, #7/8 for PXXP2, #9/10 for PXXP3, #11/12 for W160R, and #61/62 for R39K (for primers see Table 3).

The following PCR setup was performed:

Component	Final concentration
Primer A	0.5 µM
Primer B	0.5 µM
Template vector	1 ng μl^{-1}
dNTP mix	200 µM each
Phusion HF Buffer	1x
Phusion Hot Start II DNA Polymerase	0.02 U μl^{-1}

Step	Temperature [°C]	Time [s]	Cycles
1	98	60	1
2	98	30	18
3	60	50	
4	72	30 per kb	
5	72	600	1
6	4	-	

The PCR reaction was then purified using the QIAquick PCR Purification Kit (Qiagen) and subjected to *DpnI* (Thermo Fisher) digestion for 1 h at 37 °C. Plasmids were then transformed into *E. coli* TOP10 (Thermo Fisher) as described in section 2.1.2, purified, and mutations were subsequently verified by sequencing (Eurofins Genomics, Ebersberg, Germany). Mutagenesis of NLS1 and NLS2 within SASH1 has already been performed by Alexandra Gnann (Technical University of Munich, Germany) using primers #57-60 (Table 3).

2.1.5 Sequencing of CRISPR/Cas9-targeted *SASH1* locus

Primers #29/30 and #31/32 (Table 3) were used to amplify the *SASH1* locus targeted by guide RNA 1 and guide RNA 2, respectively. Isolated genomic DNA (see section 2.1.1) was used as template for PCR using the Phusion Hot Start II DNA polymerase system (Thermo Fisher) according to the manufacturer's protocol. The PCR product was then sequenced by NGS CRISPR amplicon sequencing (CCIB DNA Core, Massachusetts General Hospital, Boston, USA).

2.1.6 Genotyping of *Crkl*^{flox} mice

DNA extraction and purification were performed as described in section 2.1.1. To identify the presence of the wild type or the *Crkl*$^{tm1c(EUCOMM)Hmgu}$ (Jackson Laboratory) allele, the following PCR reaction was performed using the JumpStart REDTaq system (Sigma-Aldrich):

Component	Volume [µl]
Primer #33 [25 µM] [1]	1
Primer #34 [25 µM] [1]	1
Genomic DNA	1
JumpStart Reaction Mix	12.5
ddH_2O	to 25

[1] see Table 3

Step	Temperature [°C]	Time [s]	Cycles
1	94	120	1
2	94	30	
3	60	30	35
4	72	20	
5	72	120	1
6	4	-	

The PCR product was then subjected to agarose gel electrophoresis to analyze the specific banding pattern.

2.1.7 RNA isolation, reverse transcription and qPCR

Isolation of RNA was performed with the RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol. A total amount of 1 µg RNA was reverse transcribed using the RevertAid H Minus Reverse Transcriptase system (Thermo Fisher) with random hexamer and Oligo$(dT)_{18}$ primers:

Component	Amount
RNA	1 µg
Random hexamer	0.2 µg
Oligo(dT)$_{18}$	0.5 µg
ddH_2O	to 12.5 µl

The mixture was then incubated for 10 min at 70 °C, before being cooled down on ice. Afterwards, the following reaction was performed:

Component	Amount
Mixture from heating step	12.5 µl
Reaction buffer (5x)	4 µl
RNase inhibitor	20 U
dNTP mix	1 mM each
RevertAid H Minus RT	200 U
ddH_2O	to 20 µl

The reaction was incubated for 10 min at room temperature, subsequently for 60 min at 42 °C, and then for 5 min at 95 °C. The cDNA was diluted 1:10 or 1:100 with ddH_2O for further use. cDNA of tumor samples was already prepared by Dr. Ulrich Nitsche, Alexander Balmert and Sabine Bauer (Technical University of Munich, Germany) as described before (Nitsche et al. 2012; Franke et al. 2019). Finally, gene expression was determined by qPCR using the LightCycler 480 II system (Roche) with primers #35-54 (Table 3) and corresponding universal probe library (UPL) probes. *HPRT* expression was used as internal reference transcript. cDNA obtained from HEK293 cells or pooled cDNA from human mucosa samples was used for normalization.

The following qPCR setup was carried out in 96 well plates:

Component	Volume per well
Gene expression Mastermix Abi	10 µl
Primer A [20 µM]	0.2 µl
Primer B [20 µM]	0.2 µl
UPL probe	0.2 µl
Diluted cDNA	5 µl
ddH_2O	to 20 µl

The qPCR program "monocolour hydrolysis probe UPL" from the manufacturer was used. The data was analyzed using the LightCycler 480 v1.5 software.

2.1.8 RNA sequencing

RNA sequencing was performed in cooperation with Dr. Thomas Engleitner (Technical University of Munich, Germany), using RNA isolated as described in section 2.1.7. Library preparation was carried out as described before for bulk 3'-sequencing of poly(A)-RNA (Parekh et al. 2016). Using Maxima RT polymerase (Thermo Fisher), cDNA was generated with oligo-dT primer that contain barcodes, unique molecular identifiers (UMIs) and an adapter. Template switch oligos (TSOs) were used to extend 5'-ends of cDNAs. All samples were then pooled, before full-length cDNA was amplified using TSO-site and adapter binding primers. Fragmentation of cDNA was then performed using the Nextera XT kit (Illumina). Final amplification of 3'-end fragments was performed using primers containing Illumina P5 and P7 overhangs. Additionally, the P5 and P7 sites were altered, allowing sequencing of the cDNA in the first read and barcodes and UMIs in the second read to achieve a better cluster recognition in comparison to previously described methods (Parekh et al. 2016). NextSeq 500 (Illumina) was used for sequencing of the library using 75 cycles for the cDNA in the first read and 16 cycles for the UMIs and barcodes in the second read. Gencode gene annotations vM18 was used with the human reference genome GRCh38.p12 (https://www.gencodegenes.org/). To map raw sequencing data to the reference genome, Dropseq tools v1.12 was utilized (Macosko et al. 2015). The resulting UMI filtered countmatrix was then analyzed using R v3.4.4, and differential expression using DESeq2 v1.18.1 (Love, Huber, and Anders 2014).

Parametric fit using the genotype label as covariate was employed to estimate dispersion of the data. Further, the Wald test was used to identify differentially regulated genes upon CRK family-deficiency in HCT116 cells. Afterwards, shrunken log2 fold changes were calculated, whereby the Shrinkage-type argument of the lfcShrink function was set to 'normal'. Differential regulation of a gene was assumed, if absolute log2 fold change >1 and adjusted p-value <0.05. Finally, GSEA pre-ranked analysis was performed using GSEA v3.0 and MsigDB v6.2 within the Hallmark gene set collection (Kuleshov et al. 2016).

2.2 Cell biological methods

2.2.1 Cell culture

Human cell lines HEK293, Huh-7, H1650, PANC-1, SW480 and HCT116 were cultured at 37 °C, 7% CO_2, 95% humidity in Dulbecco's Modified Eagle Medium (DMEM) (Thermo Fisher) containing 100 U ml^{-1} Penicillin (Biochrom), 100 µg ml^{-1} Streptomycin (Biochrom), 2 mM L-glutamine (Thermo Fisher) and 10% (v/v) fetal calf serum (FCS) (Biochrom). For passaging, cells were washed with PBS (Thermo Fisher) and incubated with trypsin solution (Sigma-Aldrich) for 10 min at 37 °C. Cells were frozen in 10% (v/v) DMSO/90% (v/v) FCS, and cryovials were kept in liquid nitrogen for long-term storage.

2.2.2 Calcium phosphate transfection

HEK293 cells were transfected by calcium phosphate precipitation. Briefly, vector DNA, ddH_2O and $CaCl_2$ were mixed as shown below, before BES buffer was added dropwise during vortexing:

Component	Volume [µl] 6-well plate	Volume [µl] 100 mm dish
Vector DNA [1 µg $µl^{-1}$]	5	15
ddH_2O	130	345
$CaCl_2$ [2.5 M]	15	40
BES buffer (2X)	150	400

The mixture was incubated for 10 min at room temperature, before being added dropwise to cells at 60-70% confluency. Medium was refreshed 24 h after transfection. Two days post transfection, cells were either subjected to selection to acquire stably transfected cells or lysed for analysis. HEK293 cells transfected with pmRFP-N2 vectors were selected with 400 µg ml^{-1} G418 (Thermo Fisher) for at least 3 weeks.

2.2.3 Lipid-mediated transfection

Lipid-mediated transfection was performed using FuGENE HD (Promega) for HCT116 and PANC-1, or Attractene (Qiagen) for SW480 cells. For 6-well plates, 2 µg vector DNA was mixed with 100 µl Opti-MEM (Thermo Fisher) and 8 µl FuGENE HD or 7.5 µl Attractene. After vortexing, the mixture was incubated for 10 min at room temperature, and subsequently added dropwise to cells at 60-70% confluency. Two days post transfection, cells were lysed for analysis or subjected to selection.

2.2.4 Generation of retroviral particles and transduction

The vector pLenti CMV Puro LUC (w168-1) was a gift from Eric Campeau & Paul Kaufman (Addgene plasmid #17477; http://n2t.net/addgene:17477; RRID:Addgene_17477) (Campeau et al. 2009). A total number of $2x10^6$ HEK293T cells were seeded on a 100 mm dish. After overnight incubation, cells were transfected by calcium phosphate precipitation with 20 µg pLenti CMV Puro LUC (w168-1) as transfer vector, 15 µg psPAX2 as packaging plasmid and 6 µg pMD2.G as envelope plasmid. Medium was aspirated 6 h post transfection, before 6 ml fresh medium was added. After incubation for further 48 h, medium was collected, cleared by centrifugation at 1000 g for 5 min, and the supernatant was passaged through a 0.45 µm filter. Finally, HCT116 cells at 50% confluency were transduced either with retroviral particles based on pLenti CMV Puro LUC (w168-1) as described above (for CRK family protein-deficient and parental cells), or with premade LPP-hLUC-Lv201-025-C (GeneCopoeia) particles (for SASH1-deficient, compound SASH1/CRKL-deficient, and parental cells). After 24 h incubation, transduced cells were selected with 1 µg ml^{-1} puromycin for at least 3 weeks. Bioluminescence was measured with a FLUOstar OPTIMA luminometer to verify equivalent luciferase expression.

2.2.5 Generation of gene knockout cell lines

HCT116, PANC-1 or SW480 cells were transfected with pSpCas9(BB)-2A-Puro (PX459) V2.0 containing the appropriate guide RNA. Two days later, HCT116 cells were selected with 5 µg ml^{-1}, and SW480 or PANC-1 cells with 3 µg ml^{-1} puromycin for 72 h. Cells were allowed to recover for 48 h in refreshed medium, and then seeded at a low density. When colonies were visible without magnification, single colonies were isolated and analyzed for absence of the target protein via immunoblotting. Loss of target protein was verified during several passages to ensure a complete loss.

2.2.6 Cell proliferation assay

Cell proliferation was measured with the XTT Cell Proliferation Kit II (Roche) according to the manufacturer's protocol. Briefly, 1000 cells per well were seeded in 96-well plates and cultured in DMEM containing 2% or 10% (v/v) FCS. Colorimetric XTT assays were performed directly after seeding, as well as at indicated time points. Reagents were incubated for 4 h and the increase in absorbance was quantified by subtracting 0 h from 4 h values. Cell proliferation index was then calculated as the increase in absorbance from the measurement directly after seeding.

2.2.7 Transwell migration and invasion assays

Transwell migration (8.0 µm Transwell Permeable Supports, Corning) and invasion (8.0 µm Matrigel Invasion Chamber, Corning) assays were used to analyze single cell migration and invasion, respectively. Transwell inserts were coated with DMEM containing 10% (v/v) FCS for 2 h at 37 °C, washed with PBS, and placed in 24-well plates containing 600 µl (migration) or 750 µl (invasion) DMEM with 10% (v/v) FCS. Afterwards, $1x10^5$ cells were seeded in 100 µl DMEM without FCS for migration assays, or $2x10^5$ cells in 500 µl DMEM without FCS for invasion assays. After incubation for 20/24 h (migration) or 40/48 h (invasion) at 37 °C, inserts were washed twice with PBS, and incubated for 20 min with 3% paraformaldehyde in PBS at room temperature. Inserts were washed again, before cells at the upper site of the membrane were removed with cotton swabs. Transmigrated/invaded cells at the lower site of the membrane were permeabilized with 0.5% (v/v) Triton X-100 (Carl Roth) in PBS for 10 min at room temperature, washed with PBS, and stained with DAPI (Thermo Fisher) for further 10 min at room temperature. After a final washing step, the number of transmigrated or

invaded cells per region of interest (= one image acquired with a 20x objective in a Zeiss AxioObserver Z1 microscope) was quantified using ImageJ v1. 51 (National Institutes of Health, USA). Due to the low migrative and invasive capabilities of CRK- and/or CRKL-deficient cells, three regions of interest were counted in part II of the present thesis (see section 3.2).

2.2.8 Wound healing assay

Wound healing assays were performed to quantify collective cell migration. A total number of $1.5x10^5$ cells in 100 µl DMEM with 10% (v/v) FCS were seeded into each separated chamber of the Culture-Insert 2 Well in µ-Dish 35 mm (IBIDI). After attachment of cells overnight, the separating spacer was removed to introduce a defined gap of 500 µm. Images were acquired with an AxioObserver Z1 (Zeiss) microscope, and cells were allowed to migrate for 20 h, before images were acquired again. Wound closure was determined by subtracting the cell-free area of the two time points. The areas were measured using ImageJ v1.51 (National Institutes of Health, USA).

2.2.9 Soft agar assay

Soft agar assays were performed to quantify anchorage-independent growth. First of all, a bottom layer of 0.5 mL DMEM per well containing 10% (v/v) FCS and 0.6% (w/v) agarose was allowed to solidify in 12-well plates. Afterwards, 2500 cells per well in 0.5 mL DMEM containing 10% (v/v) FCS and 0.4% (w/v) agarose were added on top of the bottom layer. After solidification, 0.5 mL DMEM containing 10% (v/v) FCS was added on top of the solidified layers. Medium was refreshed every 3 days and cells were allowed to grow for 2 weeks. Finally, colony number was counted manually with a 5x objective using an Axiovert 100 (Zeiss) microscope.

2.2.10 3D-culture in Matrigel

To observe the morphological phenotypes of cells in a 3D-matrix, Matrigel (Sigma-Aldrich) was used as extracellular matrix. A total volume of 25 µl per well Matrigel and DMEM (1:1 mixture) containing 10% (v/v) FCS was allowed to solidify in a 96 well-plate. Next, 100 cells in 25 µl of the same ECM gel composition were added. After further solidification, 50 µl DMEM containing 10% (v/v) FCS was added on top of the layers, and cells were incubated for 5 days with medium being refreshed after 3 days. Finally, brightfield images of colonies were acquired using an AxioObserver Z1 (Zeiss) microscope.

2.2.11 Adhesion assay

Adhesion assays were performed to measure cell adhesion to fibronectin. First, 96 well plates were coated with 5 µg mL^{-1} human fibronectin (Roche) in PBS for 1 h at 37 °C. After a washing step with PBS, a total number of $1x10^4$ cells in 100 µL DMEM with 10% (v/v) FCS was added per well. Cells were allowed to adhere for different time points. Medium was aspirated, and adherent cells were washed with PBS, before they were counted under the microscope (Zeiss AxioObserver Z1). Adhesion was quantified relative to the total number of seeded cells, which was measured with additional wells without washing.

2.2.12 Chemoresistance assay

Chemoresistance assays were performed using the XTT Cell Proliferation Kit II (Roche) according to the manufacturer's protocol. Briefly, 2000 cells per well in 100 µl DMEM containing 10% (v/v) FCS were seeded in 96-well plates and allowed to attach overnight. Afterwards, medium was aspirated and 100 µl DMEM containing 10% (v/v) FCS and varying concentrations of oxaliplatin (Klinikum rechts der Isar) or 5-fluorouracil (Klinikum rechts der Isar) were added to the wells. After incubation with chemotherapeutic agents for 24 h, cells were washed with PBS and medium was refreshed. After further incubation for 2-3 days, XTT assays were carried out to determine remaining viable cells. Reagents were incubated for 4 h and absorbance was quantified by subtracting 0 h from 4 h values. The half maximal inhibitory concentration (IC50) was calculated using GraphPad Prism v6.01.

2.3 Immunological techniques and protein biochemistry

2.3.1 Cell lysis and immunoblot analysis

To generate cell lysates for immunoblot analysis, cells were washed with PBS, and lysed in ice cold RIPA buffer containing 0.1% (w/v) SDS and protease/phosphatase inhibitors (1 mM Na_3VO_4, 1 mM NaF, 1 mM Pefabloc, 2x protease inhibitor cocktail (Roche), 1 mM β-glycerophosphate, 1 mM benzamidine, 1 mM PMSF, 1 µg mL^{-1} Pepstatin). Cell lysates were subjected to sonification, incubated for 15 min at 4 °C with end over end rotation, and centrifuged for 15 min at 4 °C and 18620 g to obtain the cleared supernatant. Protein concentration was assessed using the Pierce BCA Protein Quantification Kit (Thermo Fisher) according to the manufacturer's protocol. 5x Laemmli buffer was added to the lysates to obtain a final 1x concentration of Laemmli buffer. Samples were further diluted with

1x Laemmli buffer to obtain equal protein concentrations. Samples were then boiled at 95 °C for 5 min and subjected to SDS-polyacrylamide gel electrophoresis using the PerfectBlue PAGE System (Peqlab). Proteins were then transferred from gels to 0.45 µm nitrocellulose membranes (GE Healthcare Life Sciences) using a semi-dry transblotting system (Biorad). The membrane was blocked with 5% (w/v) blotting grade milk powder (Carl Roth) in PBST for 1 h at room temperature, washed 3 times with PBST for 5 min at room temperature, and then incubated with the primary antibody (Table 4) in PBST containing 5% (w/v) blotting grade milk powder or bovine serum albumin (BSA) (Sigma-Aldrich) overnight at 4 °C. After washing 3 times with PBST for 10 min at room temperature, the membrane was incubated with the appropriate HRP-conjugated secondary antibody (Jackson Immunoresearch) in PBST containing 5% (w/v) blotting grade milk powder for 1 h at room temperature. The membrane was washed 3 times with PBST for 10 min at room temperature. Proteins were then visualized using chemiluminescent detection either with X-ray films (GE Healthcare Life Sciences) or an imaging system (Analytik Jena).

2.3.2 Co-immunoprecipitation

Co-immunoprecipitations were performed to qualitatively and quantitatively detect protein-protein interactions. HCT116 or HEK293 cells were lysed in RIPA buffer with protease/phosphatase inhibitors (1 mM Na_3VO_4, 1 mM NaF, 1 mM Pefabloc, 2x protease inhibitor cocktail (Roche), 1 mM β-glycerophosphate, 1 mM benzamidine, 1 mM PMSF, 1 µg mL^{-1} Pepstatin) as described in section 2.3.1, but without SDS and sonification. Equal concentrations and volumes of cleared lysates were incubated with 2 µg antibody (Table 4) for 2 h at 4 °C with end over end rotation. A total volume of 40 µl Protein A (for rabbit antibodies) or Protein G (for murine antibodies) Sepharose (Sigma-Aldrich) resin per sample was washed 2 times with PBS and once with RIPA buffer containing protease/phosphatase inhibitors. Washed resin was added to the samples and further incubated for 1 h at 4 °C with end over end rotation. Beads were washed 4 times with RIPA buffer containing protease/phosphatase inhibitors. For precipitation of RFP or GFP fusion proteins, RFP-trap (Chromotek rta-10) or GFP-trap (Chromotek gta-10) were used according to the manufacturer's protocol. Finally, bound proteins were eluted by addition of 1x Laemmli buffer and subjected to immunoblotting. Alternatively, corresponding bands of target and control

immunoprecipitations were cut out of the gel and analyzed via mass spectrometry as described before (Franke et al. 2019).

2.3.3 Immunofluorescence microscopy

Immunofluorescence microscopy was performed to analyze tissue cryosections or cultured cells. Cells were cultured on cover slips, which were coated with 5 µg ml^{-1} fibronectin in PBS for 1 h at 37 °C. Cryosections on glass slides were used for staining of human tumor tissues. Cells or tissues were washed with PBS, before they were fixed with 3% paraformaldehyde (Sigma-Aldrich) in PBS for 20 min at room temperature. After washing 3 times with PBS, the reaction was quenched with 50 mM NH_4Cl in PBS for 20 min at room temperature. Samples were washed again, before 0.1% (v/v) Triton X-100 (Carl Roth) in PBS was used for permeabilization at room temperature for 3 min. After further washing steps, unspecific binding sites were blocked with 2% (w/v) BSA (Sigma-Aldrich) in PBS for 1 h at room temperature. Samples were then incubated with the appropriate primary antibody (Table 4) in blocking buffer overnight at 4 °C. The next day, samples were washed 3 times and incubated with TRITC-phalloidin (Sigma-Aldrich) for F-actin staining, Alexa Fluor 488- and/or Cy3-conjugated secondary antibodies (Table 4), as well as with DAPI (Thermo Scientific), in blocking buffer for 1 h at room temperature. After further washing, samples were mounted using Prolong Gold antifade (Thermo Fisher). Finally, images were acquired using an AxioObserver Z1 (Zeiss) microscope.

2.3.4 Immunohistochemistry

Tissue samples obtained from orthotopic mouse models were fixed in formalin immediately after sacrifice and image acquisition. Formalin-fixed samples were embedded in paraffin to generate tissue sections with a thickness of 2 µm on glass slides. Sections were deparaffinized using Roticlear (Carl Roth) 3 times for 10 min at room temperature. Rehydration was performed with 100% ethanol 3 times for 2 min, followed by 96% (v/v) ethanol for 2 min, 70% (v/v) ethanol for 2 min, 50% (v/v) ethanol for 2 min, and ddH_2O for 5 min at room temperature. Tissue sections were then boiled in 10 mM citrate (pH 6.0) at 94 °C for 20 min. After cooling down to room temperature, sections were washed with PBS containing 0.1% (w/v) BSA (Sigma-Aldrich) for 5 min at room temperature. Peroxidase activity was blocked with 3% (v/v) H_2O_2 in methanol for 5 min at room temperature. Tissue sections were

washed with PBS containing 0.1% (w/v) BSA, permeabilized with 0.1% (v/v) Triton X-100 (Carl Roth) in PBS containing 0.1% (w/v) BSA for 5 min at room temperature, washed again, and blocked with 10% (v/v) goat serum in PBS containing 0.1% (w/v) BSA for 1 h at room temperature. The samples were then incubated with the primary antibody (Table 4) in PBS containing 0.1% (w/v) BSA overnight at 4 °C. After washing 3 times with PBS containing 0.1% (w/v) BSA, the appropriate HRP-coupled secondary antibody (EnVision+ System, Agilent Technologies) was added in PBS containing 0.1% (w/v) BSA and incubated for 1 h at room temperature. Samples were washed 3 times with PBS for 10 min, before the DAB chromogen mixture (Agilent Technologies) was added and incubated for 8 min at room temperature. After washing with ddH_2O, tissue sections were stained with hematoxylin for 5 min at room temperature. After a final washing step, sections were mounted with gelatin and images were acquired with an Axiolab (Zeiss) microscope.

2.3.5 Flow cytometry

Cells were washed with PBS, before Accutase (Thermo Fisher) was added for detachment. After incubation for 10 min at 37 °C, the reaction was immediately stopped by addition of DMEM containing 10% (v/v) FCS. A total number of $1x10^6$ cells were washed with PBS, before they were blocked with 100 µL PBS containing 10% (v/v) goat serum and 0.5% (w/v) BSA (Sigma-Aldrich) for 20 min at 4 °C. After washing with PBS, cells were stained with 1 µg anti-Integrin beta-1 or isotype control antibody (Table 4) in PBS with 0.5% (w/v) BSA for 30 min at 4 °C. Cells were washed 3 times with PBS, before FITC-labeled anti-mouse secondary antibody (Jackson Immunoresearch) in PBS containing 0.5% (w/v) BSA was added. After incubation for 30 min at 4 °C, two washing steps were performed. Cells were then resuspended in PBS and immediately subjected to flow cytometry using FACSCalibur (BD Biosciences) and the CellQuest Pro software (BD Biosciences). The software FlowJo v8.8.2 (FlowJo) was used for analysis. Doublets, dead cells and debris were excluded. Unspecific staining, determined with the isotype control antibody, was excluded from the data.

2.3.6 Purification of the N-terminal SH3 domain of CRKL

In order to obtain sufficient amounts of the N-terminal SH3 domain (SH3N) of human CRKL with C-terminal HIS-tag, *E. coli* One Shot BL21 Star (DE3) (Thermo Fisher) were transformed with the vector pET-21a-CRKL-SH3N that contains an IPTG-inducible expression cassette.

Bacteria were incubated in LB medium containing 100 µg ml^{-1} Ampicillin (Sigma-Aldrich) at 37 °C until an optical density of $OD_{600\ nm}$ = 0.8 was reached. Bacterial culture (total volume of 500 ml) was chilled down to room temperature, before 0.6 mM IPTG was added to induce expression of the SH3 domain. After incubation for 4 h at 21 °C, bacteria were harvested by centrifugation at 5000 g for 10 min at 4 °C. An additional washing step with PBS was performed. The sediment was freeze-thawed at -80 °C, before it was resuspended with 8 mL ice cold HIS lysis buffer and incubated for 10 min on ice. Afterwards, sonification was performed with a Sonopuls Sonicator (Bandelin) system. A total number of 6 cycles was performed, sonicating for 10 s per cycle at 70% power with 10 s breaks between each cycle. The lysate was then cleared by centrifugation for 15 min at 21000 g and 4 °C. The supernatant was transferred to 750 µl pre-washed HisPur Ni-NTA (Thermo Fisher) resin and incubated for 30 min at 4 °C with end over end rotation. Next, the sample was centrifuged for 2 min at 700 g and 4 °C, and the supernatant was discarded. The resin was washed several times with HIS wash buffer, until the absorbance at 280 nm remained stable between washing steps (determined with a NanoDrop 1000 spectrophotometer). Afterwards, 3 elution steps with 1 ml HIS elution buffer were performed. The supernatants were combined and dialyzed (MWCO of membrane: 3500) twice against PBS at 4 °C for 6 h and subsequently overnight. Finally, aliquots were frozen in liquid nitrogen.

2.3.7 Biochemical label-free assays

The affinity between SASH1 and CRKL was estimated by biochemical label-free assays (EnSpire Label-free Technology, PerkinElmer), using the immobilized N-terminal SH3 domain (SH3N) of CRKL and a peptide corresponding to the CRKL binding motif within SASH1 (residues 982-991; QPPPVPAKKS; synthesized by JPT Peptide Technologies). Activation of the assay plate (Enspire-LFB, 384-well High Sensitivity, User-Activated Biochemical Plates, Corning Epic System) was performed with 15 µl ddH_2O per well containing 50 mM N-hydroxysulfosuccinimide and 200 mM 1-Ethyl-3-(3-dimethylaminopropyl)-carbodiimide for 30 min at room temperature. The plate was then washed 4 times with 25 µL ddH_2O, before 15 µL of 25 µg ml^{-1} purified SH3N domain of CRKL (see section 2.3.6) in 20 mM sodium acetate (pH 5.0) was added to each well to immobilize the domain overnight at 4 °C. After washing 4 times with 15 µL PBS (pH 7.4) containing 0.005% (v/v) Tween-20, the plate containing the immobilized domain and PBS (pH 7.4) was equilibrated for 3 h at room temperature. The

baseline was then measured with an Enspire Multimode Plate Reader (PerkinElmer). Finally, 15 µl PBS containing either peptides or small compounds were added per well. The final read was carried out to calculate the response by subtracting the baseline from the final peak wavelength. The apparent dissociation constant of the N-terminal SH3 domain of CRKL and the SASH1 peptide was calculated using GraphPad Prism v6.01. Further, small compounds were pre-selected from a generic small molecule compound library (BioFarma-USEF Research Group, University of Santiago de Compostela, Spain) by *in silico* docking experiments (Autodock Vina, performed by Dr. Cristian R. Munteanu, University of A Coruna, Spain), using the structures of the N-terminal SH3 domains of CRKL and CRK based on RCSB PDB 2LQN and 1CKA, respectively. Pre-selected compounds (10 µM final concentration) were then used to screen for potential binders of the immobilized SH3N domain of CRKL *in vitro*. A specific high affinity peptide (HAP, residues CVDNSPPPALPPKRRRSAPS, synthesized by JPT Peptide Technologies) binding the SH3N domain of CRKL was used as positive control (Kardinal et al. 2000; Posern et al. 1998).

2.3.8 Synthesis of cell penetrating peptides

To translocate the CRK/CRKL SH3N binding high affinity peptide (HAP, residues CVDNSPPPALPPKRRRSAPS) into cells, it was dimerized to the cell penetrating peptide corresponding to residues 47-57 of HIV1 TAT via a disulfide bond (peptide dimer synthesized by GenScript). HAP additionally contained a lysine coupled to FITC at the C-terminus. The peptide was dissolved in PBS to yield a stock concentration of 800 µM.

2.4 *In vivo* analysis

2.4.1 Orthotopic mouse model

Orthotopic xenotransplantation experiments were performed to analyze the metastatic capabilities of parental, SASH1-deficient, and compound SASH1/CRKL-deficient HCT116 colorectal cancer cells. Specific pathogen-free female SCID/Beige mice were housed in mixed groups according to institutional guidelines approved by the Use Committee for Animal Care of the University of Santiago de Compostela, Spain. Mice at an age of 8 weeks were anesthetized under constant isoflurane application. A small incision was introduced in the abdominal wall to expose the caecum. Luminal content was pushed back mechanically, before the cleared tip of the caecum (approximately first 2 mm) was tied up with absorbable surgical

sutures. A total amount of $5x10^6$ cells stably expressing equivalent levels of luciferase (see section 2.2.4) were resuspended in 25 µl ice cold Matrigel, before the cell/Matrigel suspension was injected into the enclosed caecal lumen. The suspension was allowed to solidify, and the surgical wound was closed. Mice were continuously monitored and sacrificed upon weight loss >10% or signs of distress. After 34 days post injection, mice were injected intraperitoneally with 150 µg sterile D-luciferin (PerkinElmer) per g body weight. Mice were then sacrificed, and spatially resolved luminescence measurements were carried out with an IVIS Spectrum In Vivo Imaging system (PerkinElmer) before and after removal of the primary tumor to quantitatively assess tumor cell dissemination. Data was analyzed with Living Image v3.1 and ImageJ v1.51. SASH1-deficient cells failed to develop tumors attaching to the mucosa in 2 out of 6 mice, which were excluded from further analysis.

2.4.2 Hematoxylin and Eosin staining

Formalin-fixed samples were embedded in paraffin to generate tissue sections with a thickness of 2 µm on glass slides. After deparaffinization (see section 2.3.4), tissue sections were incubated with hematoxylin staining solution for 5 min at room temperature. Samples were washed with tap water for 5 min at room temperature and incubated with eosin staining solution for 1 min at room temperature. Samples were washed and dehydrated with increasing ethanol concentrations, before they were incubated with Roticlear (Carl Roth) 3 times for 3 min and mounted using Eukitt (Sigma-Aldrich). Images were acquired with an Axiolab (Zeiss) microscope.

2.4.3 Analysis of human tissue samples

Analysis of patient samples was performed as described before (Franke et al. 2019). In total, 65 patients with UICC stage III primary colorectal cancer were analyzed. Patients underwent curative surgery (R0) and gave informed consent for analysis of their tissue and data prior to surgery at the Department of Surgery, Technical University of Munich, Germany. Immediately after resection, the tissue was shock-frozen in liquid nitrogen. Clinicopathological characteristics of the patients are summarized in Table S 1 (UICC stage III colorectal cancer patient subgroup that received 5-fluorouracil as adjuvant (mono-)chemotherapy) and Table S 2 (total patient cohort with UICC stage III colorectal cancer). To avoid putative bias through poor surgical technique, patients with documented local recurrence were excluded

from this study. Supervision of the ethics committee of the Faculty of Medicine (Technical University of Munich, Germany), which approved the study (#1926/07; #5428/12), assured ethical and data protection standards of this study. Tumor-specific or distant recurrence-free survival was used as the primary endpoint for risk prediction. The IBM SPSS Statistics software v20.0 (SPSS Inc., Chicago, USA) was used for statistical evaluation. Optimal cut-off values for gene expression levels (*SASH1* 0.29; *CRKL* 0.44; *CRK* 0.6442; *CRK* family 0.7809; *ZEB1* 44.01; *CDH1* 14.66) were calculated with maximally selected log-rank statistics, using R Software v2.13.0 (R Foundation for Statistical Computing, Vienna, Austria). The R-function 'maxstat.test' was used to consider multiple test issue within these analyses. Kaplan-Meier estimates were used for survival analysis. Statistical tests were performed 2-sided. Significance level was set at 0.05. No correction of p-values was applied to adjust for multiple test issue. However, results of all statistical analyses that were performed were carefully reported to allow an informal adjustment of p-values while reviewing the data. The Cancer Genome Atlas (TCGA) data set was analyzed for prognostic association of *SASH1* and *CRKL* in colorectal cancer (n=213 fully sequenced samples with available follow-up documentation) (Cerami et al. 2012; Gao et al. 2013).

2.5 Statistical evaluation

GraphPad Prism v6.01 was used for generation of graphs and statistical analysis. All error bars represent the standard deviation. Sample sizes, statistical tests, and p-values are shown in the respective figure legends (not significant (ns): $p>0.05$; *: $p\leq0.05$; **: $p\leq0.01$; ***: $p\leq0.001$; ****: $p\leq0.0001$). Normal distribution was assessed with D'Agostino & Pearson omnibus normality test. If not otherwise indicated in the graph or graph legend, statistical tests (e.g. unpaired t-test or Mann-Whitney test) were performed to compare the respective samples with parental cells or other appropriate controls.

3. Results

3.1 Part I: The tumor suppressor SASH1 interacts with the signal adaptor CRKL, thereby inhibiting epithelial-mesenchymal transition, chemoresistance and metastasis formation

3.1.1 Loss of SASH1 induces epithelial-mesenchymal transition and cancer cell aggressiveness

The aim of this thesis was to investigate the putative impact of SASH1 on tumor progression and especially metastasis formation, as well as the underlying molecular and cellular mechanisms. Therefore, SASH1-deficient HCT116 human colorectal cancer cells were generated with the CRISPR/Cas9 system. To minimize off-target effects, two independent guide RNA sequences were used that target different regions within the coding sequence of *SASH1*. Next-generation sequencing of the genomic target areas of guide RNA 1 and 2 revealed bi-allelic, single base pair insertions in exon 1 (clone S1) and exon 2 (clone S2), respectively (Figure 5A). Due to the resulting frameshift, early stop codons were introduced (S1* and S2* in Figure 5B), leading to drastically truncated predicted SASH1 proteins that lack any known functional domain (Figure 5B). Accordingly, SASH1-deficient cells exhibited no detectable SASH1 protein levels (Figure 5C). Interestingly, the genetically independent SASH1-deficient cell clones S1 and S2 both featured a fibroblast-like appearance (data not shown, also see Figure 6A), which is a well-established feature of epithelial cells that underwent epithelial-mesenchymal transition (EMT). Therefore, EMT marker levels were assessed via immunoblotting (Figure 5C, D). SASH1-deficient cells exhibited significantly reduced protein levels of the epithelial hallmark marker E-cadherin (Figure 5D), while the mesenchymal markers ZEB1 and vimentin were strongly upregulated (Figure 5C). These changes were also observed at the transcriptional level by qRT-PCR (Figure 5E).

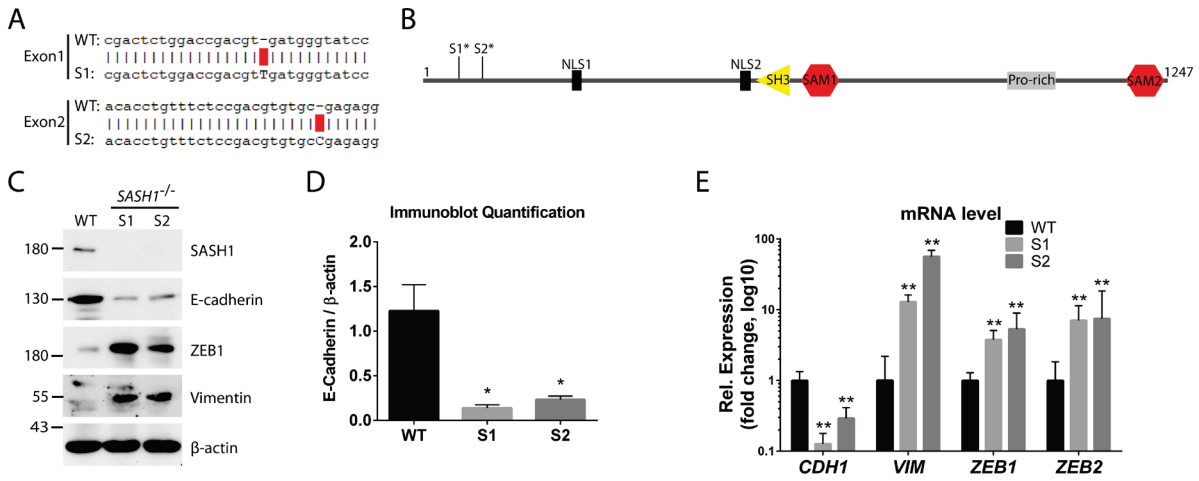

Figure 5: Loss of SASH1 induces epithelial-mesenchymal transition. (A) Genomic DNA sequences (within exon 1 and exon 2 of *SASH1*) of SASH1-deficient (S1 and S2) HCT116 cells targeted by CRISPR/Cas9 were aligned to the sequence of the parental (WT) line using ApE v2.0.55. Single base pair insertions are highlighted in red. **(B)** Primary structure and domain architecture of SASH1 according to (Martini et al. 2011). Premature stop codons of SASH1-deficient clones S1 (S1*) and S2 (S2*) are also shown (1, 1247 = residue numbers; NLS = nuclear localization signal; SH3 = Src homology 3 domain; SAM = sterile alpha motif; Pro-rich = proline-rich sequence). **(C)** Parental (WT) and SASH1-deficient (S1 and S2) HCT116 cells were analyzed for protein levels of SASH1 and EMT markers via immunoblotting. **(D)** E-cadherin protein levels were quantified relative to β-actin via immunoblot analyses (Mann-Whitney test; n=4; p=0.0286). **(E)** EMT markers were analyzed at the transcriptional level via qRT-PCR (Mann-Whitney test; n=5-6; p≤0.0079).

Next, cell morphology and E-cadherin localization was analyzed via indirect immunofluorescence microscopy. Deficiency of SASH1 induced a fibroblast-like morphology with a concomitant cell individualization and loss of intercellular adhesions (Figure 6A). SASH1-deficient cells also exhibited a complete loss of E-cadherin at the cell-cell interface, even when cells were in close contact to each other (Figure 6A). Additionally, cells underwent changes in the actin cytoskeleton, as cortical F-actin was reduced upon SASH1-deficiency, while stress fiber formation was increased (Figure 6A). Morphological alterations were also highlighted by an increase of cell size, evidenced by quantification of the cell surface area (Figure 6B), as well as by an elevated length-to-width ratio of SASH1-deficient cells (Figure 6C).

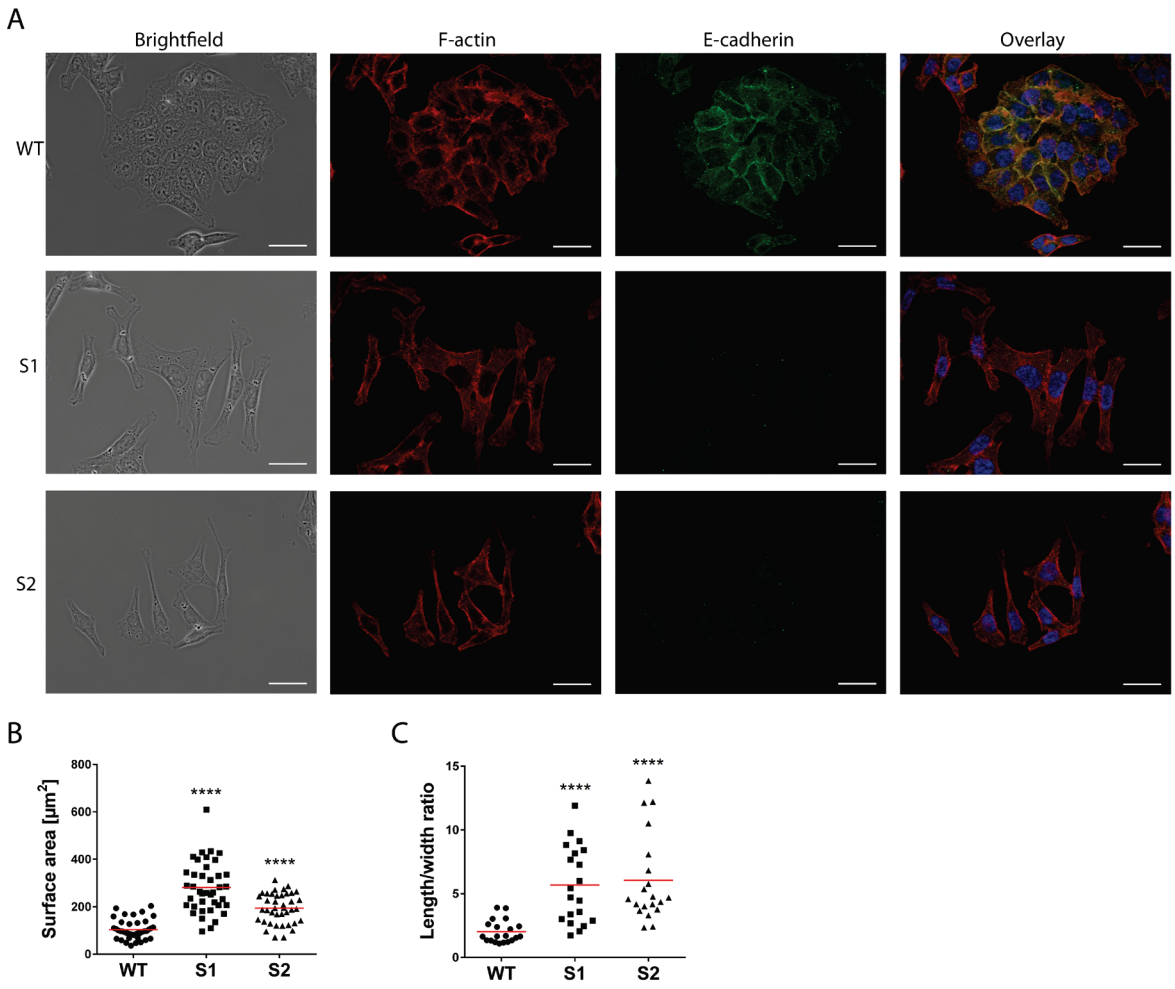

Figure 6: Loss of SASH1 induces a fibroblast-like morphology. (A) Parental (WT) and SASH1-deficient (S1 and S2) HCT116 cells were stained for E-cadherin (green), F-actin (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 µm). **(B)** Cell surface area was measured using ImageJ v1.51 (unpaired t-test; n=40 cells; p≤0.0001). **(C)** Length-to-width ratio was determined using ImageJ v1.51 (unpaired t-test; n=20 cells; p≤0.0001).

EMT is widely accepted to promote migration and invasiveness (Pino et al. 2010; Wang et al. 2016). To investigate if these phenotypes are also affected by SASH1, cell migration and invasiveness were analyzed by transwell assays. Loss of SASH1 resulted in significantly increased transwell migration (Figure 7A) and invasion through Matrigel (Figure 7B). Parental cells cultured in a 3-dimensional matrix consisting of Matrigel formed colonies with a smooth, rounded surface (Figure 7C, arrowheads). In contrast, SASH1-deficient cells generated colonies with an irregular shape, from which single cells invaded the extracellular matrix (Figure 7C, arrows). This observation is in good accordance with results from the transwell assays. Next, soft agar assays were performed to analyze anchorage-independent survival. SASH1-deficient cells exhibited a highly significantly increased number of colonies compared to the parental line (Figure 7D).

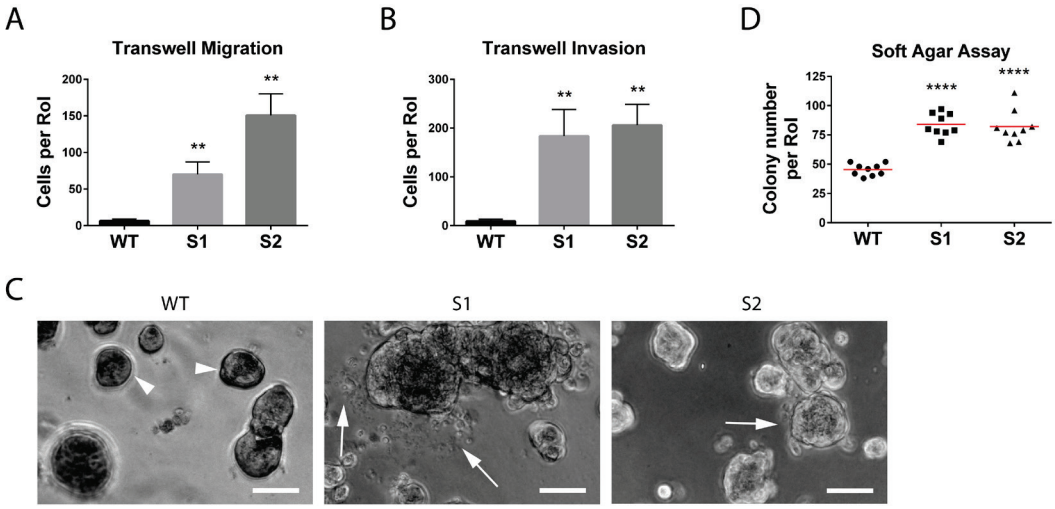

Figure 7: Cell migration, invasion and survival are increased in SASH1-deficient cells. (A) Transwell migration assays were carried out to analyze migration of parental (WT) and SASH1-deficient (S1 and S2) HCT116 cells (Mann-Whitney test; n=6; p=0.0022; RoI = region of interest). Migration was allowed for 20 h. **(B)** Invasiveness was assessed using transwell invasion assays with an incubation for 40 h (Mann-Whitney test; n=6; p=0.0022; RoI = region of interest). **(C)** Cells were cultured in a 3-dimensional extracellular matrix consisting of Matrigel for 5 days, before colony morphology was analyzed via light microscopy (size bar = 50 µm). While the parental line formed rounded colonies (arrowheads), SASH1-deficient cells grew invasively (arrows). **(D)** Cells were cultured in soft agar for 14 days, before colony number was quantified (unpaired t-test; n=9; p≤0.0001).

Taken together, these data highlight an important negative regulatory role of SASH1 in EMT. Loss of SASH1 induced mesenchymal marker expression, a fibroblast-like appearance with cell individualization, loss of E-cadherin at intercellular adhesions, as well as increased cell migration, invasiveness and survival.

3.1.2 *SASH1* is downregulated during cytokine-induced epithelial-mesenchymal transition

To verify a putative role of SASH1 in EMT with an independent approach, parental HCT116 cells were treated with TNF, a cytokine that has been described to induce EMT in this cell line (Wang et al. 2013). Treatment with TNF indeed induced *bona fide* EMT, as E-cadherin was barely detectable and cells acquired a fibroblast-like morphology (Figure 8A). Interestingly, TNF treatment resulted in an even more pronounced relative decrease of SASH1 protein levels compared to E-cadherin (Figure 8B). Treatment of cell lines originating from different tumor entities, such as SW480 (colorectal cancer), PANC-1 (pancreatic ductal adenocarcinoma), Huh7 (hepatocellular carcinoma), and H1650 (non-small cell lung cancer), with the well-known EMT-inducer TGF-β1 decreased SASH1 and E-cadherin protein levels, in accordance to the findings for TNF application (Figure 8C).

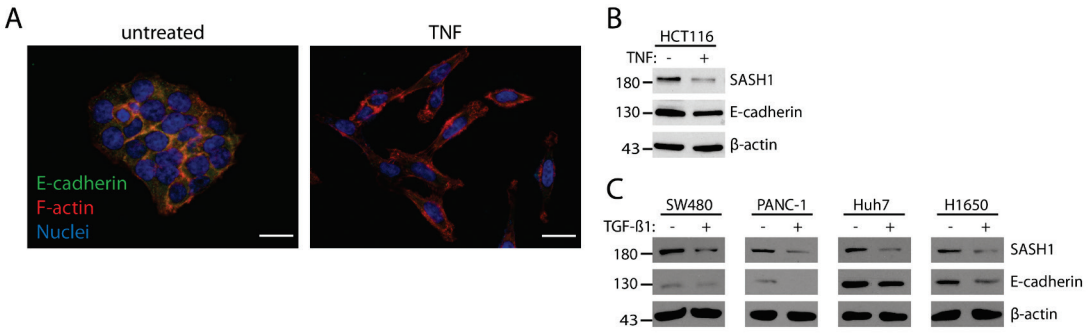

Figure 8: *SASH1* is downregulated during cytokine-induced EMT. (A) Parental HCT116 cells were treated with vehicle control or 20 ng ml^{-1} TNF for 48 h, before they were stained for E-cadherin (green), F-actin (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 µm). **(B)** Immunoblot analysis of parental HCT116 cells that were treated with vehicle or 20 ng ml^{-1} human TNF for 48 h. **(C)** Immunoblot analysis of SW480 (colorectal cancer), PANC-1 (pancreatic ductal adenocarcinoma), Huh7 (hepatocellular carcinoma), and H1650 (non-small cell lung cancer) cell lines that were treated with vehicle or 10 ng ml^{-1} human TGF-β1 for 48 h.

Next, parental HCT116 cells were transfected to transiently express full-length SASH1 with a C-terminal V5 epitope-tag under the control of the EF1-α promoter. SASH1-V5 expressing or control transfected cells were treated with TNF to induce EMT (Figure 9A). Of note, cells transfected with SASH1-V5 underwent increased cell death upon TNF treatment (data not shown), which was also reflected by increased PARP1 cleavage (Figure 9A). Forced expression of SASH1-V5 resulted in a reduced decrease of E-cadherin levels upon TNF application, as compared to control transfected cells (Figure 9B). Strikingly, cells expressing SASH1-V5 exhibited increased E-cadherin levels at the cell-cell interface, both when treated with TNF or vehicle (Figure 9C). Therefore, multiple lines of evidence point towards a crucial role of SASH1 as negative regulator of EMT. Forced expression of *SASH1* counteracted EMT, while loss of SASH1 was sufficient to induce EMT and EMT-associated aggressiveness. Further, endogenous SASH1 levels were strongly reduced upon cytokine-induced EMT.

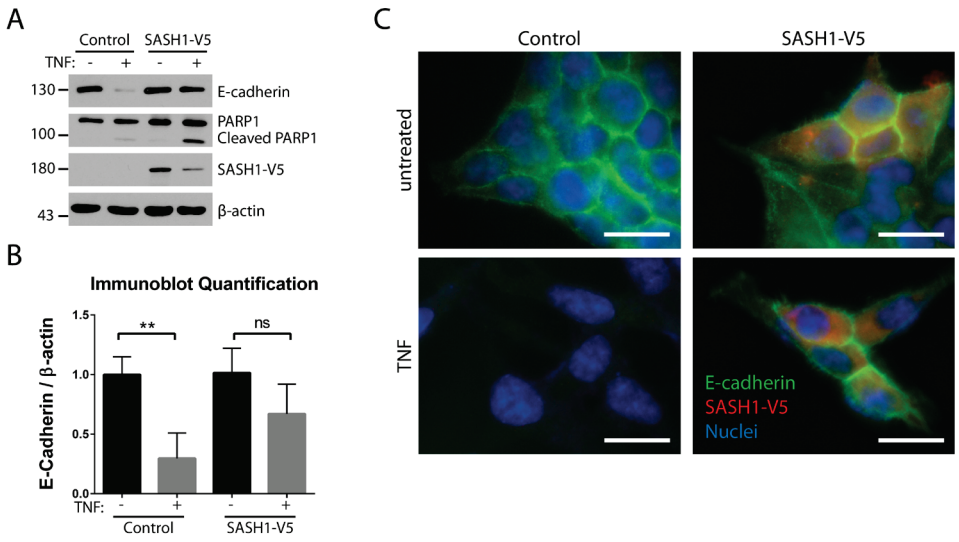

Figure 9: Forced expression of *SASH1* counteracts TNF-induced EMT. (A) Immunoblot analysis of parental HCT116 cells, which were transfected with an empty control vector or a vector encoding SASH1-V5, and subsequently treated with vehicle control or 20 ng ml^{-1} human TNF for 48 h. **(B)** E-cadherin protein levels were measured relative to β-actin via immunoblot analyses (unpaired t-test; n=3; **: p=0.0094; ns = not significant). **(C)** Cells with the indicated treatments were also stained for E-cadherin (green), V5 epitope (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 µm).

3.1.3 SASH1 interacts with the signal adaptor CRKL

The next aim was to investigate the underlying molecular mechanism of how SASH1 counteracts EMT. Therefore, orthogonal screenings were performed in order to identify potential protein-protein interaction partners of SASH1. GFP immunoprecipitations were carried out with HEK293 embryonic kidney cells transiently transfected to express either GFP as control or SASH1 with an N-terminal GFP-tag. Additionally, a yeast two-hybrid screening was performed with a placental human cDNA bank and *SASH1* cDNA as bait. The screening was conducted in cooperation with the group of Dr. Ewa Ninio (Sorbonne University, Paris, France) as described in (Franke et al. 2019). Putative interaction partners detected by the yeast two-hybrid screening were then compared to co-precipitating proteins identified by mass spectrometry (Table 1). Hits from the yeast two-hybrid screening were also searched manually in the PubMed database for a functional role in EMT, as well as for enriched domains using the STRING database analysis tool (Table 1) (von Mering et al. 2003; Szklarczyk et al. 2015; Szklarczyk et al. 2017).

Table 1: Putative SASH1 interaction partners determined by yeast two-hybrid and co-immunoprecipitation/mass spectrometry screenings. Yeast two-hybrid screens were performed in cooperation with Dr. Ewa Ninio and Dr. Henri Weidmann (Sorbonne University, Paris, France). Putative interactions of confidence A (very high confidence), B (high confidence) and C (good confidence) were considered, while candidate interactions with confidence level D to F were discarded. Additionally, co-immunoprecipitations of GFP and GFP-SASH1 were performed in HEK293 cells and analyzed via mass spectrometry. Mass spectrometric analyses were conducted in cooperation with Solenne Chardonnet (Sorbonne University, Paris, France). Further, results from GFP and GFP-SASH1 immunoprecipitations with HMEC cell lysates, which were performed by Dr. Ewa Ninio and Dr. Henri Weidmann, were considered (SAM = Sterile alpha motif; SH2/3 = Src homology 2/3 domain; PH = Pleckstrin homology domain; CH = Calponin homology domain).

GENE	Confidence	EMT-related PubMed articles	Interaction via Co-IP	Enriched SMART protein domain
AFF1	A	0		
AKAP10	C	0		
ARRB1	A	2		
ASCC2	C	0	Yes	
BAP1	C	2		
C1D	A	0		
CCNDBP1	A	1		
CNKSR1	B	0		SAM, PH
CRK	A	17		SH2, SH3
CRKL	C	6	Yes	SH2, SH3
CTNNA1	A	1		
EPS8L2	A	0		SH3
IKBKB	A	18		
JADE2	A	0		PHD zinc finger
JADE3	A	0		PHD zinc finger
KIF3B	B	0		
MCRS1	A	1	Yes	
OAS2	B	0	Yes	
PPP4R1	A	0		
RBM12B	B	0		
SECISBP2L	B	0		
SNX1	C	0		
SPG7	C	0		
SPTBN1	B	2	Yes	Spectrin repeat, CH, PH
SYNE2	B	0		Spectrin repeat, CH
TNKS2	A	0		SAM
UTRN	C	0		Spectrin repeat, CH
EXOC8	A	1		PH

Additionally, a protein-protein association network was generated using the STRING database, excluding SASH1 to avoid potential bias (von Mering et al. 2003; Szklarczyk et al. 2015; Szklarczyk et al. 2017) (Figure S 1). However, the network analysis lacked clustering and exhibited only 4 direct or indirect protein-protein associations.

The CRK family of signal adaptors, consisting of CRK and its homolog CRKL, was present in the yeast two-hybrid screen, showed a strong association in the STRING network analysis, and was furthermore also related to EMT. However, only CRKL was detected independently via immunoprecipitation/mass spectrometry (Table 1). For these reasons, the focus of the

present work was on CRKL, which we already had identified as physical interaction partner of SASH1, whereby both proteins co-localized in the nucleus, cytoplasm and especially at membrane ruffles (Fabian Christoph Franke, Master thesis 2015, Technical University of Munich, Germany). The interaction of proteins at the endogenous level was further verified by immunoprecipitation of CRKL in HEK293 and HCT116 cells, which successfully co-precipitated SASH1 (Figure 10). Immunoprecipitation of CRK showed no association with SASH1 (data not shown).

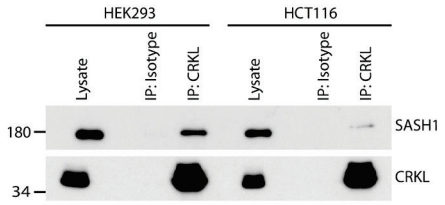

Figure 10: Endogenous SASH1 physically interacts with endogenous CRKL. Endogenous SASH1 was co-precipitated by anti-CRKL antibodies but not by isotype control antibodies in immunoprecipitation experiments with HEK293 and HCT116 cell lysates.

3.1.4 Mapping of the interaction: The N-terminal SH3 domain of CRKL binds to a specific PXXPXK motif of SASH1

To investigate the interaction between SASH1 and CRKL in detail, SASH1 deletion constructs either with N-terminal GFP-tag or C-terminal V5 epitope-tag were utilized, which were generated previously (Martini et al. 2011), as well as CRKL constructs with C-terminal RFP-tag (Figure 11A). Immunoprecipitation of endogenous CRKL was performed in HEK293 cells transiently transfected to express GFP as control, full-length GFP-SASH1, or the indicated SASH1 deletion constructs with N-terminal GFP-tag (Figure 11B). While GFP was not detected in CRKL immunoprecipitates, full-length GFP-SASH1 (residues 1-1247) showed an interaction with endogenous CRKL, as well as the deletion constructs containing the residues 1-887 or 709-1247 (Figure 11B). The deletion construct comprising the residues 1-887 co-precipitated to a lower extent (Figure 11B). Next, reciprocal immunoprecipitations were performed with HEK293 cells transiently expressing full-length SASH1-V5 (residues 1-1247), deletion constructs of SASH1 with C-terminal V5 epitope-tag, or a deletion construct with mutated nuclear localization signal (Figure 11C). V5 epitope-tag immunoprecipitations revealed co-precipitation of endogenous CRKL only for full-length SASH1 and the deletion construct containing the C-terminal part of SASH1 (residues 590-1247), while its N-terminal part did not associate with CRKL (Figure 11C). Accordingly, endogenous CRKL immunoprecipitations only

co-precipitated full-length SASH1 and the deletion construct with the residues 590-1247 (Figure 11C).

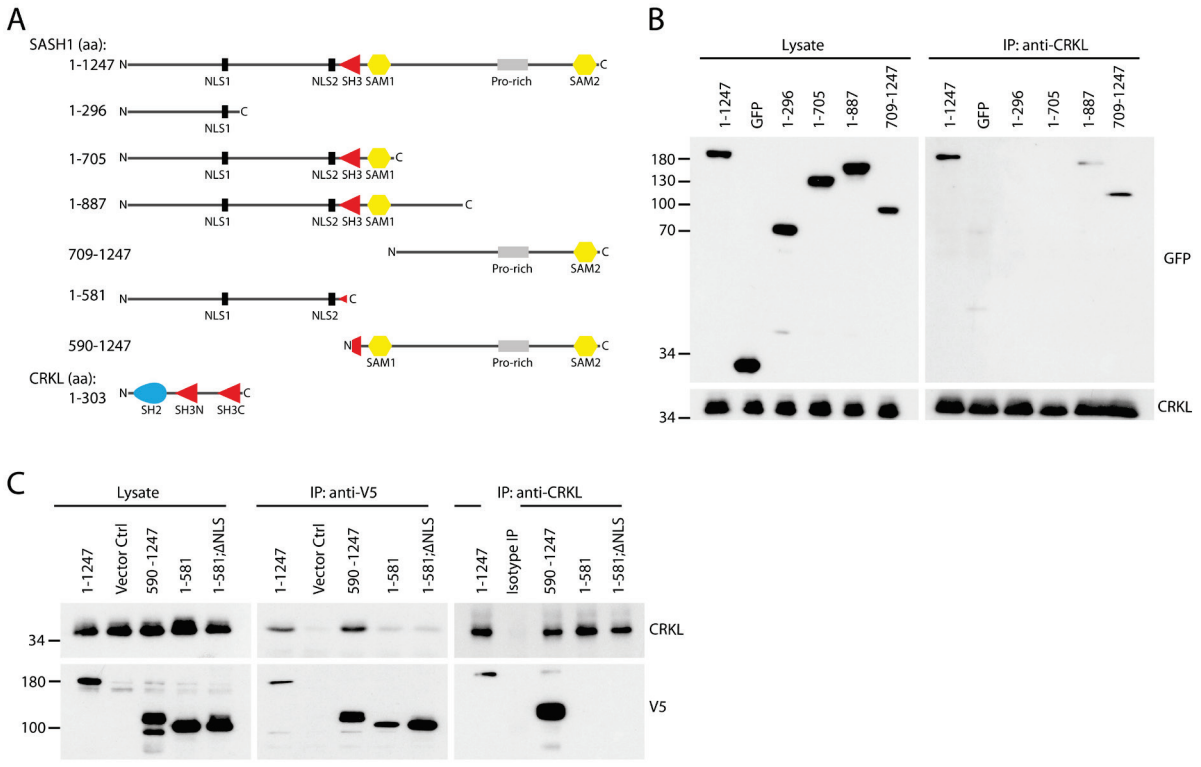

Figure 11: CRKL binds to the C-terminal part of SASH1. (A) Primary structures of full-length SASH1, deletion constructs and of full-length CRKL are shown (NLS = nuclear localization signal; SH2 = Src homology 2 domain; SH3 = Src homology 3 domain; SAM = Sterile alpha motif). The number of residues (aa = amino acids) of each construct is also indicated. **(B)** Immunoblot analysis of cell lysates and endogenous CRKL immunoprecipitations (IP), using HEK293 cells transiently transfected to express GFP alone, full-length, or deletion constructs of SASH1 with N-terminal GFP-tag. **(C)** Immunoblot analysis of cell lysates and reciprocal immunoprecipitations (IP) of V5 epitopes or endogenous CRKL. HEK293 cells were transiently transfected to express full-length, or deletion constructs of SASH1 with C-terminal V5-tag. As negative control, cells were transfected with empty vector. Additionally, a SASH1 deletion construct with point-mutated nuclear localization signal (ΔNLS) was analyzed.

Taken together, CRKL likely binds to the C-terminal part of SASH1. This region contains a proline-rich region and the second sterile alpha motif (SAM). Interestingly, the N-terminal SH3 domains of CRK proteins have been described to bind polyproline type II (PPII) motifs with the core consensus PXXPXK (Knudsen, Feller, and Hanafusa 1994; Bell and Park 2012). Analysis of the SASH1 primary structure using the RCSB Protein Data Bank (UniProt entry: O94885) showed that SASH1 contains three distinct PXXPXK motifs in the C-terminal part: residues 865-870, 984-989, and 1016-1021 (Figure 12A, P1-P3). To investigate whether these motifs are relevant for the interaction with CRKL, site-directed mutagenesis was employed, changing each PXXPXK motif to AXXAXK to abolish its function. Immunoprecipitation of endogenous CRKL was carried out with HEK293 cells transiently expressing GFP as control, wildtype or mutated full-length GFP-SASH1, in which each motif was mutated either individually (P1, P2

and P3) or altogether in one construct (P1/2/3). Strikingly, mutagenesis of the second PXXPXK motif (P2) was sufficient to completely abolish any detectable co-precipitation with endogenous CRKL (Figure 12B). Wild type SASH1, as well as variants with the first or third PXXPXK motif mutated, still co-precipitated with endogenous CRKL, while mutagenesis of all motifs in a single construct behaved similarly to mutagenesis of the second motif only (Figure 12B). Next, the N-terminal SH3 domain (SH3N) of CRKL was inactivated by mutagenesis of the residue W160, which has been reported to be crucial for domain function (Senechal et al. 1998). The hydrophobic tryptophan was exchanged to arginine (W160R) with the rationale to introduce electrostatic repulsion with the positively charged PXXPXK motifs. RFP immunoprecipitations were performed with wild type CRKL or CRKL^{W160R}, both containing a C-terminal RFP-tag (Figure 12C). The CRKL SH2 domain interaction partner GAB1 (Sakkab et al. 2000) was still co-precipitated by CRKL^{W160R}-RFP. However, co-precipitation of the CRKL SH3N interaction partner C3G (Nosaka et al. 1999; Arai et al. 1999), as well as of SASH1 was completely abolished by the W160R exchange (Figure 12C). This suggested that the interaction between SASH1 and CRKL is indeed mediated by the CRKL SH3N domain binding to the second PXXPXK motif of SASH1. Based on these data, a peptide corresponding to the second PXXPXK motif (QPPPVPAKKS) of SASH1 was probed for its ability to bind the bacterially expressed and purified CRKL SH3N domain via label-free biochemical assays. Within these assays, the change of the peak wavelength of reflected light as a result of a changed refractive index is measured as “response”. The CRKL SH3N domain was immobilized to the sensor surface, thereby allowing to assess the dissociation constant by measuring the response upon addition of the peptide in various concentrations. The SH3N domain bound the peptide with an apparent dissociation constant of K_d=7.4 µM (Figure 12D).

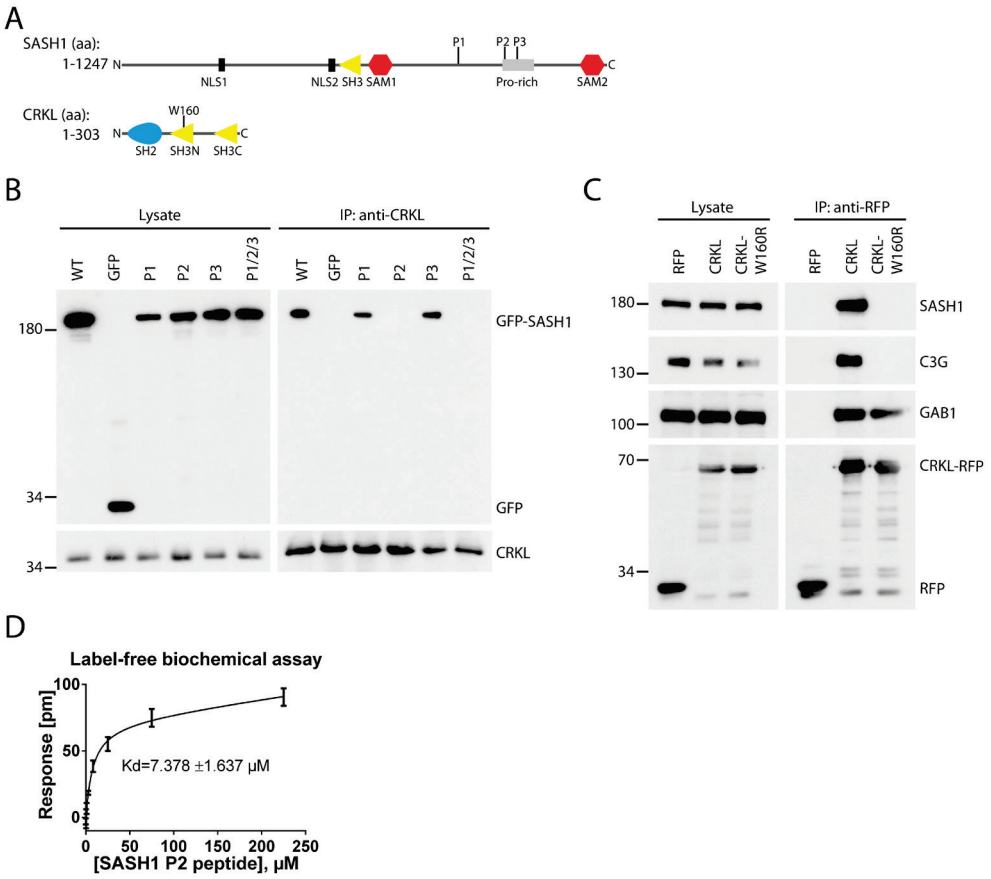

Figure 12: SASH1 binds to the N-terminal SH3 domain of CRKL through a specific PXXPXK motif. (A) Primary structures of full-length SASH1 and CRKL are shown (NLS = nuclear localization signal; SH3 = Src homology 3 domain; SAM = Sterile alpha motif; P1-P3 = PXXPXK motifs; SH2 = Src homology 2 domain; W160 = residue important for SH3N function). The number of residues (aa = amino acids) of each construct is also indicated. **(B)** Immunoblot analysis of cell lysates and endogenous CRKL immunoprecipitations with HEK293 cells transiently transfected to express GFP as control, wild type GFP-SASH1, or GFP-SASH1 variants with mutated PXXPXK motifs (P1-P3 = single motif mutated; P1/2/3 = all motifs mutated in a single construct). **(C)** Immunoblot analysis of cell lysates and RFP immunoprecipitations with HEK293 cells transiently transfected to analyze interactions with RFP as control, wild type CRKL-RFP, or CRKL^{W160R}-RFP. **(D)** The affinity between a small peptide corresponding to the P2 PXXPXK motif within SASH1 and the purified CRKL SH3N domain was estimated using label-free biochemical assays, which were performed in cooperation with Dr. Eduardo Dominguez Medina (University of Santiago de Compostela, Spain). An apparent dissociation constant of K_d=7.4 µM (n=3) was determined.

Next, the hypothesis was raised that SASH1 counteracts interactions between CRKL and its effectors by physically associating with and blocking the CRKL SH3N domain. CRKL is thought to act as signal adaptor by binding to membrane-associated, tyrosine-phosphorylated proteins via the SH2 domain, thereby recruiting effector proteins like the guanine nucleotide exchange factor C3G via its SH3N domain into close proximity to the membrane, leading to activation of small GTPases like RAP1, RAC1 or RAS (Bell and Park 2012). SASH1 could thus potentially inhibit CRKL-mediated signaling, which depends on the CRKL SH3N domain (Senechal et al. 1998; Nosaka et al. 1999; Arai et al. 1999; Cheung et al. 2011). To investigate if SASH1 can indeed counteract CRKL effector protein binding, HEK293 cells were transiently transfected with empty vector control or to express SASH1 with C-terminal V5 epitope-tag, and subjected

to immunoprecipitations with a specific antibody against endogenous CRKL (Figure 13A). Recombinant expression of SASH1-V5 significantly reduced co-precipitation of the CRKL effector C3G, known to bind to the SH3N domain of CRKL, while co-precipitation of the SH2 domain interaction partner GAB1 was not altered (Figure 13A, B).

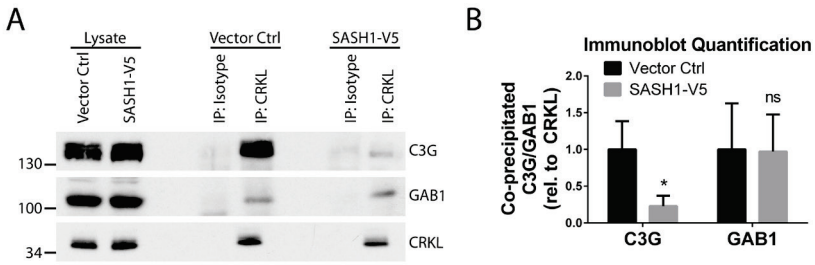

Figure 13: SASH1 counteracts CRKL effector protein binding. (A) Immunoblot analysis of cell lysates and endogenous CRKL immunoprecipitations using HEK293 cells transiently transfected with empty vector (Ctrl = Control) or with full-length SASH1-V5. **(B)** Co-precipitated C3G and GAB1 were quantified relative to precipitated endogenous CRKL via immunoblot analyses (unpaired t-test; n=3; *: p=0.0310; ns = not significant).

3.1.5 Epithelial-mesenchymal transition induced by SASH1-deficiency is dependent on CRKL

According to the current working model, SASH1 binds to the SH3N domain of CRKL, counteracting CRKL-mediated signaling and EMT. To verify this, CRKL-deficiency was introduced via CRISPR/Cas9 in parental and SASH1-deficient HCT116 cells. Deficiency of CRKL was confirmed with different antibodies (Figure 14A). The close homolog CRK (the splice variant CRK-II is recognized by the antibody) was not affected (Figure 14A). Loss of CRKL in SASH1-deficient cells (clone S1C2) strongly increased E-cadherin levels, while the mesenchymal markers vimentin and ZEB1 were reduced, essentially reflecting the phenotype of the parental line (Figure 14A). Loss of CRKL in parental cells further led to slightly decreased ZEB1 levels (Figure 14A), with cells exhibiting a cobblestone-like epithelial appearance and pronounced E-cadherin staining at intercellular adhesions (Figure 14B). Accordingly, CRKL-deficiency rescued the absence of E-cadherin at intercellular adhesions and the fibroblast-like morphology of SASH1-deficient cells (Figure 14B). The reduced cortical localization of F-actin in SASH1-deficient cells was also reverted by loss of CRKL (Figure 14B).

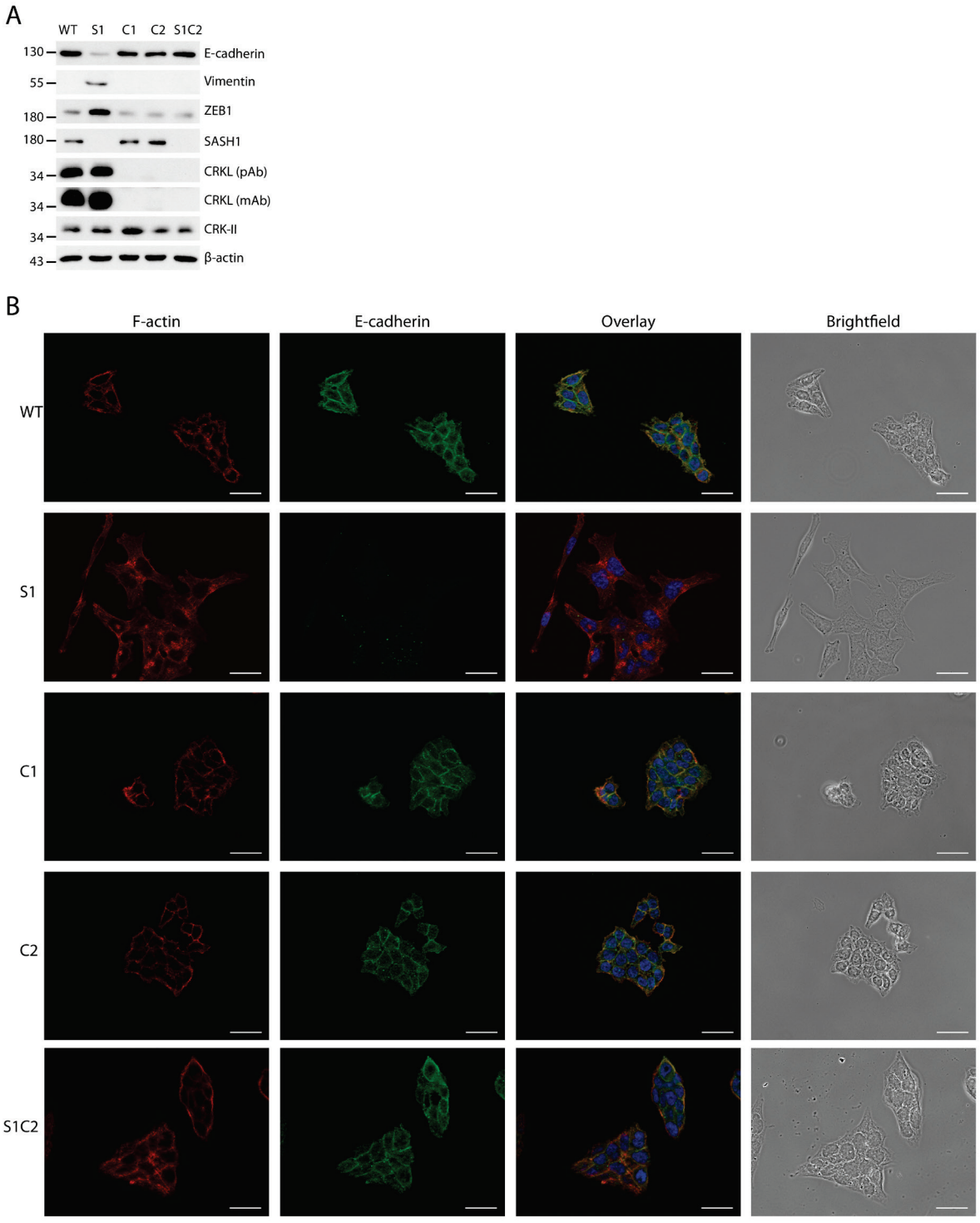

Figure 14: EMT induced by loss of SASH1 is entirely dependent on the presence of CRKL. (A) Immunoblot analysis of parental (WT), SASH1-deficient (S1), CRKL-deficient (C1 and C2), and compound SASH1/CRKL-deficient (S1C2) HCT116 cells. **(B)** The same cell lines were stained for E-cadherin (green), F-actin (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 µm).

Next, the functional impact of CRKL-deficiency on cell migration, invasiveness through Matrigel and cell survival was examined by transwell and soft agar assays. Similar to the phenotype observed for the EMT marker levels, loss of CRKL rescued the highly migratory and invasive phenotype of SASH1-deficient cells (Figure 15A, B), as well as anchorage-independent survival (Figure 15C). Loss of CRKL in parental cells reduced migration and invasion, but not survival in soft agar (Figure 15A-C). Taken together, the data demonstrated that EMT induced by loss of SASH1 is completely dependent on CRKL, as judged by both marker levels and functional phenotypes.

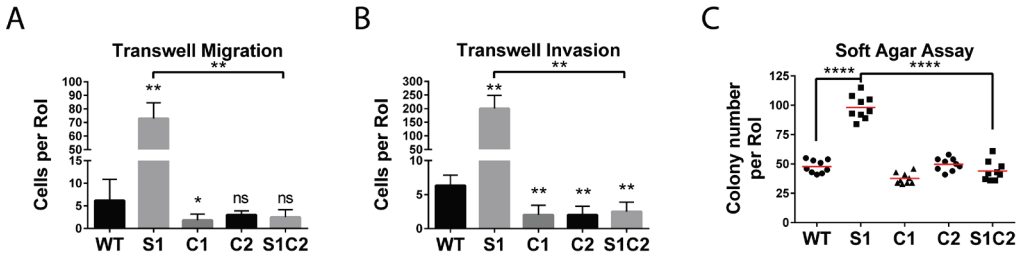

Figure 15: The aggressive phenotype of SASH1-deficient cells is dependent on CRKL. (A) Parental (WT), SASH1-deficient (S1), CRKL-deficient (C1 and C2), and compound SASH1/CRKL-deficient (S1C2) HCT116 cells were subjected to transwell migration assays (Mann-Whitney test; n=6; *: p=0.0303; **: p=0.0022; ns = not significant; RoI = region of interest). Migration was allowed for 20 h. **(B)** Invasiveness was also analyzed by transwell invasion assays with an incubation for 40 h (Mann-Whitney test; n=6; p=0.0022). **(C)** Soft agar assays were performed to analyze anchorage-independent survival (unpaired t-test; n=9; p≤0.0001).

3.1.6 SASH1 counteracts CRKL-mediated SRC signaling, which is required for EMT

The CRK family of signal adaptors is important in several signaling pathways, primarily by recruiting guanine nucleotide exchange factors to activate small GTPases (Bell and Park 2012; Birge et al. 2009). Recently, CRKL was shown to regulate invasiveness downstream of the EMT inducing transcription factor ZEB1 via integrin signaling (Ungewiss et al. 2016). To study if integrin signaling is indeed affected by loss of SASH1 and/or CRKL, cells were allowed to attach to fibronectin-coated cell culture dishes, before they were subjected to immunoblot analysis. Paxillin phosphorylation at residue Y118 and SRC family phosphorylation at residue Y419 were increased in SASH1-deficient cells after adhesion for 3 h or 24 h (Figure 16A, B). Strikingly, loss of CRKL in SASH1-deficient cells reduced phosphorylation of SRC and paxillin to the parental level (Figure 16A, B). Formation of active, phospho-paxillin containing focal adhesions was also strongly increased in SASH1-deficient cells compared to the parental line, at both timepoints of 3 h or 24 h adhesion (Figure 16C). In contrast, loss of CRKL either in parental or SASH1-deficient cells impaired the formation of active focal adhesion, which rather resembled

small focal complexes with a spot-like appearance (Figure 16C). Forced expression of CRKL with C-terminal RFP-tag in turn increased phosphorylation of SRC and its target paxillin in HEK293 cells, essentially phenocopying the effect of SASH1-deficiency (Figure 16D). This was abolished by pharmacological treatment with the Abl/SRC inhibitor Dasatinib (Figure 16D).

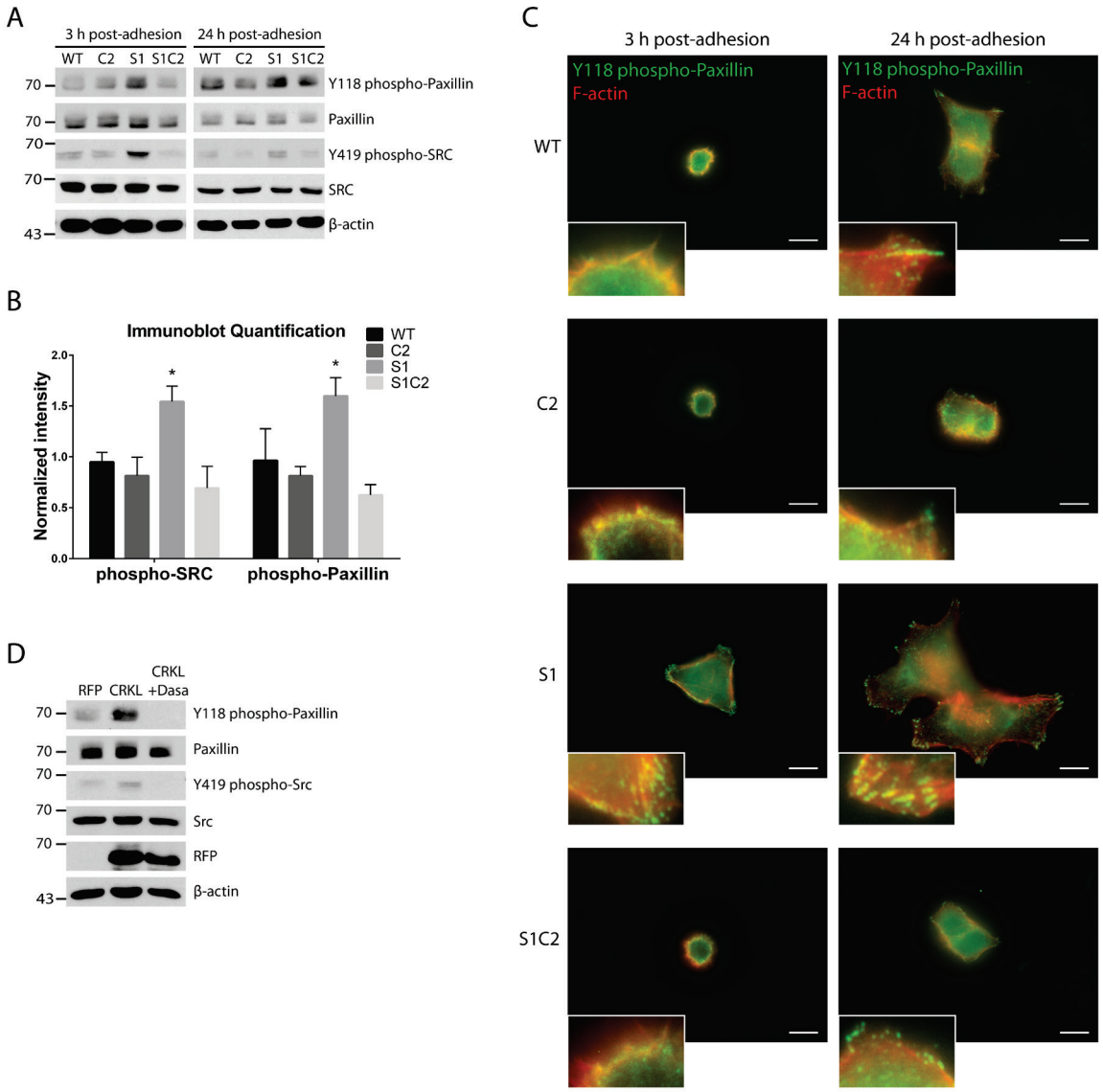

Figure 16: SASH1 counteracts CRKL-mediated SRC signaling. (A) Immunoblot analysis of parental (WT), CRKL-deficient (C2), SASH1-deficient (S1), and compound SASH1/CRKL-deficient (S1C2) HCT116 cells, which were allowed to adhere to fibronectin coated dishes for 3 h or 24 h. **(B)** Phosphorylation of paxillin at Y118 and SRC at Y419 after 3 h adhesion was quantified by immunoblot analyses (Mann-Whitney test; n=4; p=0.0286). **(C)** The same cell lines were stained for Y118 phospho-paxillin (green) and F-actin (red) for immunofluorescence microscopy (size bar = 10 µm). **(D)** Immunoblot analysis of HEK293 cells transiently transfected to express RFP or CRKL with C-terminal RFP-tag. Additionally, cells were treated with vehicle as control or 100 nM Dasatinib (Dasa) for 24 h, as indicated.

In summary, the data indicate that SASH1 inhibits CRKL-mediated activation of SRC. To investigate, whether elevated SRC activity is required for the mesenchymal phenotype of SASH1-deficient cells, SRC was inhibited pharmacologically by Dasatinib treatment. Inhibition of SRC resulted in increased levels of E-cadherin and diminished levels of ZEB1, phenocopying CRKL-deficiency (Figure 17). This suggested that EMT induced by genetic loss of SASH1 is indeed dependent on CRKL-mediated SRC signaling.

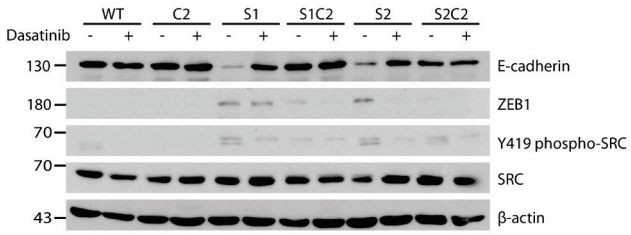

Figure 17: EMT induced by loss of SASH1 is essentially dependent on SRC activity. Immunoblot analysis of parental (WT), CRKL-deficient (C2), SASH1-deficient (S1 and S2), and compound SASH1/CRKL-deficient (S1C2 and S2C2) HCT116 cells, which were treated with vehicle as control or 100 nM Dasatinib for 72 h.

3.1.7 Loss of SASH1 promotes metastasis in a CRKL-dependent manner

To study the role of SASH1 and its novel interaction partner CRKL in metastasis formation, parental, SASH1-deficient, and compound SASH1/CRKL-deficient HCT116 cells were retrovirally transduced to stably express equivalent levels of luciferase (Figure 18A). In cooperation with Dr. Miguel Abal (University of Santiago de Compostela, Spain), cells were then injected into the tip of the caecum of immunodeficient female mice (SCID/beige background), which was previously cleared from luminal content and enclosed by surgical sutures. Using this new orthotopic model, tumor formation was observed to occur close to and within the caecum mucosa, extending to the serosa upon local invasion at the site of implantation (Figure 18B). Mice were sacrificed after 34 days, at which timepoint essentially all animals featured primary tumors, histologically confirmed as poorly differentiated colon carcinoma. The overall tissue architecture and differentiation pattern was similar for all cell lines tested (Figure 18C, left and middle panels). In accordance with *in vitro* observations, SASH1-deficient cells expressed only very low levels of E-cadherin, as compared to parental or compound SASH1/CRKL-deficient cells (Figure 18C, right panels).

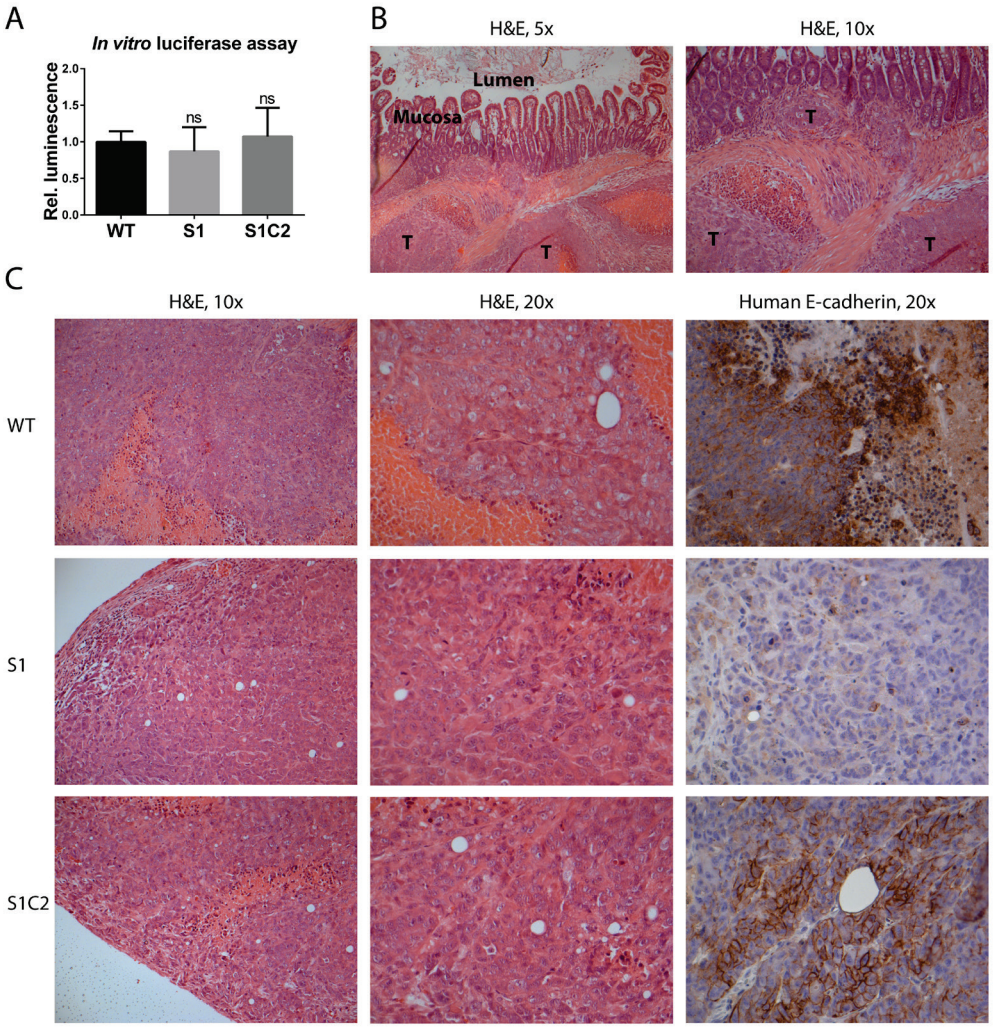

Figure 18: Tissue architecture and EMT phenotype of orthotopic xenograft mouse models for colorectal cancer. (A) Luciferase assays were performed with parental (WT), SASH1-deficient (S1), and compound SASH1/CRKL-deficient (S1C2) HCT116 cells, which were retrovirally transduced to stably express luciferase (unpaired t-test; n=8; p≥0.2795; ns = not significant). **(B)** Primary tumors generated by parental cells were analyzed by H&E staining (5x and 10x magnification; T = cluster of tumor cells). **(C)** Primary tumors spawned by the indicated cell lines, which were orthotopically implanted into immunodeficient mice (n=6 per cell line), were analyzed by H&E staining (10x and 20x magnification). Primary tumors were also subjected to immunohistochemistry with an antibody specific for human E-cadherin (20x magnification; brown staining).

Luminescence imaging was performed directly after sacrifice, both before and after mechanical removal of the primary tumor, which exhibited a very strong signal, as expected (Figure 19A). The metastatic capabilities of parental, SASH1-deficient, and compound SASH1/CRKL-deficient cells were analyzed by automated quantification (particle counting function of Image J v1.51) of the number of luminescent lesions after removal of the primary tumor (Figure 19B). Alternatively, the total luminescence signal after removal of the primary tumor was calculated relative to the total signal prior removal to assess the metastatic efficiency (Figure 19C). Strikingly, SASH1-deficient cells generated significantly more metastatic lesions compared to the parental line (Figure 19B). This increase in the metastatic capability was fully rescued by loss of CRKL (Figure 19B). Importantly, the metastatic luminescence signal relative to the total signal was significantly increased in mice injected with SASH1-deficient cells (Figure 19C). Resected distant lesions were further subjected to immunohistochemistry to ensure their origin from human cancer cells, using an antibody specifically detecting human but not murine mitochondrial epitopes (Thibaudeau et al. 2014). Parental cells primarily exhibited peritoneal carcinomatosis, characterized by large clusters of tumor cells attaching to the outer surface of organs in the peritoneum (e.g. duodenum or kidney) (Figure 19D). In contrast, SASH1-deficient cells formed *bona fide* distant metastatic lesions, as indicated by cancer cells completely enclosed by host tissue, and were frequently detected macroscopically in lymph nodes (Figure 19A, E). Interestingly, SASH1-deficient cells were also extensively dispersed as small clusters and single cells upon microscopic analysis, reflecting their presumed aggressive, mesenchymal phenotype (Figure 19E, lower panel).

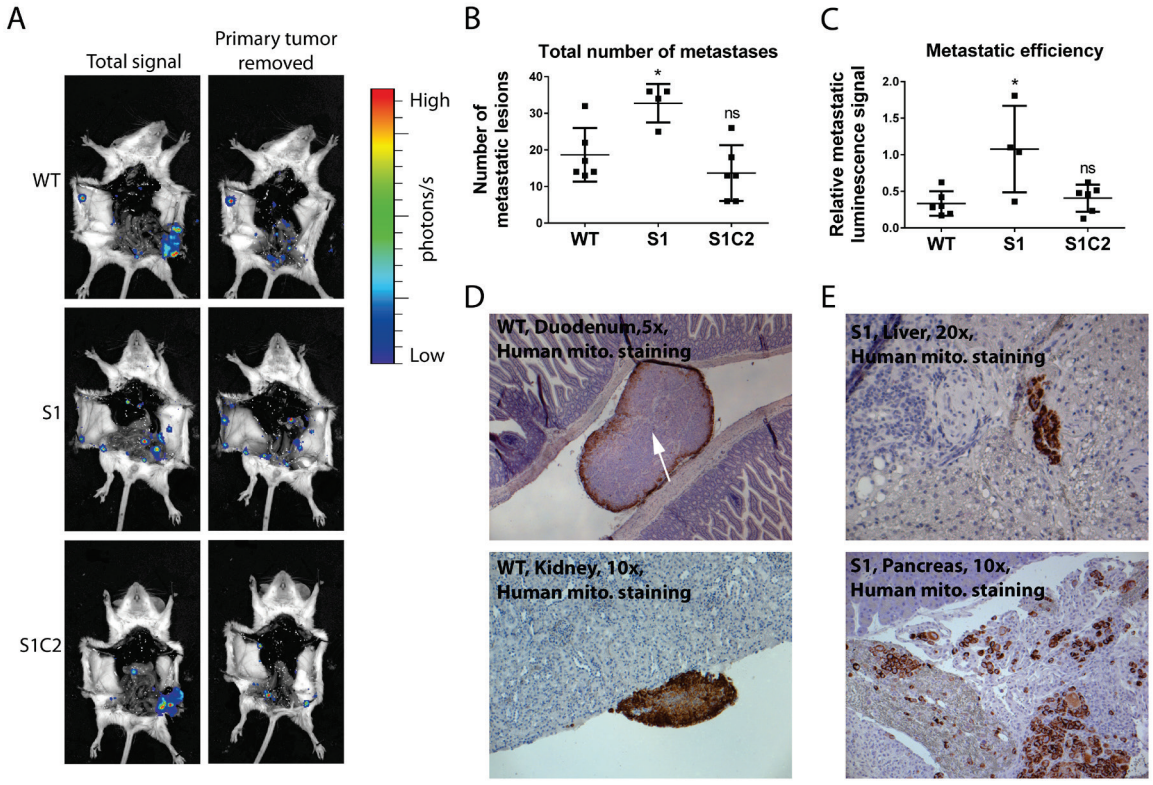

Figure 19: The highly metastatic phenotype of SASH1-deficient cells is entirely dependent on the presence of CRKL. (A) Immunodeficient mice orthotopically implanted with luciferase-expressing parental (WT), SASH1-deficient (S1), and compound SASH1/CRKL-deficient (S1C2) HCT116 cells were sacrificed after 34 days and subjected to luminescence imaging prior and after removal of the primary tumor. **(B)** The total number of metastatic lesions was quantified after removal of the primary tumor, using a particle counting function within Image J v1.51 (Mann-Whitney test; n=4-6; p=0.019; ns = not significant). **(C)** Metastatic efficiency was assessed by the luminescence signal after removal of the primary tumor relative to the total signal prior removal (Mann-Whitney test; n=4-6; p=0.0381; ns = not significant). **(D, E)** Distant lesions generated by the parental line (D) or SASH1-deficient cells (E) were subjected to immunohistochemistry for human mitochondrial epitopes (brown staining; arrow = central necrotic region).

3.1.8 SASH1 counteracts CRKL-mediated resistance against chemotherapy

Recently, EMT has been described to promote chemoresistance (Fischer et al. 2015; Zheng et al. 2015). Therefore, the impact of SASH1-deficiency or compound SASH1/CRKL-deficiency on resistance against the cytotoxic drug 5-fluorouracil, which is the most frequently used agent for treatment of colorectal cancer, was probed *in vitro*. SASH1-deficient cells exhibited an increased resistance, which was reduced upon loss of CRKL (Figure 20A). Of note, this *in vitro* phenotype was in accordance with clinical observations in patients with advanced colorectal cancer. High intratumoral *CRKL* expression was significantly associated with poor tumor-specific survival in a subset of patients with UICC stage III colorectal cancer (Table S 1), who received 5-fluorouracil as adjuvant (mono-)chemotherapy (Figure 20B). Further, analysis of an independent TCGA colorectal cancer dataset (213 colorectal cancer patients with complete sequence information and follow-up data) via the cBioPortal platform revealed a reduced overall survival of patients with alterations or aberrant expression of *SASH1* and/or *CRKL* (Figure 20C) (Cerami et al. 2012; Gao et al. 2013). *CRKL* was primarily amplified at the genomic level and/or its expression was increased at the transcriptional level, while *SASH1* was transcriptionally downregulated or genomically homodeleted. In summary, SASH1 and its novel interaction partner CRKL have prognostic significance in colorectal cancer, which could be explained by changes in EMT, invasiveness and chemoresistance.

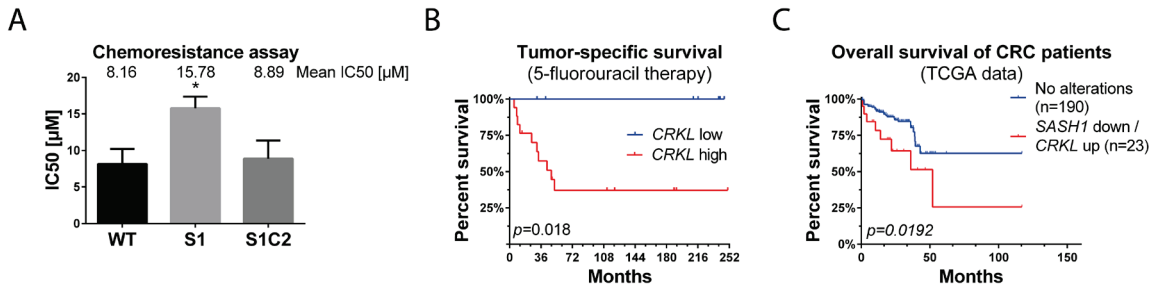

Figure 20: SASH1 and CRKL are involved in chemoresistance and have prognostic relevance in colorectal cancer. (A) Chemoresistance against 5-fluorouracil (24 h treatment and 72 h recovery) of parental (WT), SASH1-deficient (S1), and compound SASH1/CRKL-deficient (S1C2) HCT116 cells was determined by XTT assays (Mann-Whitney test; n=4; p=0.0286; IC50 = half maximal inhibitory concentration). **(B)** Kaplan-Meier analysis of UICC III colorectal cancer patients, who received adjuvant 5-fluorouracil, is shown by high or low intratumoral *CRKL* expression (Log-rank test; n=25 patients; p=0.018). Analysis was performed with the help of Dr. Dr. Ulrich Nitsche (Technical University of Munich, Germany). **(C)** Kaplan-Meier analysis of colorectal cancer patients (TCGA dataset) with alterations in *CRKL* and/or *SASH1* (23 out of 213 patients) compared to patients with no alterations was performed (Log-rank test; n=213 patients; p=0.0192).

3.1.9 Preliminary working model of SASH1 and CRKL in SRC signaling and EMT

The aim of this thesis was to investigate the underlying mechanism of how SASH1 counteracts tumor progression and potentially metastasis formation. According to the preliminary working model (Figure 21), which is based on the data described within this thesis, SASH1 physically associates with CRKL. This interaction is mediated by a specific proline-rich motif in the C-terminal part of SASH1 binding to the N-terminal SH3 domain of CRKL. This way, SASH1 counteracts binding of CRKL effectors (e.g. C3G), and thus CRKL-mediated downstream signaling. Upon loss or downregulation of SASH1, cancer cells undergo EMT. This is due to relieved inhibition of CRKL by SASH1, leading to CRKL-mediated activation of SRC, which is essential for EMT. The tyrosine kinase SRC in turn phosphorylates its substrates, e.g. paxillin, which associates with the CRKL SH2 domain upon phosphorylation (Salgia et al. 1995). This potentially represents a positive feedback loop, as binding of CRKL to paxillin and other focal adhesion components induces membrane recruitment of guanine nucleotide exchange factors (Bell and Park 2012). Upon loss of SASH1, cancer cells thus undergo EMT and acquire an aggressive, invasive phenotype, leading to increased metastasis formation and chemoresistance. In accordance with the working model, these phenotypes were dependent on the presence of CRKL. Furthermore, cytokine-induced EMT induces a downregulation of SASH1, representing a second feedback loop, shifting the cellular state towards a more pronounced mesenchymal phenotype.

Nevertheless, the mechanism of how CRKL activates the kinase SRC remained unknown. Additionally, the role of the close homolog and protein family member CRK, as well as potentially redundant effects between CRKL and CRK, remained to be addressed. Therefore, further experiments were carried out to address these points, which will be described in the following part II of the thesis (see section 3.2).

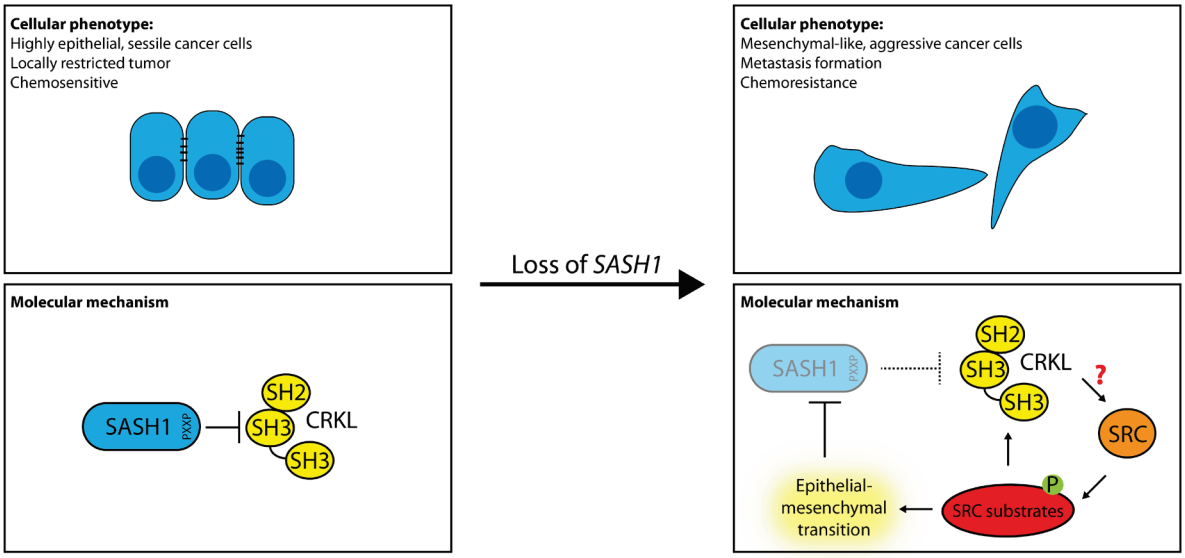

Figure 21: Preliminary working model of the role of SASH1 in metastasis formation. SASH1 binds to the N-terminal SH3 domain of CRKL, counteracting CRKL-mediated SRC-signaling. Upon loss of SASH1 (e.g. during tumor progression), CRKL is relieved from inhibition, leading to increased SRC activation, phosphorylation of SRC substrates, and induction of EMT, thereby generating highly aggressive, pro-metastatic cells with increased chemoresistance (PXXP = proline-rich motif; SH2/3 = Src homology 2/3 domain; P = phosphorylation; ? = hitherto unknown mechanism).

3.2 Part II: The CRK family of signal adaptors acts as central amplifier of SRC/FAK signaling to promote epithelial-mesenchymal transition and cancer cell aggressiveness

3.2.1 Loss of CRKL or the homolog CRK counteracts epithelial-mesenchymal transition

To investigate the role of the CRK family of signal adaptors, which comprises CRKL and the close homolog CRK (with the well-characterized splice variants CRK-I and CRK-II), in EMT either CRKL- or CRK-deficient colorectal cancer cell lines were generated using the CRISPR/Cas9 system with two independent guide RNAs per gene locus. Interestingly, loss of CRK or CRKL in HCT116 cells both resulted in a strong reduction of the mesenchymal marker ZEB1, while E-cadherin levels were upregulated (Figure 22A). Accordingly, loss of CRKL or CRK in SW480 cells, which exhibit a more mesenchymal phenotype compared to the cell line HCT116, resulted in mesenchymal-epithelial transition (MET) with an increase of E-cadherin and downregulation of ZEB1 (Figure 22B).

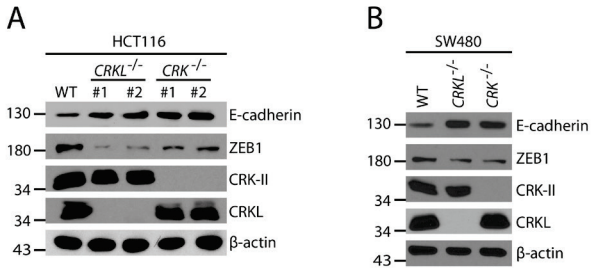

Figure 22: CRK and CRKL have redundant functions in EMT. (A) Immunoblot analysis of parental HCT116 cells (WT), two independent CRKL-deficient clones, and two independent CRK-deficient clones. **(B)** Immunoblot analysis of parental (WT), CRKL-, and CRK-deficient SW480 cells.

Compared to parental cells, which partially exhibited reduced E-cadherin levels at the cell-cell interface (Figure 23A, arrowheads), CRKL- or CRK-deficient HCT116 cells featured a pronounced cobblestone-like morphology and grew closely together (Figure 23A, left panels). In accordance with the observations described in part I (see section 3.1), TNF application resulted in a fibroblast-like morphology with a complete loss of E-cadherin at cell-cell adhesions in parental cells (Figure 23A). In CRKL- or CRK-deficient cells, however, E-cadherin was still present at intercellular adhesions, and cells still exhibited primarily a cobblestone-like morphology upon TNF treatment (Figure 23A). Parental SW480 cells displayed a fibroblast-like appearance with strongly reduced E-cadherin staining (Figure 23B, upper panel). Loss of either

CRKL or CRK reverted this phenotype, as cells expressed E-cadherin at intercellular adhesions and acquired an epithelial morphology (Figure 23B, middle and lower panel).

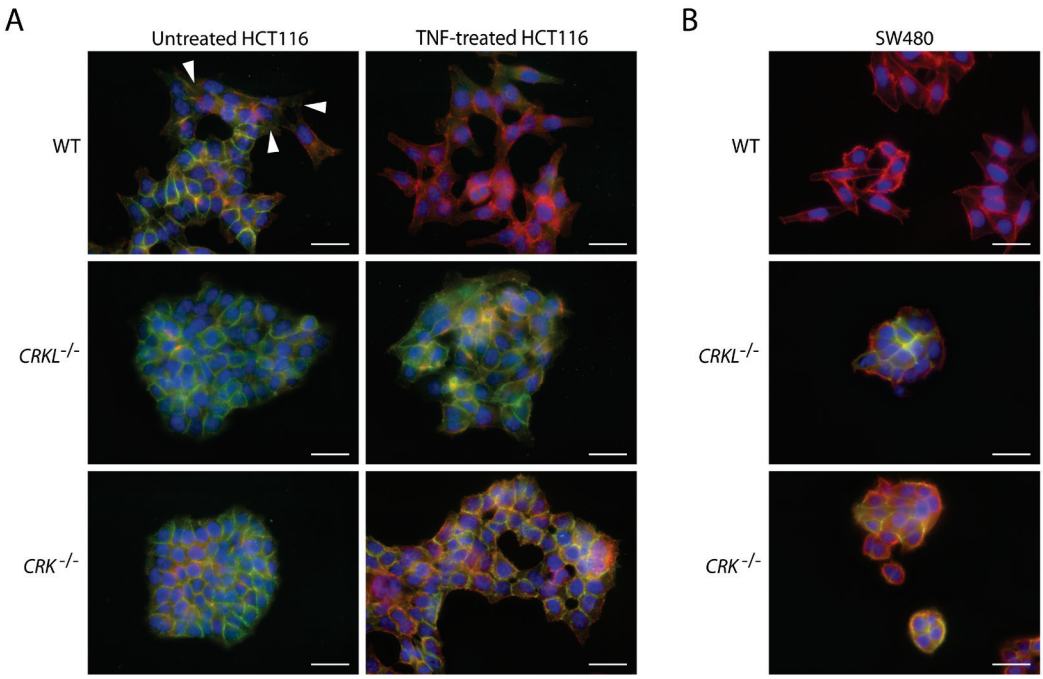

Figure 23: Loss of CRKL or CRK counteracts TNF-induced EMT. (A, B) Parental (WT), CRKL- and CRK-deficient HCT116 (A) or SW480 (B) cells were stained for E-cadherin (green), F-actin (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 µm). Additionally, HCT116 cells were treated with 20 ng ml^{-1} TNF for 72 h as indicated.

3.2.2 Complete loss of the CRK family induces a strongly pronounced epithelial phenotype

Since the data indicate that CRK and CRKL share largely redundant functions in EMT, HCT116 cells were generated that lack both CRK and CRKL, now referred to as CRK family-deficient cells. Both CRK family-deficient clones exhibited increased E-cadherin levels, while ZEB1 was barely detectable (Figure 24A). This phenotype was also observed at the transcriptional level (Figure 24B). Expression of other mesenchymal markers like *SNAI1*, *VIM* and *TWIST* was also significantly reduced, as evidenced by qRT-PCR (Figure 24B). Strikingly, loss of the whole CRK family resulted in a drastically changed cellular morphology. In contrast to parental cells, which featured a morphology typical for epithelial-derived cancer cell lines cultured *in vitro*, CRK family-deficient cells were present as large, densely packed, three-dimensional clusters (Figure 24C, upper panels). CRK family-deficient cells also displayed primarily cortical F-actin with pronounced E-cadherin at intercellular adhesions (Figure 24C, lower panels). In contrast to parental cells, formation of lamellipodia was not observed (Figure 24C, inlets).

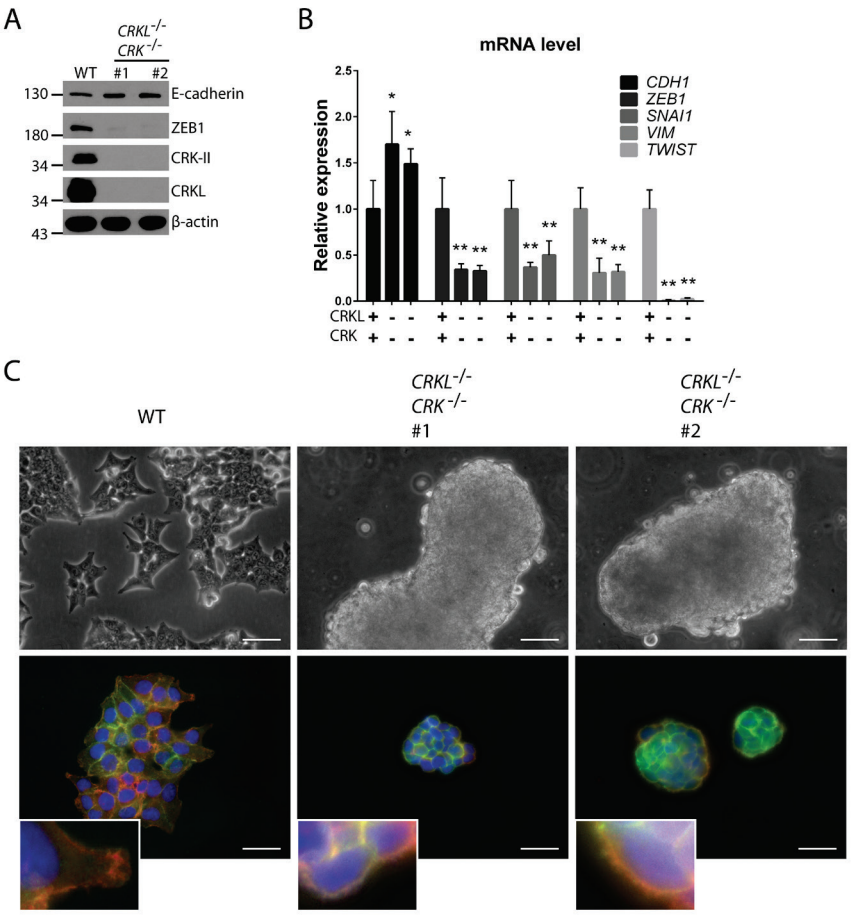

Figure 24: CRK family-deficient cells acquire a highly epithelial phenotype. (A) Immunoblot analysis of parental HCT116 (WT) and cells lacking both CRK and CRKL (CRK family-deficient). **(B)** EMT markers were analyzed at the transcriptional level via qRT-PCR (Mann-Whitney test; n=5-6; *: p≤0.05; **: p≤0.01). **(C)** Cells were analyzed by phase contrast microscopy (upper panels; size bar = 40 µm), and also stained for E-cadherin (green), F-actin (red) and nuclei (blue) for immunofluorescence microscopy (lower panels; size bar = 20 µm).

Next, cell migration and invasiveness, processes that have been linked to EMT, were investigated with parental, CRK-deficient, CRKL-deficient, and compound CRK family-deficient HCT116 cells. Loss of either CRKL or CRK drastically reduced cell migration (Figure 25A) and invasion (Figure 25B) in transwell assays. CRK family-deficiency exhibited synergistic effects, as migration and invasion were completely abolished (Figure 25A, B). To address collective cell migration, wound healing assays were performed. In accordance with single cell migration assessed by transwell assays, CRK- or CRKL-deficient cells exhibited reduced wound closure by collective migration, while complete CRK family-deficiency abolished wound healing (Figure 25C, D).

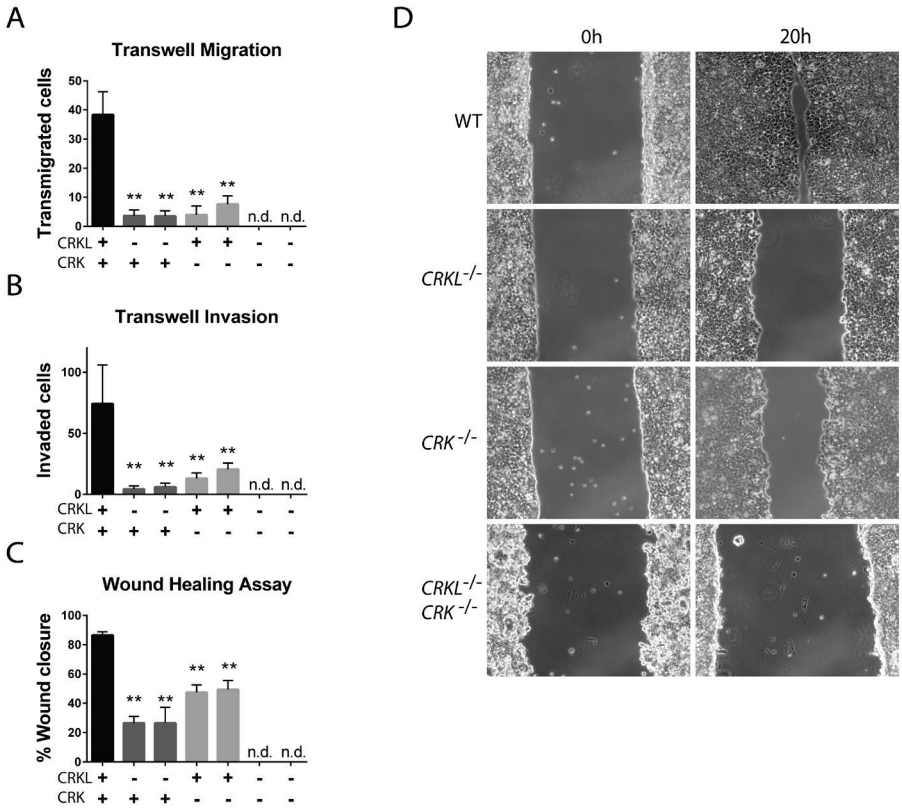

Figure 25: CRK family proteins are required for migration and invasiveness. (A) Transwell migration assays were carried out to analyze migration of parental, CRKL-, CRK-, and compound CRK family-deficient HCT116 cells (Mann-Whitney test; n=6; p=0.0022; n.d. = not detected). Migration was allowed for 24 h. **(B)** Transwell invasion assays were performed to assess invasiveness with an incubation for 48 h (Mann-Whitney test; n=6; p=0.0022; n.d. = not detected). **(C, D)** Collective cell migration was measured by wound healing assays. The cell-free gap area was quantified using ImageJ v1.51 at 0 h and 20 h after removal of the insert to calculate the wound closure (Mann-Whitney test; n=6; p=0.0022; n.d. = not detected).

In summary, CRK and CRKL have largely redundant roles in EMT, migration and invasiveness. However, loss of CRKL induced a stronger effect on invasiveness and wound healing, as compared to loss of CRK. Importantly, complete loss of CRK family proteins exerted synergistic effects, generating highly epithelial cells that lack migratory capacities and invasiveness.

3.2.3 CRK family proteins are central amplifiers of SRC/FAK kinase complex signaling

Since SASH1 counteracts EMT by inhibiting CRKL-mediated activation of SRC (see section 3.1), the impact of a complete CRK family-deficiency on SRC, as well as on other related components of integrin signaling was investigated. Complete loss of CRK family proteins resulted in drastically reduced activation of SRC family proteins, as evidenced by diminished phosphorylation of residue Y419 (Figure 26A). Surprisingly, further upstream in the pathway, autophosphorylation of FAK at residue Y397, which normally occurs early during integrin engagement and induces SRC/FAK complex formation (Schaller et al. 1994; Toutant et al.

2002), was nearly undetectable in CRK family-deficient cells (Figure 26A). Accordingly, the SRC/FAK kinase complex substrates p130CAS and paxillin were poorly phosphorylated (Figure 26A). Parental cells formed focal adhesions, which were apparent as streak-like structures that featured talin, Y118 phospho-paxillin and total paxillin (Figure 26B). In contrast, CRK family-deficient cells exhibited a rounded cellular shape accompanied by a diffuse localization of all focal adhesion components tested, including integrin beta-1, and lacked formation of streak-like focal adhesions (Figure 26B).

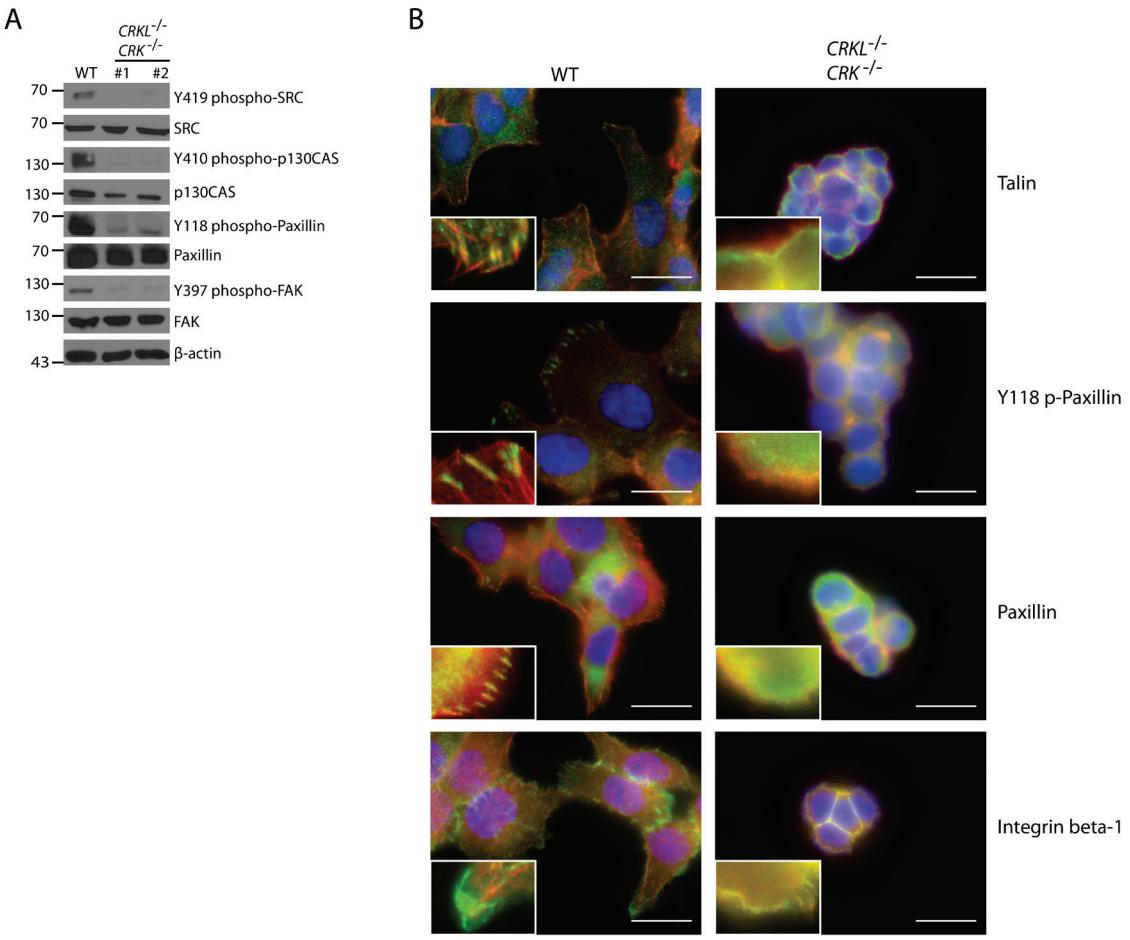

Figure 26: CRK family proteins are required for proper activation of SRC/FAK signaling and focal adhesion formation. **(A)** Immunoblot analysis of parental (WT) and CRK family-deficient HCT116 cells. **(B)** Parental and CRK family-deficient cells were also stained for focal adhesion components (as indicated, green), F-actin (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 µm).

Next, the impact of a complete CRK family-deficiency was investigated in PANC-1 pancreatic cancer cells as independent cell line. In accordance with previous observations, loss of CRK family proteins resulted in strongly decreased SRC/FAK substrate phosphorylation (Figure 27A). ZEB1 levels were also reduced, while E-cadherin levels were drastically increased upon loss of CRK family proteins (Figure 27B). Further, induction of EMT via TGF-β1 exhibited

diminished effects in CRK family-deficient cells, as E-cadherin levels were still strongly increased (Figure 27B).

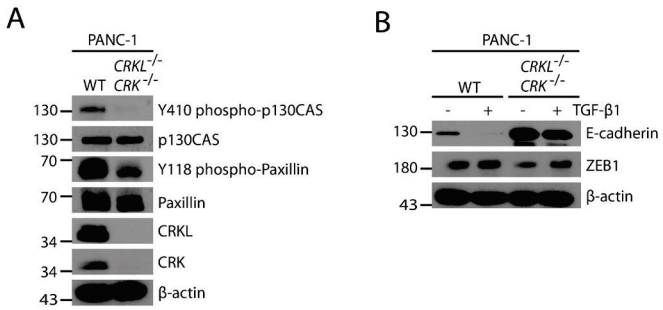

Figure 27: CRK family-deficiency exerts similar effects in PANC-1 pancreatic cancer cells. (A) Immunoblot analysis of parental (WT) and CRK family-deficient PANC-1 cells. **(B)** Immunoblot analysis of parental (WT) and CRK family-deficient PANC-1 cells, which were either left untreated or stimulated with 10 ng ml^{-1} TGF-β1 for 72 h. Immunoblot analysis was performed with the help of Widya Johannes (Technical University of Munich, Germany).

Due to the central role of CRK family proteins in activation of the SRC/FAK complex, deficiency of the kinase CSK was introduced by CRISPR/Cas9 either in the parental line or in CRK family-deficient HCT116 cells. The underlying rationale was to generate cells with strongly increased genetic activation of SRC by loss of CSK, which acts as its negative regulator (Okada 2012). Introduction of CSK-deficiency also allowed to investigate if CRK family proteins regulate SRC activity in a manner dependent on an interaction with CSK, as previous studies have suggested (Watanabe et al. 2009; Kumar et al. 2018). Loss of CSK in parental cells indeed induced a strong constitutive activation of SRC family members as indicated by phosphorylation at residue Y419, as well as of FAK at several tyrosine residues, and of the SRC/FAK substrates p130CAS and paxillin (Figure 28A, B). Strikingly, FAK remained poorly phosphorylated in CRK family-deficient cells upon loss of CSK, despite of a similar increase in SRC phosphorylation at residue Y419 (Figure 28A, B). Phosphorylation of paxillin and p130CAS was also not increased to the same extent as observed in CSK-deficient parental cells, suggesting that activation of FAK is partially uncoupled from activation of SRC specifically in CRK family-deficient cells (Figure 28A, B).

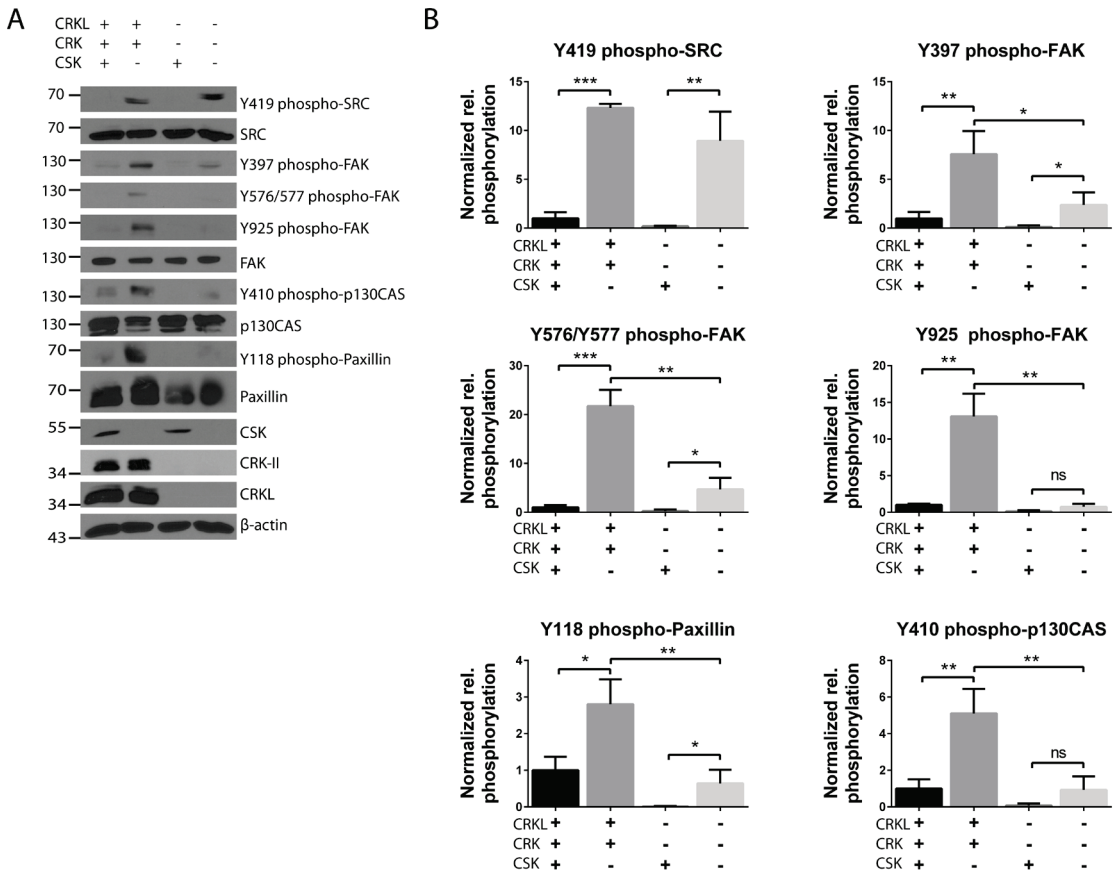

Figure 28: CRK family proteins mediate the partially independent activation of SRC and FAK kinases. (A) Immunoblot analysis of parental and CRK family-deficient HCT116 cells, either proficient or deficient for the kinase CSK. **(B)** Phosphorylation of SRC, FAK, paxillin and p130CAS was measured relative to the respective total protein via immunoblot analyses (unpaired t-test; n=3; ns = not significant: p>0.05; *: p≤0.05; **: p≤0.01; ***: p≤0.001).

Next, site-directed mutagenesis was performed to inactivate either the SH2 domain within human CRKL by an R39K exchange or the N-terminal SH3 (SH3N) domain by a W160R exchange. Disruption of domain function was confirmed by co-immunoprecipitation of stably expressed CRKL variants with C-terminal RFP-tag. The R39K exchange abolished the activity of the CRKL SH2 domain to bind GAB1 (Sakkab et al. 2000), while the W160R exchange led to inactivation of the N-terminal SH3 domain, abrogating interactions with C3G (Nosaka et al. 1999; Arai et al. 1999) or SASH1, which has been established as novel SH3N binding partner in part I (see section 3.1) of the present thesis (Figure 29A). Of note, p130CAS and paxillin, which are known to bind to the SH2 domain of CRKL upon their phosphorylation (Salgia et al. 1995; Salgia et al. 1996), only bound to wild type CRKL (Figure 29A). Stable expression of wild type CRKL in HEK293 cells promoted phosphorylation of SRC, FAK, p130CAS and paxillin, despite of an only moderate increase in total CRKL levels (Figure 29B). While forced expression of CRKL with an inactive SH2 domain did not lead to increased phosphorylation, expression of CRKL

with an inactive SH3N domain further reduced phosphorylation of SRC, FAK, p130CAS and paxillin (Figure 29B). Accordingly, the phenotype of CRK family-deficient HCT116 cells was only rescued by wild type CRKL, but not by CRKL^{R39K} or CRKL^{W160R}, as indicated by cell morphology (Figure 29C).

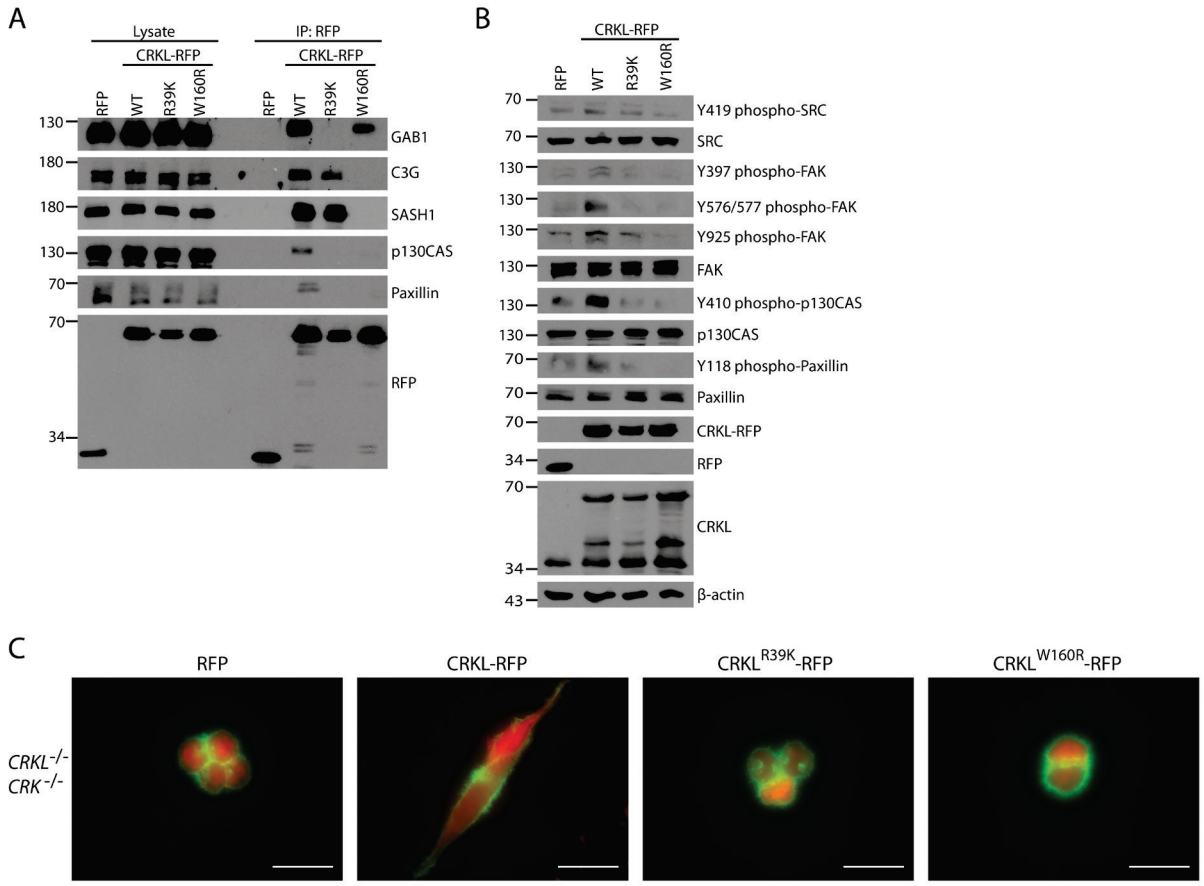

Figure 29: The SH2 and SH3N domains of CRKL are required for activation of SRC/FAK signaling. (A) Immunoblot analysis of cell lysates and RFP immunoprecipitations with HEK293 cells stably transfected to express RFP as control, wild type CRKL-RFP, CRKL^{R39K}-RFP or CRKL^{W160R}-RFP. **(B)** Immunoblot analysis of HEK293 cells stably expressing RFP as control, wild type CRKL-RFP, CRKL^{R39K}-RFP or CRKL^{W160R}-RFP. **(C)** CRK family-deficient HCT116 cells transiently expressing RFP or the indicated CRKL-RFP variants were subjected to fluorescence microscopy (size bar = 20 µm) for RFP (red) and F-actin (green).

3.2.4 CRKL activates SRC/FAK signaling in dependence of the small GTPase RAP1

Due to the relevance of SRC/FAK kinase complex signaling for integrin-mediated cell adhesion, parental and CRK family-deficient HCT116 cells were subjected to cell-matrix adhesion assays. Loss of CRK proteins significantly reduced adhesion to fibronectin (Figure 30A), despite of equivalent integrin beta-1 surface levels, as determined by flow cytometry (Figure 30B). Next, cell-matrix adhesion assays were carried out in presence or absence of EDTA or Mg^{2+} to inhibit or induce activation of integrins, respectively. While adhesion was diminished by EDTA treatment, addition of $MgCl_2$ significantly increased adhesion of parental and CRK family-deficient cells (Figure 30C). However, compared to parental cells, CRK family-deficient cells were still poorly adherent even upon activation of integrins. The small GTPase RAP1, which has been described to be activated by CRK proteins (Sakkab et al. 2000; Cheung et al. 2011), is a positive regulator of inside-out activation of integrins and integrin-mediated cell adhesion (Katagiri et al. 2000; Arai et al. 2001). Thus, CRK proteins might regulate activation of the SRC/FAK complex dependent on RAP1. To investigate this, HEK293 cells stably expressing RFP or CRKL-RFP were treated with the geranylgeranyltransferase I inhibitor GGTI-298 that counteracts RAP1 processing and thus activation. Inhibition of RAP1 activation in CRKL-RFP expressing cells strongly reduced phosphorylation of the SRC/FAK substrates p130CAS and paxillin (Figure 30D). Treatment of parental HCT116 cells with GGTI-298 decreased phosphorylation of p130CAS and paxillin (Figure 30E). Further, RAP1 inhibition resulted in increased E-cadherin and slightly decreased ZEB1 levels (Figure 30E). Treatment of parental HCT116 cells also led to a diffuse localization of talin, similar to the phenotype observed in CRK family-deficient cells (Figure 30F).

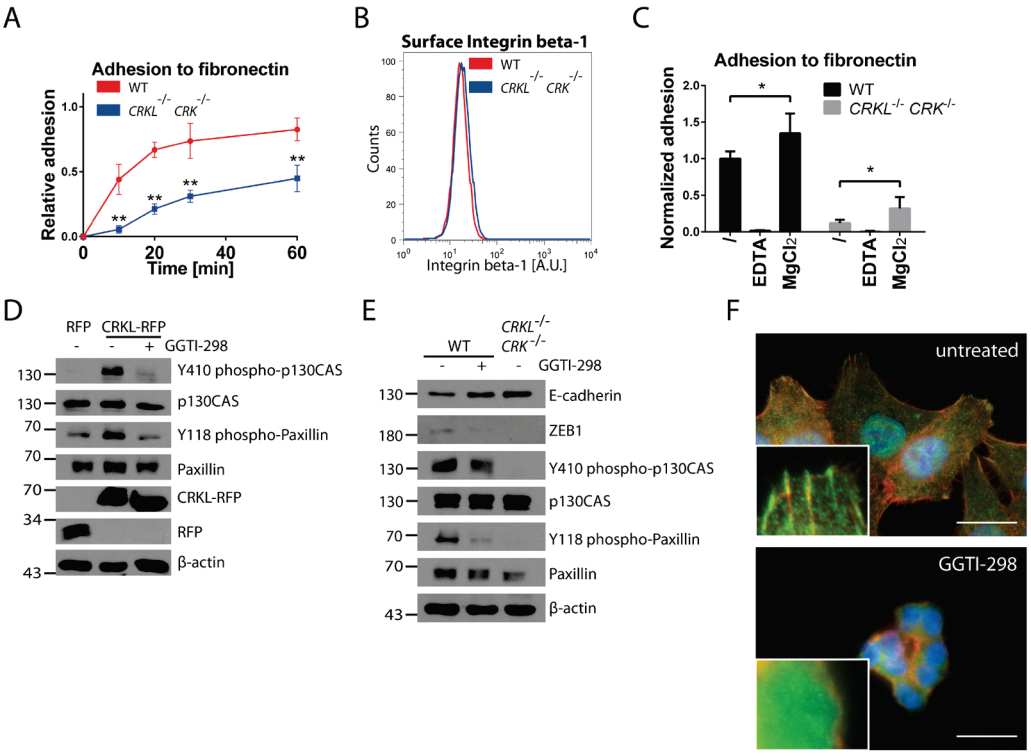

Figure 30: CRK proteins induce SRC/FAK substrate phosphorylation via RAP1. (A) Adhesion assays were performed with parental (WT) and CRK family-deficient HCT116 cells to quantify cell adhesion to fibronectin. Adherent cells were counted relative to the total number of seeded cells at different time points (Mann-Whitney test; n=6; **: p=0.0022). **(B)** Total integrin beta-1 levels of parental (WT) and CRK family-deficient HCT116 cells were quantified by flow cytometry. **(C)** Adhesion assays were performed with parental (WT) and CRK family-deficient HCT116 cells, allowing adhesion to fibronectin for 20 min in the presence or absence of 5 mM EDTA or 5 mM $MgCl_2$/2 mM EGTA (Mann-Whitney test; n=6; p=0.026). **(D)** Immunoblot analysis of HEK293 cells stably expressing RFP or CRKL-RFP. Additionally, CRKL-RFP expressing cells were treated with DMSO as vehicle control or 10 µM GGTI-298 for 48 h. **(E)** Immunoblot analysis of parental (WT) and CRK family-deficient HCT116 cells. Parental cells were either treated with DMSO as vehicle control or 10 µM GGTI-298 for 72 h. **(F)** Parental HCT116 cells treated with DMSO or 10 µM GGTI-298 for 72 h were also stained for immunofluorescence microscopy (size bar = 20 µm) for talin (green), F-actin (red) and nuclei (blue).

3.2.5 Loss of CRK family proteins abrogates the ability to undergo epithelial-mesenchymal transition

The data obtained so far highlighted that CRK family proteins act as central amplifiers of SRC/FAK signaling. Of note, CSK-deficiency induced sustained activation of the SRC/FAK complex only in CRK family-proficient cells. This was also reflected by the EMT phenotype. While loss of CSK was sufficient to induce a *bona fide* EMT in the parental line with strongly reduced E-cadherin and elevated ZEB1 protein levels, the epithelial phenotype of CRK family-deficient cells was essentially unaffected by loss of CSK (Figure 31A). Loss of CSK in single CRKL- or CRK-deficient HCT116 cells supported this observation, since CSK-deficiency did not drastically alter E-cadherin or ZEB1 levels (Figure 31B). CSK-deficient parental cells exhibited a fibroblast-like appearance with loss of E-cadherin at cell-cell contacts (Figure 31C, upper right panel). These cells also exhibited formation of multipolar lamellipodia and putative

tail retractions, indicating a motile phenotype (Figure 31C, upper right panel). This phenotype was strongly attenuated in CRKL- or CRK-deficient cells, and completely absent in compound CRK family-deficient cells (Figure 31C). Furthermore, CRKL-, CRK-, and CRK family-deficient cells still exhibited E-cadherin at intercellular contacts, reflecting their stable epithelial phenotype even in the absence of CSK, and thus presence of activated SRC (Figure 31C).

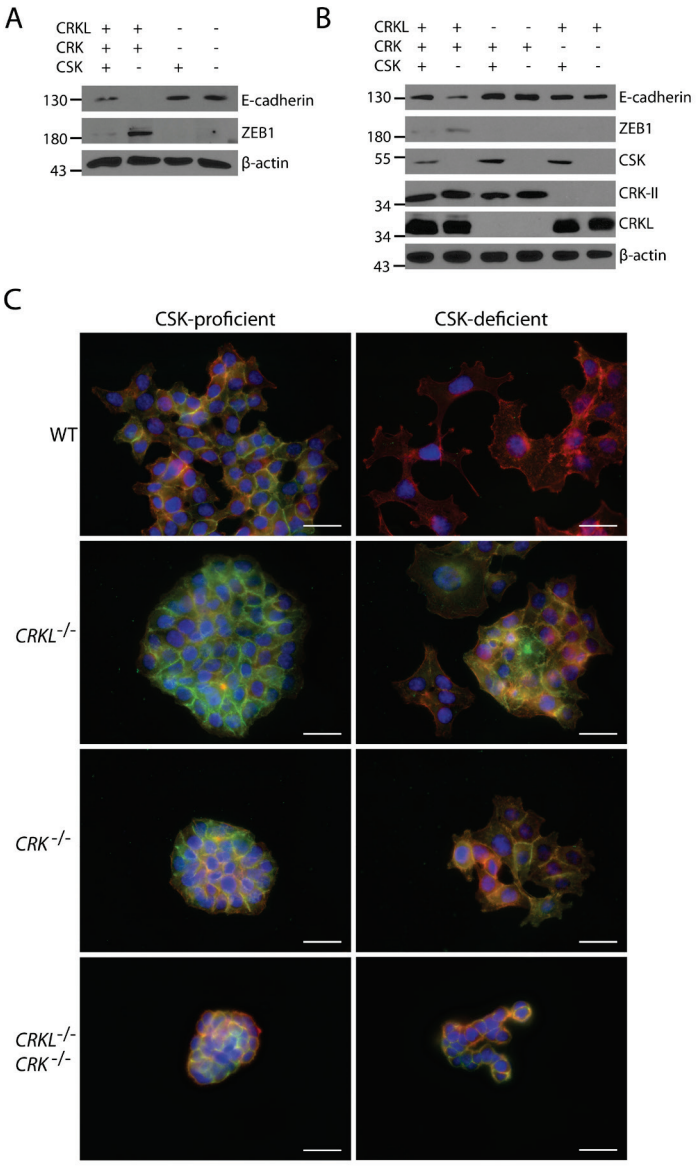

Figure 31: CRK family-deficient cells are unable to undergo EMT upon loss of CSK. (A) Immunoblot analysis of parental and CRK family-deficient HCT116 cells, either pro- or deficient for CSK. **(B)** Immunoblot analysis of parental, CRKL- or CRK-deficient cells either pro- or deficient for CSK. **(C)** Parental (WT), CRKL-, CRK- and compound CRK family-deficient cells either pro- or deficient for CSK were stained for E-cadherin (green), F-actin (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 µm).

In accordance with the mesenchymal phenotype, CSK-deficient parental cells exhibited drastically increased transwell migration (Figure 32A) and invasion (Figure 32B). This increase was strongly attenuated in CSK-deficient cells lacking either CRKL or CRK (Figure 32A, B). Strikingly, loss of CSK did not rescue the migration and invasion defects of CRK family-deficient cells (Figure 32A, B). To investigate if CRK family-deficient cells exhibit any remaining migratory capability, transwell migration assays were performed with extended incubation times. Nevertheless, cells deficient for the CRK family completely lacked migration, even if SRC was activated by loss of CSK (Figure 32C).

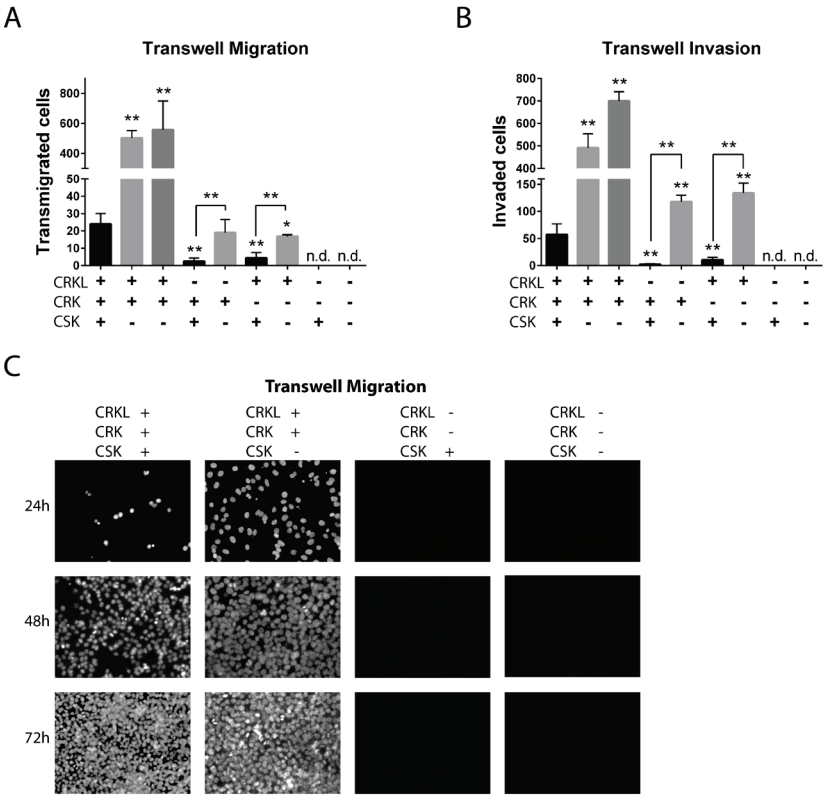

Figure 32: CSK-deficiency generates a highly migratory and invasive phenotype only in CRK family-proficient cells. (A) Transwell migration assays (24 h incubation) were carried out to analyze migration of parental, CRKL-, CRK-, and compound CRK family-deficient HCT116 cells, either proficient or deficient for CSK (Mann-Whitney test; n=6; *: p=0.0411; **: p=0.0022; n.d. = not detected). **(B)** Transwell invasion assays (48 h incubation) were performed to assess invasiveness (Mann-Whitney test; n=6; p=0.0022; n.d. = not detected). **(C)** Transwell migration assays were performed for several incubation timepoints (as indicated on the left) with parental and CRK family-deficient HCT116 cells that were either proficient or deficient for CSK. Nuclei of transmigrated cells were stained with DAPI.

3.2.6 Loss of CRK family proteins negatively regulates c-MYC and diminishes proliferation, survival and chemoresistance

To study if complete loss of CRK family proteins also affects other pathophysiological processes, parental, CRKL-, CRK-, and CRK family-deficient HCT116 cells were analyzed by RNA sequencing. While loss of either CRKL or CRK alone resulted in only minor changes on the transcriptional level (data not shown), complete CRK family-deficiency induced profound transcriptome-wide changes, and the expression of 549 genes was significantly altered (data not shown). Gene set enrichment analysis revealed several hallmark gene sets that were negatively (Table S 3) or positively (Table S 4) enriched in CRK family-deficient cells. Of note, several downregulated hallmarks were related to cell proliferation and cell cycle progression (Figure 33A). The hallmark “MYC_Targets_V1” was further displayed as heatmap analysis, demonstrating that almost all genes within this gene set were downregulated in two independent CRK family-deficient clones (Figure 33B). Upregulated hallmark gene sets in turn were in accordance with the expected functional role of the CRK family, e.g. “Apical_Junction”, but also included further processes like “Hypoxia” or “KRAS_Signaling_Up/Down“ (Figure 33C).

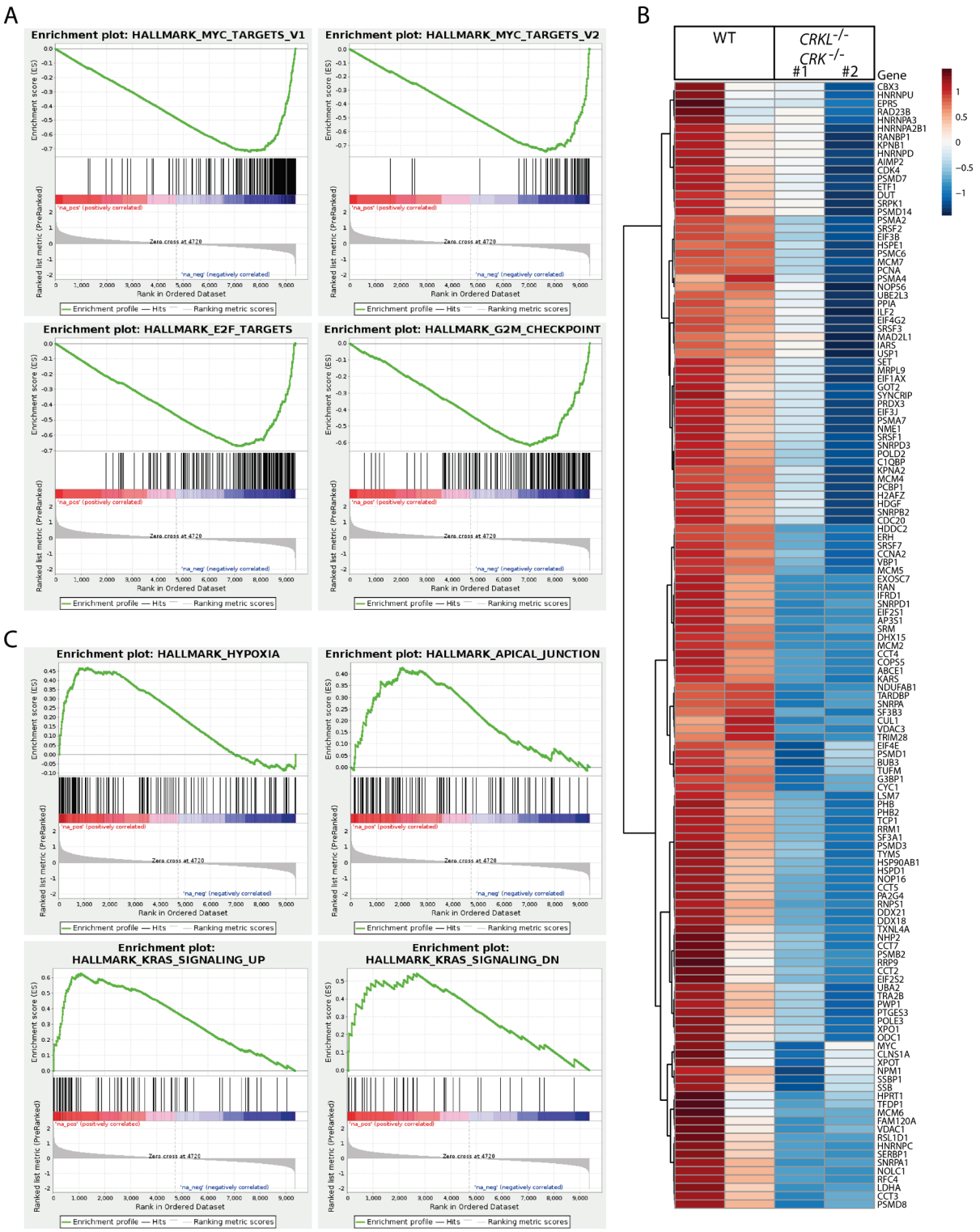

Figure 33: Loss of CRK family proteins downregulates gene sets related to cell proliferation and cell cycle progression. **(A)** RNA sequencing was performed in collaboration with Dr. Roland Rad and Dr. Thomas Engleitner (Technical University of Munich, Germany) to study the effects of CRK family-deficiency in HCT116 cells. Enrichment plots of selected hallmarks are shown, which were downregulated in CRK family-deficient cells. **(B)** Heatmap analysis of "MYC_Targets_V1" hallmark was carried out (red = upregulated; blue = downregulated). **(C)** Enrichment plots for selected hallmark gene sets are shown, which were upregulated in CRK family-deficient cells.

Due to the prevalence of prominent downregulated hallmark gene sets in the RNAseq data, the focus was on cell proliferation and the transcription factor c-MYC. In accordance with the bioinformatical analysis of the RNAseq data, CRK family-deficient cells exhibited significantly reduced c-MYC protein levels (Figure 34A, B). In addition to the reduced c-MYC levels, CRK family-deficient cells featured reduced activation loop phosphorylation of the kinases ERK1/2 at residues T202/Y204 (Figure 34A). Importantly, cell proliferation (Figure 34C), as well as cell survival in soft agar (Figure 34D), were significantly reduced upon loss of CRK family proteins. Besides hallmark gene sets associated with cell proliferation, the gene set “DNA_Repair" was also found to be downregulated in CRK family-deficient cells (Table S 3), while the gene set “Apoptosis” was upregulated (Table S 4). Therefore, the resistance of parental and CRK family-deficient HCT116 cells towards 5-fluorouracil and oxaliplatin was investigated. Loss of CRK family proteins indeed resulted in a significantly decreased chemoresistance against both chemotherapeutic agents (Figure 34E, F).

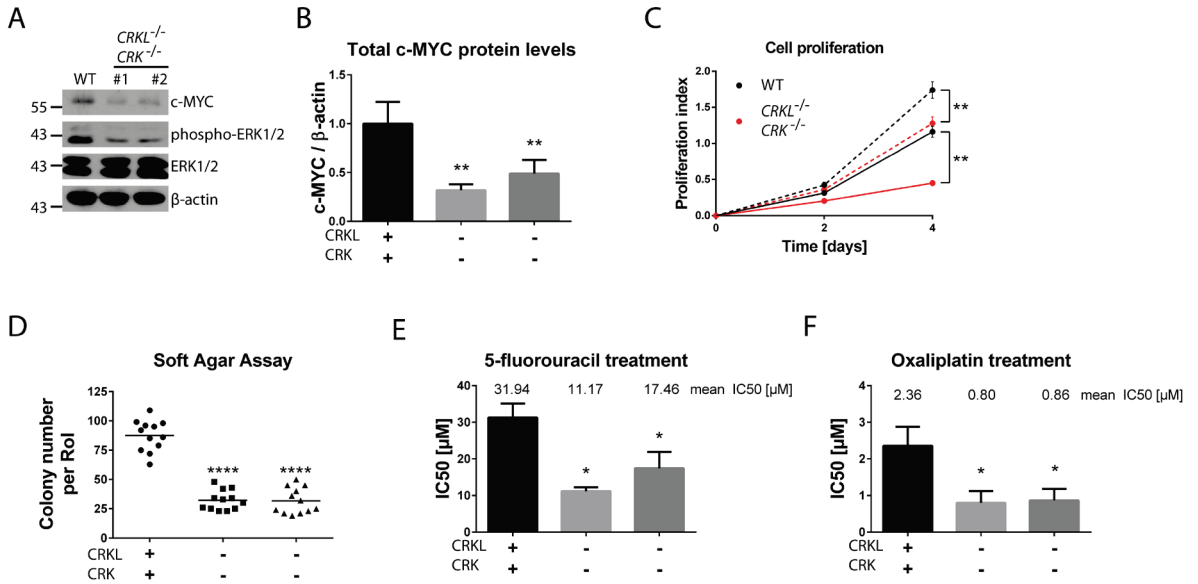

Figure 34: Loss of CRK family proteins also affects ERK1/2 phosphorylation, c-MYC levels, and thus related cellular phenotypes. (A) Immunoblot analysis of parental (WT) and CRK family-deficient HCT116 cells. **(B)** c-MYC protein levels were quantified relative to β-actin via immunoblot analyses (Mann-Whitney test; n=6; p=0.0022). **(C)** Cell proliferation of parental and CRK family-deficient cells was assessed by XTT assays with cells cultured either in 2% (solid lines) or 10% (dashed lines) FCS (Mann-Whitney test; n=4-8; 2 independent clones combined; p=0.004). **(D)** Soft agar assays were performed to analyze anchorage-independent survival (unpaired t-test; n=12; p≤0.0001; RoI = region of interest). **(E)** Resistance against 5-fluorouracil (24 h treatment and 48 h recovery) was assessed by XTT assays (Mann-Whitney test; n=4; p≤0.05; IC50 = half maximal inhibitory concentration). **(F)** Resistance against oxaliplatin (24 h treatment and 48 h recovery) was assessed by XTT assays (Mann-Whitney test; n=4; p≤0.05; IC50 = half maximal inhibitory concentration).

3.2.7 Expression of *CRK* family genes has prognostic relevance in locally advanced colorectal cancer

To investigate the clinical relevance of CRK family proteins in colorectal cancer, expression levels of *CRK* and *CRKL* in primary tumors were assessed by qRT-PCR. Cryosamples from patients with locally advanced (UICC stage III) cancer from the Department of Surgery (Klinikum rechts der Isar, Technical University of Munich, Germany) were analyzed. Additionally, expression of the EMT markers *ZEB1* and *CDH1* was quantified. Intratumoral expression of *CRK* significantly correlated with *CRKL* expression in individual patient tissues (Figure 35A). To assess *CRK* family expression, *CRK* and *CRKL* were normalized to their respective means, before the mean of *CRK* and *CRKL* was calculated. *CRK* family expression directly correlated with expression of the mesenchymal marker *ZEB1*, while the epithelial marker *CDH1* showed no significant correlation (Figure 35B, C). Next, fresh-frozen tissue slides of the same cohort were stained for ZEB1 and CRKL. ZEB1 staining was increased in nuclei of cancer cells in de-differentiated, invasive regions (Figure 35D, panels b, f, arrows), whereas more differentiated tumor regions exhibited only low ZEB1 levels (Figure 35D, panels j, n). Interestingly, CRKL staining was partially increased in cancer cells with high ZEB1 levels, with some indication of nuclear staining (Figure 35D, panels c, g, arrowheads), while it was rather uniformly stained in central regions (Figure 35D, panels k, o).

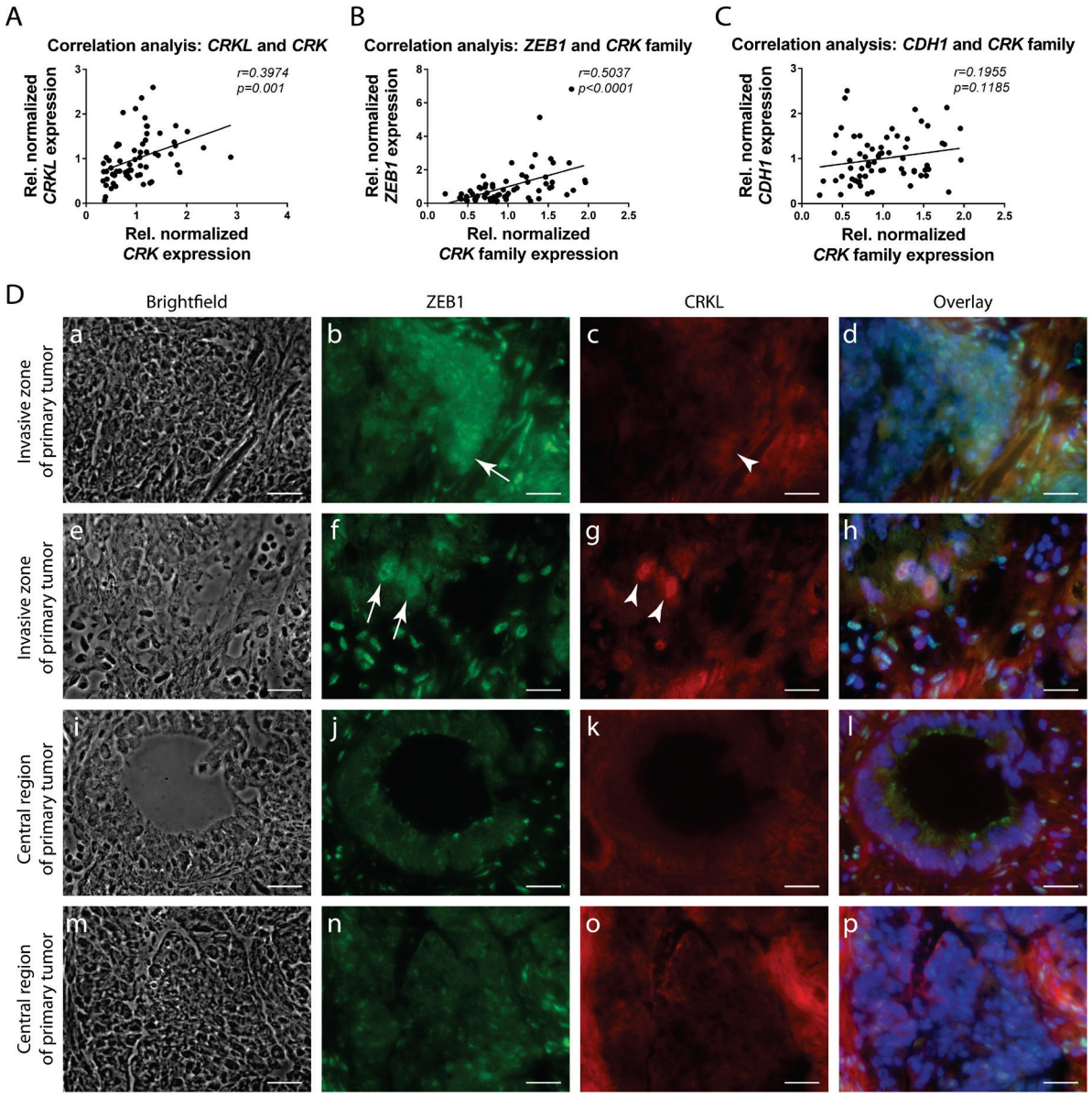

Figure 35: *CRK* family expression correlates with the EMT transcription factor *ZEB1* in colorectal cancer. (A) *CRKL*, *CRK*, *CDH1* and *ZEB1* expression was determined via qRT-PCR of tumor samples of UICC III colorectal cancer patients with the help of Benjamin Owen Slusarenko (Technical University of Munich, Germany). Correlation analysis between intratumoral expression of *CRKL* and *CRK* (Pearson r; n=65; r=0.3974; p=0.001). **(B)** Expression of *CRK* family was determined by calculating the mean of normalized *CRK* and *CRKL* expression. Correlation analysis between expression of *CRK* family and *ZEB1* (Pearson r; n=65; r=0.5037; p≤0.0001). **(C)** Correlation analysis between expression of *CRK* family and *CDH1* (Pearson r; n=65; r=0.1955; p=0.1185). **(D)** Cryosections of primary tumors were stained for ZEB1 (green), CRKL (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 µm; arrows = invasive tumor cells with high ZEB1 staining; arrowheads = invasive tumor cells with high CRKL staining).

Finally, Kaplan-Meier survival analysis revealed that high *CRKL* and high *CRK* expression were significantly associated with poor distant recurrence-free survival, with *CRKL* expression having a higher prognostic significance (Figure 36A, B). *CRK* family expression, however, was also of high prognostic significance, highlighting a clinical relevance of CRK family adaptor proteins in advanced colorectal cancer (Figure 36C). While *CDH1* expression was not of prognostic significance, high *ZEB1* expression was associated with strongly reduced distant recurrence-free survival, which is in accordance with the direct correlation between *ZEB1* and *CRK* family expression (Figure 36D, E). No further correlation with clinical or pathological parameters was observed by multivariable regression analysis with post-operative recurrence as relevant test parameter (data not shown).

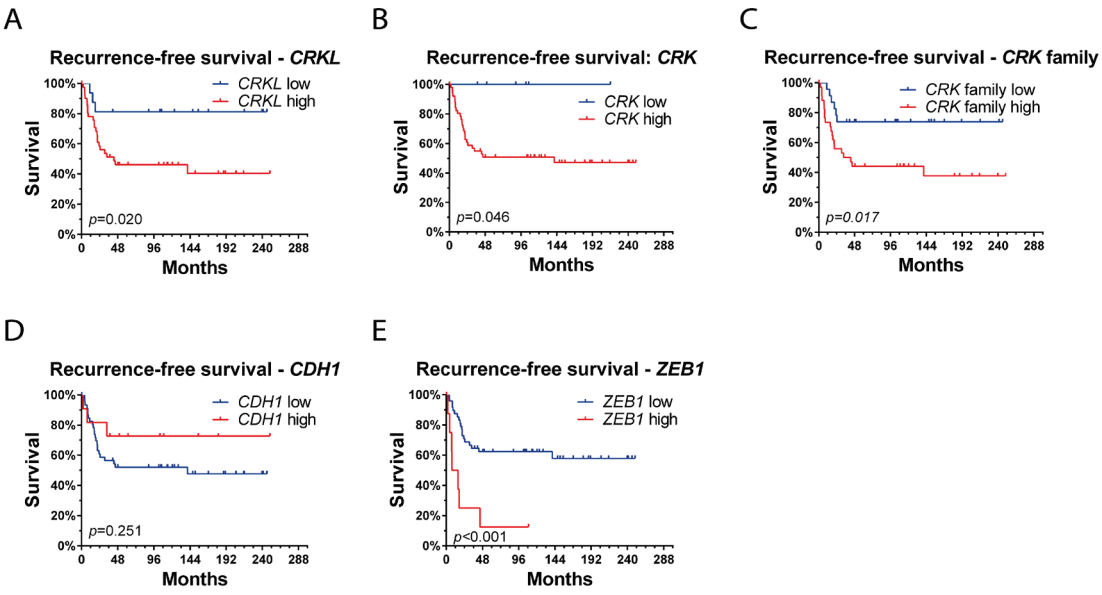

Figure 36: *CRK* family expression has prognostic significance in colorectal cancer. (A-E) Kaplan-Meier analysis was performed with the help of Dr. Dr. Ulrich Nitsche (Technical University of Munich, Germany), showing the impact of high or low expression of *CRKL* (A), *CRK* (B), *CRK* family (C), *CDH1* (D) and *ZEB1* (E) on distant recurrence-free survival (Log-rank test; n=65; p-values indicated in the graphs).

3.2.8 Outlook: Generation of an inducible *Crkl*-deficient mouse model for colorectal cancer

The present study demonstrated a significant clinical role of CRK proteins in colorectal cancer. To further investigate the role of CRK proteins in an *in vivo* setting, the establishment of a genetic mouse model with inducible *Crkl*-deficiency, and potentially compound *Crk* family-deficiency, is planned. To generate a mouse model with inducible Crkl-deficiency, C57BL/6N mice carrying the "conditional ready" allele *Crkl*$^{tm1c(EUCOMM)Hmgu}$ (*Crkl* flanked by loxP sites on chromosome 16, generated by the EUCOMM mutagenesis consortium, in short *Crkl*^{flox}) are currently propagated and will be crossed with the transgenic line pVillin-Cre-ER^{T2} (el Marjou et al. 2004) to allow inducible *Crkl*-deficiency in the intestinal epithelia, and with *Apc*$^{+/1638N}$ mice, which spontaneously develop intestinal tumors (Figure 37A) (Fodde et al. 1994; Holtorf et al. 2018; Buchert et al. 2015). In this project, *Crkl*^{flox} genotyping was optimized to distinguish homo- and heterozygote alleles during crossing (Figure 37B).

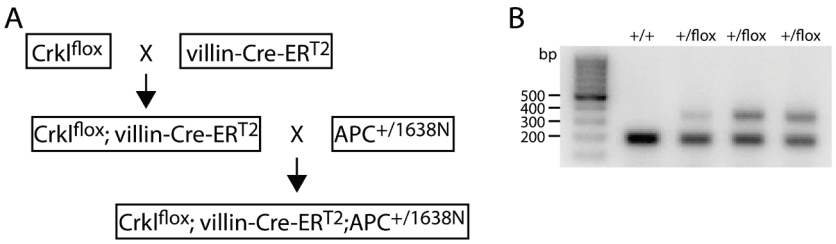

Figure 37: Crossing scheme and genotyping for the generation of an inducible *Crkl*-deficient mouse model for colorectal cancer. (A) Planned crossing and interbreeding scheme with mice represented by the indicated alleles. **(B)** Agarose gel DNA analysis for PCR-based genotyping of the *Crkl*^{flox} allele (flox: approx. 300 bp; wild type approx. 200 bp).

3.3 Part III: Development of small molecule CRK family inhibitors

3.3.1 Label-free biochemical screening for CRKL inhibitors

Currently, EMT is believed to be a driver of metastasis and chemoresistance, supported by data from this thesis. Further, this thesis demonstrated that CRK proteins play a crucial role in EMT, highlighting them as promising therapeutic targets. Due to the dominant negative effect of the inactive N-terminal SH3 domain of CRKL (see Figure 29B) on SRC/FAK signaling, the N-terminal SH3 domains (SH3N) of CRK proteins, which are highly similar regarding their primary and tertiary structure, were selected as target structures (Figure 38A, B).

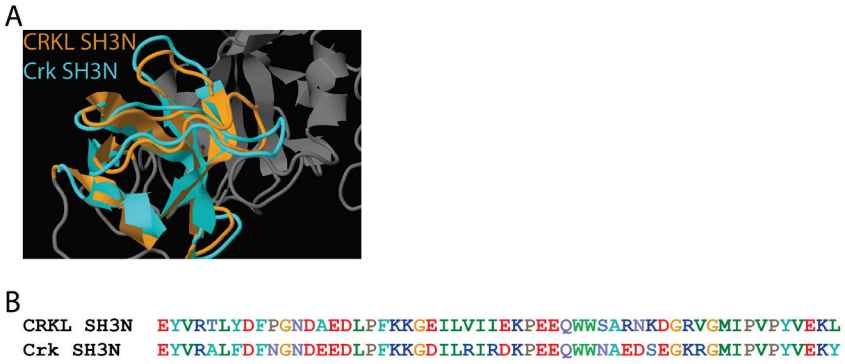

Figure 38: Structural comparison of CRKL and Crk N-terminal SH3 domains. (A) Alignment of tertiary structures of human CRKL (partial sequence from PBD ID: 2LQN) and murine Crk (partial sequence from PDB ID: 1CKA) N-terminal SH3 domains was performed using the RCSB PDB tool Pairwise Structure Alignment v1.1 and jCE alignment algorithm (RMSD = 1.96 Å). **(B)** Alignment of primary structure of the same sequence was performed with BioEdit v7.0.5.3.

To identify compounds that bind the SH3N domains, biochemical label-free assays were selected as screening approach in cooperation with Dr. Eduardo Dominguez Medina (University of Santiago de Compostela, Spain). The SH3N domain of CRKL was purified from recombinant bacterial expression using a C-terminal HIS-tag and Ni^{2+}-NTA-based affinity purification. The SH3N domain contains several surface exposed lysine residues (Figure 39A), allowing efficient immobilization on EDC/NHS-activated sensor plates. Immobilization of the purified domain was tested at various pH conditions and concentrations (Figure 39B). Since the theoretical isoelectric point of the domain is pI=5.59 (calculated with ExPASy ProtParam tool), a moderately low pH (pH=5) was selected for immobilization of 25 µg ml^{-1} of the domain to the negatively charged sensor surface. To analyze if the immobilized domain remains active, a synthetic peptide with high affinity towards the SH3N domains of the CRK family (high affinity peptide = HAP; residues CVDNSPPPALPPKRRRSAPS), which was described earlier, was

used (Posern et al. 1998; Kardinal et al. 2000). Addition of the peptide to the sensor plate with the immobilized CRKL SH3N domain resulted in a typical dose-response curve with EC_{50}=3.294 µM, demonstrating that the domain was still active (Figure 39C). Next, the purified wild type domain was compared with the inactive CRKL-SH3N^{W160R} variant. While addition of HAP peptide to the wild type domain elicited a strong response, addition to the CRKL^{W160R} SH3N domain resulted in a response similar to wells without any immobilized domain, indicating specificity of the interaction (Figure 39D). To assess the quality of the screening system, the Z-factor was calculated using vehicle as negative and HAP as positive control according to (Zhang, Chung, and Oldenburg 1999). With Z=0.8 the assay was appropriate for screenings (Figure 39E).

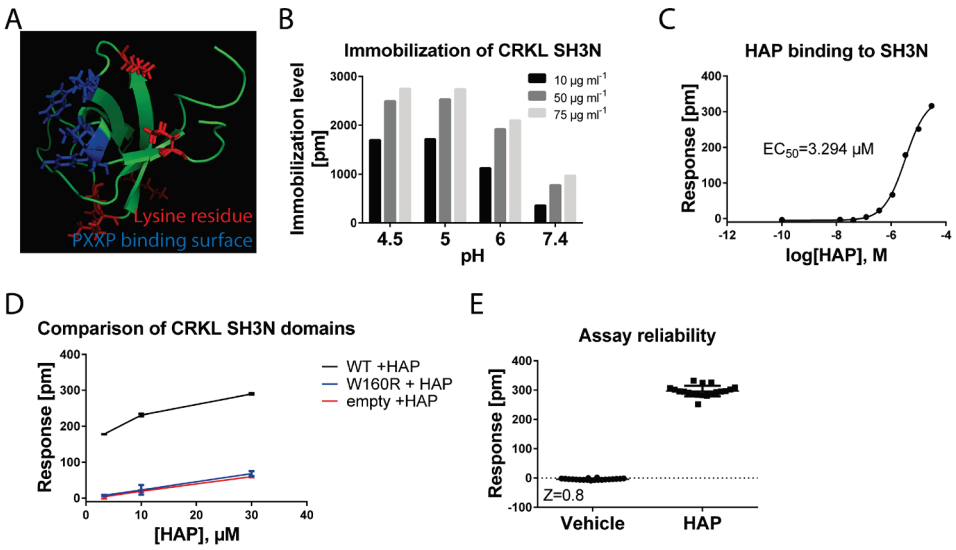

Figure 39: Development of a screening system for small molecule ligands of the SH3N domain of CRKL. (A) Structural analysis (PyMOL viewer) of the SH3N domain of human CRKL highlighting the presence of surface exposed lysine residues (partial sequence from PBD ID: 2LQN). **(B)** Immobilization of the SH3N domain of human CRKL to label-free sensor plates at different pH-values and protein concentrations (pm = picometer). **(C)** Dose-response curve of high affinity peptide (HAP) added to the immobilized domain. **(D)** HAP was added at the indicated concentrations to the wild type CRKL SH3N domain, inactive W160R variant or to wells without immobilized protein (empty). **(E)** The Z-factor was calculated based on measurements with vehicle control or 30 µM HAP added to the immobilized wild type CRKL SH3N domain.

After establishment of the binding assay, screenings were performed with compounds of a generic small molecule compound library, which was developed and supplied by BioFarma-USEF Research Group (University of Santiago de Compostela, Spain) in cooperation with Dr. Eduardo Dominguez Medina (University of Santiago de Compostela, Spain). Compounds were selected from *in silico* docking experiments (data not shown) with the CRKL SH3N (from PDB ID: 2LQN) as target. Docking was performed by Dr. Cristian R. Munteanu (University of A Coruna, Spain). However, in contrast to the HAP peptide, a subsequent *in vitro* screening showed no positive response upon addition of the selected compounds (Figure 40A). Further *in silico* docking experiments using both SH3N domains of CRK (from PDB ID: 1CKA) and CRKL (from PDB ID: 2LQN) as targets were performed (data not shown), but the second *in vitro* screening still showed no binding of pre-selected compounds (Figure 40B).

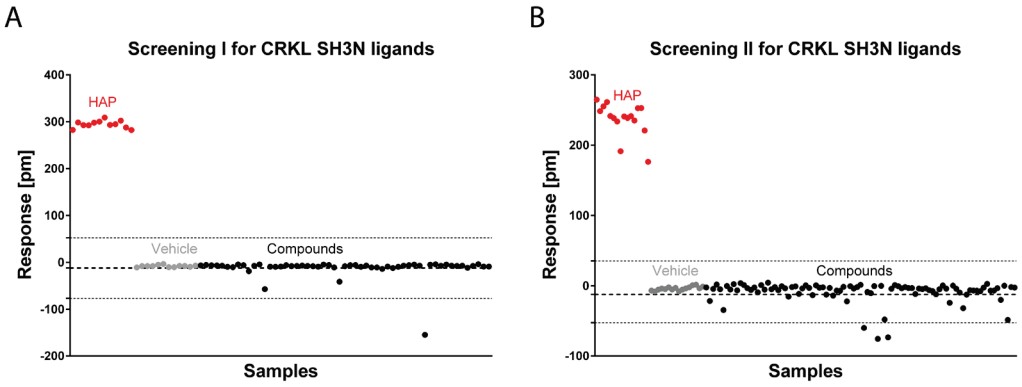

Figure 40: Screening of preselected small compounds revealed no interaction with the N-terminal SH3 domain of CRKL. (A, B) Preselected small molecules were screened by biochemical label-free assays for ligands of the immobilized SH3N domain of CRKL (dashed line = mean of compounds ± 3x standard deviation of compounds).

3.3.2 Peptide-mediated inhibition of CRK proteins

Due to the lack of any available CRK-targeting compounds, the focus was on peptide-mediated inhibition of CRK family proteins. To investigate, if interactions between the SH3N domain of CRK proteins and effectors like C3G can be inhibited by HAP *in vitro* and putatively also *in vivo*, endogenous CRKL was precipitated in the presence of various HAP concentrations. A dose-dependent inhibitory effect of HAP on C3G co-precipitation was observed, while GAB1 bound independently to CRKL (Figure 41).

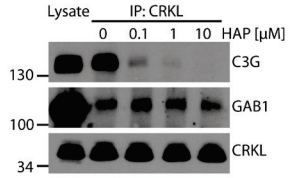

Figure 41: HAP efficiently blocks interactions with the SH3N domain of CRKL. Immunoblot analysis of cell lysate and endogenous CRKL immunoprecipitations using HEK293 cells. Immunoprecipitations were performed in the presence of different HAP concentrations.

Next, the sequence corresponding to HAP was fused to the C-terminus of GFP (Figure 42A). The aim was to mimic a potential cell-permeable CRK family inhibitor by transient transfection of the plasmid pEGFP-HAP encoding the fusion protein. In contrast to GFP alone, GFP-HAP associated with both CRK and CRKL in immunoprecipitation experiments, indicating that the C-terminal HAP is still active and a *pan*-CRK inhibitor (Figure 42B).

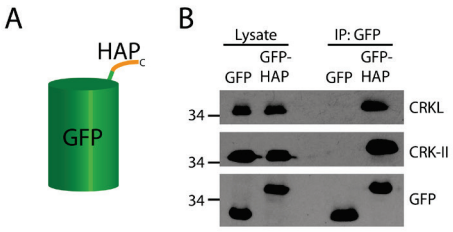

Figure 42: The GFP-HAP fusion protein physically associates with CRKL and CRK. (A) Putative structure of the GFP-HAP fusion protein (C = C-terminus). **(B)** Immunoblot analysis of cell lysates and GFP immunoprecipitations with HEK293 cells transiently transfected to express either GFP or GFP-HAP.

Strikingly, transient expression of GFP-HAP in HEK293 cells resulted in impaired phosphorylation of the SRC/FAK substrates p130CAS and paxillin (Figure 43A). Further, GFP-HAP induced the acquisition of a rounded morphology with a significantly reduced cell size, in good accordance to the observations on CRK family-deficient cells (Figure 43B, C). Importantly, GFP-HAP counteracted CRKL-mediated phosphorylation of SRC, FAK, paxillin and p130CAS (Figure 43D). This indicated that peptides, which bind the SH3N domains of CRK family proteins, counteract CRK family-mediated signaling. For this reason, HAP as synthetic small peptide was dimerized to the HIV1-TAT_{47-57} cell penetrating peptide via a disulfide bond (Figure 43E). The rationale was to trap the HAP peptide inside the cell due to the reducing environment and disulfide cleavage (Kardinal et al. 2000). To detect translocation of the peptide into the cell, a lysine residue coupled to FITC was added to the C-terminus of HAP. Low concentrations (0.1-1 µM) of the peptide dimer showed no translocation, as evidenced by fluorescence microscopy (data not shown). However, a concentration of 10 µM was sufficient to penetrate CRKL-RFP expressing HEK293 cells (Figure 43F). Preliminary data indicated, that HEK293 cells treated with the peptide dimer featured a rounded morphology, similar to GFP-HAP expressing cells. Nevertheless, the peptide dimer primarily localized to specific spots within the cells, suggesting uptake via endocytosis (Figure 43F, inlets). This was also observed at higher concentrations up to 80 µM (data not shown).

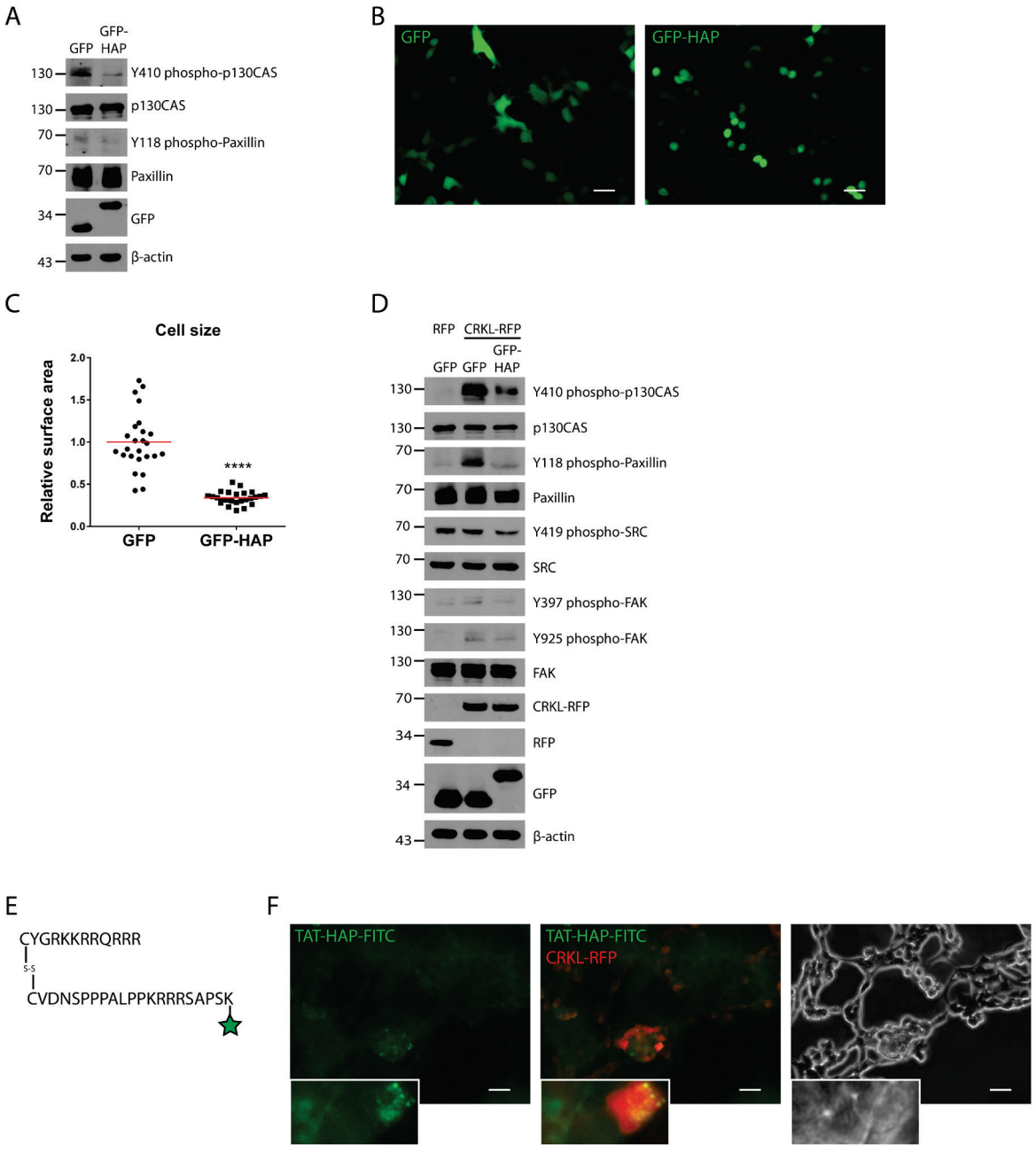

Figure 43: CRK proteins are potential druggable targets to counteract SRC/FAK signaling. (A) Immunoblot analysis of HEK293 cells transiently transfected to express GFP or GFP-HAP. Cells were lysed 48 h post transfection. **(B)** Cell morphology of HEK293 cells transiently transfected with GFP or GFP-HAP was analyzed by GFP fluorescence (size bar = 50 µm). **(C)** Cell surface area was measured using ImageJ v1.51 (unpaired t-test; n=25 cells; p≤0.0001). **(D)** Immunoblot analysis of HEK293 cells stably expressing RFP or CRKL-RFP. Additionally, cells were transiently transfected to express GFP or GFP-HAP. Cells were lysed 48 h post transfection. **(E)** Sequence of HIV1 TAT_{47-57} dimerized to HAP with C-terminal lysine coupled to FITC (green asterisk). **(F)** Fluorescence microsopic analysis of HEK293 cells expressing CRKL-RFP (red). Cells were treated with 10 µM of the peptide dimer (green; TAT-HAP-FITC) for 24 h (size bar = 50 µm).

3.4 Excursus: Role of the tumor suppressor DUSP5 in colorectal cancer

A complementary project within this thesis was to investigate the effects of the dual specificity phosphatase DUSP5 on tumor progression and metastasis formation, which had been suggested to be regulated by SASH1 by unpublished data from our cooperation partner Dr. Henri Weidmann (Sorbonne University, Paris, France). Further, unpublished data from our own group had shown that *DUSP5* expression can suppress metastasis *in vivo* in an orthotopic colorectal cancer model. Therefore, DUSP5 with C-terminal FLAG-tag was transiently expressed in HEK293 cells. In contrast to cells that were transfected with empty control vector or a vector encoding "phosphatase-dead" mutant DUSP5^{C263S}, forced expression of wild type *DUSP5* counteracted serum-induced phosphorylation of ERK1/2 at residues T202/Y204 (Figure 44A). However, when transfected cells were analyzed under normal growth conditions, ERK1/2 phosphorylation was not affected by DUSP5 (Figure 44B).

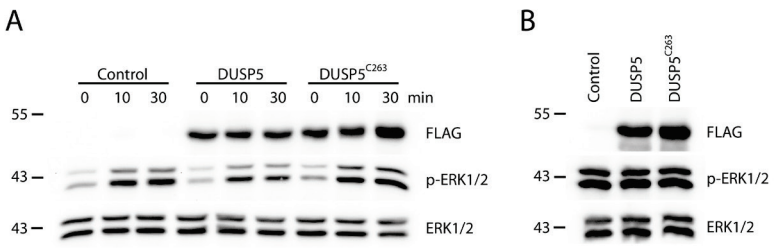

Figure 44: Forced expression of *DUSP5* counteracts serum-induced ERK1/2 phosphorylation. (A) Immunoblot analysis of HEK293 cells transiently transfected with empty control vector or vectors encoding DUSP5-Flag or DUSP5^{C263S}-FLAG. Cells were serum starved overnight and re-stimulated with 10% fetal calf serum for the indicated time points, before they were lysed 48 h post transfection (p-ERK1/2 = T202/Y204 phospho-ERK1/2). **(B)** Immunoblot analysis of similarly transfected HEK293 cells, but under normal growth conditions (cultured with 10% fetal calf serum).

Next, cell proliferation assays were performed with HEK293 cells transiently transfected with vector control or *DUSP5*. However, in accordance with ERK1/2 phosphorylation under normal growth conditions, cell proliferation was not found to be altered, both when assessed with cells cultured in reduced serum (2%) or under standard growth conditions (10% serum) (Figure 45).

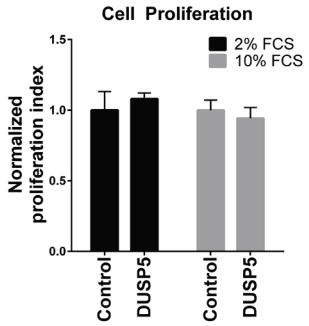

Figure 45: Forced expression of *DUSP5* has no impact on cell proliferation. Cell proliferation of HEK293 cells, which were transfected with empty vector or to transiently express *DUSP5*, was assessed by XTT assays with cells cultured either in 2% or 10% fetal calf serum (FCS) for 4 days (data normalized to each respective control; Mann-Whitney test; n=4; not significant; p=0.4857).

4. Discussion

4.1 Role of the tumor suppressor SASH1 in EMT and metastasis formation

Cancer is a major health problem worldwide. In the United States, for instance, cancer is the second leading cause of death (Siegel, Miller, and Jemal 2018). While locally restricted, benign tumors are rarely lethal, they potentially progress to highly metastatic, life-threatening diseases. During tumor progression, cancer cells acquire malignant traits that endow them with the capabilities to grow invasively, enter the systemic circulation and spawn lesions distant to the primary tumor. Given that the formation of metastases accounts for most cancer-related deaths, it is essential to study key players and molecular mechanisms underlying tumor progression and metastatic spread (Mehlen and Puisieux 2006; Gupta and Massague 2006). The scaffold protein SASH1 is a candidate tumor suppressor with clinical relevance in various solid tumor entities (Rimkus et al. 2006; Nitsche et al. 2012; Alcock et al. 2003; Zeller et al. 2003; Burgess, Bolderson, Saunus, et al. 2016; Zhou et al. 2018; Yang et al. 2015; Meng et al. 2013). Downregulation of *SASH1* has been associated with tumor progression and metachronous metastasis formation in colorectal cancer, as shown by previous work from our own group (Rimkus et al. 2006; Nitsche et al. 2012). Nevertheless, the underlying molecular and cellular mechanisms remained poorly understood, and a causal relationship between SASH1 and metastasis formation was lacking. Within the present thesis, the role of SASH1 as metastasis suppressor in colorectal cancer was verified experimentally by several orthogonal and independent approaches *in vitro*, and by an *in vivo* metastasis mouse model, which might be considered the most convincing and clinically relevant. In contrast to most studies that utilize forced expression of *SASH1* to study its function, a complete deficiency of SASH1 was introduced within the present thesis by the CRISPR/Cas9 system for the first time, to the best of our knowledge (Chen et al. 2012; Lin et al. 2012; Zong et al. 2016; He et al. 2016; Yang et al. 2012; Pan and Liu 2016; Sun et al. 2015). The underlying rational was to experimentally resemble the frequently downregulated or lost expression of *SASH1* in clinical cancer specimens (Zhou et al. 2018; Nitsche et al. 2012; Rimkus et al. 2006). As a central finding of the present thesis, SASH1-deficient HCT116 colorectal cancer cells generated significantly more metastases *in vivo* compared to parental cells, when they were orthotopically engrafted in in the caecum of immunodeficient mice. Further, the pattern of metastases was also changed dramatically. Parental cells primarily exhibited local dissemination, while SASH1-deficient cells generated metastases, which were completely

enclosed by host tissue, e.g. in the liver and pancreas. This implicates different routes for dissemination. Peritoneal carcinomatosis occurs due to the outgrowth of the primary tumor into the peritoneal cavity, from which tumor cells detach and spread locally, and is associated with a dismal prognosis (Sugarbaker 2014). This is in accordance with findings within this thesis, that distant lesions generated by parental cells were primarily attached to the outer surface of organs in the peritoneal cavity, e.g. to the kidney or intestine. In sharp contrast to this, disseminated SASH1-deficient cells were also found within hepatic capillaries, highlighting a hematogenous route for metastasis. This is especially important, since the formation of life-threatening distant metastases requires the presence of tumor cells within the systemic circulation (Gupta and Massague 2006). In summary, loss of SASH1 generated cells that disseminate with a higher efficiency and with a shift from local outgrowth to systemic spread. Therefore, SASH1 can indeed be considered as tumor suppressor counteracting metastasis formation. In accordance with this, expression of *SASH1* is more and more downregulated during colorectal cancer progression, which also correlates with metastasis formation in human patients (Rimkus et al. 2006). However, our recent data indicated that *SASH1* may be far more complex as previously anticipated, since increased *SASH1* expression was actually associated with poor distant metastasis-free survival in patients with locally advanced UICC stage III colorectal cancer (Sperlich et al. 2018). Of note, a similar observation was reported in HER2-positive and triple-negative breast cancer (Burgess, Bolderson, Saunus, et al. 2016). Nevertheless, it should be mentioned that *SASH1* expression is drastically downregulated in the vast majority of cases of colorectal cancer, including locally advanced UICC stage III (Rimkus et al. 2006; Sperlich et al. 2018). A major difference between these patient subgroups in breast or colorectal cancer is the presence or absence of cytotoxic drugs in the therapy regimes. Whereas stage II colorectal cancer patients do not receive any routine chemotherapy, stage III patients are routinely treated with adjuvant therapy. Thus, whereas *SASH1* expression can be considered a positive prognostic parameter due to its tumor suppressive roles, it is currently not clear whether it is also predictive for adjuvant or neoadjuvant therapy, since it may have additional roles in drug response. Furthermore, completely extinguished *SASH1* expression could have additional detrimental effects on tumor progression. Alternatively, the observed very low levels of *SASH1* expression might be due to an altered tumor stroma, which was part of the tissue samples

that were analyzed, and may feature low or absent expression of *SASH1*. Further experiments with clinical specimens and use of histological data are required to address this.

In contrast to primary tumors, which were surprisingly similar regarding their tissue architecture and growth pattern, SASH1-deficient colorectal cancer cells spawned distant metastases with an overall different histology compared to their parental counterpart. Disseminated parental cells grew as large, densely packed clusters, while SASH1-deficient cells localized to distant tissue parenchyma as extensively dispersed small clusters or even single isolated cells. Single cell invasion programs are typical for carcinoma cells that underwent EMT (Friedl and Wolf 2003). In accordance with that, loss of SASH1 was sufficient to induce EMT *in vitro*, as demonstrated by various markers, altered cell morphology and acquisition of a highly aggressive and invasive phenotype. The epithelial hallmark gene *CDH1* encodes for the cell-cell adhesion molecule E-cadherin, which is a central component of epithelial adherens junctions (Hartsock and Nelson 2008). Loss of E-cadherin in turn, a hallmark of EMT, is mediated by EMT promoting transcription factors, including ZEB1 and ZEB2 (Batlle et al. 2000; Comijn et al. 2001; Aigner et al. 2007). Expression of *ZEB1* and *ZEB2* was increased in SASH1-deficient cells, while E-cadherin levels and accordingly intercellular adhesions were diminished. Furthermore, expression of genes encoding mesenchymal markers was induced by loss of SASH1. Morphologically, SASH1-deficient cells were individualized and featured a fibroblast-like appearance. In contrast to that, parental cells exhibited a cobblestone-like morphology and grew as clusters of cells that were attached together by intercellular adhesions. This transition in cellular morphology and growth pattern is typical for cells that underwent EMT (Dongre and Weinberg 2019). In accordance, recombinant expression of *SASH1* in parental HCT116 cells led to a more pronounced epithelial morphology and counteracted EMT induced by the pro-inflammatory cytokine TNF, reported to trigger EMT (Bates and Mercurio 2003; Wang et al. 2013; Ricciardi et al. 2015). An antagonistic effect of recombinant expression of *SASH1* on EMT has also been described in various cell lines of other cancer entities (He et al. 2016; Zong et al. 2016; Pan and Liu 2016). Cytokine-mediated induction of EMT resulted in drastically reduced SASH1 levels in various cell lines, highlighting a reciprocal antagonism between SASH1 and EMT. Reduced SASH1 levels upon hypoxia- or TGF-β1-induced EMT have also been described independently in additional cancer cell lines (Zong et al. 2016; Pan and Liu 2016). Due to the apparent independence from the EMT-inducing signal, *SASH1* is likely downregulated by general EMT promoting transcription factors. In support of this, luciferase

reporter assays performed by our group revealed a negative regulation of the activity of the *SASH1* promoter by EMT (Franke et al. 2019).

The data highlight a crucial role of SASH1 as negative regulator of EMT. Accordingly, loss of SASH1 promoted several EMT-related phenotypes. Most importantly, SASH1-deficient cells exhibited increased cell motility and invasiveness. Cancer cell invasion is fundamentally required for the initiation and progression of the metastatic cascade (Gupta and Massague 2006; Valastyan and Weinberg 2011). Further, anchorage-independent survival was increased upon SASH1-deficiency. Avoidance of anoikis, i.e. programmed cell death upon detachment from the (physiologically correct) ECM, in turn is important during tumor cell circulation and seeding at the metastatic site with a foreign microenvironment (Taddei et al. 2012). Therefore, EMT mediated by loss of SASH1 leads to the acquisition of several malignant traits, which are required for an efficient metastatic cascade. This likely explains why SASH1-deficient cells were highly metastatic *in vivo*, according to the data presented here. Clinically, EMT has been implicated as a relevant transdifferentiation program that mediates tumor progression and metastasis formation in several carcinoma entities (Dongre and Weinberg 2019). In human colorectal cancer, high levels of ZEB2 at the invasive front have been shown to correlate with tumor progression and poor survival (Kahlert et al. 2011). Additionally, elevated levels of the EMT promoting transcription factor SLUG and the EMT marker N-cadherin were associated with tumor progression, formation of distant metastases and poor survival (Shioiri et al. 2006; Yan et al. 2015). In the present study, the EMT promoting transcription factor ZEB1 was identified as prognostic factor for patients with colorectal cancer (UICC stage III), whereby high expression of *ZEB1* was associated with poor distant recurrence-free survival. This has also been described independently (Guo et al. 2017). Further, patients with progressive colorectal cancer exhibited increased levels of EMT-like circulating tumor cells (Satelli et al. 2015). In summary, EMT is clinically associated with tumor progression and metastasis formation in colorectal cancer and various other carcinoma entities. This is in accordance with data presented here, since loss of SASH1 was sufficient to induce EMT, thereby generating invasive cancer cells with increased anchorage-independent survival, which efficiently spawn distant metastases via the hematogenous route. SASH1 thus acts as metastasis suppressor by counteracting EMT and EMT-associated pro-metastatic traits.

4.2 SASH1 counteracts EMT and metastasis formation by inhibition of CRKL-mediated SRC signaling

To investigate the molecular mechanism underlying the tumor suppressive properties of SASH1, but also to identify its putative physiological interaction partners, orthogonal screening approaches for protein-protein interaction partners of SASH1 were performed. Several proteins were detected by both yeast two-hybrid and co-immunoprecipitation/mass spectrometry screens, which were carried out in part by our cooperation partners (Dr. Ewa Ninio, Sorbonne University, Paris, France). Among these proteins, CRK family signal adaptors exhibited the highest number of manually curated EMT-related and PubMed-listed publications. The adaptor protein CRK was identified by yeast two-hybrid screens, but not in co-immunoprecipitations of GFP-SASH1 in HEK293 cells. In contrast to that, the closely related homolog CRKL was detected in both screenings. The focus was thus on CRKL, and the interaction between CRKL and SASH1 was successfully verified by immunoprecipitation of endogenous proteins, as well as protein deletion construct mapping and mutagenesis. The interaction between SASH1 and CRKL is mediated by a specific PXXPXK motif (residues 984-989) within the C-terminal region of SASH1 that binds to the N-terminal SH3 (SH3N) domain of CRKL with a dissociation constant in the low micromolar range (7.4 µM), which is typically observed for SH3 domain-mediated interactions (Nguyen et al. 2000). A contribution of the other PXXPXK motifs should not be completely ruled out, however, since a SASH1 deletion construct (residues 1-887) containing only the first motif (residues 865-870) associated with CRKL, albeit only weakly. Of note, the SH3N domains of CRK and CRKL exhibit highly similar ligand binding preferences (the aforementioned PXXPXK motif) (Bell and Park 2012). Therefore, lack of CRK binding could be due to the altered regulation of CRK compared to CRKL. In contrast to CRKL, the SH3N domain of CRK is negatively regulated by a phosphorylation-dependent autoinhibition (Jankowski et al. 2012). While co-precipitation of SASH1 and CRK was not detected within this thesis, CRK could not be ruled out as putative SASH1 interaction partner at the current time.

CRKL acts as signal adaptor by binding to phosphorylated, membrane-associated receptor tyrosine kinase and integrin complexes via the SH2 domain, thereby recruiting guanine nucleotide exchange factors via the SH3N domain. This way, membrane-associated small GTPase like RAP1, RAC1 and RAS are activated (Feller 2001; Bell and Park 2012). By physically associating with the SH3N domain of CRKL, SASH1 counteracted interactions between CRKL

and guanine nucleotide exchange factors. This highlights a potential mechanism of how SASH1 inhibits CRKL-mediated signaling and potentially EMT. Indeed, EMT induced by loss of SASH1 was entirely dependent on the presence of CRKL. Loss of CRKL completely rescued the more-mesenchymal phenotype of SASH1-deficient cells. Cell migration, invasiveness and anchorage-independent survival were drastically reduced. In accordance with that, CRKL has been shown to foster cell migration and invasiveness in various cancer cell lines (Li et al. 2003; Fu et al. 2015; Ji et al. 2016; Ungewiss et al. 2016; Cai, Wang, and Yang 2017; Guo, Zhao, et al. 2018). CRKL mediates activation of the kinases PI3K/AKT and ERK1/2, which play a rather pleiotropic role in cancer, but were repeatedly associated with cell invasiveness (Nosaka et al. 1999; Cheung et al. 2011; Lin et al. 2015; Cheng, Guo, Yang, and Han 2015; Zhang, Gao, et al. 2017). However, CRK proteins have also been described to mediate activation of the tyrosine kinase SRC (Sabe, Shoelson, and Hanafusa 1995; Watanabe et al. 2006; Watanabe et al. 2009; Ungewiss et al. 2016). Integrin-dependent SRC/FAK signaling represents a pathway crucially important for invasiveness and has also been linked to EMT (Brakebusch et al. 2002; Patel et al. 2016). Importantly, SASH1-deficient cells exhibited increased activating phosphorylation of SRC and its substrate paxillin. Activation of SRC is pivotal for the formation of focal adhesions, which also play a role in the metastatic cascade (Li, Okura, and Imamoto 2002; Barbazan et al. 2017). In accordance with the increased phosphorylation of SRC, loss of SASH1 resulted in pronounced formation of active, phospho-paxillin positive focal adhesions. Similar to the EMT phenotype, loss of CRKL reversed the increased SRC and paxillin phosphorylation observed in SASH1-deficient cells. Even further, pharmacological inhibition of SRC phenocopied CRKL-deficiency, as the EMT phenotype of cells lacking SASH1 was rescued.

Based on these data, the following working model was proposed. SASH1 directly binds and thereby inhibits the SH3N domain of CRKL, counteracting recruitment of guanine nucleotide exchange factors and subsequent activation of small GTPases. Loss of SASH1 relieves CRKL from inhibition, thus inducing CRKL-mediated activation of SRC and induction of EMT. Downstream of CRKL-mediated SRC activation and EMT, SASH1-deficient cells feature strongly enhanced cell motility, invasiveness, chemoresistance and anchorage-independent survival. These phenotypes in turn enable SASH1-deficient cells to overcome the manifold barriers within the metastatic cascade. SASH1-deficient cells are therefore highly metastatic and disseminate efficiently via the hematogenous route. This model is supported by independent large-scale clinical data from the public TCGA data base. Genomic losses of *SASH1* and/or

amplifications of *CRKL* (frequently accompanied by elevated mRNA expression) were primarily observed in primary tumors of colorectal cancer patients. These alterations were significantly associated with reduced overall survival. Strikingly, also in line with the proposed model, the highly metastatic phenotype of SASH1-deficient cells was entirely dependent on the presence of its novel interaction partner CRKL.

4.3 Role of the CRK family of signal adaptors in EMT, invasiveness and SRC/FAK signaling

The signal adaptor CRK has been described as inducer of EMT in several studies (Matsumoto et al. 2015; Elmansuri et al. 2016; Kumar et al. 2017). The data in this thesis indicated a similar role of the homolog CRKL, since the mesenchymal phenotype of SASH1-deficient cells was rescued by CRKL-deficiency. The focus was thus on a potential functional redundancy of CRKL and CRK, as well as on the relevant signaling pathways that mediate EMT downstream of CRK proteins. Loss of either CRKL or CRK resulted in a more pronounced epithelial phenotype in the colorectal cancer cell lines HCT116 and SW480, suggesting that CRKL and CRK have indeed overlapping functions in EMT. This can be explained by essentially identical ligand binding preferences of the SH2 and SH3N domains of CRK proteins (Feller 2001; Bell and Park 2012). However, loss of CRKL exhibited stronger effects on cell migration and invasiveness compared to loss of CRK. Due to the aforementioned domain organization differences of CRK and CRKL, interactions und thus signaling is modulated differently (Jankowski et al. 2012). Accordingly, CRKL but not CRK-II transforms fibroblasts, highlighting different biological activities (Matsuda et al. 1992). Li and colleagues found that *Crkl*-deficient murine embryonic fibroblasts exhibit severe haptotaxis defects, despite the presence of elevated Crk levels (Li et al. 2003). This suggests that CRKL has unique functions in integrin-induced cell migration, which is likely due to the different domain organization. This would also explain the additional effects on migration and invasion observed in this thesis.

Ungewiss and colleagues described that downregulation of *CRKL* by RNA interference reduced cancer cell invasiveness, but not EMT marker expression (Ungewiss et al. 2016). This is in obvious contrast to data from this thesis. Firstly, loss of either CRKL or its homolog CRK led to a pronounced epithelial phenotype in HCT116 and SW480 cells, also affecting EMT marker expression. Secondly, CRKL-deficiency rescued the EMT-like phenotypes of cells lacking its novel interaction partner SASH1. While SASH1-deficient cells featured diminished levels of the epithelial marker E-cadherin, the mesenchymal markers ZEB1 and vimentin were strongly

elevated. Loss of CRKL in SASH1-deficient cells in turn induced increased E-cadherin levels and reduced mesenchymal marker levels, phenocopying the parental line. Thirdly, complete deficiency of the CRK family resulted in the acquisition of a highly epithelial phenotype with drastically reduced EMT marker expression in colorectal and pancreatic cancer cells. Finally, cells deficient for CRKL and/or CRK were unable to undergo EMT either induced by CSK-deficiency or by cytokine administration. This demonstrates that CRKL and CRK are crucial positive regulators of EMT. Cell-line or tissue-specific differences, as well as different experimental approaches (shRNA-mediated downregulation of *CRKL* in lung cancer cells by Ungewiss and colleagues or complete deficiency of CRKL and/or CRK in the present thesis), might explain why Ungewiss and colleagues found no change in EMT marker expression upon downregulation of *CRKL* in lung cancer cell lines (Ungewiss et al. 2016). However, data from our own group indicated that loss of CRKL resulted in a pronounced epithelial phenotype also in A549 human non-small cell lung cancer cells, indicating a multi-tissue role of CRKL in EMT (Maximilian Ehrenfeld, Master thesis 2018, Technical University of Munich, Germany). Accordingly, the signal adaptor CRK has been implicated in EMT in various cancer entities, including bladder cancer, triple negative breast cancer and lung cancer (Matsumoto et al. 2015; Elmansuri et al. 2016; Kumar et al. 2017). Thus, both CRKL and CRK likely play a role in EMT in various solid cancer entities.

As consequence of this functional redundancy, compound deficiency of the complete CRK family resulted in synergistic effects. In contrast to cells lacking either CRKL or CRK alone, loss of the CRK family resulted in a drastically altered cellular morphology, with cells completely lacking lamellipodia formation and growing exclusively as large, densely packed clusters with pronounced E-cadherin staining at intercellular adhesions. Further, while single loss of CRKL or CRK only led to reduced migration and invasiveness, CRK family-deficient cells completely lacked collective and individual cell migration, even after prolonged incubation for several days. Additionally, invasiveness was completely absent, while anchorage-independent survival and chemoresistance were reduced. As discussed above, cancer cell invasion and anchorage-independent survival is critical to accomplish the metastatic cascade (Gupta and Massague 2006; Taddei et al. 2012). While *in vivo* data for cells lacking the complete CRK family are yet unavailable, CRKL-deficiency alone was already sufficient to fully reverse the highly metastatic phenotype of SASH1-deficient cells. Thus, it is tempting to speculate that metastasis formation is completely abolished by the synergistic effect of a complete CRK

family-deficiency. Nevertheless, orthotopic xenograft mouse models, as well as genetically engineered mouse models, would be required to investigate this topic further.

Due to the importance of SRC activation in CRKL-dependent EMT upon loss of SASH1, a detailed analysis of CRK proteins in SRC signaling was performed. In general, CRK proteins have been implicated in signaling downstream of integrins and receptor tyrosine kinases by recruitment of guanine nucleotide exchange factors and subsequent activation of small GTPases (Feller 2001; Bell and Park 2012). Interestingly, data from the present thesis demonstrated that CRK proteins also mediate activation of the kinases SRC and FAK, which were previously thought to act mainly upstream in the pathway. Activation of these upstream kinases by CRK proteins occurred in a manner dependent on small GTPases. As molecular mechanism of how CRK proteins mediate EMT and metastasis formation, the following working model is proposed (Figure 46). Upon initial engagement of integrins by the ECM, scaffolds cluster at the cytoplasmic site of the plasma membrane (Liu, Calderwood, and Ginsberg 2000). Upon clustering, focal adhesion kinase (FAK) is activated by autophosphorylation at Y397, which is then bound by the SH2 domain of SRC (Schaller et al. 1994; Toutant et al. 2002). Once the SRC/FAK complex is formed, SRC phosphorylates FAK at several sites, including Y576/577 and Y925, which leads to activation of the SRC/FAK kinase complex (Calalb, Polte, and Hanks 1995; Schlaepfer and Hunter 1996; Ciccimaro, Hevko, and Blair 2006). Next, the active SRC/FAK complex phosphorylates the integrin-associated scaffolds p130CAS and paxillin (Figure 46A). At this point, the signal adaptors CRK and CRKL come into play by binding via their SH2 domains to phosphotyrosine sites of p130CAS and paxillin (Birge et al. 1993; Sakai et al. 1994; Salgia et al. 1995; Schaller and Parsons 1995; Salgia et al. 1996; Lamorte et al. 2003; Li et al. 2003; Mitra, Hanson, and Schlaepfer 2005; Mitra and Schlaepfer 2006). The SH3N domains of CRK family proteins thereby recruit guanine nucleotide exchange factors like SOS1, C3G and DOCK180 into close proximity to the membrane, which then induce local activation of small GTPases like RAS, RAC1 and RAP1 (Figure 46B) (Sakkab et al. 2000; Sakakibara 2002; Cheung et al. 2011; Bell and Park 2012). This way, CRK adaptor proteins are crucial for downstream signaling. However, the small GTPase RAP1 has been described to mediate integrin activation by talin recruitment (Arai et al. 1999; Han et al. 2006; Lagarrigue, Kim, and Ginsberg 2016). Therefore, downstream of the initial integrin engagement event, further integrin molecules are activated in a manner dependent on CRK family proteins. Locally activated integrins efficiently bind to the ECM,

which promotes clustering of integrins and integrin-associated scaffold proteins. This induces further activation of SRC and FAK, closing the positive feedback circuit by providing additional bindings sites for CRK proteins (Figure 46C). Therefore, CRK family proteins mediate a novel small GTPase-dependent positive feedback loop, which induces formation of focal adhesions and sustained activation of SRC and FAK.

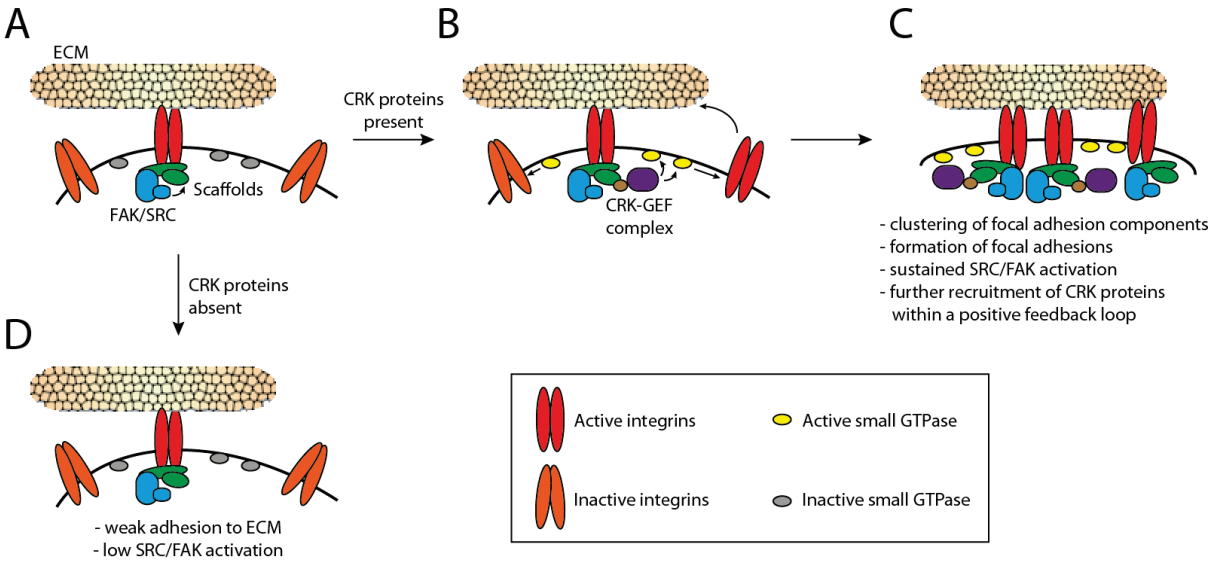

Figure 46: CRK family proteins act as central components in a small GTPase-dependent positive feedback loop to promote clustering of integrin-associated scaffolds, focal adhesion formation and sustained SRC/FAK activation. (A) Initial engagement of integrins by the extracellular matrix (ECM) leads to recruitment and activation of the SRC/FAK kinase complex. **(B)** CRK proteins are recruited to scaffold proteins, which were phosphorylated by the SRC/FAK complex, and mediate local activation of small GTPases like RAP1. RAP1 in turn induces integrin activation by talin recruitment. **(C)** This leads to local clustering of active integrin complexes, focal adhesion formation, sustained activation of SRC/FAK kinases, and further recruitment of CRK proteins within a positive feedback loop. **(D)** In the absence of CRK family proteins, the positive feedback loop is lacking, allowing only weak cell-matrix adhesions and low activation of the SRC/FAK complex (red/orange = active/inactive integrins; green = scaffolds p130CAS and paxillin; blue = kinases FAK and SRC; grey = inactive small GTPases; brown = CRK and CRKL; purple = guanine nucleotide exchange factors; yellow = activated small GTPases).

In accordance with this model, pharmacological inhibition of RAP1 reversed the effects of recombinant *CRKL* expression on SRC/FAK substrate phosphorylation. RAP1 inhibition also resulted in a diffuse localization of talin and a more pronounced epithelial marker expression. The compound GGTI-298, which was used here to inhibit RAP1, is a geranylgeranyltransferase I inhibitor that counteracts RAP1 activation (McSherry et al. 2011; Lilja et al. 2017; Zhang, Wang, et al. 2017; Molina-Ortiz et al. 2014; Yang et al. 2017). Despite GGTI-298 being described as specific RAP1 inhibitor, the results should be taken with caution, since other small GTPases might potentially be inhibited as well. Nevertheless, CRK proteins clearly mediate a small GTPase-dependent positive feedback loop. Besides RAP1, small GTPases such as RAC1 likely also play a role. Further research, e.g. with RAP1 targeted by RNA interference, will

reveal if this process is fully dependent on RAP1 or potentially also on other small GTPases that are activated by CRK proteins. Nevertheless, silencing of RAP1 has been shown to reduce autophosphorylation of FAK (Carmona et al. 2009). Further, RAP1 has been described to mediate clustering of integrins (Katagiri et al. 2003; Eppler, Quast, and Kolanus 2017). These independent data are in support of the proposed RAP1-dependent mechanism.

When CRK proteins are lacking, initial engagement of integrins still occurs according to the model. This explains the not completely abolished phosphorylation of integrin-associated signaling proteins in CRK family-deficient cells, as well as their ability to adhere to fibronectin, albeit only weakly. However, activation of the proposed small GTPase-dependent positive feedback loop by CRK proteins is lacking. Accordingly, clustering of focal adhesion components like paxillin and talin, and potentially of integrins was diminished in CRK family-deficient cells. Due to impaired clustering, the initial event of FAK autophosphorylation was also drastically reduced, as well as activation of SRC and SRC/FAK substrate phosphorylation (Figure 46D). Even artificial activation of integrins poorly increased adhesion of CRK family-deficient cells, presumably due to this clustering defect. The model also explains why loss of CSK, which resulted in massive SRC activation in both parental and CRK family-deficient cells, poorly rescued the SRC/FAK signaling defects. Even when SRC is artificially activated, CRK family proteins are still required to mediate local clustering of focal adhesion components, and thus activation of FAK (by autophosphorylation), as well as formation of the active SRC/FAK complex. Loss of CSK in CRK-proficient cells in turn induced strong phosphorylation of FAK and SRC/FAK substrates. Therefore, specifically in the absence of CRK family proteins, activation of FAK is at least partially uncoupled from SRC activity. It has been proposed that CRK proteins mediate SRC activation by binding its negative regulatory kinase CSK (Sabe, Shoelson, and Hanafusa 1995; Watanabe et al. 2009; Kumar et al. 2018). However, the failure of CSK-deficiency to rescue SRC/FAK signaling defects of cells lacking CRK family proteins demonstrates that the mechanism is not fully dependent on CSK. While a CSK-dependent mechanism for SRC activation by CRK proteins could occur independently, the data from this thesis indicated that CRK proteins mainly mediate the spatiotemporal clustering of integrin-associated scaffolds to induce sustained SRC/FAK activation via positive feedback signaling.

Loss of CRK family proteins generated highly epithelial cells that completely lack migration, invasiveness and the ability to undergo EMT. This can be explained by the crucial role of CRK proteins as central positive feedback inducers of SRC/FAK signaling. SRC and FAK cooperate to promote cell migration, invasion and metastasis formation via various signaling pathways and mechanisms, including the aforementioned activation of small GTPases via CRK family proteins (Summy and Gallick 2003; Frame 2004; Mitra, Hanson, and Schlaepfer 2005). The importance of SRC family kinases in EMT is highlighted by the observation that loss of CSK, a negative regulator of SRC, was sufficient to induce EMT in parental cells. Further, inhibition of SRC phenocopied CRKL-deficiency, and reversed the mesenchymal phenotype of SASH1-deficient cells. Besides inducing various oncogenic signaling pathways, SRC disrupts E-cadherin-mediated adherens junctions, thereby shifting cell-cell adhesion programs to cell-matrix adhesion and ultimately promoting EMT (Behrens et al. 1993; Frame 2004). Strikingly, integrin signaling and phosphorylation of FAK are required for SRC-induced disruption of E-cadherin-mediated cell-cell junctions (Avizienyte et al. 2002). Therefore, by uncoupling FAK activation from SRC activity, loss of CRK family proteins renders cells unable to undergo EMT induced by loss of CSK or cytokines. In summary, the required cooperation of FAK and SRC to induce EMT, migration and invasion explains why loss of CSK did not rescue the phenotypes of CRK family-deficient cells. Clinically, elevated specific activity of SRC has been shown to be associated with poor survival in colon carcinoma (Aligayer et al. 2002). Genetic alterations at codon 531 of SRC have been described as a potential mechanism of activation in human colorectal cancer, but this is still highly controversially discussed (Irby et al. 1999; Daigo et al. 1999; Laghi et al. 2001; Aligayer et al. 2002). Increased levels of CRK family proteins could thus represent a novel mechanism of SRC activation in human colorectal cancer. This could be investigated by assessing expression levels of CRK family genes, as well as the specific SRC activity, in clinical specimens.

4.4 Revised model of SASH1, CRK proteins and DUSP5 in tumor progression

As described above, downstream of initial integrin engagement by the ECM, integrin-associated scaffolds are clustered, leading to activation of FAK and SRC, which form the active SRC/FAK complex. This complex phosphorylates the scaffolds p130CAS and paxillin. Here, CRK proteins play a crucial role by inducing a positive feedback loop likely via the small GTPase RAP1 and subsequent local activation of integrins. This allows clustering of integrin-associated complexes, formation of focal contacts/adhesions and stable cell-ECM adhesion, thus promoting sustained activation of SRC and FAK. Downstream of this increased SRC/FAK activation, further binding sites for CRK proteins are generated to induce more and more positive feedback signaling. This signaling promotes EMT, which in turn induces transcriptional downregulation of *SASH1*. This provides a second feedback loop, since SASH1 is a negative regulator of CRKL. In accordance with this model, loss of SASH1 relieves CRKL from inhibition, inducing SRC/FAK positive feedback signaling and EMT, which confer the observed highly aggressive, pro-metastatic phenotype. Loss of CSK as negative regulator of SRC in turn enhances recruitment of CRK proteins to integrin-associated scaffolds, thereby inducing EMT. For this reason, CRK family proteins are required for EMT induction and cancer cell aggressiveness induced by loss of CSK or SASH1 (Figure 47). This is also reflected clinically, as expression of *CRK* family genes directly correlated with expression of the EMT promoting transcription factor *ZEB1* in locally advanced human colorectal cancer. Both proteins were especially enriched in de-differentiated, invasive cancer cells within the primary tumor, supporting their involvement in EMT and invasiveness.

However, the phenotype of CRK family-deficient cells was not limited to regulation of EMT, cell migration and adhesion. Unbiased analysis of the transcriptome by RNAseq-approach revealed several key processes that were significantly regulated by the CRK family, as indicated by gene set enrichment analysis (GSEA). Of note, these alterations predicted *in silico* could be verified experimentally, such as reduced cell proliferation and reduced phosphorylation of the mitogen-activated protein kinases (MAPKs) ERK1/2, as well as decreased c-MYC protein levels and transcriptome-wide MYC target gene expression in CRK family-deficient cells. CRK proteins have been described to activate the small GTPase RAS via recruitment of the guanine nucleotide exchange factor SOS1 (Cheung et al. 2011; Bell and Park 2012). Following RAS activation, a kinase cascade is activated that culminates in phosphorylation and activation of ERK1/2, which phosphorylates various targets to promote

proliferation and other potentially malignant traits (Yoon and Seger 2006; Shaul and Seger 2007). A prominent target of ERK1/2 is the transcription factor c-MYC, which is stabilized upon ERK1/2-dependent phosphorylation (Junttila and Westermarck 2008). The proto-oncogene c-MYC acts as transcription factor, thereby regulating various processes, including E2F transcription factors and cell cycle (Chen, Liu, and Qing 2018). Accordingly, not only MYC target genes, but also E2F targets and cell cycle-related genes were drastically reduced upon CRK family-deficiency (GSEA). While direct evidence of an ERK1/2-dependent mechanism for the increased c-MYC levels is still lacking, it is tempting to speculate that CRK family proteins promote proliferation via this pathway (Figure 47).

Besides SASH1 and CSK, which primarily regulate SRC/FAK signaling, the nuclear dual-specificity phosphatase DUSP5 is a negative feedback regulator of ERK1/2 activation (Figure 47) (Mandl, Slack, and Keyse 2005; Keyse 2008). Furthermore, previous unpublished data from our group indicate that DUSP5 has a tumor- and metastasis-suppressive role in colorectal cancer (Dr. Klaus-Peter Janssen, Fabian Christoph Franke and Dr. Miguel Abal, unpublished data). This was also analyzed and partially verified in the present thesis, since forced expression of the phosphatase in HEK293 cells counteracted serum-induced ERK1/2 phosphorylation. Further, exchange of the catalytic cysteine to serine (C263S) abrogated this effect of DUSP5, which is in accordance with data from others (Talipov et al. 2016). However, the effect of DUSP5 on ERK1/2 phosphorylation was rather small, which also explains why forced expression of *DUSP5* did not affect cell proliferation. Tögel and colleagues also found that DUSP5 fails to regulate ERK1/2 phosphorylation, as well as intestinal cell proliferation and tumorigenesis (Togel et al. 2018). DUSP5 has been shown to anchor and inactivate ERK2 in the nucleus (Mandl, Slack, and Keyse 2005). Surprisingly, while ERK is inactivated in the nucleus, cytoplasmic ERK activation has recently been shown to be increased and prolonged by DUSP5 due to reduced negative feedback signaling to RAF and MEK (Kidger et al. 2017). Instead of counteracting total ERK1/2 phosphorylation, DUSP5 therefore rather regulates ERK1/2 in a spatial manner. This way, DUSP5 may even exert oncogenic functions. For instance, increased DUSP5 levels drive $BRAF^{V600E}$-induced proliferation and transformation due to relieved negative feedback signaling and prevention of senescence (Kidger et al. 2017).

Increased ERK1/2 phosphorylation, c-MYC protein levels and MYC target gene expression likely contribute profoundly to cell proliferation, and thus to the highly aggressive, pro-metastatic and chemoresistant phenotype downstream of CRK family proteins

(Figure 47). The clinical data of *CRK* family expression in locally advanced colorectal cancer highlights these oncogenic functions of CRK family adaptors. High intratumoral expression of *CRKL* was associated with a poor survival of patients who received 5-fluorouracil as adjuvant chemotherapy. Accordingly, the chemoresistant *in vitro* phenotype of SASH1-deficient cells was fully rescued by loss of CRKL. Further, elevated expression of *CRK*, *CRKL* and *CRK* family were significantly associated with reduced distant recurrence-free survival in patients.

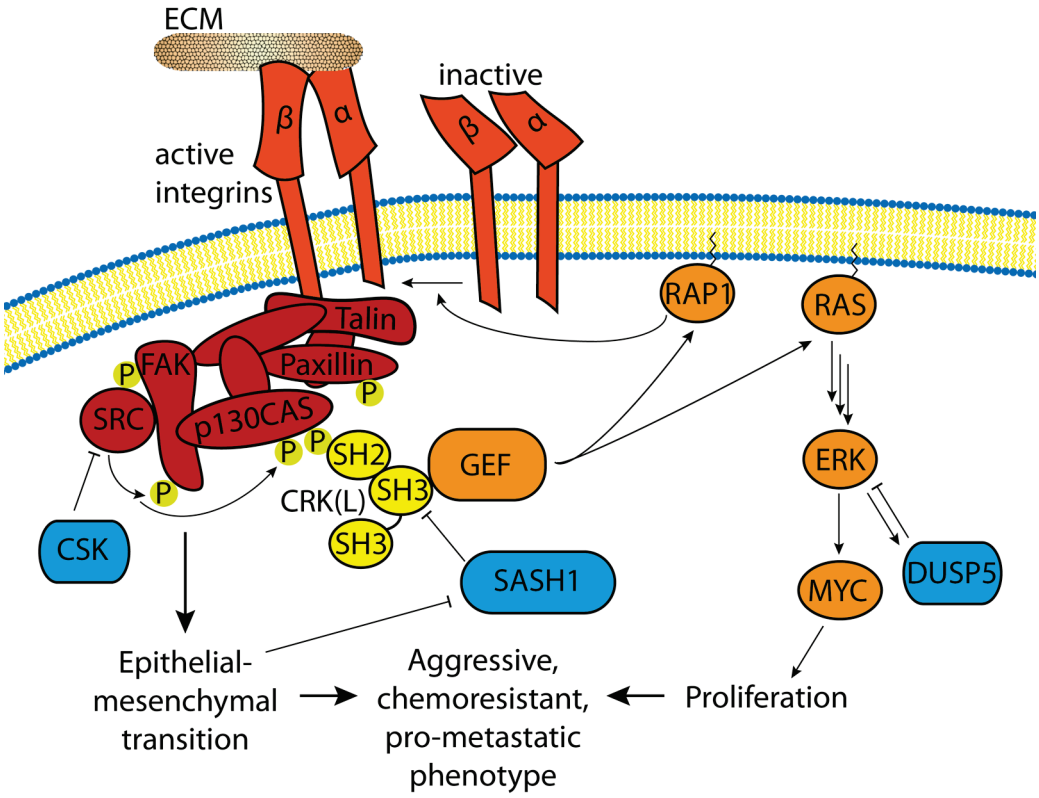

Figure 47: Revised, final working model of SASH1, DUSP5 and CRK proteins in cancer aggressiveness. Upon initial engagement of integrins, scaffold proteins cluster inside the cell, leading to autophosphorylation of FAK and subsequent activation of the SRC/FAK kinase complex. The scaffold proteins p130CAS and paxillin are phosphorylated by SRC/FAK kinases, providing binding sites for CRK family proteins. This way, CRK proteins mediate the recruitment of GEFs into close proximity to the membrane, inducing local activation of small GTPases like RAP1 or RAS. While activated RAS leads to activation of ERK1/2 and potentially increased levels of c-MYC, the small GTPase RAP1 mediates local integrin activation via talin recruitment. This represents a positive feedback loop, since SRC/FAK kinases are activated downstream of integrin signaling, providing further binding sites for CRK family proteins. Sustained activation of SRC/FAK kinases in turn induces EMT, and together with MYC a highly aggressive, pro-metastatic and chemoresistant phenotype. Negative regulators of these pathways include CSK, SASH1 and DUSP5. As a second feedback mechanism, *SASH1* is transcriptionally downregulated upon EMT induction, leading to increased binding of GEFs by CRK family proteins (ECM = extracellular matrix; P = phosphorylation; GEF = guanine nucleotide exchange factor).

4.5 Therapeutic targeting of CRK family proteins to counteract EMT and metastasis formation

The central role of CRK and CRKL in SRC/FAK signaling and ERK activation highlights them as promising targets for anticancer therapy. Inhibition of CRK proteins would have the potential to overcome EMT and malignant phenotypes such as proliferation, invasiveness and chemoresistance. Given that CRK family proteins exerted profound effects not only in colorectal but also in pancreatic cancer cells, potential CRK family inhibitors could be beneficial in various cancer entities. However, in contrast to kinases and other "druggable" proteins, which can be targeted by small molecular compounds specifically interacting with their ligand binding pockets, inhibitors of protein-protein interactions are required to counteract CRK family function. There are several major difficulties in developing protein-protein interaction inhibitors: protein-protein interactions are mediated by large areas, deep pockets are lacking, the presence of non-continuous binding sites and the general lack of natural ligands (Ivanov, Khuri, and Fu 2013). Nevertheless, targeting protein-protein interactions is feasible, e.g. as demonstrated by counteracting the MDM2/p53 interaction by peptidomimetics or nonpeptidic small compounds (Khoury and Domling 2012; Ivanov, Khuri, and Fu 2013).

Forced expression of $CRKL^{W160R}$ further decreased SRC/FAK signaling compared to control transfected cells, suggesting a dominant negative effect of CRKL with inactive SH3N domain. This might be due to the SH2 domain, which remains active in the $CRKL^{W160R}$ mutant, binding and blocking upstream phosphotyrosine residues, thereby downregulating SRC/FAK signaling. Likely due to the reduced phosphorylation of paxillin and p130CAS, the SH2 domain of $CRKL^{W160R}$ also failed to co-precipitate these proteins. Due to the dominant-negative effect, the SH3N domain of CRKL was selected as target structure for developing protein-protein interaction inhibitors. Of note, the SH3N domains of CRKL and CRK have a highly similar structure and share the same ligand binding preferences (Feller 2001; Bell and Park 2012). Therefore, potentially identified small compounds likely target both CRK and CRKL.

Unfortunately, screenings for *in silico* pre-selected small compound binders of this domain remained unsuccessful, despite of an optimized assay with excellent performance (Z-factor = 0.8) and specific binding of a high affinity peptide as positive control (Zhang, Chung, and Oldenburg 1999). The failure of these assays to identify SH3N binding lead compounds

highlights the difficulty of targeting protein-protein interactions. The compound NSC-603169 (PubChem CID: 6326622), which was identified by *in silico* docking experiments to bind both CRKL and CRK SH3N domains, was not available at the time of the actual experiments and could thus not be tested. This compound is inactive in NCI human tumor cell line growth inhibition assays (PubChem) and has an elongated structure with a central hydrophobic region flanked by positively charged moieties. This could potentially mimic the PXXPXK motif as CRK family SH3N ligands. Further research is necessary to investigate this and test the compound for CRK family inhibition.

Inhibition of SH3 domains via small compounds has already been successfully demonstrated. For instance, the SRC SH3 domain has been shown to be inhibited by a benzoquinoline derivate (Atatreh et al. 2008). Others reported the ability of 2-aminoquinoline derivates to inhibit the Tec SH3 domain (Inglis et al. 2004; Inglis et al. 2005). Based on these structures, inhibitors of the N-terminal SH3 domains of CRK proteins could be developed in a rational drug discovery approach. Further, non-natural peptidomimetics could be employed to mimic the PXXP-containing ligands in future research (Nguyen et al. 2000).

Peptides that target the SH3N domain of CRK proteins have already been described (Posern et al. 1998; Kardinal et al. 2000). Indeed, GFP with such a SH3N domain binding high affinity peptide (HAP) fused to the C-terminus co-precipitated both CRK-II and CRKL, highlighting that both proteins could likely inhibited by a single small compound. This is highly important due the redundant functions of CRKL and CRK in EMT. Further, transient expression of GFP-HAP fusion proteins efficiently inhibited CRK protein-mediated SRC/FAK signaling. This indicated that transient targeting of CRK family-mediated signaling is feasible. Additionally, HAP was delivered across the membrane via coupling to the HIV1 TAT_{47-57} cell-penetrating peptide. Nevertheless, the efficiency was low and the peptide was primarily taken up via endocytosis, as evidenced by a spot like, vesicular intracellular localization. Cell penetrating peptides are already in various clinical trials, and might be additional tools to study and develop CRK family inhibitors (Guidotti, Brambilla, and Rossi 2017). Further emphasis should thus be put on the optimization of intracellular delivery of HAP by cell penetrating peptides, e.g. by the use of modified variants such as cyclic TAT peptides or incorporation of D-arginine residues (Ma et al. 2012; Nischan et al. 2015).

4.6 Conclusion and outlook

This thesis aimed to identify a causal relationship between the putative tumor suppressor SASH1 and metastasis formation in the context of colorectal cancer, as well as to investigate the underlying molecular mechanisms.

The role of SASH1 as tumor suppressor could be successfully verified, as loss of SASH1 induced distant metastases *in vivo* in a xenograft mouse model for colorectal cancer. Mechanistically, SASH1 was found to physically associate with the signal adaptor and proto-oncogene CRKL, thereby counteracting CRKL-mediated SRC signaling and EMT. Accordingly, the highly metastatic and aggressive phenotype of SASH1-deficient cells was rescued by genetic loss of CRKL or pharmacological inhibition of SRC. In addition to CRKL, the close homolog CRK was also shown to be a positive regulator of EMT, indicating functional redundancy. CRK family adaptors were demonstrated to act as central positive feedback inducers of SRC and FAK kinases, which cooperate to induce EMT. This positive feedback loop involves a novel small GTPase-dependent mechanism. Further, a hitherto unknown role of CRK proteins as positive regulators of c-MYC was identified. CRK family-deficient cells lacked migration, invasiveness and the ability to undergo EMT. Cell proliferation, survival and chemoresistance were also reduced, highlighting CRK proteins as promising therapeutic targets. This was also supported by clinical findings, since elevated expression of *CRK* family genes were associated with poor survival. While initial screenings for small compound inhibitors of the N-terminal SH3 domains of CRK proteins failed, transient inhibition of CRK proteins was accomplished using recombinant proteins and cell-penetrating peptides.

Besides the identification of small compounds, peptide-based or peptidomimetic inhibitors should be employed to target the CRK family. While the mechanism of CRK proteins in SRC/FAK signaling and EMT could be clarified, their role as positive regulators of c-MYC needs to be addressed in further studies. Recently, CRK was shown to promote immune evasion (Kumar et al. 2017). Given the importance of immunotherapy for future cancer treatment, the role of the CRK family in immune escape should also be addressed. In the present thesis, CRKL was identified as novel interaction partner of SASH1 by orthogonal approaches. However, SASH1 contains multiple protein-protein interactions domains, and SASH1 likely exerts functions in addition to inhibition of CRKL and potentially also CRK. It is also conceivable that SASH1 is recruited to integrin- or receptor tyrosine kinase-associated complexes by

association with CRKL, where it modulates downstream signaling pathways by yet unknown mechanisms. This could be investigated with special fluorescence microscopy techniques or analysis of SASH1-containing protein complexes, e.g. by co-immunoprecipitation of intracellularly crosslinked proteins. Further research should also focus on the function of the protein-protein interaction domains of SASH1. This could be performed by specifically identifying association partners of isolated domains by mass spectrometry. Results from the yeast two-hybrid screening in addition to the co-immunoprecipitation/mass spectrometry data described in this thesis should also provide valuable input for further research.

Besides the functional role of SASH1 as tumor suppressor, the present thesis highlighted CRK family proteins as promising new targets to counteract EMT, invasiveness, proliferation, chemoresistance, metastasis formation, and minimal residual disease. These findings might allow the development of future strategies of therapeutic intervention, e.g. by small molecule compounds, which could help to control metastatic spread and minimal residual disease in cancer patients, especially in combination with more conventional multi-modal approaches.

5. References

Acevedo, C. M., M. Henriquez, M. R. Emmert-Buck, and R. F. Chuaqui. 2002. 'Loss of heterozygosity on chromosome arms 3p and 6q in microdissected adenocarcinomas of the uterine cervix and adenocarcinoma in situ', *Cancer*, 94: 793-802.

Aigner, K., B. Dampier, L. Descovich, M. Mikula, A. Sultan, M. Schreiber, W. Mikulits, T. Brabletz, D. Strand, P. Obrist, W. Sommergruber, N. Schweifer, A. Wernitznig, H. Beug, R. Foisner, and A. Eger. 2007. 'The transcription factor ZEB1 (deltaEF1) promotes tumour cell dedifferentiation by repressing master regulators of epithelial polarity', *Oncogene*, 26: 6979-88.

Akagi, T., K. Murata, T. Shishido, and H. Hanafusa. 2002. 'v-Crk activates the phosphoinositide 3-kinase/AKT pathway by utilizing focal adhesion kinase and H-Ras', *Mol Cell Biol*, 22: 7015-23.

Akalay, I., B. Janji, M. Hasmim, M. Z. Noman, F. Andre, P. De Cremoux, P. Bertheau, C. Badoual, P. Vielh, A. K. Larsen, M. Sabbah, T. Z. Tan, J. H. Keira, N. T. Hung, J. P. Thiery, F. Mami-Chouaib, and S. Chouaib. 2013. 'Epithelial-to-mesenchymal transition and autophagy induction in breast carcinoma promote escape from T-cell-mediated lysis', *Cancer Res*, 73: 2418-27.

Al-Mehdi, A. B., K. Tozawa, A. B. Fisher, L. Shientag, A. Lee, and R. J. Muschel. 2000. 'Intravascular origin of metastasis from the proliferation of endothelium-attached tumor cells: a new model for metastasis', *Nat Med*, 6: 100-2.

Alcock, H. E., T. J. Stephenson, J. A. Royds, and D. W. Hammond. 2003. 'Analysis of colorectal tumor progression by microdissection and comparative genomic hybridization', *Genes Chromosomes Cancer*, 37: 369-80.

Aligayer, H., D. D. Boyd, M. M. Heiss, E. K. Abdalla, S. A. Curley, and G. E. Gallick. 2002. 'Activation of Src kinase in primary colorectal carcinoma: an indicator of poor clinical prognosis', *Cancer*, 94: 344-51.

Antoku, S., and B. J. Mayer. 2009. 'Distinct roles for Crk adaptor isoforms in actin reorganization induced by extracellular signals', *J Cell Sci*, 122: 4228-38.

Arai, A., Y. Nosaka, E. Kanda, K. Yamamoto, N. Miyasaka, and O. Miura. 2001. 'Rap1 is activated by erythropoietin or interleukin-3 and is involved in regulation of beta1 integrin-mediated hematopoietic cell adhesion', *J Biol Chem*, 276: 10453-62.

Arai, A., Y. Nosaka, H. Kohsaka, N. Miyasaka, and O. Miura. 1999. 'CrkL activates integrin-mediated hematopoietic cell adhesion through the guanine nucleotide exchange factor C3G', *Blood*, 93: 3713-22.

Atatreh, N., C. Stojkoski, P. Smith, G. W. Booker, C. Dive, A. D. Frenkel, S. Freeman, and R. A. Bryce. 2008. 'In silico screening and biological evaluation of inhibitors of Src-SH3 domain interaction with a proline-rich ligand', *Bioorg Med Chem Lett*, 18: 1217-22.

Avizienyte, E., A. W. Wyke, R. J. Jones, G. W. McLean, M. A. Westhoff, V. G. Brunton, and M. C. Frame. 2002. 'Src-induced de-regulation of E-cadherin in colon cancer cells requires integrin signalling', *Nat Cell Biol*, 4: 632-8.

Bao, W., H. Qiu, T. Yang, X. Luo, H. Zhang, and X. Wan. 2013. 'Upregulation of TrkB promotes epithelial-mesenchymal transition and anoikis resistance in endometrial carcinoma', *PLoS One*, 8: e70616.

Barbazan, J., L. Alonso-Alconada, N. Elkhatib, S. Geraldo, V. Gurchenkov, A. Glentis, G. van Niel, R. Palmulli, B. Fernandez, P. Viano, T. Garcia-Caballero, R. Lopez-Lopez, M. Abal, and D. M. Vignjevic. 2017. 'Liver Metastasis Is Facilitated by the Adherence of Circulating Tumor Cells to Vascular Fibronectin Deposits', *Cancer Res*, 77: 3431-41.

Barghorn, A., E. J. Speel, B. Farspour, P. Saremaslani, S. Schmid, A. Perren, J. Roth, P. U. Heitz, and P. Komminoth. 2001. 'Putative tumor suppressor loci at 6q22 and 6q23-q24 are involved in the malignant progression of sporadic endocrine pancreatic tumors', *Am J Pathol*, 158: 1903-11.

Bates, R. C., and A. M. Mercurio. 2003. 'Tumor necrosis factor-alpha stimulates the epithelial-to-mesenchymal transition of human colonic organoids', *Mol Biol Cell*, 14: 1790-800.

Batlle, E., E. Sancho, C. Franci, D. Dominguez, M. Monfar, J. Baulida, and A. Garcia De Herreros. 2000. 'The transcription factor snail is a repressor of E-cadherin gene expression in epithelial tumour cells', *Nat Cell Biol*, 2: 84-9.

Beer, S., A. B. Simins, A. Schuster, and B. Holzmann. 2001. 'Molecular cloning and characterization of a novel SH3 protein (SLY) preferentially expressed in lymphoid cells', *Biochim Biophys Acta*, 1520: 89-93.

Behrens, J., L. Vakaet, R. Friis, E. Winterhager, F. Van Roy, M. M. Mareel, and W. Birchmeier. 1993. 'Loss of epithelial differentiation and gain of invasiveness correlates with tyrosine phosphorylation of the E-cadherin/beta-catenin complex in cells transformed with a temperature-sensitive v-SRC gene', *J Cell Biol*, 120: 757-66.

Bell, E. S., and M. Park. 2012. 'Models of crk adaptor proteins in cancer', *Genes Cancer*, 3: 341-52.

Birge, R. B., J. E. Fajardo, C. Reichman, S. E. Shoelson, Z. Songyang, L. C. Cantley, and H. Hanafusa. 1993. 'Identification and characterization of a high-affinity interaction between v-Crk and tyrosine-phosphorylated paxillin in CT10-transformed fibroblasts', *Mol Cell Biol*, 13: 4648-56.

Birge, R. B., C. Kalodimos, F. Inagaki, and S. Tanaka. 2009. 'Crk and CrkL adaptor proteins: networks for physiological and pathological signaling', *Cell Commun Signal*, 7: 13.

Bockhorn, M., R. K. Jain, and L. L. Munn. 2007. 'Active versus passive mechanisms in metastasis: do cancer cells crawl into vessels, or are they pushed?', *Lancet Oncol*, 8: 444-8.

Bos, J. L. 1989. 'ras oncogenes in human cancer: a review', *Cancer Res*, 49: 4682-9.

Bos, J. L., K. de Bruyn, J. Enserink, B. Kuiperij, S. Rangarajan, H. Rehmann, J. Riedl, J. de Rooij, F. van Mansfeld, and F. Zwartkruis. 2003. 'The role of Rap1 in integrin-mediated cell adhesion', *Biochem Soc Trans*, 31: 83-6.

Bos, J. L., E. R. Fearon, S. R. Hamilton, M. Verlaan-de Vries, J. H. van Boom, A. J. van der Eb, and B. Vogelstein. 1987. 'Prevalence of ras gene mutations in human colorectal cancers', *Nature*, 327: 293-7.

Bourcy, M., M. Suarez-Carmona, J. Lambert, M. E. Francart, H. Schroeder, C. Delierneux, N. Skrypek, E. W. Thompson, G. Jerusalem, G. Berx, M. Thiry, S. Blacher, B. G. Hollier, A. Noel, C. Oury, M. Polette, and C. Gilles. 2016. 'Tissue Factor Induced by Epithelial-Mesenchymal Transition Triggers a Procoagulant State That Drives Metastasis of Circulating Tumor Cells', *Cancer Res*, 76: 4270-82.

Brabletz, T., A. Jung, S. Reu, M. Porzner, F. Hlubek, L. A. Kunz-Schughart, R. Knuechel, and T. Kirchner. 2001. 'Variable beta-catenin expression in colorectal cancers indicates tumor progression driven by the tumor environment', *Proc Natl Acad Sci U S A*, 98: 10356-61.

Brady, C. A., and L. D. Attardi. 2010. 'p53 at a glance', *J Cell Sci*, 123: 2527-32.

Brakebusch, C., D. Bouvard, F. Stanchi, T. Sakai, and R. Fassler. 2002. 'Integrins in invasive growth', *J Clin Invest*, 109: 999-1006.

Buchert, M., F. Rohde, M. Eissmann, N. Tebbutt, B. Williams, C. W. Tan, A. Owen, Y. Hirokawa, A. Gnann, G. Orend, G. Orner, R. H. Dashwood, J. K. Heath, M. Ernst, and K. P. Janssen. 2015. 'A hypermorphic epithelial beta-catenin mutation facilitates intestinal tumorigenesis in mice in response to compounding WNT-pathway mutations', *Dis Model Mech*, 8: 1361-73.

Burgess, J. T., E. Bolderson, M. N. Adams, A. M. Baird, S. D. Zhang, K. A. Gately, K. Umezawa, K. J. O'Byrne, and D. J. Richard. 2016. 'Activation and cleavage of SASH1 by caspase-3 mediates an apoptotic response', *Cell Death Dis*, 7: e2469.

Burgess, J. T., E. Bolderson, J. M. Saunus, S. D. Zhang, L. E. Reid, A. M. McNicol, S. R. Lakhani, K. Cuff, K. Richard, D. J. Richard, and K. J. O'Byrne. 2016. 'SASH1 mediates sensitivity of breast cancer cells to chloropyramine and is associated with prognosis in breast cancer', *Oncotarget*, 7: 72807-18.

Cai, L., H. Wang, and Q. Yang. 2017. 'CRKL overexpression promotes cell proliferation and inhibits apoptosis in endometrial carcinoma', *Oncol Lett*, 13: 51-56.

Calalb, M. B., T. R. Polte, and S. K. Hanks. 1995. 'Tyrosine phosphorylation of focal adhesion kinase at sites in the catalytic domain regulates kinase activity: a role for Src family kinases', *Mol Cell Biol*, 15: 954-63.

Campeau, E., V. E. Ruhl, F. Rodier, C. L. Smith, B. L. Rahmberg, J. O. Fuss, J. Campisi, P. Yaswen, P. K. Cooper, and P. D. Kaufman. 2009. 'A versatile viral system for expression and depletion of proteins in mammalian cells', *PLoS One*, 4: e6529.

Carmeliet, P., and R. K. Jain. 2011. 'Principles and mechanisms of vessel normalization for cancer and other angiogenic diseases', *Nat Rev Drug Discov*, 10: 417-27.

Carmona, G., S. Gottig, A. Orlandi, J. Scheele, T. Bauerle, M. Jugold, F. Kiessling, R. Henschler, A. M. Zeiher, S. Dimmeler, and E. Chavakis. 2009. 'Role of the small GTPase Rap1 for integrin activity regulation in endothelial cells and angiogenesis', *Blood*, 113: 488-97.

Carnero, A., and M. Lleonart. 2016. 'The hypoxic microenvironment: A determinant of cancer stem cell evolution', *Bioessays*, 38 Suppl 1: S65-74.

Cerami, E., J. Gao, U. Dogrusoz, B. E. Gross, S. O. Sumer, B. A. Aksoy, A. Jacobsen, C. J. Byrne, M. L. Heuer, E. Larsson, Y. Antipin, B. Reva, A. P. Goldberg, C. Sander, and N. Schultz. 2012. 'The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data', *Cancer Discov*, 2: 401-4.

Chambers, A. F., A. C. Groom, and I. C. MacDonald. 2002. 'Dissemination and growth of cancer cells in metastatic sites', *Nat Rev Cancer*, 2: 563-72.

Chen, E. G., Y. Chen, L. L. Dong, and J. S. Zhang. 2012. 'Effects of SASH1 on lung cancer cell proliferation, apoptosis, and invasion in vitro', *Tumour Biol*, 33: 1393-401.

Chen, H., H. Liu, and G. Qing. 2018. 'Targeting oncogenic Myc as a strategy for cancer treatment', *Signal Transduct Target Ther*, 3: 5.

Chen, L., D. L. Gibbons, S. Goswami, M. A. Cortez, Y. H. Ahn, L. A. Byers, X. Zhang, X. Yi, D. Dwyer, W. Lin, L. Diao, J. Wang, J. Roybal, M. Patel, C. Ungewiss, D. Peng, S. Antonia, M. Mediavilla-Varela, G. Robertson, M. Suraokar, J. W. Welsh, B. Erez, Wistuba, II, L. Chen, D. Peng, S. Wang, S. E. Ullrich, J. V. Heymach, J. M. Kurie, and F. X. Qin. 2014. 'Metastasis is regulated via microRNA-200/ZEB1 axis control of tumour cell PD-L1 expression and intratumoral immunosuppression', *Nat Commun*, 5: 5241.

Cheng, S., J. Guo, Q. Yang, and L. Han. 2015. 'Crk-like adapter protein is required for TGF-beta-induced AKT and ERK-signaling pathway in epithelial ovarian carcinomas', *Tumour Biol*, 36: 915-9.

Cheng, S., J. Guo, Q. Yang, and X. Yang. 2015. 'Crk-like adapter protein regulates CCL19/CCR7-mediated epithelial-to-mesenchymal transition via ERK signaling pathway in epithelial ovarian carcinomas', *Med Oncol*, 32: 47.

Cheung, H. W., J. Du, J. S. Boehm, F. He, B. A. Weir, X. Wang, M. Butaney, L. V. Sequist, B. Luo, J. A. Engelman, D. E. Root, M. Meyerson, T. R. Golub, P. A. Janne, and W. C. Hahn. 2011. 'Amplification of CRKL induces transformation and epidermal growth factor receptor inhibitor resistance in human non-small cell lung cancers', *Cancer Discov*, 1: 608-25.

Ciccimaro, E., J. Hevko, and I. A. Blair. 2006. 'Analysis of phosphorylation sites on focal adhesion kinase using nanospray liquid chromatography/multiple reaction monitoring mass spectrometry', *Rapid Commun Mass Spectrom*, 20: 3681-92.

Claudio, J. O., Y. X. Zhu, S. J. Benn, A. H. Shukla, C. J. McGlade, N. Falcioni, and A. K. Stewart. 2001. 'HACS1 encodes a novel SH3-SAM adaptor protein differentially expressed in normal and malignant hematopoietic cells', *Oncogene*, 20: 5373-7.

Clevers, H. 2011. 'The cancer stem cell: premises, promises and challenges', *Nat Med*, 17: 313-9.

Comijn, J., G. Berx, P. Vermassen, K. Verschueren, L. van Grunsven, E. Bruyneel, M. Mareel, D. Huylebroeck, and F. van Roy. 2001. 'The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion', *Mol Cell*, 7: 1267-78.

Condeelis, J. S., J. Wyckoff, and J. E. Segall. 2000. 'Imaging of cancer invasion and metastasis using green fluorescent protein', *Eur J Cancer*, 36: 1671-80.

Courcet, J. B., S. C. Elalaoui, L. Duplomb, M. Tajir, J. B. Riviere, J. Thevenon, N. Gigot, N. Marle, B. Aral, Y. Duffourd, A. Sarasin, V. Naim, E. Courcet-Degrolard, M. H. Aubriot-Lorton, L. Martin, J. E. Abrid, C. Thauvin, A. Sefiani, P. Vabres, and L. Faivre. 2015. 'Autosomal-recessive SASH1 variants associated with a new genodermatosis with pigmentation defects, palmoplantar keratoderma and skin carcinoma', *Eur J Hum Genet*, 23: 957-62.

Crouin, C., M. Arnaud, F. Gesbert, J. Camonis, and J. Bertoglio. 2001. 'A yeast two-hybrid study of human p97/Gab2 interactions with its SH2 domain-containing binding partners', *FEBS Lett*, 495: 148-53.

Daigo, Y., Y. Furukawa, T. Kawasoe, H. Ishiguro, M. Fujita, S. Sugai, S. Nakamori, G. J. Liefers, R. A. Tollenaar, C. J. van de Velde, and Y. Nakamura. 1999. 'Absence of genetic alteration at codon 531 of the human c-src gene in 479 advanced colorectal cancers from Japanese and Caucasian patients', *Cancer Res*, 59: 4222-4.

Dauphinee, S. M., A. Clayton, A. Hussainkhel, C. Yang, Y. J. Park, M. E. Fuller, J. Blonder, T. D. Veenstra, and A. Karsan. 2013. 'SASH1 is a scaffold molecule in endothelial TLR4 signaling', *J Immunol*, 191: 892-901.

Denkert, C., W. D. Schmitt, S. Berger, A. Reles, S. Pest, A. Siegert, W. Lichtenegger, M. Dietel, and S. Hauptmann. 2002. 'Expression of mitogen-activated protein kinase phosphatase-1 (MKP-1) in primary human ovarian carcinoma', *Int J Cancer*, 102: 507-13.

Derynck, R., B. P. Muthusamy, and K. Y. Saeteurn. 2014. 'Signaling pathway cooperation in TGF-beta-induced epithelial-mesenchymal transition', *Curr Opin Cell Biol*, 31: 56-66.

Deryugina, E. I., and J. P. Quigley. 2006. 'Matrix metalloproteinases and tumor metastasis', *Cancer Metastasis Rev*, 25: 9-34.

Dhillon, A. S., S. Hagan, O. Rath, and W. Kolch. 2007. 'MAP kinase signalling pathways in cancer', *Oncogene*, 26: 3279.

Doehn, U., C. Hauge, S. R. Frank, C. J. Jensen, K. Duda, J. V. Nielsen, M. S. Cohen, J. V. Johansen, B. R. Winther, L. R. Lund, O. Winther, J. Taunton, S. H. Hansen, and M. Frodin. 2009. 'RSK is a principal effector of the RAS-ERK pathway for eliciting a coordinate promotile/invasive gene program and phenotype in epithelial cells', *Mol Cell*, 35: 511-22.

Dongre, A., and R. A. Weinberg. 2019. 'New insights into the mechanisms of epithelial-mesenchymal transition and implications for cancer', *Nat Rev Mol Cell Biol*, 20: 69-84.

Downward, J. 2003. 'Targeting RAS signalling pathways in cancer therapy', *Nat Rev Cancer*, 3: 11-22.

Drasin, D. J., T. P. Robin, and H. L. Ford. 2011. 'Breast cancer epithelial-to-mesenchymal transition: examining the functional consequences of plasticity', *Breast Cancer Res*, 13: 226.

Duan, H., Z. Yan, W. Chen, Y. Wu, J. Han, H. Guo, and J. Qiao. 2017. 'TET1 inhibits EMT of ovarian cancer cells through activating Wnt/beta-catenin signaling inhibitors DKK1 and SFRP2', *Gynecol Oncol*, 147: 408-17.

Dubois, F., F. Vandermoere, A. Gernez, J. Murphy, R. Toth, S. Chen, K. M. Geraghty, N. A. Morrice, and C. MacKintosh. 2009. 'Differential 14-3-3 affinity capture reveals new downstream targets of phosphatidylinositol 3-kinase signaling', *Mol Cell Proteomics*, 8: 2487-99.

el Marjou, F., K. P. Janssen, B. H. Chang, M. Li, V. Hindie, L. Chan, D. Louvard, P. Chambon, D. Metzger, and S. Robine. 2004. 'Tissue-specific and inducible Cre-mediated recombination in the gut epithelium', *Genesis*, 39: 186-93.

Elmansuri, A. Z., M. A. Tanino, R. Mahabir, L. Wang, T. Kimura, H. Nishihara, I. Kinoshita, H. Dosaka-Akita, M. Tsuda, and S. Tanaka. 2016. 'Novel signaling collaboration between TGF-beta and adaptor protein Crk facilitates EMT in human lung cancer', *Oncotarget*, 7: 27094-107.

Eppler, F. J., T. Quast, and W. Kolanus. 2017. 'Dynamin2 controls Rap1 activation and integrin clustering in human T lymphocyte adhesion', *PLoS One*, 12: e0172443.

Fathers, K. E., S. Rodrigues, D. Zuo, I. V. Murthy, M. Hallett, R. Cardiff, and M. Park. 2010. 'CrkII transgene induces atypical mammary gland development and tumorigenesis', *Am J Pathol*, 176: 446-60.

Feller, S. M. 2001. 'Crk family adaptors-signalling complex formation and biological roles', *Oncogene*, 20: 6348-71.

Fidler, I. J. 2003. 'The pathogenesis of cancer metastasis: the 'seed and soil' hypothesis revisited', *Nat Rev Cancer*, 3: 453-8.

Fischer, K. R., A. Durrans, S. Lee, J. Sheng, F. Li, S. T. Wong, H. Choi, T. El Rayes, S. Ryu, J. Troeger, R. F. Schwabe, L. T. Vahdat, N. K. Altorki, V. Mittal, and D. Gao. 2015. 'Epithelial-to-mesenchymal transition is not required for lung metastasis but contributes to chemoresistance', *Nature*, 527: 472-6.

Fodde, R., W. Edelmann, K. Yang, C. van Leeuwen, C. Carlson, B. Renault, C. Breukel, E. Alt, M. Lipkin, P. M. Khan, and et al. 1994. 'A targeted chain-termination mutation in the mouse Apc gene results in multiple intestinal tumors', *Proc Natl Acad Sci U S A*, 91: 8969-73.

Frame, M. C. 2004. 'Newest findings on the oldest oncogene; how activated src does it', *J Cell Sci*, 117: 989-98.

Franke, F. C., J. Muller, M. Abal, E. D. Medina, U. Nitsche, H. Weidmann, S. Chardonnet, E. Ninio, and K. P. Janssen. 2019. 'The Tumor Suppressor SASH1 Interacts With the Signal Adaptor CRKL to Inhibit Epithelial-Mesenchymal Transition and Metastasis in Colorectal Cancer', *Cell Mol Gastroenterol Hepatol*, 7: 33-53.

Fre, S., S. K. Pallavi, M. Huyghe, M. Lae, K. P. Janssen, S. Robine, S. Artavanis-Tsakonas, and D. Louvard. 2009. 'Notch and Wnt signals cooperatively control cell proliferation and tumorigenesis in the intestine', *Proc Natl Acad Sci U S A*, 106: 6309-14.

Friedl, P., and K. Wolf. 2003. 'Tumour-cell invasion and migration: diversity and escape mechanisms', *Nat Rev Cancer*, 3: 362-74.

Frose, J., M. B. Chen, K. E. Hebron, F. Reinhardt, C. Hajal, A. Zijlstra, R. D. Kamm, and R. A. Weinberg. 2018. 'Epithelial-Mesenchymal Transition Induces Podocalyxin to Promote Extravasation via Ezrin Signaling', *Cell Rep*, 24: 962-72.

Fu, L., Q. Dong, C. Xie, Y. Wang, and Q. Li. 2015. 'CRKL protein overexpression enhances cell proliferation and invasion in pancreatic cancer', *Tumour Biol*, 36: 1015-22.

Furukawa, T., M. Sunamura, F. Motoi, S. Matsuno, and A. Horii. 2003. 'Potential tumor suppressive pathway involving DUSP6/MKP-3 in pancreatic cancer', *Am J Pathol*, 162: 1807-15.

Furukawa, T., T. Yatsuoka, E. M. Youssef, T. Abe, T. Yokoyama, S. Fukushige, E. Soeda, M. Hoshi, Y. Hayashi, M. Sunamura, M. Kobari, and A. Horii. 1998. 'Genomic analysis of DUSP6, a dual specificity MAP kinase phosphatase, in pancreatic cancer', *Cytogenet Cell Genet*, 82: 156-9.

Gao, D., L. T. Vahdat, S. Wong, J. C. Chang, and V. Mittal. 2012. 'Microenvironmental regulation of epithelial-mesenchymal transitions in cancer', *Cancer Res*, 72: 4883-9.

Gao, J., B. A. Aksoy, U. Dogrusoz, G. Dresdner, B. Gross, S. O. Sumer, Y. Sun, A. Jacobsen, R. Sinha, E. Larsson, E. Cerami, C. Sander, and N. Schultz. 2013. 'Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal', *Sci Signal*, 6: pl1.

Garcia-Guzman, M., F. Dolfi, K. Zeh, and K. Vuori. 1999. 'Met-induced JNK activation is mediated by the adapter protein Crk and correlates with the Gab1 - Crk signaling complex formation', *Oncogene*, 18: 7775-86.

Gassmann, P., J. Haier, K. Schluter, B. Domikowsky, C. Wendel, U. Wiesner, R. Kubitza, R. Engers, S. W. Schneider, B. Homey, and A. Muller. 2009. 'CXCR4 regulates the early extravasation of metastatic tumor cells in vivo', *Neoplasia*, 11: 651-61.

Gerlinger, M., A. J. Rowan, S. Horswell, M. Math, J. Larkin, D. Endesfelder, E. Gronroos, P. Martinez, N. Matthews, A. Stewart, P. Tarpey, I. Varela, B. Phillimore, S. Begum, N. Q. McDonald, A. Butler, D. Jones, K. Raine, C. Latimer, C. R. Santos, M. Nohadani, A. C. Eklund, B. Spencer-Dene, G. Clark, L. Pickering, G. Stamp, M. Gore, Z. Szallasi, J. Downward, P. A. Futreal, and C. Swanton. 2012. 'Intratumor heterogeneity and branched evolution revealed by multiregion sequencing', *N Engl J Med*, 366: 883-92.

Gregory, P. A., A. G. Bert, E. L. Paterson, S. C. Barry, A. Tsykin, G. Farshid, M. A. Vadas, Y. Khew-Goodall, and G. J. Goodall. 2008. 'The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1', *Nat Cell Biol*, 10: 593-601.

Grivennikov, S. I., F. R. Greten, and M. Karin. 2010. 'Immunity, inflammation, and cancer', *Cell*, 140: 883-99.

Guidotti, G., L. Brambilla, and D. Rossi. 2017. 'Cell-Penetrating Peptides: From Basic Research to Clinics', *Trends Pharmacol Sci*, 38: 406-24.

Guo, C., S. Liu, and M. Z. Sun. 2014. 'The role of CT10 regulation of kinase-like in cancer', *Future Oncol*, 10: 2687-97.

Guo, C., J. Ma, G. Deng, Y. Qu, L. Yin, Y. Li, Y. Han, C. Cai, H. Shen, and S. Zeng. 2017. 'ZEB1 Promotes Oxaliplatin Resistance through the Induction of Epithelial - Mesenchymal Transition in Colon Cancer Cells', *J Cancer*, 8: 3555-66.

Guo, C., D. Zhao, Q. Zhang, S. Liu, and M. Z. Sun. 2018. 'miR-429 suppresses tumor migration and invasion by targeting CRKL in hepatocellular carcinoma via inhibiting Raf/MEK/ERK pathway and epithelial-mesenchymal transition', *Sci Rep*, 8: 2375.

Guo, F., H. Gong, H. Zhao, J. Chen, Y. Zhang, L. Zhang, X. Shi, A. Zhang, H. Jin, J. Zhang, and Y. He. 2018. 'Mutation status and prognostic values of KRAS, NRAS, BRAF and PIK3CA in 353 Chinese colorectal cancer patients', *Sci Rep*, 8: 6076.

Gupta, G. P., and J. Massague. 2006. 'Cancer metastasis: building a framework', *Cell*, 127: 679-95.

Gupta, P. B., C. L. Chaffer, and R. A. Weinberg. 2009. 'Cancer stem cells: mirage or reality?', *Nat Med*, 15: 1010-2.

Haigis, K. M. 2017. 'KRAS Alleles: The Devil Is in the Detail', *Trends Cancer*, 3: 686-97.

Han, G., D. Wu, Y. Yang, Z. Li, J. Zhang, and C. Li. 2015. 'CrkL meditates CCL20/CCR6-induced EMT in gastric cancer', *Cytokine*, 76: 163-9.

Han, J., C. J. Lim, N. Watanabe, A. Soriani, B. Ratnikov, D. A. Calderwood, W. Puzon-McLaughlin, E. M. Lafuente, V. A. Boussiotis, S. J. Shattil, and M. H. Ginsberg. 2006. 'Reconstructing and deconstructing agonist-induced activation of integrin alphaIIbbeta3', *Curr Biol*, 16: 1796-806.

Hanahan, D., and R. A. Weinberg. 2011. 'Hallmarks of cancer: the next generation', *Cell*, 144: 646-74.

Hartsock, A., and W. J. Nelson. 2008. 'Adherens and tight junctions: structure, function and connections to the actin cytoskeleton', *Biochim Biophys Acta*, 1778: 660-9.

Hasegawa, H., E. Kiyokawa, S. Tanaka, K. Nagashima, N. Gotoh, M. Shibuya, T. Kurata, and M. Matsuda. 1996. 'DOCK180, a major CRK-binding protein, alters cell morphology upon translocation to the cell membrane', *Mol Cell Biol*, 16: 1770-6.

Haynes, J., J. Srivastava, N. Madson, T. Wittmann, and D. L. Barber. 2011. 'Dynamic actin remodeling during epithelial-mesenchymal transition depends on increased moesin expression', *Mol Biol Cell*, 22: 4750-64.

Hazan, R. B., G. R. Phillips, R. F. Qiao, L. Norton, and S. A. Aaronson. 2000. 'Exogenous expression of N-cadherin in breast cancer cells induces cell migration, invasion, and metastasis', *J Cell Biol*, 148: 779-90.

He, P., H. X. Zhang, C. Y. Sun, C. Y. Chen, and H. Q. Jiang. 2016. 'Overexpression of SASH1 Inhibits the Proliferation, Invasion, and EMT in Hepatocarcinoma Cells', *Oncol Res*, 24: 25-32.

Hemmeryckx, B., A. van Wijk, A. Reichert, V. Kaartinen, R. de Jong, P. K. Pattengale, I. Gonzalez-Gomez, J. Groffen, and N. Heisterkamp. 2001. 'Crkl enhances leukemogenesis in BCR/ABL P190 transgenic mice', *Cancer Res*, 61: 1398-405.

Heng, H. H., S. W. Bremer, J. B. Stevens, S. D. Horne, G. Liu, B. Y. Abdallah, K. J. Ye, and C. J. Ye. 2013. 'Chromosomal instability (CIN): what it is and why it is crucial to cancer evolution', *Cancer Metastasis Rev*, 32: 325-40.

Heng, H. H., J. B. Stevens, S. W. Bremer, K. J. Ye, G. Liu, and C. J. Ye. 2010. 'The evolutionary mechanism of cancer', *J Cell Biochem*, 109: 1072-84.

Hollstein, M., D. Sidransky, B. Vogelstein, and C. C. Harris. 1991. 'p53 mutations in human cancers', *Science*, 253: 49-53.

Holtorf, A., A. Conrad, B. Holzmann, and K. P. Janssen. 2018. 'Cell-type specific MyD88 signaling is required for intestinal tumor initiation and progression to malignancy', *Oncoimmunology*, 7: e1466770.

Howe, E. N., D. R. Cochrane, and J. K. Richer. 2011. 'Targets of miR-200c mediate suppression of cell motility and anoikis resistance', *Breast Cancer Res*, 13: R45.

Hsu, D. S., H. Y. Lan, C. H. Huang, S. K. Tai, S. Y. Chang, T. L. Tsai, C. C. Chang, C. H. Tzeng, K. J. Wu, J. Y. Kao, and M. H. Yang. 2010. 'Regulation of excision repair cross-complementation group 1 by Snail contributes to cisplatin resistance in head and neck cancer', *Clin Cancer Res*, 16: 4561-71.

Huang, S. M., M. K. Hancock, J. L. Pitman, A. P. Orth, and N. Gekakis. 2009. 'Negative regulators of insulin signaling revealed in a genome-wide functional screen', *PLoS One*, 4: e6871.

Inglis, S., R. Jones, D. Fritz, C. Stojkoski, G. Booker, and S. Pyke. 2005. 'Synthesis of 5-, 6- and 7-substituted-2-aminoquinolines as SH3 domain ligands', *Org Biomol Chem*, 3: 2543-57.

Inglis, S. R., C. Stojkoski, K. M. Branson, J. F. Cawthray, D. Fritz, E. Wiadrowski, S. M. Pyke, and G. W. Booker. 2004. 'Identification and specificity studies of small-molecule ligands for SH3 protein domains', *J Med Chem*, 47: 5405-17.

Irby, R. B., W. Mao, D. Coppola, J. Kang, J. M. Loubeau, W. Trudeau, R. Karl, D. J. Fujita, R. Jove, and T. J. Yeatman. 1999. 'Activating SRC mutation in a subset of advanced human colon cancers', *Nat Genet*, 21: 187-90.

Irie, H. Y., R. V. Pearline, D. Grueneberg, M. Hsia, P. Ravichandran, N. Kothari, S. Natesan, and J. S. Brugge. 2005. 'Distinct roles of Akt1 and Akt2 in regulating cell migration and epithelial-mesenchymal transition', *J Cell Biol*, 171: 1023-34.

Ivanov, A. A., F. R. Khuri, and H. Fu. 2013. 'Targeting protein-protein interactions as an anticancer strategy', *Trends Pharmacol Sci*, 34: 393-400.

Iwahara, T., T. Akagi, Y. Fujitsuka, and H. Hanafusa. 2004. 'CrkII regulates focal adhesion kinase activation by making a complex with Crk-associated substrate, p130Cas', *Proc Natl Acad Sci U S A*, 101: 17693-8.

Jankowski, W., T. Saleh, M. T. Pai, G. Sriram, R. B. Birge, and C. G. Kalodimos. 2012. 'Domain organization differences explain Bcr-Abl's preference for CrkL over CrkII', *Nat Chem Biol*, 8: 590-6.

Janssen, K. P., P. Alberici, H. Fsihi, C. Gaspar, C. Breukel, P. Franken, C. Rosty, M. Abal, F. El Marjou, R. Smits, D. Louvard, R. Fodde, and S. Robine. 2006. 'APC and oncogenic KRAS are synergistic in enhancing Wnt signaling in intestinal tumor formation and progression', *Gastroenterology*, 131: 1096-109.

Janssen, K. P., F. el-Marjou, D. Pinto, X. Sastre, D. Rouillard, C. Fouquet, T. Soussi, D. Louvard, and S. Robine. 2002. 'Targeted expression of oncogenic K-ras in intestinal epithelium causes spontaneous tumorigenesis in mice', *Gastroenterology*, 123: 492-504.

Ji, H., B. Li, S. Zhang, Z. He, Y. Zhou, and L. Ouyang. 2016. 'Crk-like adapter protein is overexpressed in cervical carcinoma, facilitates proliferation, invasion and chemoresistance, and regulates Src and Akt signaling', *Oncol Lett*, 12: 3811-17.

Joyce, J. A., and J. W. Pollard. 2009. 'Microenvironmental regulation of metastasis', *Nat Rev Cancer*, 9: 239-52.

Junttila, M. R., and J. Westermarck. 2008. 'Mechanisms of MYC stabilization in human malignancies', *Cell Cycle*, 7: 592-6.

Jurasz, P., D. Alonso-Escolano, and M. W. Radomski. 2004. 'Platelet--cancer interactions: mechanisms and pharmacology of tumour cell-induced platelet aggregation', *Br J Pharmacol*, 143: 819-26.

Kahlert, C., S. Lahes, P. Radhakrishnan, S. Dutta, C. Mogler, E. Herpel, K. Brand, G. Steinert, M. Schneider, M. Mollenhauer, C. Reissfelder, F. Klupp, J. Fritzmann, C. Wunder, A. Benner, M. Kloor, C. Huth, P. Contin, A. Ulrich, M. Koch, and J. Weitz. 2011. 'Overexpression of ZEB2 at the invasion front of colorectal cancer is an independent prognostic marker and regulates tumor invasion in vitro', *Clin Cancer Res*, 17: 7654-63.

Kalluri, R., and R. A. Weinberg. 2009. 'The basics of epithelial-mesenchymal transition', *J Clin Invest*, 119: 1420-8.

Kardinal, C., B. Konkol, A. Schulz, G. Posern, H. Lin, K. Adermann, M. Eulitz, Z. Estrov, M. Talpaz, R. B. Arlinghaus, and S. M. Feller. 2000. 'Cell-penetrating SH3 domain blocker peptides inhibit proliferation of primary blast cells from CML patients', *FASEB J*, 14: 1529-38.

Katagiri, K., M. Hattori, N. Minato, Sk Irie, K. Takatsu, and T. Kinashi. 2000. 'Rap1 is a potent activation signal for leukocyte function-associated antigen 1 distinct from protein kinase C and phosphatidylinositol-3-OH kinase', *Mol Cell Biol*, 20: 1956-69.

Katagiri, K., A. Maeda, M. Shimonaka, and T. Kinashi. 2003. 'RAPL, a Rap1-binding molecule that mediates Rap1-induced adhesion through spatial regulation of LFA-1', *Nat Immunol*, 4: 741-8.

Kessenbrock, K., V. Plaks, and Z. Werb. 2010. 'Matrix metalloproteinases: regulators of the tumor microenvironment', *Cell*, 141: 52-67.

Keyse, S. M. 2000. 'Protein phosphatases and the regulation of mitogen-activated protein kinase signalling', *Curr Opin Cell Biol*, 12: 186-92.

———. 2008. 'Dual-specificity MAP kinase phosphatases (MKPs) and cancer', *Cancer Metastasis Rev*, 27: 253-61.

Khoury, K., and A. Domling. 2012. 'P53 mdm2 inhibitors', *Curr Pharm Des*, 18: 4668-78.

Kidger, A. M., L. K. Rushworth, J. Stellzig, J. Davidson, C. J. Bryant, C. Bayley, E. Caddye, T. Rogers, S. M. Keyse, and C. J. Caunt. 2017. 'Dual-specificity phosphatase 5 controls the localized inhibition, propagation, and transforming potential of ERK signaling', *Proc Natl Acad Sci U S A*, 114: E317-E26.

Kiesslich, T., M. Pichler, and D. Neureiter. 2013. 'Epigenetic control of epithelial-mesenchymal-transition in human cancer', *Mol Clin Oncol*, 1: 3-11.

Kim, C. A., and J. U. Bowie. 2003. 'SAM domains: uniform structure, diversity of function', *Trends Biochem Sci*, 28: 625-8.

Kim, D. H., T. Xing, Z. Yang, R. Dudek, Q. Lu, and Y. H. Chen. 2017. 'Epithelial Mesenchymal Transition in Embryonic Development, Tissue Repair and Cancer: A Comprehensive Overview', *J Clin Med*, 7.

Kim, Y., M. C. Kugler, Y. Wei, K. K. Kim, X. Li, A. N. Brumwell, and H. A. Chapman. 2009. 'Integrin alpha3beta1-dependent beta-catenin phosphorylation links epithelial Smad signaling to cell contacts', *J Cell Biol*, 184: 309-22.

Klein, C. A. 2009. 'Parallel progression of primary tumours and metastases', *Nat Rev Cancer*, 9: 302-12.

Knudsen, B. S., S. M. Feller, and H. Hanafusa. 1994. 'Four proline-rich sequences of the guanine-nucleotide exchange factor C3G bind with unique specificity to the first Src homology 3 domain of Crk', *J Biol Chem*, 269: 32781-7.

Kobashigawa, Y., M. Sakai, M. Naito, M. Yokochi, H. Kumeta, Y. Makino, K. Ogura, S. Tanaka, and F. Inagaki. 2007. 'Structural basis for the transforming activity of human cancer-related signaling adaptor protein CRK', *Nat Struct Mol Biol*, 14: 503-10.

Kohno, M., and J. Pouyssegur. 2006. 'Targeting the ERK signaling pathway in cancer therapy', *Ann Med*, 38: 200-11.

Kondoh, K., and E. Nishida. 2007. 'Regulation of MAP kinases by MAP kinase phosphatases', *Biochim Biophys Acta*, 1773: 1227-37.

Kucharska, A., L. K. Rushworth, C. Staples, N. A. Morrice, and S. M. Keyse. 2009. 'Regulation of the inducible nuclear dual-specificity phosphatase DUSP5 by ERK MAPK', *Cell Signal*, 21: 1794-805.

Kuleshov, M. V., M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann, M. G. McDermott, C. D. Monteiro, G. W. Gundersen, and A. Ma'ayan. 2016. 'Enrichr: a comprehensive gene set enrichment analysis web server 2016 update', *Nucleic Acids Res*, 44: W90-7.

Kumar, S., V. Davra, A. E. Obr, K. Geng, T. L. Wood, M. S. De Lorenzo, and R. B. Birge. 2017. 'Crk adaptor protein promotes PD-L1 expression, EMT and immune evasion in a murine model of triple-negative breast cancer', *Oncoimmunology*, 7: e1376155.

Kumar, S., B. Lu, V. Davra, P. Hornbeck, K. Machida, and R. B. Birge. 2018. 'Crk Tyrosine Phosphorylation Regulates PDGF-BB-inducible Src Activation and Breast Tumorigenicity and Metastasis', *Mol Cancer Res*, 16: 173-83.

Kumarswamy, R., G. Mudduluru, P. Ceppi, S. Muppala, M. Kozlowski, J. Niklinski, M. Papotti, and H. Allgayer. 2012. 'MicroRNA-30a inhibits epithelial-to-mesenchymal transition by targeting Snai1 and is downregulated in non-small cell lung cancer', *Int J Cancer*, 130: 2044-53.

Kurochkina, N., and U. Guha. 2013. 'SH3 domains: modules of protein-protein interactions', *Biophys Rev*, 5: 29-39.

Kurrey, N. K., S. P. Jalgaonkar, A. V. Joglekar, A. D. Ghanate, P. D. Chaskar, R. Y. Doiphode, and S. A. Bapat. 2009. 'Snail and slug mediate radioresistance and chemoresistance by antagonizing p53-mediated apoptosis and acquiring a stem-like phenotype in ovarian cancer cells', *Stem Cells*, 27: 2059-68.

Lagarrigue, F., C. Kim, and M. H. Ginsberg. 2016. 'The Rap1-RIAM-talin axis of integrin activation and blood cell function', *Blood*, 128: 479-87.

Laghi, L., P. Bianchi, O. Orbetegli, L. Gennari, M. Roncalli, and A. Malesci. 2001. 'Lack of mutation at codon 531 of SRC in advanced colorectal cancers from Italian patients', *Br J Cancer*, 84: 196-8.

Lai, S., and S. Pelech. 2016. 'Regulatory roles of conserved phosphorylation sites in the activation T-loop of the MAP kinase ERK1', *Mol Biol Cell*, 27: 1040-50.

Lamorte, L., S. Rodrigues, V. Sangwan, C. E. Turner, and M. Park. 2003. 'Crk associates with a multimolecular Paxillin/GIT2/beta-PIX complex and promotes Rac-dependent relocalization of Paxillin to focal contacts', *Mol Biol Cell*, 14: 2818-31.

Lamouille, S., J. Xu, and R. Derynck. 2014. 'Molecular mechanisms of epithelial-mesenchymal transition', *Nat Rev Mol Cell Biol*, 15: 178-96.

Levental, K. R., H. Yu, L. Kass, J. N. Lakins, M. Egeblad, J. T. Erler, S. F. Fong, K. Csiszar, A. Giaccia, W. Weninger, M. Yamauchi, D. L. Gasser, and V. M. Weaver. 2009. 'Matrix crosslinking forces tumor progression by enhancing integrin signaling', *Cell*, 139: 891-906.

Li, L., D. L. Guris, M. Okura, and A. Imamoto. 2003. 'Translocation of CrkL to focal adhesions mediates integrin-induced migration downstream of Src family kinases', *Mol Cell Biol*, 23: 2883-92.

Li, L., M. Okura, and A. Imamoto. 2002. 'Focal adhesions require catalytic activity of Src family kinases to mediate integrin-matrix adhesion', *Mol Cell Biol*, 22: 1203-17.

Lilja, J., T. Zacharchenko, M. Georgiadou, G. Jacquemet, N. De Franceschi, E. Peuhu, H. Hamidi, J. Pouwels, V. Martens, F. H. Nia, M. Beifuss, T. Boeckers, H. J. Kreienkamp, I. L. Barsukov, and J. Ivaska. 2017. 'SHANK proteins limit integrin activation by directly interacting with Rap1 and R-Ras', *Nat Cell Biol*, 19: 292-305.

Lin, F., X. Chengyao, L. Qingchang, D. Qianze, W. Enhua, and W. Yan. 2015. 'CRKL promotes lung cancer cell invasion through ERK-MMP9 pathway', *Mol Carcinog*, 54 Suppl 1: E35-44.

Lin, S., J. Zhang, J. Xu, H. Wang, Q. Sang, Q. Xing, and L. He. 2012. 'Effects of SASH1 on melanoma cell proliferation and apoptosis in vitro', *Mol Med Rep*, 6: 1243-8.

Lin, T., A. Ponn, X. Hu, B. K. Law, and J. Lu. 2010. 'Requirement of the histone demethylase LSD1 in Snai1-mediated transcriptional repression during epithelial-mesenchymal transition', *Oncogene*, 29: 4896-904.

Liu, S., D. A. Calderwood, and M. H. Ginsberg. 2000. 'Integrin cytoplasmic domain-binding proteins', *J Cell Sci*, 113 (Pt 20): 3563-71.

Love, M. I., W. Huber, and S. Anders. 2014. 'Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2', *Genome Biol*, 15: 550.

Ma, Y., C. Gong, Y. Ma, F. Fan, M. Luo, F. Yang, and Y. H. Zhang. 2012. 'Direct cytosolic delivery of cargoes in vivo by a chimera consisting of D- and L-arginine residues', *J Control Release*, 162: 286-94.

Macosko, E. Z., A. Basu, R. Satija, J. Nemesh, K. Shekhar, M. Goldman, I. Tirosh, A. R. Bialas, N. Kamitaki, E. M. Martersteck, J. J. Trombetta, D. A. Weitz, J. R. Sanes, A. K. Shalek, A. Regev, and S. A. McCarroll. 2015. 'Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets', *Cell*, 161: 1202-14.

Magi-Galluzzi, C., R. Mishra, M. Fiorentino, R. Montironi, H. Yao, P. Capodieci, K. Wishnow, I. Kaplan, P. J. Stork, and M. Loda. 1997. 'Mitogen-activated protein kinase phosphatase 1 is overexpressed in prostate cancers and is inversely related to apoptosis', *Lab Invest*, 76: 37-51.

Makrodouli, E., E. Oikonomou, M. Koc, L. Andera, T. Sasazuki, S. Shirasawa, and A. Pintzas. 2011. 'BRAF and RAS oncogenes regulate Rho GTPase pathways to mediate migration and invasion properties in human colon cancer cells: a comparative study', *Mol Cancer*, 10: 118.

Mandl, M., D. N. Slack, and S. M. Keyse. 2005. 'Specific inactivation and nuclear anchoring of extracellular signal-regulated kinase 2 by the inducible dual-specificity protein phosphatase DUSP5', *Mol Cell Biol*, 25: 1830-45.

Mani, S. A., W. Guo, M. J. Liao, E. N. Eaton, A. Ayyanan, A. Y. Zhou, M. Brooks, F. Reinhard, C. C. Zhang, M. Shipitsin, L. L. Campbell, K. Polyak, C. Brisken, J. Yang, and R. A. Weinberg. 2008. 'The epithelial-mesenchymal transition generates cells with properties of stem cells', *Cell*, 133: 704-15.

Martin, M. D., G. J. Kremers, K. W. Short, J. V. Rocheleau, L. Xu, D. W. Piston, L. M. Matrisian, and D. L. Gorden. 2010. 'Rapid extravasation and establishment of breast cancer micrometastases in the liver microenvironment', *Mol Cancer Res*, 8: 1319-27.

Martini, M., A. Gnann, D. Scheikl, B. Holzmann, and K. P. Janssen. 2011. 'The candidate tumor suppressor SASH1 interacts with the actin cytoskeleton and stimulates cell-matrix adhesion', *Int J Biochem Cell Biol*, 43: 1630-40.

Maschler, S., G. Wirl, H. Spring, D. V. Bredow, I. Sordat, H. Beug, and E. Reichmann. 2005. 'Tumor cell invasiveness correlates with changes in integrin expression and localization', *Oncogene*, 24: 2032-41.

Matsuda, M., Y. Hashimoto, K. Muroya, H. Hasegawa, T. Kurata, S. Tanaka, S. Nakamura, and S. Hattori. 1994. 'CRK protein binds to two guanine nucleotide-releasing proteins for the Ras family and modulates nerve growth factor-induced activation of Ras in PC12 cells', *Mol Cell Biol*, 14: 5495-500.

Matsuda, M., S. Tanaka, S. Nagata, A. Kojima, T. Kurata, and M. Shibuya. 1992. 'Two species of human CRK cDNA encode proteins with distinct biological activities', *Mol Cell Biol*, 12: 3482-9.

Matsumoto, R., M. Tsuda, L. Wang, N. Maishi, T. Abe, T. Kimura, M. Tanino, H. Nishihara, K. Hida, Y. Ohba, N. Shinohara, K. Nonomura, and S. Tanaka. 2015. 'Adaptor protein CRK induces epithelial-mesenchymal transition and metastasis of bladder cancer cells through HGF/c-Met feedback loop', *Cancer Sci*, 106: 709-17.

Mayer, B. J., M. Hamaguchi, and H. Hanafusa. 1988. 'A novel viral oncogene with structural similarity to phospholipase C', *Nature*, 332: 272-5.

McGranahan, N., and C. Swanton. 2017. 'Clonal Heterogeneity and Tumor Evolution: Past, Present, and the Future', *Cell*, 168: 613-28.

McSherry, E. A., K. Brennan, L. Hudson, A. D. Hill, and A. M. Hopkins. 2011. 'Breast cancer cell migration is regulated through junctional adhesion molecule-A-mediated activation of Rap1 GTPase', *Breast Cancer Res*, 13: R31.

Mehlen, P., and A. Puisieux. 2006. 'Metastasis: a question of life or death', *Nat Rev Cancer*, 6: 449-58.

Mendez, M. G., S. Kojima, and R. D. Goldman. 2010. 'Vimentin induces changes in cell shape, motility, and adhesion during the epithelial to mesenchymal transition', *FASEB J*, 24: 1838-51.

Meng, Q., M. Zheng, H. Liu, C. Song, W. Zhang, J. Yan, L. Qin, and X. Liu. 2013. 'SASH1 regulates proliferation, apoptosis, and invasion of osteosarcoma cell', *Mol Cell Biochem*, 373: 201-10.

Mitra, S. K., D. A. Hanson, and D. D. Schlaepfer. 2005. 'Focal adhesion kinase: in command and control of cell motility', *Nat Rev Mol Cell Biol*, 6: 56-68.

Mitra, S. K., and D. D. Schlaepfer. 2006. 'Integrin-regulated FAK-Src signaling in normal and cancer cells', *Curr Opin Cell Biol*, 18: 516-23.

Miyoshi, A., Y. Kitajima, K. Sumi, K. Sato, A. Hagiwara, Y. Koga, and K. Miyazaki. 2004. 'Snail and SIP1 increase cancer invasion by upregulating MMP family in hepatocellular carcinoma cells', *Br J Cancer*, 90: 1265-73.

Molina-Ortiz, P., S. Polizzi, E. Ramery, S. Gayral, C. Delierneux, C. Oury, S. Iwashita, and S. Schurmans. 2014. 'Rasa3 controls megakaryocyte Rap1 activation, integrin signaling and differentiation into proplatelet', *PLoS Genet*, 10: e1004420.

Morel, A. P., M. Lievre, C. Thomas, G. Hinkal, S. Ansieau, and A. Puisieux. 2008. 'Generation of breast cancer stem cells through epithelial-mesenchymal transition', *PLoS One*, 3: e2888.

Muda, M., A. Theodosiou, C. Gillieron, A. Smith, C. Chabert, M. Camps, U. Boschert, N. Rodrigues, K. Davies, A. Ashworth, and S. Arkinstall. 1998. 'The mitogen-activated protein kinase phosphatase-3 N-terminal noncatalytic region is responsible for tight substrate binding and enzymatic specificity', *J Biol Chem*, 273: 9323-9.

Muralidharan, V., K. Dutta, J. Cho, M. Vila-Perello, D. P. Raleigh, D. Cowburn, and T. W. Muir. 2006. 'Solution structure and folding characteristics of the C-terminal SH3 domain of c-Crk-II', *Biochemistry*, 45: 8874-84.

Nagaharu, K., X. Zhang, T. Yoshida, D. Katoh, N. Hanamura, Y. Kozuka, T. Ogawa, T. Shiraishi, and K. Imanaka-Yoshida. 2011. 'Tenascin C induces epithelial-mesenchymal transition-like change accompanied by SRC activation and focal adhesion kinase phosphorylation in human breast cancer cells', *Am J Pathol*, 178: 754-63.

Negrini, S., V. G. Gorgoulis, and T. D. Halazonetis. 2010. 'Genomic instability--an evolving hallmark of cancer', *Nat Rev Mol Cell Biol*, 11: 220-8.

Nelson, W. J., and R. Nusse. 2004. 'Convergence of Wnt, beta-catenin, and cadherin pathways', *Science*, 303: 1483-7.

Nguyen, J. T., M. Porter, M. Amoui, W. T. Miller, R. N. Zuckermann, and W. A. Lim. 2000. 'Improving SH3 domain ligand selectivity using a non-natural scaffold', *Chem Biol*, 7: 463-73.

Nieman, M. T., R. S. Prudoff, K. R. Johnson, and M. J. Wheelock. 1999. 'N-cadherin promotes motility in human breast cancer cells regardless of their E-cadherin expression', *J Cell Biol*, 147: 631-44.

Nischan, N., H. D. Herce, F. Natale, N. Bohlke, N. Budisa, M. C. Cardoso, and C. P. Hackenberger. 2015. 'Covalent attachment of cyclic TAT peptides to GFP results in protein delivery into live cells with immediate bioavailability', *Angew Chem Int Ed Engl*, 54: 1950-3.

Nitsche, U., H. Friess, A. Agha, M. Angele, R. Eckel, W. Heitland, K. W. Jauch, D. Krenz, N. C. Nussler, H. G. Rau, R. Ruppert, G. Schubert-Fritschle, D. Wilhelm, J. Werner, and J. Engel. 2016. 'Prognosis of mucinous and signet-ring cell colorectal cancer in a population-based cohort', *J Cancer Res Clin Oncol*, 142: 2357-66.

Nitsche, U., R. Rosenberg, A. Balmert, T. Schuster, J. Slotta-Huspenina, P. Herrmann, F. G. Bader, H. Friess, P. M. Schlag, U. Stein, and K. P. Janssen. 2012. 'Integrative marker analysis allows risk assessment for metastasis in stage II colon cancer', *Ann Surg*, 256: 763-71; discussion 71.

Nosaka, Y., A. Arai, N. Miyasaka, and O. Miura. 1999. 'CrkL mediates Ras-dependent activation of the Raf/ERK pathway through the guanine nucleotide exchange factor C3G in hematopoietic cells stimulated with erythropoietin or interleukin-3', *J Biol Chem*, 274: 30154-62.

Nowell, P. C. 1976. 'The clonal evolution of tumor cell populations', *Science*, 194: 23-8.

Okada, M. 2012. 'Regulation of the SRC family kinases by Csk', *Int J Biol Sci*, 8: 1385-97.

Pan, J, and Y Liu. 2016. 'SASH1 inhibits hypoxia-induced epithelial-to-mesenchymal transition via suppression of the PI3K/Akt pathway in human pancreatic cancer cells', *Int J Clin Exp Pathol*, 9: 750-57.

Pantel, K., and R. H. Brakenhoff. 2004. 'Dissecting the metastatic cascade', *Nat Rev Cancer*, 4: 448-56.

Parekh, S., C. Ziegenhain, B. Vieth, W. Enard, and I. Hellmann. 2016. 'The impact of amplification on differential expression analyses by RNA-seq', *Sci Rep*, 6: 25533.

Park, S. M., A. B. Gaur, E. Lengyel, and M. E. Peter. 2008. 'The miR-200 family determines the epithelial phenotype of cancer cells by targeting the E-cadherin repressors ZEB1 and ZEB2', *Genes Dev*, 22: 894-907.

Patel, A., H. Sabbineni, A. Clarke, and P. R. Somanath. 2016. 'Novel roles of Src in cancer cell epithelial-to-mesenchymal transition, vascular permeability, microinvasion and metastasis', *Life Sci*, 157: 52-61.

Payne, D. M., A. J. Rossomando, P. Martino, A. K. Erickson, J. H. Her, J. Shabanowitz, D. F. Hunt, M. J. Weber, and T. W. Sturgill. 1991. 'Identification of the regulatory phosphorylation sites in pp42/mitogen-activated protein kinase (MAP kinase)', *EMBO J*, 10: 885-92.

Peng, D. H., C. Ungewiss, P. Tong, L. A. Byers, J. Wang, J. R. Canales, P. A. Villalobos, N. Uraoka, B. Mino, C. Behrens, Wistuba, II, R. I. Han, C. A. Wanna, M. Fahrenholtz, K. J. Grande-Allen, C. J. Creighton, and D. L. Gibbons. 2016. 'ZEB1 induces LOXL2-mediated collagen stabilization and deposition in the extracellular matrix to drive lung cancer invasion and metastasis', *Oncogene*.

Peng, J. M., R. Bera, C. Y. Chiou, M. C. Yu, T. C. Chen, C. W. Chen, T. R. Wang, W. L. Chiang, S. P. Chai, Y. Wei, H. Wang, M. C. Hung, and S. Y. Hsieh. 2018. 'Actin cytoskeleton remodeling drives epithelial-mesenchymal transition for hepatoma invasion and metastasis in mice', *Hepatology*, 67: 2226-43.

Peng, L., H. Wei, and L. Liren. 2014. 'Promoter methylation assay of SASH1 gene in hepatocellular carcinoma', *J BUON*, 19: 1041-7.

Pino, M. S., H. Kikuchi, M. Zeng, M. T. Herraiz, I. Sperduti, D. Berger, D. Y. Park, A. J. Iafrate, L. R. Zukerberg, and D. C. Chung. 2010. 'Epithelial to mesenchymal transition is impaired in colon cancer cells with microsatellite instability', *Gastroenterology*, 138: 1406-17.

Posern, G., J. Zheng, B. S. Knudsen, C. Kardinal, K. B. Muller, J. Voss, T. Shishido, D. Cowburn, G. Cheng, B. Wang, G. D. Kruh, S. K. Burrell, C. A. Jacobson, D. M. Lenz, T. J. Zamborelli, K. Adermann, H. Hanafusa, and S. M. Feller. 1998. 'Development of highly selective SH3 binding peptides for

Crk and CRKL which disrupt Crk-complexes with DOCK180, SoS and C3G', *Oncogene*, 16: 1903-12.

Powell, S. M., N. Zilz, Y. Beazer-Barclay, T. M. Bryan, S. R. Hamilton, S. N. Thibodeau, B. Vogelstein, and K. W. Kinzler. 1992. 'APC mutations occur early during colorectal tumorigenesis', *Nature*, 359: 235-7.

Prosser, S., E. Sorokina, P. Pratt, and A. Sorokin. 2003. 'CrkIII: a novel and biologically distinct member of the Crk family of adaptor proteins', *Oncogene*, 22: 4799-806.

Psaila, B., and D. Lyden. 2009. 'The metastatic niche: adapting the foreign soil', *Nat Rev Cancer*, 9: 285-93.

Ran, F. A., P. D. Hsu, J. Wright, V. Agarwala, D. A. Scott, and F. Zhang. 2013. 'Genome engineering using the CRISPR-Cas9 system', *Nat Protoc*, 8: 2281-308.

Rankin, E. B., and A. J. Giaccia. 2016. 'Hypoxic control of metastasis', *Science*, 352: 175-80.

Reddy, K. B., S. M. Nabha, and N. Atanaskova. 2003. 'Role of MAP kinase in tumor progression and invasion', *Cancer Metastasis Rev*, 22: 395-403.

Ricciardi, M., M. Zanotto, G. Malpeli, G. Bassi, O. Perbellini, M. Chilosi, F. Bifari, and M. Krampera. 2015. 'Epithelial-to-mesenchymal transition (EMT) induced by inflammatory priming elicits mesenchymal stromal cell-like immune-modulatory properties in cancer cells', *Br J Cancer*, 112: 1067-75.

Rimkus, C., M. Martini, J. Friederichs, R. Rosenberg, D. Doll, J. R. Siewert, B. Holzmann, and K. P. Janssen. 2006. 'Prognostic significance of downregulated expression of the candidate tumour suppressor gene SASH1 in colon cancer', *Br J Cancer*, 95: 1419-23.

Rosen, M. K., T. Yamazaki, G. D. Gish, C. M. Kay, T. Pawson, and L. E. Kay. 1995. 'Direct demonstration of an intramolecular SH2-phosphotyrosine interaction in the Crk protein', *Nature*, 374: 477-9.

Saad, S., S. R. Stanners, R. Yong, O. Tang, and C. A. Pollock. 2010. 'Notch mediated epithelial to mesenchymal transformation is associated with increased expression of the Snail transcription factor', *Int J Biochem Cell Biol*, 42: 1115-22.

Sabe, H., S. E. Shoelson, and H. Hanafusa. 1995. 'Possible v-Crk-induced transformation through activation of Src kinases', *J Biol Chem*, 270: 31219-24.

Sahlgren, C., M. V. Gustafsson, S. Jin, L. Poellinger, and U. Lendahl. 2008. 'Notch signaling mediates hypoxia-induced tumor cell migration and invasion', *Proc Natl Acad Sci U S A*, 105: 6392-7.

Sakai, R., A. Iwamatsu, N. Hirano, S. Ogawa, T. Tanaka, H. Mano, Y. Yazaki, and H. Hirai. 1994. 'A novel signaling molecule, p130, forms stable complexes in vivo with v-Crk and v-Src in a tyrosine phosphorylation-dependent manner', *EMBO J*, 13: 3748-56.

Sakakibara, A. 2002. 'Novel function of Chat in controlling cell adhesion via Cas-Crk-C3G-pathway-mediated Rap1 activation', *Journal of Cell Science*, 115: 4915-24.

Sakkab, D., M. Lewitzky, G. Posern, U. Schaeper, M. Sachs, W. Birchmeier, and S. M. Feller. 2000. 'Signaling of hepatocyte growth factor/scatter factor (HGF) to the small GTPase Rap1 via the large docking protein Gab1 and the adapter protein CRKL', *J Biol Chem*, 275: 10772-8.

Salgia, R., E. Pisick, M. Sattler, J. L. Li, N. Uemura, W. K. Wong, S. A. Burky, H. Hirai, L. B. Chen, and J. D. Griffin. 1996. 'p130CAS forms a signaling complex with the adapter protein CRKL in hematopoietic cells transformed by the BCR/ABL oncogene', *J Biol Chem*, 271: 25198-203.

Salgia, R., N. Uemura, K. Okuda, J. L. Li, E. Pisick, M. Sattler, R. de Jong, B. Druker, N. Heisterkamp, L. B. Chen, and et al. 1995. 'CRKL links p210BCR/ABL with paxillin in chronic myelogenous leukemia cells', *J Biol Chem*, 270: 29145-50.

Sanchez-Perez, I., M. Martinez-Gomariz, D. Williams, S. M. Keyse, and R. Perona. 2000. 'CL100/MKP-1 modulates JNK activation and apoptosis in response to cisplatin', *Oncogene*, 19: 5142-52.

Sanchez-Tillo, E., A. Lazaro, R. Torrent, M. Cuatrecasas, E. C. Vaquero, A. Castells, P. Engel, and A. Postigo. 2010. 'ZEB1 represses E-cadherin and induces an EMT by recruiting the SWI/SNF chromatin-remodeling protein BRG1', *Oncogene*, 29: 3490-500.

Satelli, A., A. Mitra, Z. Brownlee, X. Xia, S. Bellister, M. J. Overman, S. Kopetz, L. M. Ellis, Q. H. Meng, and S. Li. 2015. 'Epithelial-mesenchymal transitioned circulating tumor cells capture for detecting tumor progression', *Clin Cancer Res*, 21: 899-906.

Saxena, M., M. A. Stephens, H. Pathak, and A. Rangarajan. 2011. 'Transcription factors that mediate epithelial-mesenchymal transition lead to multidrug resistance by upregulating ABC transporters', *Cell Death Dis*, 2: e179.

Schaller, M. D., J. D. Hildebrand, J. D. Shannon, J. W. Fox, R. R. Vines, and J. T. Parsons. 1994. 'Autophosphorylation of the focal adhesion kinase, pp125FAK, directs SH2-dependent binding of pp60src', *Mol Cell Biol*, 14: 1680-8.

Schaller, M. D., and J. T. Parsons. 1995. 'pp125FAK-dependent tyrosine phosphorylation of paxillin creates a high-affinity binding site for Crk', *Mol Cell Biol*, 15: 2635-45.

Schlaepfer, D. D., and T. Hunter. 1996. 'Evidence for in vivo phosphorylation of the Grb2 SH2-domain binding site on focal adhesion kinase by Src-family protein-tyrosine kinases', *Mol Cell Biol*, 16: 5623-33.

Schnekenburger, M., C. Florean, M. Dicato, and M. Diederich. 2016. 'Epigenetic alterations as a universal feature of cancer hallmarks and a promising target for personalized treatments', *Curr Top Med Chem*, 16: 745-76.

Senechal, K., J. Halpern, and C. L. Sawyers. 1996. 'The CRKL adaptor protein transforms fibroblasts and functions in transformation by the BCR-ABL oncogene', *J Biol Chem*, 271: 23255-61.

Senechal, K., C. Heaney, B. Druker, and C. L. Sawyers. 1998. 'Structural requirements for function of the Crkl adapter protein in fibroblasts and hematopoietic cells', *Mol Cell Biol*, 18: 5082-90.

Seubert, B., B. Grunwald, J. Kobuch, H. Cui, F. Schelter, S. Schaten, J. T. Siveke, N. H. Lim, H. Nagase, N. Simonavicius, M. Heikenwalder, T. Reinheckel, J. P. Sleeman, K. P. Janssen, P. A. Knolle, and A. Kruger. 2015. 'Tissue inhibitor of metalloproteinases (TIMP)-1 creates a premetastatic niche in the liver through SDF-1/CXCR4-dependent neutrophil recruitment in mice', *Hepatology*, 61: 238-48.

Shaul, Y. D., and R. Seger. 2007. 'The MEK/ERK cascade: from signaling specificity to diverse functions', *Biochim Biophys Acta*, 1773: 1213-26.

Shellman, Y. G., K. A. Lambert, A. Brauweiler, P. Fain, R. A. Spritz, M. Martini, K. P. Janssen, N. F. Box, T. Terzian, M. Rewers, A. Horvath, C. A. Stratakis, W. A. Robinson, S. E. Robinson, D. A. Norris, K. B. Artinger, and T. R. Pacheco. 2015. 'SASH1 Is Involved in an Autosomal Dominant Lentiginous Phenotype', *J Invest Dermatol*, 135: 3192-4.

Sheyu, L., L. Hui, Z. Junyu, X. Jiawei, W. Honglian, S. Qing, Z. Hengwei, G. Xuhui, X. Qinghe, and H. Lin. 2013. 'Promoter methylation assay of SASH1 gene in breast cancer', *J BUON*, 18: 891-8.

Shioiri, M., T. Shida, K. Koda, K. Oda, K. Seike, M. Nishimura, S. Takano, and M. Miyazaki. 2006. 'Slug expression is an independent prognostic parameter for poor survival in colorectal carcinoma patients', *Br J Cancer*, 94: 1816-22.

Shridhar, V., J. Staub, B. Huntley, W. Cliby, R. Jenkins, H. I. Pass, L. Hartmann, and D. I. Smith. 1999. 'A novel region of deletion on chromosome 6q23.3 spanning less than 500 Kb in high grade invasive epithelial ovarian cancer', *Oncogene*, 18: 3913-8.

Siegel, R. L., K. D. Miller, and A. Jemal. 2018. 'Cancer statistics, 2018', *CA Cancer J Clin*, 68: 7-30.

Small, G. W., Y. Y. Shi, L. S. Higgins, and R. Z. Orlowski. 2007. 'Mitogen-activated protein kinase phosphatase-1 is a mediator of breast cancer chemoresistance', *Cancer Res*, 67: 4459-66.

Smit, M. A., T. R. Geiger, J. Y. Song, I. Gitelman, and D. S. Peeper. 2009. 'A Twist-Snail axis critical for TrkB-induced epithelial-mesenchymal transition-like transformation, anoikis resistance, and metastasis', *Mol Cell Biol*, 29: 3722-37.

Sperlich, A., A. Balmert, D. Doll, S. Bauer, F. Franke, G. Keller, D. Wilhelm, A. Mur, M. Respondek, H. Friess, U. Nitsche, and K. P. Janssen. 2018. 'Genetic and immunological biomarkers predict metastatic disease recurrence in stage III colon cancer', *BMC Cancer*, 18: 998.

Srikantan, V., I. A. Sesterhenn, L. Davis, G. R. Hankins, F. A. Avallone, J. R. Livezey, R. Connelly, F. K. Mostofi, D. G. McLeod, J. W. Moul, S. C. Chandrasekharappa, and S. Srivastava. 1999. 'Allelic loss on chromosome 6Q in primary prostate cancer', *Int J Cancer*, 84: 331-5.

Sriram, G., C. Reichman, A. Tunceroglu, N. Kaushal, T. Saleh, K. Machida, B. Mayer, Q. Ge, J. Li, P. Hornbeck, C. G. Kalodimos, and R. B. Birge. 2011. 'Phosphorylation of Crk on tyrosine 251 in the RT loop of the SH3C domain promotes Abl kinase transactivation', *Oncogene*, 30: 4645-55.

Strilic, B., and S. Offermanns. 2017. 'Intravascular Survival and Extravasation of Tumor Cells', *Cancer Cell*, 32: 282-93.

Sugarbaker, P. H. 2014. 'Colorectal cancer: prevention and management of metastatic disease', *Biomed Res Int*, 2014: 782890.

Summy, J. M., and G. E. Gallick. 2003. 'Src family kinases in tumor progression and metastasis', *Cancer Metastasis Rev*, 22: 337-58.

Sun, D., R. Zhou, H. Liu, W. Sun, A. Dong, and H. Zhang. 2015. 'SASH1 inhibits proliferation and invasion of thyroid cancer cells through PI3K/Akt signaling pathway', *Int J Clin Exp Pathol*, 8: 12276-83.

Szklarczyk, D., A. Franceschini, S. Wyder, K. Forslund, D. Heller, J. Huerta-Cepas, M. Simonovic, A. Roth, A. Santos, K. P. Tsafou, M. Kuhn, P. Bork, L. J. Jensen, and C. von Mering. 2015. 'STRING v10: protein-protein interaction networks, integrated over the tree of life', *Nucleic Acids Res*, 43: D447-52.

Szklarczyk, D., J. H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, A. Santos, N. T. Doncheva, A. Roth, P. Bork, L. J. Jensen, and C. von Mering. 2017. 'The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible', *Nucleic Acids Res*, 45: D362-D68.

Taddei, M. L., E. Giannoni, T. Fiaschi, and P. Chiarugi. 2012. 'Anoikis: an emerging hallmark in health and diseases', *J Pathol*, 226: 380-93.

Talipov, M. R., J. Nayak, M. Lepley, R. D. Bongard, D. S. Sem, R. Ramchandran, and R. Rathore. 2016. 'Critical Role of the Secondary Binding Pocket in Modulating the Enzymatic Activity of DUSP5 toward Phosphorylated ERKs', *Biochemistry*, 55: 6187-95.

Tanaka, S., T. Morishita, Y. Hashimoto, S. Hattori, S. Nakamura, M. Shibuya, K. Matuoka, T. Takenawa, T. Kurata, K. Nagashima, and et al. 1994. 'C3G, a guanine nucleotide-releasing protein expressed ubiquitously, binds to the Src homology 3 domains of CRK and GRB2/ASH proteins', *Proc Natl Acad Sci U S A*, 91: 3443-7.

ten Hoeve, J., C. Morris, N. Heisterkamp, and J. Groffen. 1993. 'Isolation and chromosomal localization of CRKL, a human crk-like gene', *Oncogene*, 8: 2469-74.

Theys, J., B. Jutten, R. Habets, K. Paesmans, A. J. Groot, P. Lambin, B. G. Wouters, G. Lammering, and M. Vooijs. 2011. 'E-Cadherin loss associated with EMT promotes radioresistance in human tumor cells', *Radiother Oncol*, 99: 392-97.

Thibaudeau, L., A. V. Taubenberger, B. M. Holzapfel, V. M. Quent, T. Fuehrmann, P. Hesami, T. D. Brown, P. D. Dalton, C. A. Power, B. G. Hollier, and D. W. Hutmacher. 2014. 'A tissue-engineered humanized xenograft model of human breast cancer metastasis to bone', *Dis Model Mech*, 7: 299-309.

Tian, X. J., H. Zhang, and J. Xing. 2013. 'Coupled reversible and irreversible bistable switches underlying TGFbeta-induced epithelial to mesenchymal transition', *Biophys J*, 105: 1079-89.

Tien, Y. W., K. J. Chang, Y. M. Jeng, P. H. Lee, M. S. Wu, J. T. Lin, and S. M. Hsu. 2001. 'Tumor angiogenesis and its possible role in intravasation of colorectal epithelial cells', *Clin Cancer Res*, 7: 1627-32.

Timmerman, L. A., J. Grego-Bessa, A. Raya, E. Bertran, J. M. Perez-Pomares, J. Diez, S. Aranda, S. Palomo, F. McCormick, J. C. Izpisua-Belmonte, and J. L. de la Pompa. 2004. 'Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation', *Genes Dev*, 18: 99-115.

Togel, L., R. Nightingale, R. Wu, A. C. Chueh, S. Al-Obaidi, I. Luk, M. Davalos-Salas, F. Chionh, C. Murone, D. D. Buchanan, Z. Chatterton, O. M. Sieber, D. Arango, N. C. Tebbutt, D. Williams, A. S. Dhillon, and J. M. Mariadason. 2018. 'DUSP5 is methylated in CIMP-high colorectal cancer but is not a major regulator of intestinal cell proliferation and tumorigenesis', *Sci Rep*, 8: 1767.

Toutant, M., A. Costa, J. M. Studler, G. Kadare, M. Carnaud, and J. A. Girault. 2002. 'Alternative splicing controls the mechanisms of FAK autophosphorylation', *Mol Cell Biol*, 22: 7731-43.

Tripathi, S. C., H. L. Peters, A. Taguchi, H. Katayama, H. Wang, A. Momin, M. K. Jolly, M. Celiktas, J. Rodriguez-Canales, H. Liu, C. Behrens, Wistuba, II, E. Ben-Jacob, H. Levine, J. J. Molldrem, S. M. Hanash, and E. J. Ostrin. 2016. 'Immunoproteasome deficiency is a feature of non-small cell

lung cancer with a mesenchymal phenotype and is associated with a poor outcome', *Proc Natl Acad Sci U S A*, 113: E1555-64.

Tsai, J. H., and J. Yang. 2013. 'Epithelial-mesenchymal plasticity in carcinoma metastasis', *Genes Dev*, 27: 2192-206.

Tsuda, M., and S. Tanaka. 2012. 'Roles for crk in cancer metastasis and invasion', *Genes Cancer*, 3: 334-40.

Tsuji, T., S. Ibaragi, K. Shima, M. G. Hu, M. Katsurano, A. Sasaki, and G. F. Hu. 2008. 'Epithelial-mesenchymal transition induced by growth suppressor p12CDK2-AP1 promotes tumor cell local invasion but suppresses distant colony growth', *Cancer Res*, 68: 10377-86.

Uchida, T., A. Nakao, N. Nakano, A. Kuramasu, H. Saito, K. Okumura, C. Ra, and H. Ogawa. 2001. 'Identification of Nash1, a novel protein containing a nuclear localization signal, a sterile alpha motif, and an SH3 domain preferentially expressed in mast cells', *Biochem Biophys Res Commun*, 288: 137-41.

Ueda, K., H. Arakawa, and Y. Nakamura. 2003. 'Dual-specificity phosphatase 5 (DUSP5) as a direct transcriptional target of tumor suppressor p53', *Oncogene*, 22: 5586-91.

Ungewiss, C., Z. H. Rizvi, J. D. Roybal, D. H. Peng, K. A. Gold, D. H. Shin, C. J. Creighton, and D. L. Gibbons. 2016. 'The microRNA-200/Zeb1 axis regulates ECM-dependent beta1-integrin/FAK signaling, cancer cell invasion and metastasis through CRKL', *Sci Rep*, 6: 18652.

Valastyan, S., and R. A. Weinberg. 2011. 'Tumor metastasis: molecular insights and evolving paradigms', *Cell*, 147: 275-92.

Verdugo, R. A., T. Zeller, M. Rotival, P. S. Wild, T. Munzel, K. J. Lackner, H. Weidmann, E. Ninio, D. A. Tregouet, F. Cambien, S. Blankenberg, and L. Tiret. 2013. 'Graphical modeling of gene expression in monocytes suggests molecular mechanisms explaining increased atherosclerosis in smokers', *PLoS One*, 8: e50888.

Vogelstein, B., E. R. Fearon, S. R. Hamilton, S. E. Kern, A. C. Preisinger, M. Leppert, Y. Nakamura, R. White, A. M. Smits, and J. L. Bos. 1988. 'Genetic alterations during colorectal-tumor development', *N Engl J Med*, 319: 525-32.

von Holleben, M., A. Gohla, K. P. Janssen, B. M. Iritani, and S. Beer-Hammer. 2011. 'Immunoinhibitory adapter protein Src homology domain 3 lymphocyte protein 2 (SLy2) regulates actin dynamics and B cell spreading', *J Biol Chem*, 286: 13489-501.

von Mering, C., M. Huynen, D. Jaeggi, S. Schmidt, P. Bork, and B. Snel. 2003. 'STRING: a database of predicted functional associations between proteins', *Nucleic Acids Res*, 31: 258-61.

Wada, T., and J. M. Penninger. 2004. 'Mitogen-activated protein kinases in apoptosis regulation', *Oncogene*, 23: 2838-49.

Wang, H., H. S. Wang, B. H. Zhou, C. L. Li, F. Zhang, X. F. Wang, G. Zhang, X. Z. Bu, S. H. Cai, and J. Du. 2013. 'Epithelial-mesenchymal transition (EMT) induced by TNF-alpha requires AKT/GSK-3beta-mediated stabilization of snail in colorectal cancer', *PLoS One*, 8: e56664.

Wang, Y., J. Liu, X. Ying, P. C. Lin, and B. P. Zhou. 2016. 'Twist-mediated Epithelial-mesenchymal Transition Promotes Breast Tumor Cell Invasion via Inhibition of Hippo Pathway', *Sci Rep*, 6: 24606.

Wang, Z., P. S. Reinach, F. Zhang, K. S. Vellonen, A. Urtti, H. Turner, and J. M. Wolosin. 2010. 'DUSP5 and DUSP6 modulate corneal epithelial cell proliferation', *Mol Vis*, 16: 1696-704.

Watanabe, T., M. Tsuda, Y. Makino, S. Ichihara, H. Sawa, A. Minami, N. Mochizuki, K. Nagashima, and S. Tanaka. 2006. 'Adaptor molecule Crk is required for sustained phosphorylation of Grb2-associated binder 1 and hepatocyte growth factor-induced cell motility of human synovial sarcoma cell lines', *Mol Cancer Res*, 4: 499-510.

Watanabe, T., M. Tsuda, Y. Makino, T. Konstantinou, H. Nishihara, T. Majima, A. Minami, S. M. Feller, and S. Tanaka. 2009. 'Crk adaptor protein-induced phosphorylation of Gab1 on tyrosine 307 via Src is important for organization of focal adhesions and enhanced cell migration', *Cell Res*, 19: 638-50.

Weidmann, H., Z. Touat-Hamici, H. Durand, C. Mueller, S. Chardonnet, C. Pionneau, F. Charlotte, K. P. Janssen, R. Verdugo, F. Cambien, S. Blankenberg, L. Tiret, T. Zeller, and E. Ninio. 2015. 'SASH1, a new potential link between smoking and atherosclerosis', *Atherosclerosis*, 242: 571-9.

Weigelt, B., J. L. Peterse, and L. J. van 't Veer. 2005. 'Breast cancer metastasis: markers and models', *Nat Rev Cancer*, 5: 591-602.

Widmann, C., S. Gibson, M. B. Jarpe, and G. L. Johnson. 1999. 'Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human', *Physiol Rev*, 79: 143-80.

Woodcock, S. A., C. Rooney, M. Liontos, Y. Connolly, V. Zoumpourlis, A. D. Whetton, V. G. Gorgoulis, and A. Malliri. 2009. 'SRC-induced disassembly of adherens junctions requires localized phosphorylation and degradation of the rac activator tiam1', *Mol Cell*, 33: 639-53.

Wu, Z. Q., X. Y. Li, C. Y. Hu, M. Ford, C. G. Kleer, and S. J. Weiss. 2012. 'Canonical Wnt signaling regulates Slug activity and links epithelial-mesenchymal transition with epigenetic Breast Cancer 1, Early Onset (BRCA1) repression', *Proc Natl Acad Sci U S A*, 109: 16654-9.

Xie, L., B. K. Law, A. M. Chytil, K. A. Brown, M. E. Aakre, and H. L. Moses. 2004. 'Activation of the Erk pathway is required for TGF-beta1-induced EMT in vitro', *Neoplasia*, 6: 603-10.

Xing, Q. H., M. T. Wang, X. D. Chen, G. Y. Feng, H. Y. Ji, J. D. Yang, J. J. Gao, W. Qin, X. Q. Qian, S. N. Wu, and L. He. 2003. 'A gene locus responsible for dyschromatosis symmetrica hereditaria (DSH) maps to chromosome 6q24.2-q25.2', *Am J Hum Genet*, 73: 377-82.

Yamamura, T., S. Tsukikawa, K. Yamada, and S. Yamaguchi. 2001. 'Morphologic analysis of microvessels in colorectal tumors with respect to the formation of liver metastases', *J Surg Oncol*, 78: 259-64.

Yan, X., L. Liu, H. Li, L. Huang, M. Yin, C. Pan, H. Qin, and Z. Jin. 2016. 'Dual specificity phosphatase 5 is a novel prognostic indicator for patients with advanced colorectal cancer', *Am J Cancer Res*, 6: 2323-33.

Yan, X., L. Yan, S. Liu, Z. Shan, Y. Tian, and Z. Jin. 2015. 'N-cadherin, a novel prognostic biomarker, drives malignant progression of colorectal cancer', *Mol Med Rep*, 12: 2999-3006.

Yanagi, H., L. Wang, H. Nishihara, T. Kimura, M. Tanino, T. Yanagi, S. Fukuda, and S. Tanaka. 2012. 'CRKL plays a pivotal role in tumorigenesis of head and neck squamous cell carcinoma through the regulation of cell adhesion', *Biochem Biophys Res Commun*, 418: 104-9.

Yang, F., K. M. Waters, J. H. Miller, M. A. Gritsenko, R. Zhao, X. Du, E. A. Livesay, S. O. Purvine, M. E. Monroe, Y. Wang, D. G. Camp, 2nd, R. D. Smith, and D. L. Stenoien. 2010. 'Phosphoproteomics profiling of human skin fibroblast cells reveals pathways and proteins affected by low doses of ionizing radiation', *PLoS One*, 5: e14152.

Yang, L., M. Liu, Z. Gu, J. Chen, Y. Yan, and J. Li. 2012. 'Overexpression of SASH1 related to the decreased invasion ability of human glioma U251 cells', *Tumour Biol*, 33: 2255-63.

Yang, L., H. Zhang, Q. Yao, Y. Yan, R. Wu, and M. Liu. 2015. 'Clinical Significance of SASH1 Expression in Glioma', *Dis Markers*, 2015: 383046.

Yang, Z., H. M. Kirton, M. Al-Owais, J. Thireau, S. Richard, C. Peers, and D. S. Steele. 2017. 'Epac2-Rap1 Signaling Regulates Reactive Oxygen Species Production and Susceptibility to Cardiac Arrhythmias', *Antioxid Redox Signal*, 27: 117-32.

Yao, D., C. Dai, and S. Peng. 2011. 'Mechanism of the mesenchymal-epithelial transition and its relationship with metastatic tumor formation', *Mol Cancer Res*, 9: 1608-20.

Yilmaz, M., and G. Christofori. 2009. 'EMT, the cytoskeleton, and cancer cell invasion', *Cancer Metastasis Rev*, 28: 15-33.

Yook, J. I., X. Y. Li, I. Ota, E. R. Fearon, and S. J. Weiss. 2005. 'Wnt-dependent regulation of the E-cadherin repressor snail', *J Biol Chem*, 280: 11740-8.

Yoon, S., and R. Seger. 2006. 'The extracellular signal-regulated kinase: multiple substrates regulate diverse cellular functions', *Growth Factors*, 24: 21-44.

Yuan, Y., Y. C. Jiang, C. K. Sun, and Q. M. Chen. 2016. 'Role of the tumor microenvironment in tumor progression and the clinical applications (Review)', *Oncol Rep*, 35: 2499-515.

Zeisberg, M., and E. G. Neilson. 2009. 'Biomarkers for epithelial-mesenchymal transitions', *J Clin Invest*, 119: 1429-37.

Zeller, C., B. Hinzmann, S. Seitz, H. Prokoph, E. Burkhard-Goettges, J. Fischer, B. Jandrig, L. E. Schwarz, A. Rosenthal, and S. Scherneck. 2003. 'SASH1: a candidate tumor suppressor gene on chromosome 6q24.3 is downregulated in breast cancer', *Oncogene*, 22: 2972-83.

Zhang, G., D. Fenyo, and T. A. Neubert. 2008. 'Screening for EphB signaling effectors using SILAC with a linear ion trap-orbitrap mass spectrometer', *J Proteome Res*, 7: 4715-26.

Zhang, G., D. S. Spellman, E. Y. Skolnik, and T. A. Neubert. 2006. 'Quantitative phosphotyrosine proteomics of EphB2 signaling by stable isotope labeling with amino acids in cell culture (SILAC)', *J Proteome Res*, 5: 581-8.

Zhang, J., X. Gao, F. Schmit, G. Adelmant, M. J. Eck, J. A. Marto, J. J. Zhao, and T. M. Roberts. 2017. 'CRKL Mediates p110beta-Dependent PI3K Signaling in PTEN-Deficient Cancer Cells', *Cell Rep*, 20: 549-57.

Zhang, J. H., T. D. Chung, and K. R. Oldenburg. 1999. 'A Simple Statistical Parameter for Use in Evaluation and Validation of High Throughput Screening Assays', *J Biomol Screen*, 4: 67-73.

Zhang, Y. L., R. C. Wang, K. Cheng, B. Z. Ring, and L. Su. 2017. 'Roles of Rap1 signaling in tumor cell migration and invasion', *Cancer Biol Med*, 14: 90-99.

Zhang, Y., K. Weber-Matthiesen, R. Siebert, P. Matthiesen, and B. Schlegelberger. 1997. 'Frequent deletions of 6q23-24 in B-cell non-Hodgkin's lymphomas detected by fluorescence in situ hybridization', *Genes Chromosomes Cancer*, 18: 310-3.

Zhao, Y., X. Li, X. Sun, Y. Zhang, and H. Ren. 2012. 'EMT phenotype is induced by increased Src kinase activity via Src-mediated caspase-8 phosphorylation', *Cell Physiol Biochem*, 29: 341-52.

Zheng, L., U. Baumann, and J. L. Reymond. 2004. 'An efficient one-step site-directed and site-saturation mutagenesis protocol', *Nucleic Acids Res*, 32: e115.

Zheng, X., J. L. Carstens, J. Kim, M. Scheible, J. Kaye, H. Sugimoto, C. C. Wu, V. S. LeBleu, and R. Kalluri. 2015. 'Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer', *Nature*, 527: 525-30.

Zhou, D., Z. Wei, S. Deng, T. Wang, M. Zai, H. Wang, L. Guo, J. Zhang, H. Zhong, L. He, and Q. Xing. 2013. 'SASH1 regulates melanocyte transepithelial migration through a novel Galphas-SASH1-IQGAP1-E-Cadherin dependent pathway', *Cell Signal*, 25: 1526-38.

Zhou, N., C. Liu, X. Wang, Q. Mao, Q. Jin, and P. Li. 2018. 'Downregulated SASH1 expression indicates poor clinical prognosis in gastric cancer', *Hum Pathol*, 74: 83-91.

Zhu, Y., Y. Ma, H. Peng, L. Gong, M. Xiao, L. Xiang, D. He, and K. Cao. 2019. 'MiR-130b promotes the progression of oesophageal squamous cell carcinoma by targeting SASH1', *J Cell Mol Med*, 23: 93-103.

Zong, W., C. Yu, P. Wang, and L. Dong. 2016. 'Overexpression of SASH1 Inhibits TGF-beta1-Induced EMT in Gastric Cancer Cells', *Oncol Res*, 24: 17-23.

6. Appendix

6.1 Supplementary data

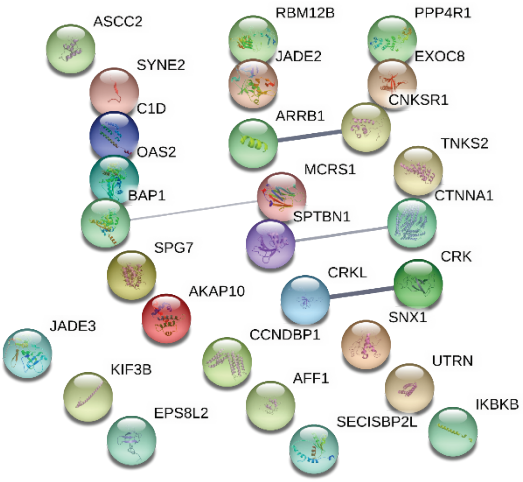

Figure S 1: Protein-protein association network of putative physical interaction partners of SASH1. The network was generated with the STRING database v11.0 using yeast two-hybrid hits as input. Line thickness indicates the strength of data support (protein-protein interaction p=0.601).

Table S 1: Clinical and pathological parameters of the UICC stage III colorectal cancer patient subgroup that received 5-fluorouracil as adjuvant (mono-)chemotherapy. All patients underwent tumor resection at the Department of Surgery (Klinikum rechts der Isar, Technical University of Munich, Germany). The study was approved by the Ethics committee of the Faculty of Medicine of the Technical University of Munich (1926/07), and all patients gave informed consent prior to the use of the tissue.

Gender (n)	Male	17
	Female	8
Age (years)	Median	65
Grading (n)	G1/2	15
	G3/4	10
T-stage (n)	T2	1
	T3	16
	T4	8
N-stage (n)	N1	16
	N2	9

Table S 2: Clinical and pathological parameters of the patient cohort with UICC stage III colorectal cancer. All patients underwent tumor resection at the Department of Surgery (Klinikum rechts der Isar, Technical University of Munich, Germany). The study was approved by the Ethics committee of the Faculty of Medicine of the Technical University of Munich (1926/07), and all patients gave informed consent prior to the use of the tissue.

Gender (n)	Male	44
	Female	21
Age (years)	Median	67
Grading (n)	G1/2	37
	G3/4	28
T-stage (n)	T2	6
	T3	43
	T4	16
N-stage (n)	N1	45
	N2	20

Table S 3: Hallmarks downregulated in CRK family-deficient cells (GSEA analysis).

Pathway Name	Pathway Size	Enrichment Score	Normalized Enrichment Score	p-value	q-value	Familywise-Error Rate p-value
HALLMARK_MYC_TARGETS_V1	183	-0.771	-3.664	0.000	0.000	0.000
HALLMARK_E2F_TARGETS	194	-0.717	-3.443	0.000	0.000	0.000
HALLMARK_G2M_CHECKPOINT	187	-0.662	-3.146	0.000	0.000	0.000
HALLMARK_MYC_TARGETS_V2	57	-0.798	-3.127	0.000	0.000	0.000
HALLMARK_MTORC1_SIGNALING	174	-0.614	-2.879	0.000	0.000	0.000
HALLMARK_UNFOLDED_PROTEIN_RESPONSE	102	-0.533	-2.314	0.000	0.000	0.000
HALLMARK_OXIDATIVE_PHOSPHORYLATION	185	-0.480	-2.267	0.000	0.000	0.000
HALLMARK_DNA_REPAIR	130	-0.474	-2.157	0.000	0.000	0.000
HALLMARK_UV_RESPONSE_UP	125	-0.329	-1.494	0.000	0.024	0.232
HALLMARK_REACTIVE_OXIGEN_SPECIES_PATHWAY	36	-0.401	-1.397	0.057	0.049	0.465
HALLMARK_PI3K_AKT_MTOR_SIGNALING	79	-0.338	-1.394	0.030	0.045	0.471
HALLMARK_FATTY_ACID_METABOLISM	117	-0.301	-1.342	0.037	0.060	0.608
HALLMARK_SPERMATOGENESIS	63	-0.314	-1.251	0.129	0.112	0.850
HALLMARK_PEROXISOME	78	-0.238	-0.988	0.473	0.510	1.000
HALLMARK_PROTEIN_SECRETION	86	-0.155	-0.650	0.995	0.985	1.000

Table S 4: Hallmarks upregulated in CRK family-deficient cells (GSEA analysis).

Pathway Name	Pathway Size	Enrichment Score	Normalized Enrichment Score	p-value	q-value	Familywise-Error Rate p-value
HALLMARK_KRAS_SIGNALING_UP	93	0.622	2.465	0.000	0.000	0.000
HALLMARK_COAGULATION	61	0.618	2.309	0.000	0.000	0.000
HALLMARK_TNFA_SIGNALING_VIA_NFKB	148	0.537	2.296	0.000	0.000	0.000
HALLMARK_COMPLEMENT	116	0.509	2.112	0.000	0.000	0.000
HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	111	0.483	1.987	0.000	0.001	0.003
HALLMARK_APICAL_SURFACE	24	0.620	1.917	0.002	0.001	0.003
HALLMARK_KRAS_SIGNALING_DN	60	0.516	1.899	0.000	0.001	0.004
HALLMARK_HYPOXIA	144	0.447	1.886	0.000	0.001	0.006
HALLMARK_MYOGENESIS	104	0.451	1.837	0.000	0.001	0.010
HALLMARK_IL2_STAT5_SIGNALING	133	0.418	1.777	0.000	0.002	0.022
HALLMARK_INTERFERON_ALPHA_RESPONSE	56	0.487	1.772	0.002	0.002	0.023
HALLMARK_ANGIOGENESIS	20	0.599	1.719	0.009	0.005	0.051
HALLMARK_APICAL_JUNCTION	117	0.415	1.718	0.000	0.005	0.052
HALLMARK_HEME_METABOLISM	145	0.403	1.715	0.000	0.004	0.055
HALLMARK_IL6_JAK_STAT3_SIGNALING	46	0.482	1.713	0.005	0.004	0.055
HALLMARK_WNT_BETA_CATENIN_SIGNALING	31	0.526	1.708	0.013	0.004	0.056
HALLMARK_ESTROGEN_RESPONSE_LATE	149	0.389	1.654	0.000	0.006	0.085
HALLMARK_ESTROGEN_RESPONSE_EARLY	156	0.374	1.609	0.001	0.009	0.141
HALLMARK_XENOBIOTIC_METABOLISM	124	0.381	1.591	0.002	0.011	0.172
HALLMARK_APOPTOSIS	119	0.365	1.516	0.009	0.022	0.321
HALLMARK_HEDGEHOG_SIGNALING	20	0.527	1.510	0.048	0.022	0.338
HALLMARK_ANDROGEN_RESPONSE	85	0.368	1.449	0.025	0.036	0.529
HALLMARK_NOTCH_SIGNALING	27	0.461	1.440	0.049	0.038	0.557
HALLMARK_P53_PATHWAY	171	0.327	1.426	0.014	0.041	0.612
HALLMARK_ALLOGRAFT_REJECTION	74	0.366	1.403	0.035	0.049	0.681
HALLMARK_INTERFERON_GAMMA_RESPONSE	113	0.335	1.378	0.045	0.058	0.767
HALLMARK_UV_RESPONSE_DN	108	0.324	1.321	0.060	0.086	0.893

Pathway Name	Pathway Size	Enrichment Score	Normalized Enrichment Score	p-value	q-value	Familywise-Error Rate p-value
HALLMARK_GLYCOLYSIS	165	0.297	1.295	0.052	0.100	0.927
HALLMARK_TGF_BETA_SIGNALING	45	0.364	1.275	0.140	0.112	0.951
HALLMARK_CHOLESTEROL_HOMEOSTASIS	62	0.294	1.077	0.319	0.365	1.000
HALLMARK_BILE_ACID_METABOLISM	65	0.278	1.036	0.421	0.433	1.000
HALLMARK_INFLAMMATORY_RESPONSE	89	0.248	0.985	0.482	0.532	1.000
HALLMARK_ADIPOGENESIS	156	0.213	0.919	0.657	0.665	1.000
HALLMARK_MITOTIC_SPINDLE	187	0.200	0.888	0.739	0.712	1.000
HALLMARK_PANCREAS_BETA_CELLS	17	0.307	0.861	0.644	0.749	1.000

6.2 Materials

Table 2: Important buffer compositions.

BES buffer (2X)	50 mM BES (N,N-bis-(2-hydroxyethyl)-2-aminoethane sulfonic acid), 410 mM NaCl, 1.5 mM Na_2HPO_4, pH 6.96
HIS elution buffer	250 mM imidazole in PBS (pH 7.4)
HIS lysis buffer	10 mM imidazole in PBS (pH 7.4), 1X protease inhibitor cocktail (Roche), 1 mM PMSF, 0.5 mg mL^{-1} lysozyme, 50 U ml^{-1} Benzonase
HIS wash buffer	25 mM imidazole in PBS (pH 7.4)
Laemmli buffer (5X)	1.5 M Tris-HCl (pH 6.8), 10% (w/v) SDS, 50% (v/v) glycerol, 77 mg ml^{-1} DTT, 0.01 % (w/v) Bromophenol blue
PBST	PBS, 0.1% (v/v) Tween-20
RIPA buffer	50 mM Tris-HCl (pH 7.4), 150 mM NaCl, 1% (v/v) NP-40, 0.25% (v/v) sodium deoxycholate, 1 mM EDTA (pH 8.0), optional: 0.1% (v/v) SDS

Table 3: Oligonucleotides.

Oligo ID	Purpose	Oligonucleotide sequence
#1	Cloning of CRKL	5’-CAAGGTACCATGTCCTCCGCCAGG-3’
#2		5’-GCCCGGGACTCGTTTTCATCTGGGTTTTG-3’
#3	Cloning of SH3N	5’-GCACATATGACAGCAGAAGATAACCTG-3’
#4		5’-CACTCGAGTGAGGATCTCACAAGC-3’
#5	Mut. of PXXP1	5’-GATTGTAGCAGAAGTGGCACAGAAGACGACCGCC-3’
#6		5’-CTGTGCCACTTCTGCTACAATCTGAGGGGGCGG-3’
#7	Mut. of PXXP2	5’-CAGCCTGCACCTGTTGCAGCCAAAAAGAGCAGAGAACG-3’
#8		5’-CTTTTTGGCTGCAACAGGTGCAGGCTGTGATGGAATTTTGG-3’
#9	Mut. of PXXP3	5’-GATGCGGCATGCCTGGCAGTGAAAAGGGGCAGCCCCGC-3’
#10		5’-CTTTTCACTGCCAGGCATGCCGCATCGGGACTGGGGAG-3’
#11	Mut. of W160R	5’-GAACAGCGGTGGAGTGCCCGGAACAAGG-3’
#12		5’-GCACTCCACCGCTGTTCTTCAGGCTTCTC-3’
#13	SASH1 guide RNA 1	5’-CACCGCGACTCTGGACCGACGTGAT-3’
#14		5’-AAACATCACGTCGGTCCAGAGTCGC-3’
#15	SASH1 guide RNA 2	5’-CACCGTTTCTCCGACGTGTGCGAG-3’
#16		5’-AAACCTCGCACACGTCGGAGAAAC-3’
#17	CRKL guide RNA 1	5’-CACCGTCCGAGGAGTCGAACCTGG-3’
#18		5’-AAACCCAGGTTCGACTCCTCGGAC-3’

Oligo ID	Purpose	Oligonucleotide sequence
#19	CRKL guide RNA 2	5’-CACCGCGAGGAGTCGAACCTGGCGG-3’
#20		5’-AAACCCGCCAGGTTCGACTCCTCGC-3’
#21	CRK guide RNA 1	5’-CACCGGCGGGCAACTTCGACTCGG-3’
#22		5’-AAACCCGAGTCGAAGTTGCCCGCC-3’
#23	CRK guide RNA 2	5’-CACCGTGGTGCTCGAGTCCCGCACC-3’
#24		5’-AAACGGTGCGGGACTCGAGCACCAC-3’
#25	CSK guide RNA 1	5’-CACCGCAATACATTCTGTACCGGA-3’
#26		5’-AAACTCCGGTACAGAATGTATTGC-3’
#27	CSK guide RNA 2	5’-CACCGAGTGCCGTGGAAGTTGTACT-3’
#28		5’-AAACAGTACAACTTCCACGGCACTC-3’
#29	CRISPR/Cas9 target locus amplification	5’-CACATCCGAGGCGTTCTC-3’
#30		5’-TGTGTGTATTTAGCCCCCTAAG-3’
#31	CRISPR/Cas9 target locus amplification	5’-GCGCAGCAGTATGCAGATTA-3’
#32		5’-GTGCAGCAGCCTACAGATTG-3’
#33	Genotyping of *Crkl*^{flox}	5’-TGTTTATCAGCCAGGAGAGGA-3’
#34		5’-GCTCTCGTCCTGCATAGGTC-3’
#35	*CDH1* qPCR (UPL #84)	5’-TGGAGGAATTCTTGCTTTGC-3’
#36		5’-CGCTCTCCTCCGAAGAAAC-3’
#37	*CRKL* qPCR (UPL #77)	5’-GGCCCTGCTGGAGTTTTAC-3’
#38		5’-TTGGTGGGCTTGGATACCT-3’
#39	*SASH1* qPCR (UPL #25)	5’-CTTGGCACAGGACTGAGGA-3’
#40		5’-GGTCAAAGAGAACCGCACTAA-3’
#41	*VIM* qPCR (UPL #39)	5’-GACCAGCTAACCAACGACAAA-3’
#42		5’-GAAGCATCTCCTCCTGCAAT-3’
#43	*ZEB1* qPCR (UPL #34)	5’-TTTTTCCTGAGGCACCTGAA-3’
#44		5’-TGAAAATGCATCTGGTGTTCC-3’
#45	*ZEB2* qPCR (UPL #06)	5’-GCTCGCACTACAATGCATCA-3’
#46		5’-GGGAAATTGATGAATAGCGAAA-3’
#47	*CRK* qPCR (UPL #81)	5’-AACAGGCAGCGCTACTCAAG-3’
#48		5’-TCCAGCCCAGTGGTTCAT-3’
#49	*TWIST1* qPCR (UPL #50)	5’-TCCAGAGAAGGAGAAAATGGAC-3’
#50		5’-TTTCCAAGAAAATCTTTGGCATA-3’
#51	*SNAI1* qPCR (UPL #30)	5’-AGGATCTCCAGGCTCGAAAG-3’
#52		5’-TCGGATGTGCATCTTGAGG-3’
#53	*HPRT* qPCR (UPL #22)	5’-GACCAGTCAACAGGGGACAT-3’
#54		5’-GTGTCAATTATATCTTCCACAATCAAG-3’
#55	Cloning of HAP	5’-AATTCTGCGTGGATAACAGCCCGCCGCCGGCGCTGCCGCCGAAACGCCGCCGCAGCGCGCCGAGCTAGG-3’
#56		5’-GATCCCTAGCTCGGCGCGCTGCGGCGGCGTTTCGGCGGCAGCGCCGGCGGCGGGCTGTTATCCACGCAG-3’

Oligo ID	Purpose	Oligonucleotide sequence
#57	Mutagenesis of NLS1	5’-CCACTCGCAGAGTCGCTGCGGCACTAATTAGGGTGG-3’
#58		5’-CCACCCTAATTAGTGCCGCAGCGACTCTGCGAGTGG-3’
#59	Mutagenesis of NLS2	5’-CGCATCTCTCTTGGGGCAGCGGTGAAATCAGTG-3’
#60		5’-CACTGATTTCACCGCTGCCCCAAGAGAGATGCG-3’
#61	Mut. of R39K	5’-CTCGTCAAGGATTCTTCCACCTGCCCTG-3’
#62		5’-GAAGAATCCTTGACGAGGAACATACCGTGGC-3’

Table 4: Primary and secondary antibodies.

Primary antibodies

C3G	Bethyl A301-965A
c-MYC	SCBT sc-40
CRK-II	SCBT sc-289
CRKL (mAb)	EMD Millipore 05-414
CRKL (pAb)	SCBT sc-319
CSK	Abcam ab125005
DUSP5	SCBT sc-393801
E-cadherin	CST #3195 and Abcam ab40772
ERK	CST #4695
FAK	CST #13009
FLAG	Sigma F1804
GAB1	Bethyl A303-288A
GFP	Chromotek 3H9
Human mitochondria	Abcam ab92824
Integrin beta-1	Abcam ab24693
p130CAS	CST #13846
PARP1	Abcam ab137653
Paxillin	CST #12065
RFP	Chromotek 5F8
SASH1	Bethyl A302-265A
SRC	CST #2123

T202/Y204 phospho-ERK1/2	CST #9101
Talin	Abcam ab17333
V5 epitope	Thermo Scientific MA5-15253
Vimentin	Abcam ab92547
Y118 phospho-Paxillin	CST #2541
Y397 phospho-FAK	CST #8556
Y410 phospho-p130CAS	CST #4011
Y419 phospho-SRC	CST #6943
Y576/Y577 phospho-FAK	CST #3281
Y925 phospho-FAK	CST #3284
ZEB1	Sigma HPA027524
β-actin	CST #3700

Secondary antibodies

Cy3 AffiniPure Goat Anti-Mouse IgG (H+L)	Jackson 115-165-003
FITC AffiniPure Rat Anti-Mouse IgG (H+L)	Jackson 415-095-100
Goat anti-Mouse IgG (H+L) Alexa Fluor 488	Thermo Scientific A32723
Goat anti-Rabbit IgG (H+L) Alexa Fluor 488	Thermo Scientific A11008
Peroxidase-AffiniPure Goat Anti-Mouse IgG (H+L)	Jackson 115-035-003
Peroxidase-AffiniPure Goat Anti-Rabbit IgG (H+L)	Jackson 111-035-144
Peroxidase-AffiniPure Mouse Anti-Rat IgG (H+L)	Jackson 212-036-102

Table 5: Important chemicals.

1-Ethyl-3-(3-dimethylaminopropyl)carbodiimide	Sigma-Aldrich, Stammheim, Germany
5-fluorouracil	Pharmacy, MRI, Munich, Germany
Acrylamide solution	BioRad, Munich, Germany
Adenosine triphosphate	New England Biolabs, Ipswich, USA
Ammonium persulfate	Sigma-Aldrich, Stammheim, Germany
Ampicillin	Sigma-Aldrich, Stammheim, Germany
Attractene	Qiagen, Hilden, Germany
Benzamidine	Sigma-Aldrich, Stammheim, Germany
Blotting grade milk powder	Carl Roth, Karlsruhe, Germany
Bovine serum albumin (BSA)	Sigma-Aldrich, Stammheim, Germany
Bromophenol blue	Merck, Darmstadt, Germany
Crystal violet solution	Sigma-Aldrich, Stammheim, Germany
DAB chromogen	Agilent Technologies, Santa Clara, USA
DAPI	Thermo Fisher, Waltham, USA
Deoxynucleotides	Fermentas, St. Leon-Rot, Germany
D-luciferine	PerkinElmer, Waltham, USA
DMSO	Merck, Darmstadt, Germany
DpnI restriction enzyme	Thermo Fisher, Waltham, USA
Dulbecco's Modified Eagle Medium	Thermo Fisher, Waltham, USA
Dulbecco's Phosphate Buffered Saline (PBS)	Thermo Fisher, Waltham, USA
ECL substrate	Pierce, Bonn, Germany
EnVision+ System	Agilent Technologies, Santa Clara, USA
Ethidium bromide	Carl Roth, Karlsruhe, Germany
Eukitt	Sigma-Aldrich, Stammheim, Germany
FastDigest restriction enzymes	Thermo Fisher, Waltham, USA
Fetal calf serum (FCS)	Biochrom, Berlin, Deutschland
Fibronectin (human)	Roche, Mannheim, Germany
FuGENE HD	Promega, Madison, USA

Gel Loading Dye, blue 6x	Fermentas, St. Leon-Rot, Germany
GeneRuler 1 kb DNA ladder	Fermentas, St. Leon-Rot, Germany
GeneRuler 100 bp DNA ladder	Fermentas, St. Leon-Rot, Germany
HAP peptide	JPT Peptides Technologies, Berlin, Germany
HAP-CPP peptide dimer	Genscript, Piscataway, USA
HisPur Ni-NTA Resin	Thermo Fisher, Waltham, USA
IPTG	Sigma-Aldrich, Stammheim, Germany
JumpStart REDTaq ReadyMix Reaction Mix	Sigma-Aldrich, Stammheim, Germany
Kanamycin	Sigma-Aldrich, Stammheim, Germany
LB medium	Institute for Medical Microbiology, Immunology and Hygiene, Munich, Germany
L-glutamine	Biochrom, Berlin, Deutschland
Matrigel	Sigma-Aldrich, Stammheim, Germany
N,N-Bis(2-hydroxyethyl)-2-aminoethanesulfonic acid	Sigma-Aldrich, Stammheim, Germany
N-hydroxysulfosuccinimide	Thermo Fisher, Waltham, USA
Nonidet P40	Thermo Fisher, Waltham, USA
Oligo(dT)18 Primer	Fermentas, St. Leon-Rot, Germany
Opti-MEM	Thermo Fisher, Waltham, USA
Oxaliplatin	Pharmacy, MRI, Munich, Germany
PageRuler Prestained Protein Ladder	Thermo Fisher, Waltham, USA
Paraformaldehyde	Sigma-Aldrich, Stammheim, Germany
PBS	Biochrom, Berlin, Germany
Pefabloc	Roche, Mannheim, Germany
Penicillin/Streptomycin	Biochrom, Berlin, Germany
Pepstatin A	Invitrogen, Karlsruhe, Germany
Phusion Hot start II DNA polymerase	Thermo Fisher, Waltham, USA
Plasmid-Safe ATP-Dependent DNase	Lucigen (via Biozym Scientific), Hessisch Oldendorf, Germany
PMSF (Phenylmethanesulfonyl fluoride)	Sigma-Aldrich, Stammheim, Germany

ProLong Gold Antifade	Thermo Fisher, Waltham, USA
Protease Inhibitor Cocktail (PIC)	Roche, Mannheim, Germany
Protein A/G Sepharose	Sigma-Aldrich, Stammheim, Germany
Proteinase K	Sigma-Aldrich, Stammheim, Germany
Puromycin	Sigma-Aldrich, Stammheim, Germany
Random Hexamer Primer	Fermentas, St. Leon-Rot, Germany
Retroviral particles (LPP-hLUC-Lv201-025-C)	GeneCopoeia, Rockville, USA
RevertAid H Minus Reverse Transcriptase	Thermo Fisher, Waltham, USA
RiboLock RNase Inhibitor	Thermo Fisher, Waltham, USA
Roticlear	Carl Roth, Karlsruhe, Germany
SOC medium	Institute for Medical Microbiology, Immunology and Hygiene, Munich, Germany
Sodium deoxycholate	Sigma-Aldrich, Stammheim, Germany
Sodium dodecyl sulfate (SDS)	Carl Roth, Karlsruhe, Germany
Sodium fluoride	Sigma-Aldrich, Stammheim, Germany
Sodium orthovanadate	Sigma-Aldrich, Stammheim, Germany
ß-Glycerophosphate	Sigma-Aldrich, Stammheim, Germany
ß-Mercaptoethanol	Merck, Darmstadt, Germany
T4 Ligase	New England Biolabs, Ipswich, USA
T4 Polynucleotide Kinase	New England Biolabs, Ipswich, USA
T7 Ligase	New England Biolabs, Ipswich, USA
Tango buffer	Thermo Fisher, Waltham, USA
TEMED	Bio-Rad, Munich, Germany
TGF-beta1 (human)	Stemcell technologies, Vancouver, Canada
Tissue-Tek	Sakura, Zoeterwoude, Netherlands
TNF (human)	Thermo Fisher, Waltham, USA
TRITC-phalloidin	Sigma-Aldrich, Stammheim, Germany
Triton X-100	Carl Roth, Karlsruhe, Germany
Trypan blue	Biochrom, Berlin, Germany

Trypsin	Sigma-Aldrich, Stammheim, Germany
Tween 20	Merck, Darmstadt, Germany

Table 6: Important consumables.

Amersham Hyperfilm	GE Healthcare Life Sciences, Munich, Germany
Cell culture plates	Corning, Corning, USA
Culture-Insert 2 Well in µ-Dish 35 mm	Ibidi, Martinsried, Germany
Enspire-LFB, 384-well High Sensitivity, User-Activated Biochemical Plates, Corning Epic System	Corning, Corning, USA
Spectra/Por membrane tubing (MWCO: 3500)	Carl Roth, Karlsruhe, Germany
Whatman Blotting Paper	GE Healthcare Life Sciences, Munich, Germany
8.0 µm Transwell Permeable Supports #3422	Corning, Corning, USA
8.0 µm Matrigel Invasion Chamber #354480	Corning, Corning, USA
Amersham Protran nitrocellulose membrane (0.45 µm)	GE Healthcare Life Sciences, Munich, Germany

Table 7: Commercial kits.

Cell Proliferation Kit II (XTT)	Sigma-Aldrich, Stammheim, Germany
LightCycler 480 Probes Master	Roche, Mannheim, Germany
Pierce BCA Protein Assay Kit	Thermo Fisher, Waltham, USA
QIAGEN Plasmid Mini/Maxi Kit	Qiagen, Hilden, Germany
QIAprep Spin Miniprep Kit	Qiagen, Hilden, Germany
QIAquick Gel Extraction Kit	Qiagen, Hilden, Germany
QIAquick PCR Purification Kit	Qiagen, Hilden, Germany
RNeasy Mini Kit	Qiagen, Hilden, Germany
Universal ProbeLibrary Set	Sigma-Aldrich, Stammheim, Germany

Table 8: Bacteria, cell lines and mouse model.

E. coli One Shot BL21 Star DE3	Thermo Fisher, Waltham, USA
E. coli TOP10	Thermo Fisher, Waltham, USA
H1650	ATCC (CRL-5883)
HCT116	ATCC (CCL-247)
HEK293	ATCC (CRL-1573)
Huh-7	JCRB (JCRB0403)
PANC-1	ATCC (CRL-1469)
SCID/Beige mice	Barcelona Biomedical Research Park, Barcelona, Spain
SW480	ATCC (CCL-228)

Table 9: Technical devices.

Axiolab	Zeiss, Jena, Germany
AxioObserver Z1 (including ApoTome)	Zeiss, Jena, Germany
Axiovert 100	Zeiss, Jena, Germany
BBD 6220 cell culture incubator	Heraeus, Hanau, Germany
Biometra Power Pack P25T	Analytik Jena, Jena, Deutschland
Bioruptor Sonicator	Diagenode, Seraing, Belgium
Chemoluminescence imager	Analytik Jena, Jena, Germany
Cryostat CM3050S	Leica, Nussloch, Deutschland
Enspire Multimode Plate Reader	PerkinElmer, Waltham, MA
FACSCalibur	BD Biosciences, San Jose, USA
Fluoroskan Ascent FL	Thermo Fisher, Waltham, USA
FLUOstar OPTIMA	BMG Labtech GMBH, Ortenberg, Germany
GelDoc XR Imaging System	Biorad, Munich, Germany
Genesys 5 Spectrophotometer	Spectronic, Helsingborg, Sweden
HERA Safe Cleanbench	Thermo Scientific, Waltham, USA
Inova 4230 refrigerated incubator shaker	New Brunswick Scientific, Edison, USA
IVIS Spectrum In Vivo Imaging System	PerkinElmer, Waltham, MA

LightCycler 480 II system	Roche, Penzberg, Germany
Mithras LB 940 Microplate reader	Berthold Technologies, Wildbad, Germany
NanoDrop 1000 Spectrophotometer	Thermo Fisher, Waltham, USA
PerfectBlue PAGE System	Peqlab (via VWR) Darmstadt, Germany
Purelab Ultra MK II	ELGA LabWater, Lane End, United Kingdom
Sonopuls Sonicator system	Bandelin, Berlin, Germany
T3 Thermocycler	Analytik Jena, Jena, Germany
Trans-Blot SD semi-dry transfer cell	Biorad, Munich, Germany
X-Ray Film Processor Optimax	Protec, Oberstenfeld, Germany

6.3 List of figures

6.4 List of tables

7. Publications

Research articles

Franke, F. C., B. O. Slusarenko, T. Engleitner, W. Johannes, M Laschinger, R. Rad, U. Nitsche, and K.P. Janssen. 2019. 'Novel role for CRK adaptor proteins as essential components of SRC/FAK signaling for epithelial-mesenchymal transition and cancer aggressiveness', Manuscript submitted.

Franke, F. C., J. Muller, M. Abal, E. D. Medina, U. Nitsche, H. Weidmann, S. Chardonnet, E. Ninio, and K. P. Janssen. 2019. 'The Tumor Suppressor SASH1 Interacts With the Signal Adaptor CRKL to Inhibit Epithelial-Mesenchymal Transition and Metastasis in Colorectal Cancer', Cell Mol Gastroenterol Hepatol, 7: 33-53.

Sperlich, A., A. Balmert, D. Doll, S. Bauer, F. Franke, G. Keller, D. Wilhelm, A. Mur, M. Respondek, H. Friess, U. Nitsche, and K. P. Janssen. 2018. 'Genetic and immunological biomarkers predict metastatic disease recurrence in stage III colon cancer', BMC Cancer, 18: 998.

Scientific talks

Franke et al. 2019. 'The tumor suppressor SASH1 interacts with CRK-family proteins to inhibit epithelial-mesenchymal transition and metastasis formation', 20th International AEK Cancer Congress, 27th February – 1st March 2019, Heidelberg.

Franke et al. 2018. 'The tumor suppressor SASH1 associates with CRK-family proteins to counteract epithelial-mesenchymal transition', International Meeting of the German Society for Cell Biology, 17th – 19th September 2018, Leipzig.

Franke et al. 2017. 'The tumor suppressor SASH1 interacts with the signal adaptor CRKL to inhibit epithelial-mesenchymal transition and metastasis formation', 3rd AEK Autumn School, 6th – 8th November 2017, Berlin.

Poster presentations

Franke et al. 2018. 'The tumor suppressor SASH1 associates with CRK-family proteins to counteract epithelial-mesenchymal transition', International Meeting of the German Society for Cell Biology, 17th – 19th September 2018, Leipzig.

Franke et al. 2017. 'The tumor suppressor SASH1 interacts with the signal adaptor CRKL to inhibit epithelial-mesenchymal transition and metastasis formation', 3rd AEK Autumn School, 6th – 8th November 2017, Berlin.

Franke et al. 2016. 'Dual-specificity phosphatase 5: a negative regulator of MAP-Kinase signaling in colorectal cancer', ISREC-SCCL Symposium 2016, 8th – 11th September 2016, Lausanne.

8. Acknowledgement

First of all, I would like to express my great gratitude to Dr. Klaus-Peter Janssen, who supported me during my thesis and provided an excellent scientific environment with the freedom to develop and conduct project ideas on my own. Working in his laboratory was an exciting and encouraging experience. Thank you very much for this great opportunity! Thank you, Dr. Karl Kramer and Dr. Melanie Laschinger, for your great support and valuable discussions. It really helped me in my scientific development. Thank you, Dr. Bernhard Holzmann, for your great advices and swift answers to any question. Furthermore, I want to acknowledge the German Cancer Aid for the financial support that allowed me to conduct research on this highly interesting topic.

I also want to express my gratitude to all the colleagues in the laboratory for the very personal and friendly working atmosphere, especially to Christoph, Andreas and Markus for the unforgettable and great time! Thank you Widya, Felicitas, Anja and Gabi for your indispensable support in the laboratory. I also want to thank Miguel, Eduardo, Elia, Lorena and Alex for the great time I had during my research stay in Santiago de Compostela. A big thank you! Thank you, Miguel, Eduardo and Cristian, for your help with the *in vivo* model and drug screenings, which were really interesting parts of my research. Further, I want to thank Dr. Roland Rad, Dr. Ewa Ninio, Dr. Henri Weidmann, Solenne Chardonnet and Dr. Thomas Engleitner for the exciting cooperations.

Finally, I want to express my great thankfulness to my family, friends and my wife Carina for the continuous encouragement and support.