

Novel role for CRK adaptor proteins as essential components of SRC/FAK signaling for epithelial–mesenchymal transition and colorectal cancer aggressiveness

Fabian C. Franke¹, Benjamin O. Slusarenko¹, Thomas Engleitner², Widya Johannes¹, Melanie Laschinger¹, Roland Rad², Ulrich Nitsche¹ and Klaus-Peter Janssen¹ ¹

¹Department of Surgery, School of Medicine, Klinikum Rechts der Isar, Technical University of Munich, Munich, Germany

²Department of Medicine II, School of Medicine, Institute of Molecular Oncology and Functional Genomics, Technical University of Munich, TranslaTUM Cancer Center, Munich, Germany

Epithelial–mesenchymal transition (EMT) is a cell plasticity process required for metastasis and chemoresistance of carcinoma cells. We report a crucial role of the signal adaptor proteins CRK and CRKL in promoting EMT and tumor aggressiveness, as well as resistance against chemotherapy in colorectal and pancreatic carcinoma. Genetic loss of either CRKL or CRK partially counteracted EMT in three independent cancer cell lines. Strikingly, complete loss of the CRK family shifted cells strongly toward the epithelial phenotype. Cells exhibited greatly increased E-cadherin and grew as large, densely packed clusters, completely lacked invasiveness and the ability to undergo EMT induced by cytokines or genetic activation of SRC. Furthermore, CRK family-deficiency significantly reduced cell survival, proliferation and chemoresistance, as well as ERK1/2 phosphorylation and c-MYC protein levels. In accordance, MYC-target gene expression was identified as novel hallmark process positively regulated by CRK family proteins. Mechanistically, CRK proteins were identified as pivotal amplifiers of SRC/FAK signaling at focal adhesions, mediated through a novel positive feedback loop depending on RAP1. Expression of the CRK family and the EMT regulator ZEB1 was significantly correlated in samples from colorectal cancer patients, especially in invasive regions. Further, high expression of CRK family genes was significantly associated with reduced survival in locally advanced colorectal cancer, as well as in pan-cancer datasets from the TCGA project. Thus, CRK family adaptor proteins are promising therapeutic targets to counteract EMT, chemoresistance, metastasis formation and minimal residual disease. As proof of concept, CRK family-mediated oncogenic signaling was successfully inhibited by a peptide-based inhibitor.

Introduction

The metastatic cascade is a highly complex process during which cancer cells detach from the primary tumor, invade adjacent tissue, enter the blood or lymph circulatory system,

Additional Supporting Information may be found in the online version of this article.

Key words: colorectal cancer, invasion, metastasis, SASH1, focal adhesion

Abbreviations: CRC: colorectal cancer; CRK: V-Crk Avian Sarcoma Virus CT10 Oncogene Homolog; CRKL: V-Crk Avian Sarcoma Virus CT10 Oncogene Homolog-Like; CSK: C-terminal SRC kinase; EMT: epithelial–mesenchymal transition; SASH1: SAM-and SH3-domain containing 1; SH2: Src-homology 2 domain; SH3: Src-homology 3 domain. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

DOI: 10.1002/ijc.32955

History: Received 20 Aug 2019; Accepted 8 Jan 2020; Online 9 Mar 2020

Correspondence to: Klaus-Peter Janssen, E-mail: klaus-peter.janssen@tum.de

extravasate and finally colonize distant organs.¹ Due to systemic dissemination, metastasis is largely incurable and the main cause of cancer-related death. Epithelial–mesenchymal transition (EMT) is a transdifferentiation process that is hijacked by cancer cells to acquire a highly motile and invasive phenotype with increased cell survival.² Hence, EMT is believed to aggravate the formation of metastasis.^{2,3} Epithelial cells lose many of their epithelial characteristics during EMT, for example, the intercellular adhesion protein E-cadherin is strongly downregulated, promoting cell individualization and the acquisition of a fibroblast-like morphology.^{3,4} Downregulation of epithelial factors and induction of a mesenchymal expression profile is mediated by EMT-promoting transcription factors like ZEB1, SNAIL and TWIST1.⁴ Since EMT is believed to foster tumor progression, especially regarding metastasis formation, identification of key players that could be targeted to reduce or reverse EMT is important to inhibit the systemic spread of cancer cells.

We have recently described the signal adaptor CRKL (v-CRK avian sarcoma virus CT10-homolog-like) as a novel interaction partner of the tumor and metastasis suppressor SASH1.⁵ The CRK protein family of signal adaptors consists of CRKL, CRK-I and CRK-II, which share a high sequence and overall structural

What's new?

Epithelial-mesenchymal transition (EMT) is an essential prerequisite for metastatic spread and resistance against chemotherapy, which cause major clinical problems in colorectal and other carcinomas. CRK adaptor proteins are known as downstream effectors of SRC/FAK kinases. Here, the authors demonstrate that CRK proteins mediate the EMT-associated phenotypes of colorectal cancer aggressiveness, act as novel feedback amplifiers of SRC/FAK kinase signaling, and induce c-MYC signaling. The findings highlight CRK family adaptor proteins as promising therapeutic targets to counteract EMT, chemoresistance, metastasis formation, and minimal residual disease. Moreover, as proof of concept, CRK family-mediated oncogenic signaling was successfully inhibited by a peptide-based inhibitor.

homology.⁶ SASH1 physically associates with the N-terminal SH3 domain of CRKL, counteracting the activation of the tyrosine kinase SRC that promotes EMT, and thereby blocks metastasis formation *in vivo*.⁵ However, the molecular mechanism underlying this novel pro-metastatic cascade, as well as the role of the close homolog CRK (v-CRK avian sarcoma virus CT10-homolog) are essentially unknown. CRK proteins act as signal adaptors in a diverse range of pathways, for example, downstream of HGF/c-Met or integrins.^{7–11} This is mediated by the SH2 domains of CRK proteins binding to phosphotyrosine residues of transmembrane receptor-associated scaffold proteins like p130CAS or paxillin.^{6,12} Thus, CRK proteins recruit guanine nucleotide exchange factors as effectors *via* their N-terminal SH3 domains into close proximity to the plasma membrane, thereby activating membrane-bound small GTPases like RAP1 and RAC1.⁶ The aim of our study was to address the potential redundancy of CRKL and CRK in EMT, and to elucidate the underlying mechanism of how CRK family proteins promote the shift toward a mesenchymal state.

Materials and Methods**Cell culture and constructs**

Human colorectal cancer cell lines HCT116 (RRID: CVCL_0291) and SW480 (RRID: CVCL_0546), PANC-1 pancreatic cancer cells (CVCL_0480) and HEK293 (CVCL_0045) embryonic kidney cells (ATCC, Rockville, MD) were cultured in DMEM (Invitrogen, Karlsruhe, Germany) containing 10% fetal bovine serum (Biochrom, Berlin, Germany), tested for mycoplasma infection by PCR every 6 weeks, and authenticated by multiplex short tandem repeat analysis (Promega) as stated earlier,¹³ most recently in June 2019. All experiments were performed with mycoplasma-free cells. To avoid contamination and phenotypic changes, cells were kept as frozen stocks and cultured consecutively for 4 weeks maximum. HCT116 cells were transfected with FuGENE HD (Promega, Madison, WI) and SW480 cells with Attractene (Qiagen, Hilden, Germany) according to the manufacturer's protocol. HEK293 cells were transfected by calcium phosphate precipitation as described.¹⁴ Mutagenesis of CRKL-RFP, which was previously cloned into pmRFP-N2, was performed using the following primers: R39K: 5'-CTCGTCAAGGATTC TTCCACCTGCCCTG-3' and 5'-GAAGAATCCTTGACGAGG AACATAACCGTGGC-3'; W160R: 5'-GAACAGCGGTGGAGTG CCCGGAACAAGG-3' and 5'-GCACTCCACCGCTGTTCTTC AGGCTTCTC-3'.⁵ GFP with C-terminal high-affinity peptide

(HAP; residues CVDNSPPALPPKRRRSAPS^{15,16}) was generated by cloning the oligonucleotides 5'-AATTCTGCGTGGATAAC AGCCCGCCGCCGCGCTGCCCGCAAACGCCGCCGCAG CGCGCCGAGCTAGG-3' and 5'-GATCCCTAGCTCGGCGCG CTGCGGCGGCGTTTCGGCGGCAGCGCCGGCGGCGGGCT GTTATCCACGCAG-3' into pEGFP-C2, using EcoRI and BamHI (Thermo Fisher) for restriction digest of the plasmid. CRKL- and/or CRK-deficient cells were generated with the CRISPR/Cas9-system, using two different guide RNAs per gene, and cloned into PX459 vector: CRK (gRNA1: 5'-GGCGGGCAA CTTCGACTCGG-3'; gRNA2: 5'-TGGTGCTCGAGTCCCGCA CC-3'); CRKL (gRNA1: 5'-GTCCGAGGAGTCGAACCTGG-3'; gRNA2: 5'-CGAGGAGTCGAACCTGGCGG-3'); CSK (gRNA1: 5'-GCAATACATTCTGTACCGGA-3'; gRNA2: 5'-AGTGCC GTGGAAGTTGTA-3'). The vector pSpCas9(BB)-2A-Puro (PX459) V2.0 was a gift from Feng Zhang (Addgene plasmid #62988).¹⁷ Then, 48 hr after transfection of PX459 containing the corresponding guide RNA sequences, transfected cells were selected by 3 µg/ml (SW480) or 5 µg/ml (HCT116) puromycin for 72 hr. Single colonies were isolated and analyzed for the absence of the target protein during several passages to ensure a complete deficiency. Nontargeting control clones were generated with an empty PX459 vector, using the same transfection and selection procedure. CRKL-deficient cells were described before.⁵ EMT was induced by application of 20 ng/ml TNF (Gibco, Waltham, MA) or 10 ng/ml TGF-β1 (Stemcell Technologies, Cologne, Germany) for 72 hr.

Immunoblotting and co-immunoprecipitation

Immunoblotting and co-immunoprecipitation were performed as described^{5,14} with the following antibodies: E-cadherin (Abcam ab40772), β-actin (CST #3700), GAB1 (Bethyl A303-288A), C3G (Bethyl A301-965A), c-MYC (SCBT #sc-40), ERK1/2 (CST #4695), Thr202/Tyr204 phospho-ERK1/2 (CST #9101), SASH1 (Novus NBP1-26650), RFP (Chromotek 5F8), CRK-II (SCBT sc-289), CRK (Merck Millipore MABC172), CRKL (SCBT sc-319), CSK (Abcam ab125005), ZEB1 (Sigma HPA027524), Y118 phospho-paxillin (CST #2541), paxillin (CST #12065), Y410 phospho-p130CAS (CST #4011), p130CAS (CST #13846), Y416 phospho-Src family (CST #6943), Src (CST #2123), Y397 phospho-FAK (CST #8556), Y576/577 phospho-FAK (CST #3281), Y925 phospho-FAK (CST #3284), FAK (CST #13009) and RFP-trap (Chromotek rta-10).

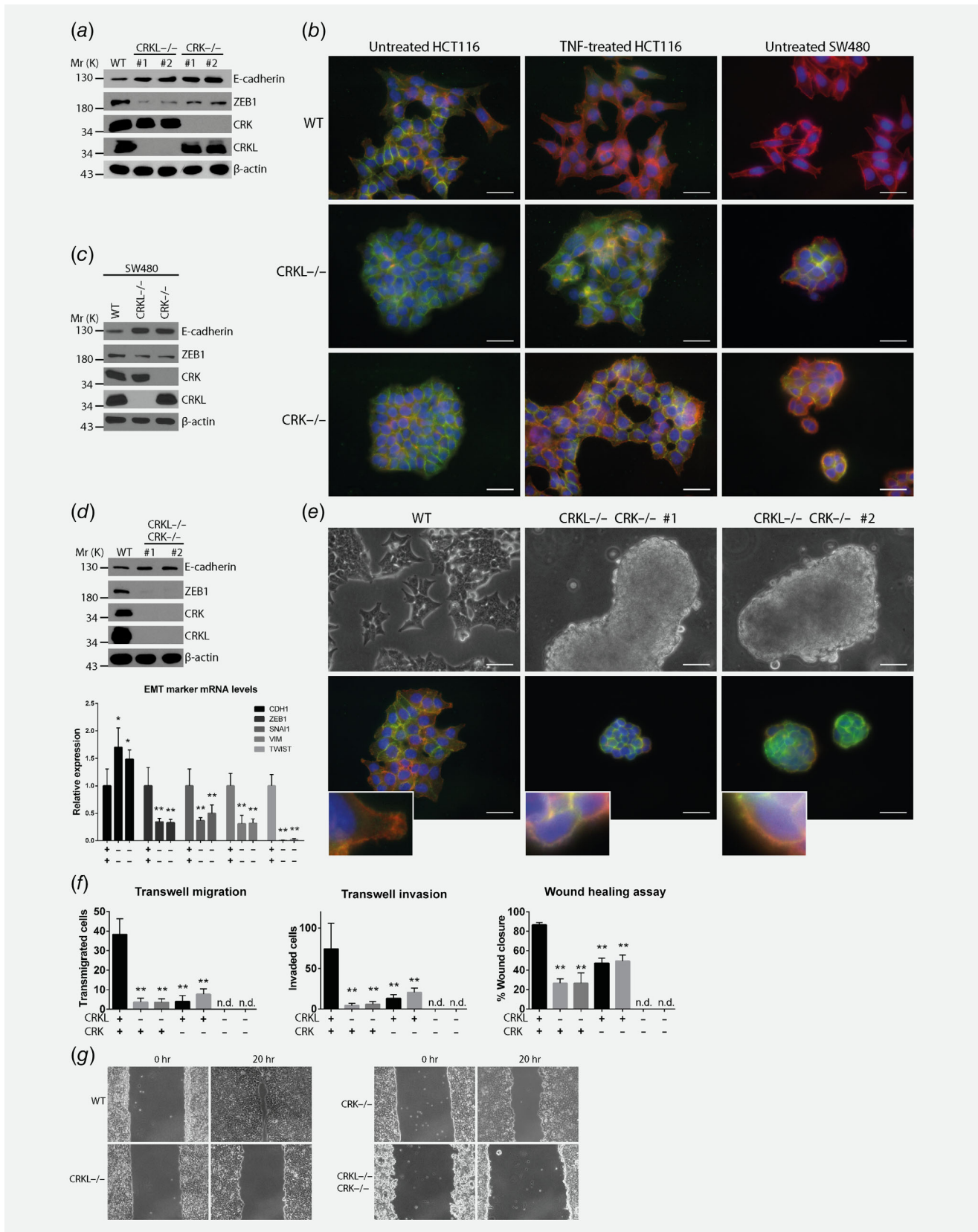


Figure 1. Legend on next page.

Immunofluorescence microscopy

Immunofluorescence microscopy was performed with a Zeiss AxioObserver Z1 microscope (Carl Zeiss, Jena, Germany) as described⁵ with the following antibodies: E-cadherin (Abcam, Cambridge, MA; ab40772), ZEB1 (Sigma, St. Louis, MO; HPA027524), Y118 phospho-paxillin (CST #2541), paxillin (CST #12065), Talin (Abcam ab17333), integrin beta-1 (Abcam ab24693), CRKL (EMD Millipore 05-414), DAPI for DNA-staining and TRITC-phalloidin for F-actin (both from Sigma). To analyze patient tissue samples, cryosections of 7 μ m were generated with a cryotome (Leica CM3050 S, Leica Microsystems, Wetzlar, Germany) and stained with hematoxylin and eosin as published.⁵

Transwell assays

Transwell migration (8.0 μ m Transwell Permeable Supports, Costar, #3422) and invasion assays (8.0 μ m Matrigel Invasion Chamber, Corning, #354480) were carried out as described.⁵ Migration was allowed for 24 hr and invasion for 48 hr. For transiently transfected HEK293 cells, inserts were additionally coated with 5 μ g/ml human fibronectin for 1 hr at 37°C.

Wound healing assays

Wound healing assays were performed using culture-insert “2 Well in μ -Dish 35 mm” (ibidi #81176, Planegg, Germany). Then, 1.5×10^5 cells in DMEM containing 10% fetal bovine serum were seeded into each chamber and allowed to adhere overnight. The separating chambers were removed and images acquired immediately, and again after 20 hr incubation, with a Zeiss AxioObserver Z1 microscope. Wound closure was determined by subtracting the cell-free area of both time points, measured using ImageJ v.1.51 (National Institutes of Health).

Reverse transcription and qRT-PCR

Isolation and reverse transcription of RNA was performed with the RNeasy Kit (Qiagen) as described.^{18,19} qRT-PCR was performed with the LightCycler 480 II system (Roche). *HPRT* was used as internal reference transcript and a pool of human colon mucosa cDNA for normalization. Pooled cDNA obtained from HCT116 cells was used for normalization of tumor samples.

The following primers were used: *HPRT* (UPL #22): 5'-GAC CAGTCAACAGGGGACAT-3' and 5'-GTGTCAATTATATCT TCCACAATCAAG-3'; *CDH1* (UPL #84): 5'-TGGAGGAAT TCTTGCTTTGC-3' and 5'-CGCTCTCCTCCGAAGAAAC-3'; *VIM* (UPL #39): 5'-GACCAGCTAACCAACGACAAA-3' and 5'-GAAGCATCTCCTCCTGCAAT-3'; *ZEB1* (UPL #34): 5'-TT TTTCTGAGGCACCTGAA-3' and 5'-TGAAAATGCATCT GGTGTTCC-3'; *TWIST1* (UPL #50) 5'-TCCAGAGAAGGA GAAAATGGAC-3' and 5'-TTTCCAAGAAAATCTTTGGCA TA-3'; *SNAI1* (UPL #30) 5'-AGGATCTCCAGGCTCGAAA G-3' and 5'-TCGGATGTGCATCTTGAGG-3'; *CRK* (UPL #81) 5'-AACAGGCAGCGCTACTCAAG-3' and 5'-TCCAGCCCA GTGGTTCAT-3'.

Adhesion and soft agar assays

Adhesion assays were performed in 96 well plates, coated with 5 μ g/ml human fibronectin in PBS. Briefly, 1×10^4 cells in 100 μ l DMEM with 10% fetal bovine serum were allowed to adhere for different time points. Medium was aspirated, adherent cells were washed with PBS and counted by light microscopy (Zeiss AxioObserver Z1). Adhesion was quantified relative to the total number of cells in wells without washing. Anchorage-independent growth (soft agar) was determined as described.⁵

Flow cytometry

Cells (1×10^6) were blocked in 100 μ l PBS containing 10% goat serum and 0.5% bovine serum albumin for 20 min at 4°C, washed and stained with 1 μ g anti-Integrin beta-1 antibody (Abcam ab24693), or isotype antibody for 30 min at 4°C in PBS with 0.5% BSA. After two washing steps, cells were incubated with the appropriate, FITC-labeled secondary antibody (Jackson ImmunoResearch #415-095-100) for 30 min at 4°C in PBS with 0.5% BSA. Cells were washed twice and subjected to flow cytometry using FACSCalibur (BD Becton Dickinson) and CellQuest Pro software (BD Biosciences, San Jose, CA). Dead cells, doublets and debris were excluded, and analysis was performed using FlowJo 8.8.2 (Tree Star, Ashland). Unspecific staining was determined with an isotype control antibody.

Figure 1. CRKL and CRK share redundant roles in EMT; compound loss of CRK family proteins leads to MET. (a) CRKL- or CRK-deficient HCT116 colon cancer cells were generated by CRISPR/Cas9 and independent guide RNAs (#1 and #2). E-cadherin, ZEB1, CRK, CRKL and β -actin were determined by immunoblotting. (b) Parental (WT), CRKL- or CRK-deficient HCT116 cells were stained for E-cadherin (green), F-actin (red) and nuclei (DAPI, blue) by immunofluorescence microscopy, after treatment with vehicle or 20 ng/ml TNF for 72 hr. Furthermore, CRKL- or CRK-deficient SW480 colon cancer cells were generated and stained for E-cadherin (green), F-actin (red) and nuclei (blue) and compared to parental cells (WT; size bar = 20 μ m). (c) Parental (WT), CRKL- or CRK-deficient SW480 cells were analyzed for E-cadherin, ZEB1, CRK, CRKL and β -actin protein levels via immunoblotting. (d) Parental HCT116 cells and two independent CRK family-deficient clones (#1 and #2) were analyzed for EMT markers (E-cadherin, ZEB1) by immunoblotting (upper panel), as well as for mRNA expression (lower panel) by qRT-PCR (*CDH1*, *ZEB1*, *SNAI1*, *VIM* and *TWIST*; Mann–Whitney test; $n = 5-6$; * $p \leq 0.05$; ** $p \leq 0.01$). (e) Phase-contrast images of parental and CRK family-deficient HCT116 cells (upper panels; size bar = 40 μ m). Additionally, cells were stained for E-cadherin (green), F-actin (red) and nuclei (blue; lower panels; size bar = 20 μ m). (f) Cell motility and invasiveness of parental, CRKL-, CRK- and CRK family-deficient HCT116 cells was quantified by transwell assays (Mann–Whitney test; $n = 6$; ** $p = 0.0022$), as well as by wound healing assays (Mann–Whitney test; $n = 6$; ** $p = 0.0022$). (g) Representative picture, collective cell migration analyzed by wound healing assays. [Color figure can be viewed at wileyonlinelibrary.com]

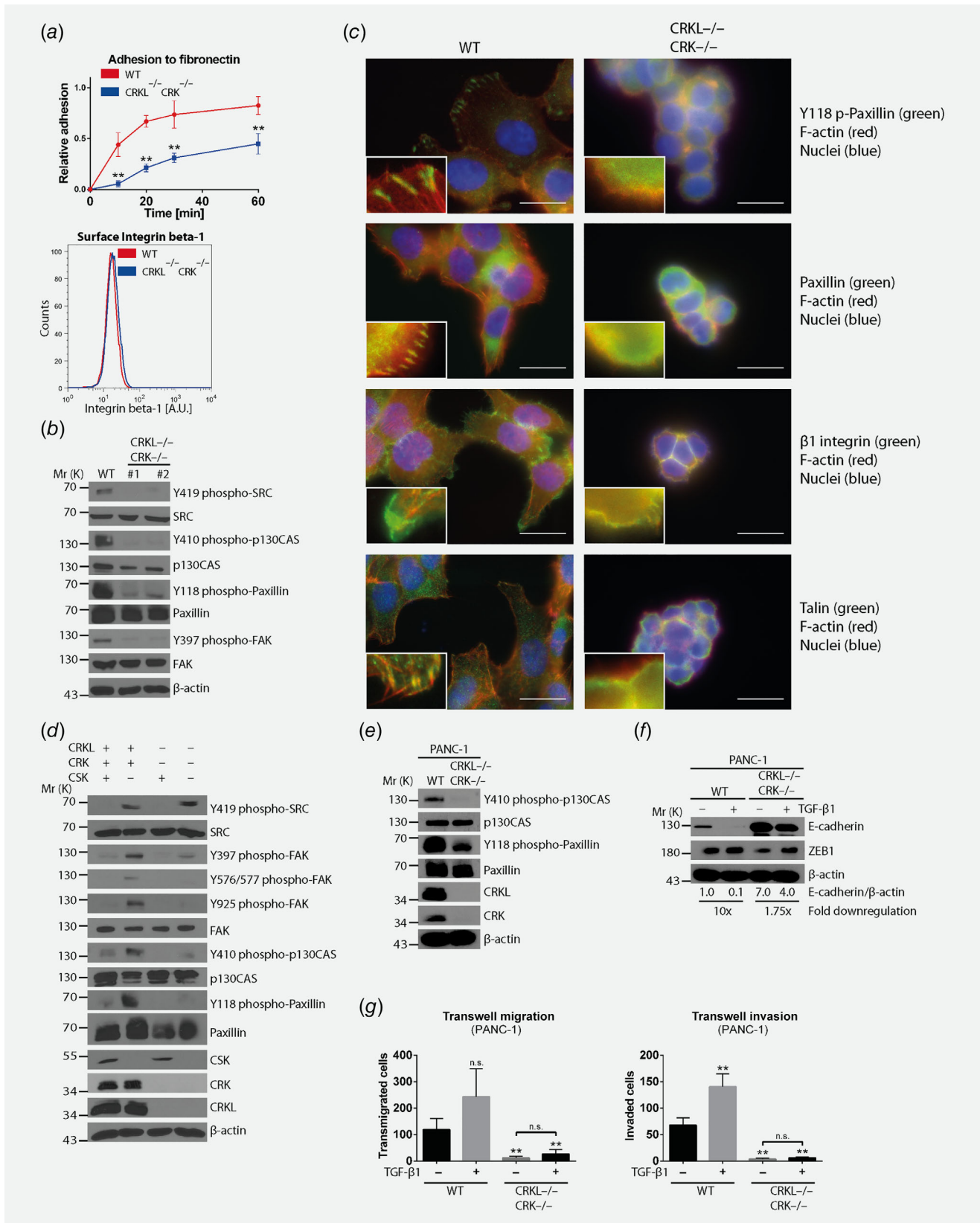


Figure 2. Legend on next page.

Chemoresistance and cell proliferation assays

A total of 2×10^3 cells were seeded per well (96-well plate) and allowed to adhere overnight. Afterward, cells were treated with varying concentrations of 5-fluorouracil and/or oxaliplatin for 24 hr. Medium was aspirated, cells washed with PBS, and incubated for 48 hr in fresh medium without chemotherapeutics. Surviving cells were quantified using the Cell Proliferation Kit II/XTT assay (Roche, Mannheim, Germany). IC50 values were calculated with GraphPad Prism 6.01. For cell proliferation, 1,000 cells were seeded and incubated in 2 or 10% fetal bovine serum for the indicated time points before XTT assays were performed.

Clinical data, TCGA analysis and statistical analysis

Sixty-five patients with Stage III (UICC/AJCC) primary colorectal cancer were analyzed, who underwent curative surgery (R0) at the Department of Surgery, TUM and gave informed consent. Ethical and data protection standards were assured by the ethics committee of the Faculty of Medicine, TUM (# 1926/07; #5428/12). Tissue was shock-frozen immediately after resection. Clinicopathological characteristics are summarized in Supporting Information Table S1. Cases with local recurrence were excluded to avoid putative bias by poor surgical technique, and recurrence-free survival was the primary endpoint. Statistical evaluation was performed using IBM SPSS 20.0 (SPSS Inc., Chicago, IL). To derive optimal cut-off values for patient stratification based on gene expression levels determined by qRT-PCR, a rational approach was chosen based on maximally selected log-rank statistics performed by R Software (see <https://cran.r-project.org/web/packages/maxstat/vignettes/maxstat.pdf>). The derived cut-off values of relative gene expression were the following: CRKL 0.44; CRK 0.6442; CRK family 0.7809; ZEB1 44.01. To consider this multiple test issue within the analyses, the R-function `maxstat.test` was used. The package `maxstat` implements both, cutpoint estimation and the test procedure above with several *p*-value approximations as well as plotting of the empirical process of the standardized statistics. In a first step, the statistics program R calculates the ideal cut-off value for a particular target by plotting absolute standardized log-rank statistics and significance bound based on the improved Bonferroni inequality. In the

second step, survival analyses are performed with the above determined cut-off value. Survival analysis was performed using Kaplan–Meier estimates. Statistical tests were performed two-sided, the significance level set at 0.05. No correction of *p* values was applied to adjust for multiple test issues. However, results of all statistical tests were thoroughly reported so that an informal adjustment of *p* values can be performed while reviewing the data. The public Cancer Genome Atlas (TCGA) data set was analyzed with the cBioPortal platform,^{20,21} the pan-cancer dataset comprising 72,175 samples from 69,835 patients in 236 studies, for genomic, mutational and transcriptomic alterations of CRKL and CRK, their co-occurrence. For association with survival, a curated, nonredundant dataset from 176 studies was used ($n = 26,607$ patients), the “HIGH” versus “LOW” status for CRK-family genes was defined as: genomic amplification or gain, and/or overexpression >twofold (HIGH), compared to shallow or deep genomic deletions, and/or average or downregulated expression (LOW). GraphPad Prism 6.01 was used for all further statistical analysis. All error bars represent standard deviation. Sample size, statistical test used and *p* values are indicated in each figure legend.

RNA sequencing and analysis

Detailed description of the sequencing methodology, number of total reads and further details can be found in the Supporting Information, as well as upon request. Briefly, two biological (i.e., independent clones) and three technical replicates (i.e., RNA prepared at different passages) were analyzed for each condition. Library preparation for bulk 3'-sequencing of poly(A)-RNA was done as described.²² Barcoded cDNA of each sample was generated with a Maxima RT polymerase (Thermo Fisher) using oligo-dT primer containing barcodes, unique molecular identifiers (UMIs) and an adapter. 5' ends of the cDNAs were extended by a template switch oligo (TSO), and after pooling of all samples, full-length cDNA was amplified with primers binding to the TSO-site and the adapter. cDNA was fragmented with the Nextera XT kit (Illumina) and 3'-end-fragments finally amplified using primers with Illumina P5 and P7 overhangs. In comparison to Parekh *et al.*, the P5 and P7 primer sites were exchanged to allow sequencing of the cDNA in read1, and barcodes and UMIs in read2, to achieve better

Figure 2. CRK proteins are required for activation of the SRC/FAK signaling complex across tumor entities. (a) Adhesion assays were performed to quantify adhesion of parental (WT) or CRK family-deficient HCT116 cells to fibronectin (Mann–Whitney test; $n = 6$; $**p \leq 0.01$). Cell surface levels of integrin beta-1 quantified by flow cytometry. (b) Activation of the SRC/FAK signaling pathway of parental and CRK family-deficient HCT116 cells was determined by immunoblotting with phosphorylation-specific antibodies. (c) Cells adhering to fibronectin-coated glass slides were stained for focal adhesion components with specific antibodies (green), F-actin (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 μm). (d) Activation of the SRC/FAK signaling pathway analyzed *via* immunoblotting of parental and CRK family-deficient cells, either CSK-proficient or CSK-deficient. (e) SRC/FAK substrate phosphorylation was analyzed *via* immunoblotting of parental and CRK family-deficient PANC-1 cells. (f) Parental and CRK family-deficient PANC-1 cells were stimulated with 10 ng/ml TGF- β 1 for 72 hr, and EMT markers analyzed *via* immunoblotting. (g) Parental and CRK family-deficient PANC-1 cells were stimulated with 10 ng/ml TGF- β 1 for 72 hr, and subsequently subjected to transwell migration (Mann–Whitney test; $n = 6$; $**p = 0.0022$; n.s. = not significant: $p \geq 0.05$) and invasion assay (Mann–Whitney test; $n = 6$; $**p = 0.0022$; n.s. = not significant: $p \geq 0.05$). Representative loading control for all samples is shown. [Color figure can be viewed at wileyonlinelibrary.com]

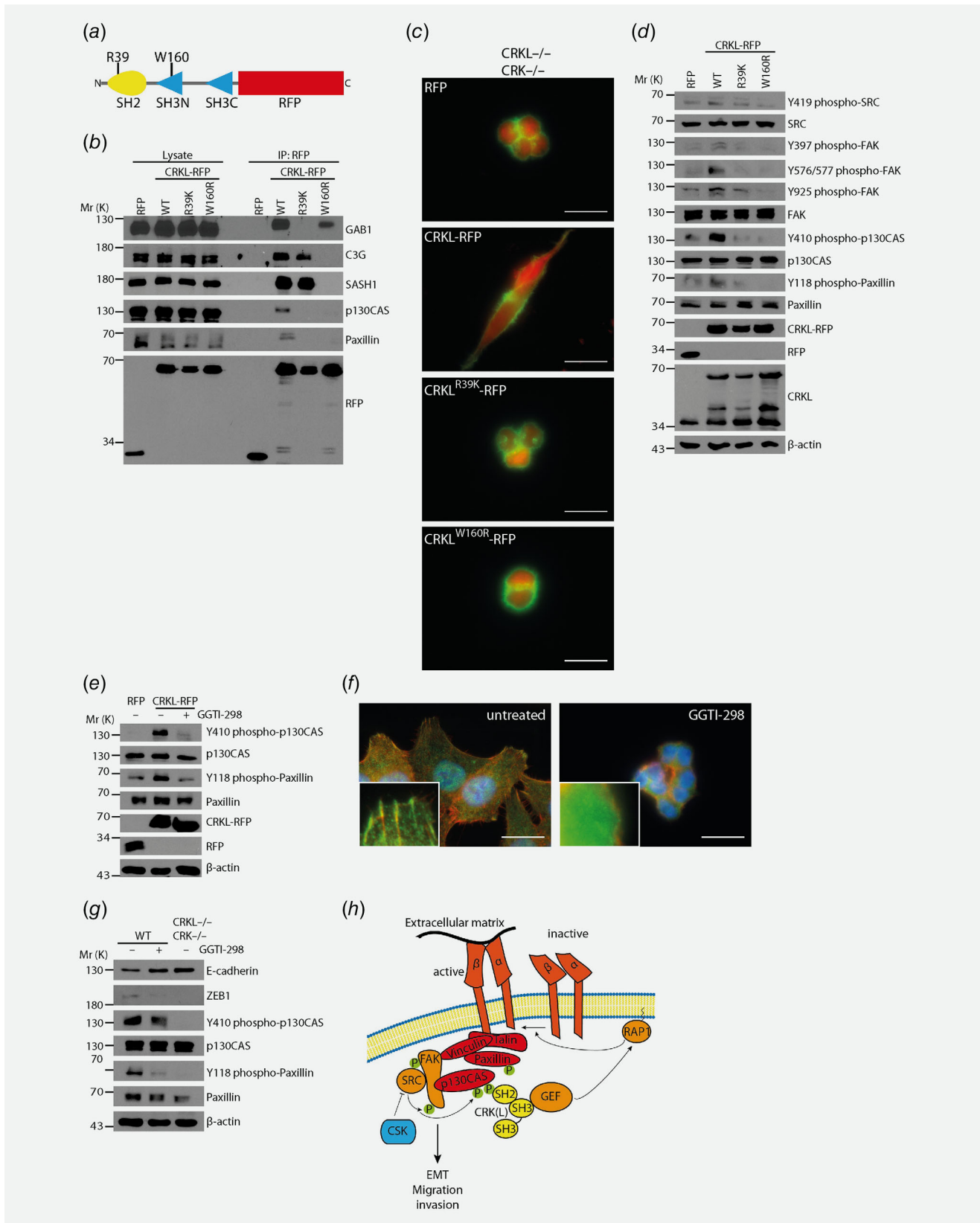


Figure 3. Legend on next page.

cluster recognition.²² The library was sequenced on a NextSeq 500 (Illumina) with 75 cycles for the cDNA in read1 and 16 cycles for the barcodes and UMIs in read2. Gencode annotations version M18 and the human reference genome GRCh38.p12 were derived from the Gencode homepage (<https://www.gencodegenes.org/>). Dropseq tools v1.12²³ was used for mapping the raw sequencing data to the reference genome. The resulting UMI filtered countmatrix was imported into R v3.4.4. Differential expression analysis was conducted with DESeq2 1.18.1.²⁴ Dispersion of the data was estimated with a parametric fit using the genotype label as covariate. The Wald test was used for determining differentially regulated genes between parental and CRK family-deficient HCT116 samples and shrunken log₂ fold changes were calculated afterward with setting the Shrinkage-type argument of the lfcShrink function to “normal”. A gene was determined to be differentially regulated if the absolute log₂ fold change was greater than 1 and the adjusted *p*-value was below 0.05. GSEA preranked analysis was conducted with GSEA v3.0 and the MsigDB v6.2 within the Hallmark gene set collection,²⁵ complete GSEA results are displayed in Supporting Information Table S2 (positively enriched) and Table S3 (negatively enriched), including all scores, *p* values and the core genes. The rank of a gene was determined by the shrunken fold-change estimation by DESeq2.

Data availability

The data that support the findings of our study are openly available in the European Nucleotide Archive (ENA, EMBL-EBI) under the accession number PRJEB36353, at the following address: <http://www.ebi.ac.uk/ena/data/view/PRJEB36353>.

Results

CRK and CRKL have redundant functions in epithelial–mesenchymal transition

CRK or CRKL deficient HCT116 human colorectal cancer cells were generated using the CRISPR/Cas9 system with two independent gRNAs per gene. Loss of either CRKL or CRK (targeting both splice variants CRK-I and CRK-II⁶) induced an increase of E-cadherin, while the EMT promoting transcription factor ZEB1 was reduced (Fig. 1a). In contrast, nontargeting control clones exhibited no alteration in EMT marker

expression (Supporting Information Fig. S1). Induction of EMT with TNF, a well-described EMT-inducing cytokine,²⁶ resulted in a fibroblast-like morphology with loss of E-cadherin at intercellular adhesions exclusively in the parental line, while CRK- or CRKL-deficient cells exhibited only minor changes (Fig. 1b). In support, loss of either CRKL or CRK in SW480 colon carcinoma cells, which have pronounced mesenchymal features, produced a highly epithelial phenotype (Figs. 1b and 1c).

Compound deficiency of CRK and CRKL promotes a highly epithelial phenotype

HCT116 cells were generated lacking CRK and CRKL (CRK family-deficient). E-cadherin protein levels were increased, while ZEB1 was strongly reduced (Fig. 1d). In accordance, *CDH1* mRNA expression was significantly increased, while the expression of mesenchymal markers *ZEB1*, *SNAIL1*, *VIM* and *TWIST* was significantly reduced (Fig. 1d). Parental cells exhibited a fibroblastic morphology, formation of lamellipodia and F-actin-rich membrane ruffles (Fig. 1e). CRK family-deficient cells formed dense, three-dimensional clusters, with pronounced E-cadherin staining at cell–cell contacts, indicating prominent intercellular adhesions, but were only loosely attached to the substratum (Fig. 1e). Lamellipodia and membrane ruffles were absent upon CRK family deficiency. Furthermore, the loss of either CRK or CRKL significantly reduced cell migration and invasion (Fig. 1f). Of note, CRKL-deficiency had a stronger impact compared to the loss of CRK. Compound CRK family-deficiency completely abolished transwell migration and invasion (Fig. 1f). Collective cell migration, assessed by wound healing assays, was also markedly impaired (Figs. 1f and 1g). Taken together, loss of CRK family proteins induced mesenchymal–epithelial transition (MET).

CRK proteins promote cell–matrix adhesion and are required for FAK/SRC complex signaling

Blunted cell migration, invasion and substrate attachment indicated that CRK proteins promote integrin signaling. Indeed, CRK family-deficient cells exhibited reduced adhesion to fibronectin, despite essentially unchanged surface expression of integrin beta-1 (Fig. 2a). CRK-family deficiency almost

Figure 3. CRK proteins promote activation of the SRC/FAK complex via the small GTPase RAP1. (a) Domain architecture of CRKL-RFP highlighting the residues R39 and W160, which are essential for domain function (SH2/3 = Src homology 2/3 domain). (b) HEK293 cells stably expressing RFP, CRKL-RFP, CRKL^{R39K}-RFP or CRKL^{W160R}-RFP were subjected to co-immunoprecipitations using anti-RFP antibody-coupled beads. Immunoprecipitates were blotted using antibodies against confirmed interaction partners. (c) CRK family-deficient HCT116 cells were transiently transfected to express RFP, CRKL-RFP, CRKL^{R39K}-RFP or CRKL^{W160R}-RFP, before they were stained for RFP (red) and F-Actin (green) for immunofluorescence microscopy (size bar = 20 μm). (d) Activation of the SRC/FAK signaling pathway was analyzed in HEK293 cells stably expressing RFP, CRKL-RFP, CRKL^{R39K}-RFP or CRKL^{W160R}-RFP via immunoblotting with phosphorylation-specific antibodies. (e) SRC/FAK substrate phosphorylation was analyzed in HEK293 cells stably expressing RFP or CRKL-RFP. CRKL-RFP expressing cells were also treated either with DMSO or 10 μM GGTI-298 for 48 hr. (f) Parental HCT116 treated either with DMSO or 10 μM GGTI-298 for 72 hr were stained for talin (green), F-actin (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 μm). (g) Parental HCT116 (WT) were subjected to immunoblot analysis after GGTI-298 treatment, and compared to CRK family-deficient cells. (h) Working model: CRK family proteins as central amplifiers of FAK/SRC signaling. [Color figure can be viewed at wileyonlinelibrary.com]

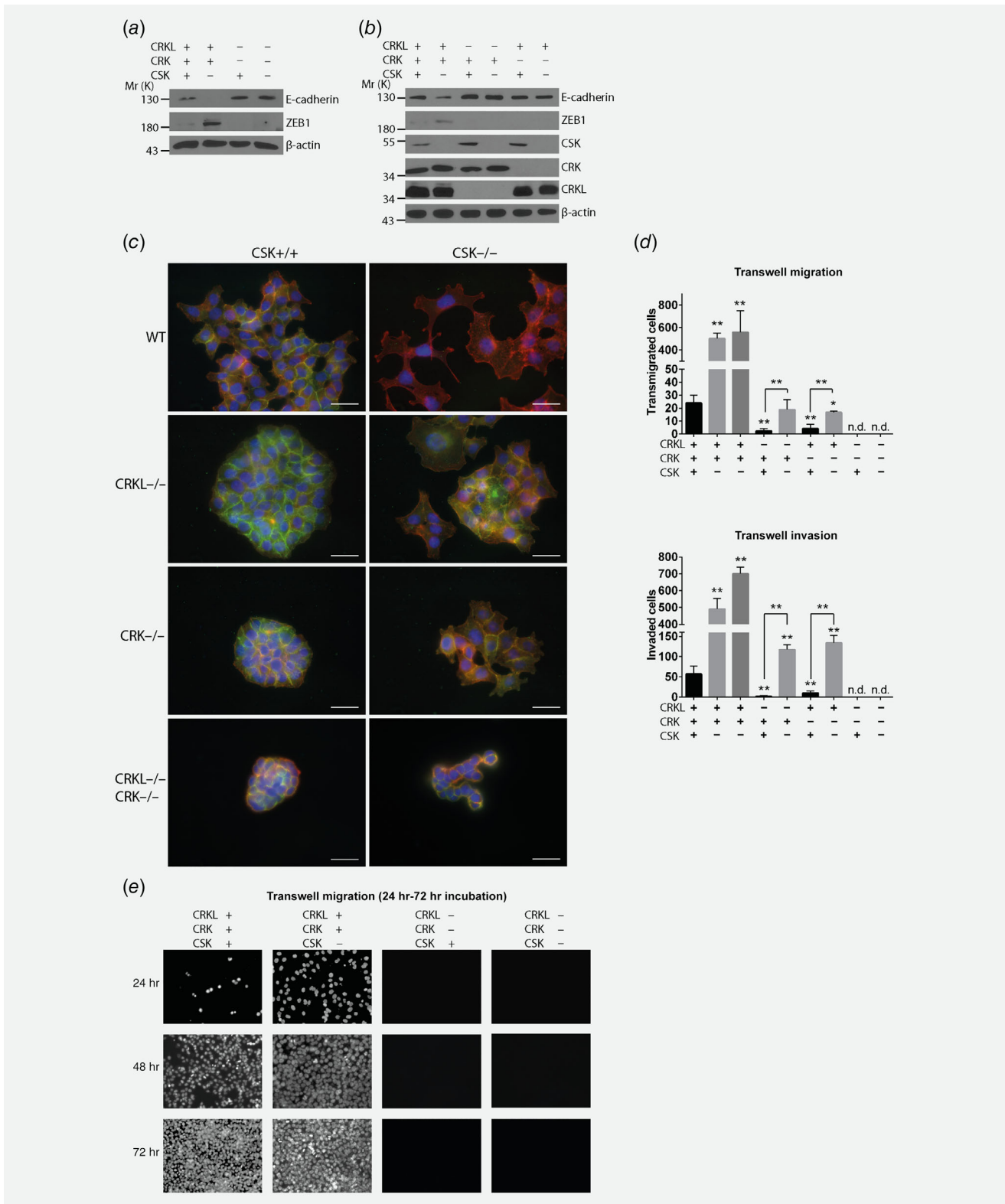


Figure 4. Loss of CRK proteins inhibits EMT induced by SRC activation. (a) Parental or CRK family-deficient HCT116 cells, either proficient or deficient for CSK, were analyzed for E-cadherin and ZEB1 levels by immunoblotting. (b) The same analysis was carried out with parental, CRKL- and CRK-deficient cells, in which CSK-deficiency was introduced. (c) Parental (WT), CRKL-, CRK- and CRK family-deficient cells were stained for E-cadherin (green), F-Actin (red) and nuclei (blue) for immunofluorescence microscopy (size bar = 20 μm). (d) Transwell migration (24 hr) and invasion (48 hr) assays were performed to quantify directional cell motility and invasiveness (Mann-Whitney test; $n = 6$; $*p \leq 0.05$; $**p \leq 0.01$). (e) Transwell migration assays for 24, 48 or 72 hr, followed by DAPI staining. [Color figure can be viewed at wileyonlinelibrary.com]

completely abolished activation-loop phosphorylation of SRC family kinases (residues Y416 in chicken and Y419 in humans; Fig. 2b). Auto-phosphorylation of FAK at Y397, an early event

during integrin signaling,²⁷ was strongly reduced (Fig. 2b). Phosphorylation of paxillin (Y118) and p130CAS (Y410), two FAK/SRC substrates, was also diminished in CRK

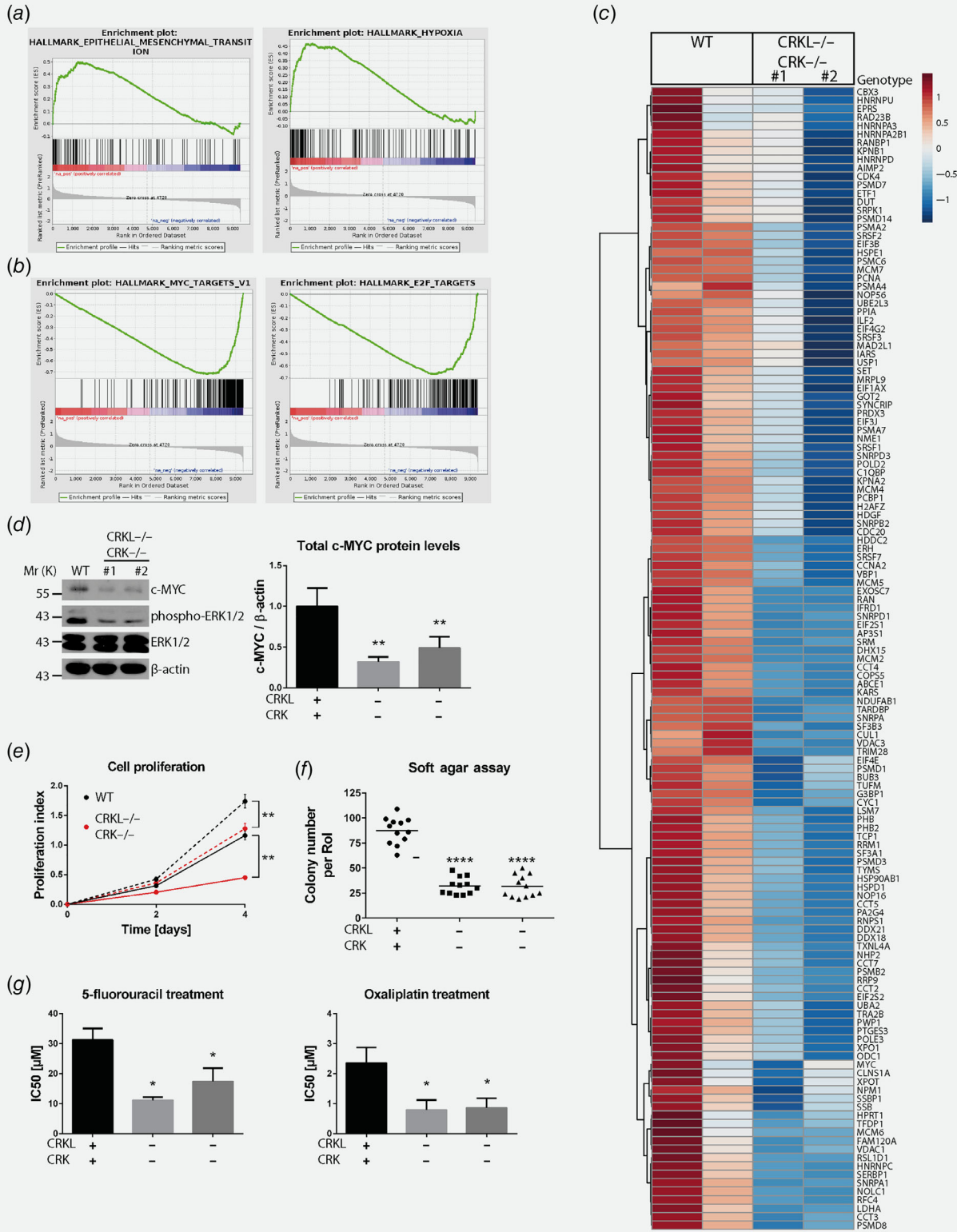


Figure 5. Legend on next page.

family-deficient cells (Fig. 2b). In accordance, only parental cells exhibited focal adhesions, positive for paxillin, Y118 phospho-paxillin and talin (Fig. 2c). Integrin beta-1 clusters were strongly reduced upon CRK family-deficiency (Fig. 2c). To dissect the role of CRK proteins in the activation of FAK/SRC signaling, CSK-deficient cells were generated. In accordance with published data, reporting the kinase CSK as a negative regulator of SRC,²⁸ phosphorylation of SRC at Y419 was significantly increased upon loss of CSK in both parental and CRK family-deficient cells (Fig. 2d and Supporting Information Fig. S2). Despite of an equivalent activation of SRC upon CSK deficiency, the parental line exhibited significantly higher levels of FAK phosphorylation at Y397, Y576/577 and Y925, compared to CRK family-deficient cells (Fig. 2d and Supporting Information Fig. S2). Accordingly, paxillin and p130CAS were poorly phosphorylated in CRK family-deficient cells despite the loss of CSK (Fig. 2d and Supporting Information Fig. S2). Thus, CRK proteins promote activation of FAK partially independent from SRC activation. A drastic reduction in SRC/FAK substrate phosphorylation upon loss of CRK family proteins was observed in the PANC-1 cells, derived from pancreatic ductal adenocarcinoma (Fig. 2e). Loss of CRK family proteins leads to a reduction of ZEB1 levels and a drastic increase of E-cadherin levels in the pancreatic cancer cells (Fig. 2f). Upon induction of EMT in parental PANC-1 cells with TGF- β 1, ZEB1 was increased, while E-cadherin was almost completely lost (Fig. 2f). This EMT-inducing effect was not completely absent but strongly diminished in cells lacking CRK-family proteins, as TGF- β 1 mediated E-cadherin down-regulation was reduced compared to the parental line, as assessed by densitometry (Fig. 2f). CRK family-deficient PANC-1 cells also exhibited strongly reduced migration and invasion capabilities (Fig. 2g). Importantly, the pretreatment effects of TGF- β 1 on transwell migration and invasion were blunted in cells lacking CRK family proteins (Fig. 2g).

CRK proteins act as central amplifiers of FAK/SRC signaling within a RAP1-dependent positive feedback loop

To investigate the role of the SH2-domain and the N-terminal SH3-domain of CRKL, inactivating point mutations were introduced (R39K for the SH2 domain; W160R for the N-SH3

domain), based on CRKL constructs with C-terminal RFP-tag (Fig. 3a). The ability of CRKL-RFP fusion protein variants to bind known interaction partners^{5,6,9,12,29} was assessed by co-immunoprecipitation experiments. The R39K replacement abolished the activity of the SH2 domain, as GAB1 showed no binding (Fig. 3b). The W160R exchange eliminated binding of C3G and SASH1 to the N-terminal SH3 domain (Fig. 3b). Importantly, only wild type CRKL-RFP interacted with p130CAS and paxillin (Fig. 3b), which are described to bind to the CRKL SH2 domain upon phosphorylation.⁶ Next, RFP controls or CRKL-RFP variants were expressed in CRK family-deficient HCT116 cells. Only wild-type CRKL rescued the parental fibroblast-like morphology (Fig. 3c). Furthermore, stable expression of CRKL-RFP fusion proteins in HEK293 cells demonstrated that only the wild-type promoted phosphorylation of SRC (Y419), FAK (Y397/Y576/Y577/Y925), p130CAS (Y410) and paxillin (Y118; Fig. 3d). Inactivation of the SH2 domain showed no effect, while inactivation of the N-terminal SH3 domain further decreased phosphorylation (Fig. 3d). Transient expression of these constructs in HCT116 cells (approximately 20–30% of cells transfected) resembled the observation in HEK293 cells, since only wild type CRKL promoted phosphorylation of the SRC/FAK substrates p130CAS and paxillin (Supporting Information Fig. S3).

These data indicate the importance of the N-terminal SH3 domain of CRKL, described to bind guanine nucleotide exchange factors to induce activation of the small GTPases RAC1, RAP1 and RAS.⁶ RAP1 plays important roles in integrin inside-out signaling.³⁰ Therefore, HEK293 cells expressing CRKL-RFP were treated with GGTI-298, an inhibitor of RAP1 activation.^{31–33} GGTI-298 strongly reduced CRKL-mediated phosphorylation of p130CAS and paxillin, indicating that RAP1 activation is required for CRKL-mediated SRC/FAK substrate phosphorylation (Fig. 3e). Accordingly, treatment of parental HCT116 with GGTI-298 leads to a rounded morphology with a diffuse localization of talin, phenocopying CRK family-deficiency (Fig. 3f). GGTI-298 application increased E-cadherin protein levels, while ZEB1 levels were reduced (Fig. 3g). Taken together, CRK proteins act as amplifiers within a RAP1-dependent positive feedback loop to induce sustained activation of FAK/SRC, and thus EMT (Fig. 3h).

Figure 5. CRK family promotes cell survival and chemoresistance. (a,b) Gene set enrichment analysis based on Hallmark gene sets was performed with RNA sequencing data of parental and CRK family-deficient HCT116 cells ($n = 2$ independent clones). Enrichment plots of selected upregulated (a) or downregulated (b) hallmarks upon CRK family-deficiency. (c) Heatmap analysis of the mRNA expression of hallmark MYC targets_V1, comparing parental cells with two independent CRK family-deficient clones (#1 and #2). (d) Parental or CRK family-deficient cells were analyzed for c-MYC levels and phosphorylation of ERK1/2 by immunoblot; quantification of c-MYC levels relative to β -actin (Mann–Whitney test; $n = 6$; $p = 0.0022$). (e) Proliferation of parental or CRK family-deficient cells was analyzed by XTT assays with cells cultured in 2% (solid line) or 10% (dashed line) fetal bovine serum (Mann–Whitney test; $n = 4–8$; two independent clones combined; $p = 0.004$). (f) Anchorage-independent growth of parental and CRK family-deficient cells was quantified by soft agar assays (unpaired t -test; $n = 12$; $p < 0.0001$). (g) The IC50 concentration of 5-fluorouracil or oxaliplatin treatment (24 hr incubation time) of parental and CRK family-deficient cells was determined by XTT assays (Mann–Whitney test; $n = 4$; $*p \leq 0.05$). Dose–response curves and IC50 data shown in Supporting Information Figure S4. Representative loading control for all samples is shown. [Color figure can be viewed at wileyonlinelibrary.com]

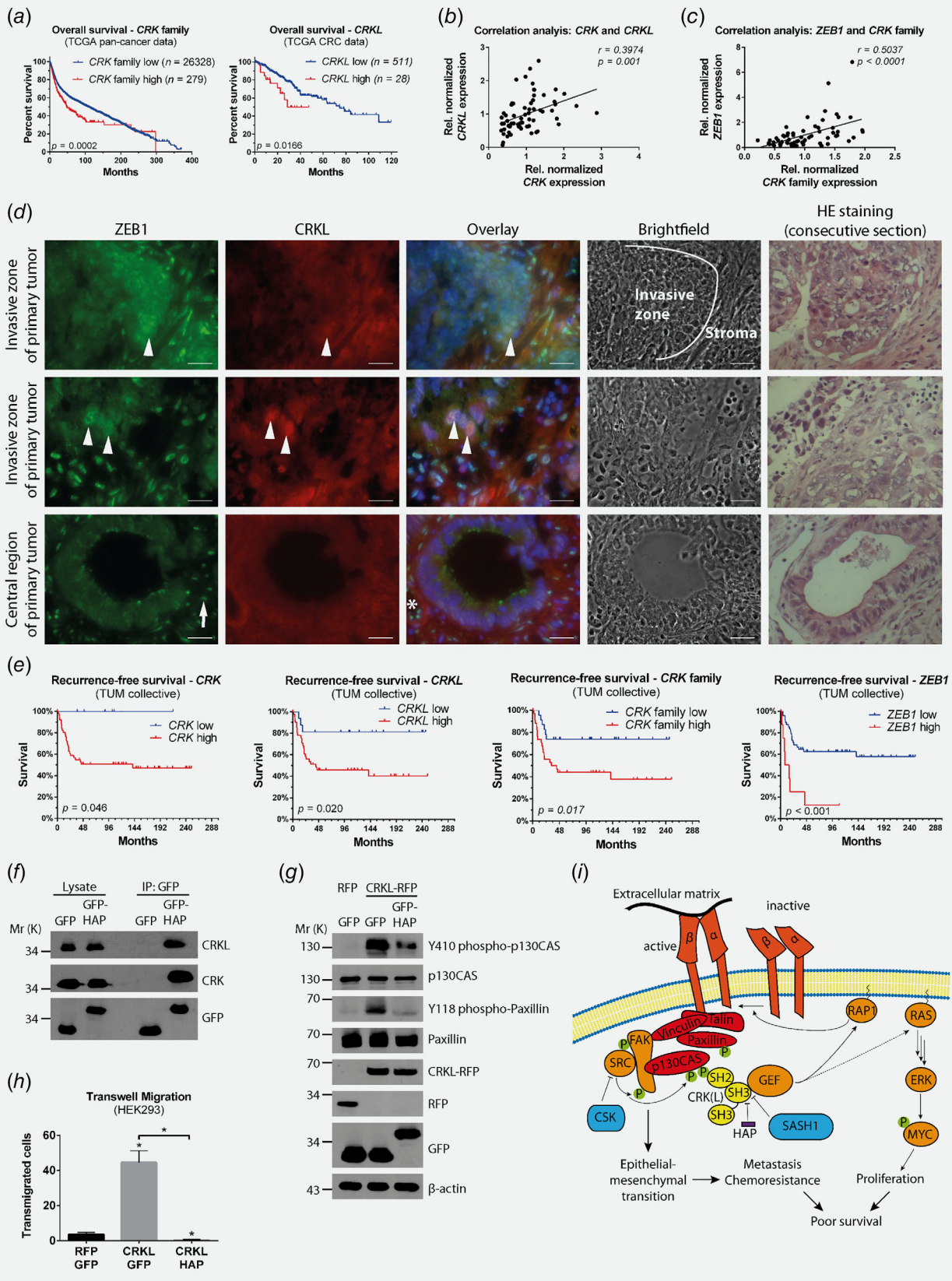


Figure 6. Legend on next page.

CRK protein-deficiency locks colorectal cancer cells in an epithelial state

While activation of SRC through deficiency of its inhibitor CSK was sufficient to induce EMT in parental cells, CRK family-deficiency rendered cells unable to undergo EMT upon loss of CSK (Fig. 4a). Moreover, treatment with TNF induced EMT only in the parental line, while treated CRK family-deficient cells rapidly underwent cell death (not shown). Furthermore, loss of either CRKL or CRK alone kept cells in a more epithelial state upon CSK-deficiency (Fig. 4b). In both the parental line and CRKL- or CRK-deficient cells, CSK-deficiency induced the formation of multipolar lamellipodia (Fig. 4c). Cell individualization and loss of E-cadherin at intercellular adhesions were only observed for CSK-deficient parental cells. Cells lacking the CRK-family showed no alteration of their highly epithelial phenotype (Fig. 4c). In accordance, CSK-deficient parental cells exhibited a striking increase in motility and invasiveness (Fig. 4d). This increase was severely diminished in CRKL- or CRK-deficient cells (Fig. 4d). Cell migration and invasiveness were completely abrogated in CRK family-deficient cells even upon loss of CSK, and after prolonged assay duration (Figs. 4d and 4e).

Compound CRK family deficiency exerts synergistic effects on MYC signaling and attenuates proliferation and chemoresistance

Transcriptomes of parental cells were compared to CRK-, CRKL- or compound-deficient cells *via* RNAseq. Some variation in gene expression was observed between replicates, likely caused by cell density, however, only compound CRK family-deficiency lead to major differences in gene expression, with 549 genes significantly altered compared to the parental line (Supporting Information Table S4). In contrast, single loss of either CRK or CRKL resulted in minor changes (CRKL: 1 deregulated gene; CRK: 6 deregulated genes; data not shown). Transcripts associated with a highly epithelial and differentiated state (e.g., *CLDN3*, *CLDN4*) were significantly upregulated in

CRK family-deficient cells (Supporting Information Table S4). Gene set enrichment analysis revealed significantly deregulated hallmark processes like “hypoxia”, “epithelial-mesenchymal transition” and “apical junction”, in accordance with our earlier findings (Fig. 5a; Supporting Information Tables S2 and S3). Importantly, hallmarks involved in cell-cycle progression (MYC Targets_V1, E2F Targets and G2M Checkpoint) were negatively enriched in CRK family-deficient cells, suggesting that CRK proteins promote cell proliferation and MYC transcriptional activity (Fig. 5b; Supporting Information Table S3). The transcriptome analysis further revealed a downregulation of genes in the “DNA repair” hallmark upon CRK family-deficiency (Supporting Information Table S3). Heatmap analysis illustrates the downregulation of the gene set “MYC Targets_V1” in two independent CRK family-deficient clones (Fig. 5c). CRK family-deficient cells exhibited significantly reduced c-MYC protein levels; of note, c-MYC protein levels were not altered in nontargeting controls (Supporting Information Fig. S1), as well as strongly reduced phosphorylation of ERK1/2 (Fig. 5d). Moreover, loss of CRK family genes lead to a significant reduction in cell proliferation, more pronounced under serum-stress conditions (2% fetal bovine serum; Fig. 5e). To validate the transcriptome data, and since EMT was described to be important for cell survival and chemoresistance,^{34,35} anchorage-independent growth and sensitivity against cytotoxic agents was assessed. CRK family-deficient cells exhibited a significantly reduced number of colonies in soft agar (Fig. 5f), and had increased sensitivity toward 5-fluorouracil or oxaliplatin (Fig. 5g and Supporting Information Fig. S4).

Intratumoral CRK family expression correlates with ZEB1 and is prognostic in colorectal cancer

To investigate their clinical relevance, the effects of aberrant expression and/or genomic alteration of *CRK* and/or *CRKL* on survival was analyzed in a TCGA pan-cancer dataset comprising 72,175 tissue samples from 236 studies in the cBioPortal platform.^{20,21} *CRKL* was altered in 1,398 of 72,175 samples

Figure 6. CRK family expression correlates with ZEB1 in patient samples, and has prognostic significance in advanced colorectal cancer. (a) Left panel: Data from the TCGA database across all cancer entities were analyzed. Patients with high levels (amplification, gain or over twofold increased expression) of *CRK* and/or *CRKL* exhibited significantly reduced overall survival (Log-rank test; *p* values indicated in the graph). Right panel: High expression of *CRKL* in a colorectal cancer (CRC) TCGA dataset was significantly associated with poor overall survival. (b,c) Expression of *CRK* family and *ZEB1* in UICC Stage III colorectal cancer samples (TUM collective) was determined by qRT-PCR. Correlation analysis between *CRK* and *CRKL* expression, as well as between *ZEB1* expression and *CRK* family expression (mean of normalized *CRK* and *CRKL*), was performed (Pearson *r*; *n* = 65; *r* and *p* values indicated in the graphs). (d) Tumors from the TUM cohort were stained for ZEB1 (green), CRKL (red) and nuclei (DAPI, blue) for immunofluorescence microscopy (size bar = 20 μm; arrowheads indicate tumor cells with high ZEB1 and CRKL staining; arrow indicates example of a stroma cell; asterisk indicates blood vessel). Hematoxylin and eosin (HE) staining of consecutive sections in the right panels. (e) Distant recurrence-free survival, dependent on high or low expression of *CRK* (low: 8; high: 57), *CRKL* (low: 17; high: 48), *CRK* family (low: 25; high: 40) or *ZEB1* (low: 54; high: 10; *n* = 64 due to failed qRT-PCR in one case) is shown (Log-rank test; *n* = 65; *p* values indicated in the graphs). (f) HEK293 cells transiently expressing GFP or GFP-HAP were subjected to co-immunoprecipitations using anti-GFP antibody-coupled beads. Immunoprecipitates were blotted to investigate interactions with CRK and CRKL. (g, h) HEK293 cells stably expressing RFP or CRKL-RFP were transfected to transiently express GFP or GFP-HAP. Phosphorylation of p130CAS and paxillin was analyzed by immunoblotting (g). Transwell assays (24 hr) were performed to analyze cell migration (h, Mann-Whitney test; *n* = 4; **p* = 0.0286). (i) Schematic model for the functional role of the CRK family in EMT, tumor progression and metastasis. [Color figure can be viewed at wileyonlinelibrary.com]

(1.9%), and *CRK* in 1,498 (2%) of cases. There was a significant co-occurrence between *CRK* and *CRKL* alterations (Log Odds ratio: 1.756, adjusted *p*-value: 0.001), as well as with the tumor suppressor *SASH1* (Log Odds ratio: 1.855, adjusted *p*-value: 0.001), and with the EMT regulator *ZEB1* (Log Odds ratio: 2.589, adjusted *p*-value: 0.001). Genomic and transcriptomic amplification of *MYC* and *CRK* family genes were significantly associated (Log Odds ratio: 1.899, adjusted *p*-value: <0.001). For *MYC*, genomic amplifications represented the most frequent type of alteration; similar findings were obtained for *CRKL* (66% of all cases), whereas 35% of cases showed amplification for *CRK*, and 43% had genomic deletions (Supporting Information Appendix S1). No significant association was observed between microsatellite instability status, mutations in *KRAS* exon 2 or *BRAF* exon 15, and the expression of either *CRK* or *CRKL*, in the TUM cohort (Supporting Information Fig. S6). Prognostic significance was assessed in a curated, non-redundant subset of cases with follow-up data (*n* = 26,607). Whereas alterations in *CRK* alone showed no significant prognostic association, *CRKL* genetic alterations were significantly associated with reduced overall survival (Supporting Information Fig. S5). Analysis of both *CRK* family members revealed that overall survival, as well as disease-free survival, were highly significantly reduced in the 702 cases with altered *CRK* family genes, as opposed to 37,500 patients without alterations (Supporting Information Fig. S5). Furthermore, the pan-cancer TCGA dataset, as well as the subset from patients with colorectal cancer (CRC), were subgrouped into high (i.e., amplification, gain or over twofold increased expression) and low (deletions, average or downregulated expression) levels of *CRK* family genes. High *CRK* family levels were highly significantly associated with poor overall survival in the pan-cancer TCGA dataset (Fig. 6a, left panel). In contrast to high *CRK* levels (Supporting Information Fig. S5), high levels of *CRKL* were also significantly associated with poor overall survival in the CRC subset (Fig. 6a, right panel). Expression of *CRK* and *ZEB1* was determined in an independent cohort of locally advanced (UICC Stage III) colorectal cancer from our clinic (Supporting Information Table S1). *CRKL* expression, which had previously been determined in the same cohort,⁵ was included to calculate *CRK* family gene expression as the mean of normalized *CRK* and *CRKL* expression. Expression of *CRK* significantly correlated with expression of *CRKL*, and *CRK* family expression with *ZEB1* in tumor tissue (Figs. 6b and 6c). This was in accordance with protein levels, and increased nuclear *ZEB1* (an EMT hallmark), was found in dedifferentiated regions (Fig. 6d, upper and middle panels, arrowheads), especially in the invasive front (Fig. 6d, upper panels). Central regions of the tumor exhibited relatively low levels of *ZEB1* with no particular enrichment of nuclear staining (Fig. 6d, lower panels). Of note, invasive cancer cells with nuclear *ZEB1* showed increased *CRKL* staining (Fig. 6d, arrowheads). Upon Kaplan–Meier analysis, mRNA expression of either *CRK* or *CRKL* was significantly associated with

reduced recurrence-free survival for distant metastasis (Fig. 6e). Importantly, patients with elevated mRNA expression of *CRK* family genes, or *ZEB1*, exhibited significantly reduced distant recurrence-free survival (Fig. 6e). Multivariable regression analysis revealed no further correlation with clinicopathological parameters (data not shown). Given their clinical relevance, we investigated if *CRK* family-mediated signaling can be targeted by small ligands. The N-terminal SH3 domains of *CRK* family adaptors were selected as target structure, since the inactivation of the *CRKL* N-terminal SH3 domain exerted inhibitory effects on SRC/FAK signaling. HAPs have been described to bind these domains, and the corresponding sequence was fused to the C-terminus of GFP.^{15,16} The HAP (residues CVDNSPPALPPKRRRSAPS) fusion protein associated with both *CRK* and *CRKL* (Fig. 6f). Transient expression of GFP-HAP counteracted *CRKL*-mediated phosphorylation of the SRC/FAK substrates p130CAS and paxillin, as well as the increased transwell migration of HEK293 cells stably expressing *CRKL*-RFP, demonstrating that inhibition of *CRK* family adaptor signaling is feasible (Figs. 6g and 6h). In summary, a working model of *CRK* proteins and their proposed novel role in EMT is proposed (Fig. 6i).

Discussion

Epithelial–mesenchymal transition (EMT) allows sessile cancer cells of epithelial origin to acquire a motile and invasive phenotype, required for metastasis formation.² Due to its reversibility and the existence of intermediate states, EMT is recognized as cell plasticity.^{3,36} While its role in metastasis is controversially discussed, EMT has been shown to render cells resistant to chemotherapeutics and to increase cell survival.^{34,35,37,38} Inhibition of EMT represents an innovative strategy by simultaneously reducing metastatic capabilities and increasing the effectiveness of chemotherapy. The present study demonstrates that the *CRK* family of signal adaptor proteins plays a crucial role as positive regulators of EMT in colorectal and, presumably, also in pancreatic cancer. Mechanistically, *CRK* proteins act as essential positive feedback amplifiers of FAK/SRC signaling *via* the small GTPase RAPI, leading to sustained FAK/SRC activation, and thus invasiveness, chemoresistance, EMT and ultimately poor patient survival.

The *CRK* family consists of two ubiquitously expressed genes: *CRK*, encoding the isoforms *CRK*-I and *CRK*-II and *CRKL*,⁶ sharing similar structures and interaction specificities, but with some functional differences.³⁹ Here we found that loss of either *CRKL* or *CRK* shifted colorectal cancer cells toward a more epithelial and less malignant phenotype, reverting the mesenchymal phenotype. We found considerable functional redundancy of *CRK* proteins in EMT, and single *CRKL*- or *CRK* deficiency showed only attenuated effects. Loss of the *CRK* family exerted synergistic inhibitory effects on most aspects of cancer cell aggressiveness and malignancy. This was reflected by transcriptome-wide changes observed only upon complete *CRK* family-deficiency, but not upon

individual loss of CRK or CRKL. In accordance, clinical data showed a significant co-occurrence of genetic and genomic alterations of *CRK* together with *CRKL* in individual tumors (TCGA datasets). *In vitro*, complete loss of CRK family proteins induced a pronounced epithelial phenotype, a concomitant loss of motility, invasiveness, anchorage-independent survival and the ability to undergo EMT, as well as a strong increase in chemosensitivity. Therefore, targeting of CRK family proteins may be a promising therapeutic strategy. Besides toxicity, a major disadvantage of current EMT-targeting therapeutics is increased proliferation of cells in a more epithelial state. Inhibition of CRK proteins would have the advantage of decreasing proliferation, likely by reduced activity of ERK1/2 and c-MYC. Activation of ERK1/2 is mediated by CRK proteins through activation of RAS, downstream of SOS1.^{6,40} RAS and the downstream MAP-kinases ERK1/2, in turn, have been described to stabilize c-MYC.⁴¹ This could explain the observed transcriptome-wide downregulation of c-MYC targets, diminished c-MYC protein levels and the reduced proliferation upon CRK family-deficiency. Genetic alterations of *MYC* and *CRK* family members co-occurred highly significantly, most frequently gene locus amplifications. Thus, inhibition of CRK family signaling might constitute an innovative mechanism to address the notoriously difficult MYC pathway.

Next, we investigated how CRK-family proteins induce the observed downstream signaling. CRK proteins require both a functional SH2-domain and the N-terminal SH3 domain.^{6,42} In accordance, only wild-type CRKL promoted signaling of SRC/FAK kinases, but not CRKL with inactivating mutation in either the SH2- or N-terminal SH3 domain. Furthermore, expression of CRKL with a point-mutated inactive N-terminal SH3 domain (W160R) led to decreased SRC/FAK activation and substrate phosphorylation, suggesting a dominant-negative effect. This could be explained by the presence of the functional SH2 domain in this mutant, binding and blocking phosphotyrosine residues of paxillin and p130CAS,⁶ while downstream signaling cannot occur due to the inactive N-terminal SH3 domain. This interaction interferes with binding of endogenous type CRK and CRKL, leading to reduced SRC/FAK activation. Forced localization of CRKL to focal adhesions was sufficient to rescue haptotaxis defects of murine embryonic fibroblasts lacking SRC, FYN and YES, highlighting the importance of CRK protein recruitment to phosphorylated integrin-associated scaffold proteins.¹⁰ Therefore, the N-terminal SH3 domains of CRK and CRKL represent a target structure to simultaneously counteract the recruitment of other SH2 domain-containing adaptors in a dominant-negative way. Given that the N-terminal SH3 domains are highly similar, both proteins could be targeted with a single antagonist, even though toxicity remains to be determined.⁶ Since CRK and CRKL also have similar ligand-binding preferences, antagonistic peptides to the N-terminal SH3 domain have been developed and could be analyzed for their ability to

inhibit EMT.¹⁶ Indeed, recombinant expression of GFP with a C-terminal high-affinity peptide, which associated with both CRKL and CRK N-terminal SH3 domains, counteracted SRC/FAK substrate phosphorylation. This highlights that transient targeting of CRK family-mediated signaling is feasible.

According to the working model (Fig. 6i), CRK proteins act as central amplifiers of SRC/FAK kinase complex signaling downstream of integrins: upon initial integrin engagement, the integrin-associated kinase FAK is auto-phosphorylated, which induces complex formation between SRC and FAK.^{27,43} After further phosphorylation of FAK by SRC, the active SRC/FAK kinase complex phosphorylates the integrin-associated scaffold proteins paxillin and p130CAS to induce CRK protein recruitment *via* their SH2 domains.^{10,12,44,45} CRK proteins then recruit guanine nucleotide exchange factors like C3G through their N-terminal SH3 domains into close proximity to the membrane, inducing activation of the membrane-associated small GTPases, for example, RAP1.^{6,29,40,46} Active, GTP-loaded RAP1 is described to induce local activation of integrins *via* talin recruitment. In accordance, both CRK family-deficiency and RAP1 inhibition induced a diffuse localization of talin, and CRK family-deficient cells exhibited strongly reduced adhesion to fibronectin. Furthermore, our data indicated that CRK proteins activate SRC/FAK signaling through RAP1, as RAP1 inhibition reversed the effect of elevated CRKL expression. These data are supported by independent reports that indicate a role of RAP1 in FAK autophosphorylation.⁴⁷ Further integrin molecules may thus be activated due to local GTP-loading of small GTPases like RAP1, downstream of CRK proteins. This potentially promotes integrin clustering and focal adhesion formation, absent in CRK family-deficient cells, but required for enhanced activation of the SRC/FAK complex. CRK proteins, which were previously regarded only as downstream effectors, can therefore rather be seen as central amplifiers to promote sustained local activation of SRC/FAK signaling. In accordance, activation of SRC *via* engineered deficiency of its inhibitor CSK, was sufficient to induce EMT in the parental line, but failed to do so in CRK family-deficient cells, despite an equivalent level of SRC activation. This indicates a novel potential role of CSK in EMT. CRK has been proposed earlier to activate SRC dependent on CSK.^{48,49} However, genetic loss of CSK did not completely rescue the SRC/FAK signaling defects of CRK family-deficient cells, indicating that the mechanism is not solely dependent on CSK. CRK proteins may rather mediate the local assembly of integrin-associated protein complexes by spatiotemporally restricted GTP loading of small GTPases like RAP1. Thus, they enhance activation of SRC and FAK partially independent from each other to induce SRC/FAK complex signaling. Therefore, CRK family-deficiency uncouples SRC activation from the activation of FAK. Increased SRC/FAK signaling downstream of CRK proteins phosphorylates paxillin and p130CAS, providing further binding sites for CRK proteins and amplifying the signaling circuit.

Our data support the clinical importance of CRK family proteins, and confirm the functional interactions between the CRK-family and the metastasis suppressor SASH1, as well as with ZEB1 and MYC, in patient samples. High expression or genomic amplification of CRK-family genes, as well as of the EMT regulator ZEB1, was significantly associated with reduced survival. Moreover, a direct correlation between intratumoral expression of ZEB1 and CRK family was observed. Furthermore, ZEB1 and CRKL were increased at protein level in the invasive margins of primary colorectal tumors. Accordingly, *in vitro* findings showed that loss of CRK family induced a strongly reduced ZEB1 expression. Similar findings for ZEB1 were reported earlier, highlighting the role of EMT in cancer cell invasion.⁵⁰

Taken together, the data highlight CRK family proteins as potential therapeutic targets. In the present study, SRC/FAK substrate phosphorylation and cell migration were inhibited by a peptide-based CRK family inhibitor in HEK293 cells.

Therefore, pharmacological targeting of CRK proteins, specifically their N-terminal SH3 domain, might be feasible and could potentially provide a new and innovative therapeutic strategy to counteract EMT, invasiveness, chemoresistance and thus metastasis formation. Further studies are required to analyze the effects on EMT and other pro-metastatic traits in colorectal cancer cells, using technologies such as cell-penetrating peptides.

Acknowledgements

Our study was supported by the Deutsche Krebshilfe e.V., No. 111822 (to K.P.J.). The authors would like to thank Anja Conrad for excellent technical assistance, and Maximilian Ehrenfeld for critical discussion.

Conflict of interest

The authors declare no competing financial interests in relation to the work described.

References

- Valastyan S, Weinberg RA. Tumor metastasis: molecular insights and evolving paradigms. *Cell* 2011;147:275–92.
- Kalluri R, Weinberg RA. The basics of epithelial-mesenchymal transition. *J Clin Invest* 2009;119:1420–8.
- Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. *Nat Rev Mol Cell Biol* 2014;15:178–96.
- Peinado H, Olmeda D, Cano A. Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat Rev Cancer* 2007;7:415–28.
- Franke FC, Muller J, Abal M, et al. The tumor suppressor SASH1 interacts with the signal adaptor CRKL to inhibit epithelial-mesenchymal transition and metastasis in colorectal cancer. *Cell Mol Gastroenterol Hepatol* 2019;7:33–53.
- Bell ES, Park M. Models of crk adaptor proteins in cancer. *Genes Cancer* 2012;3:341–52.
- Lamorte L, Royal I, Naujokas M, et al. CRK adapter proteins promote an epithelial-mesenchymal-like transition and are required for HGF-mediated cell spreading and breakdown of epithelial adherens junctions. *Mol Biol Cell* 2002;13:1449–61.
- Matsumoto R, Tsuda M, Wang L, et al. Adaptor protein CRK induces epithelial-mesenchymal transition and metastasis of bladder cancer cells through HGF/c-met feedback loop. *Cancer Sci* 2015;106:709–17.
- Arai A, Nosaka Y, Kohsaka H, et al. CRKL activates integrin-mediated hematopoietic cell adhesion through the guanine nucleotide exchange factor C3G. *Blood* 1999;93:3713–22.
- Li L, Guris DL, Okura M, et al. Translocation of CRKL to focal adhesions mediates integrin-induced migration downstream of Src family kinases. *Mol Cell Biol* 2003;23:2883–92.
- Ungewiss C, Rizvi ZH, Roybal JD, et al. The microRNA-200/Zeb1 axis regulates ECM-dependent beta1-integrin/FAK signaling, cancer cell invasion and metastasis through CRKL. *Sci Rep* 2016;6:18652.
- Lamorte L, Rodrigues S, Sangwan V, et al. Crk associates with a multimolecular Paxillin/GIT2/beta-PIX complex and promotes Rac-dependent relocalization of Paxillin to focal contacts. *Mol Biol Cell* 2003;14:2818–31.
- Geyer PE, Maak M, Nitsche U, et al. Gastric adenocarcinomas express the glycosphingolipid Gb3/CD77: targeting of gastric cancer cells with Shiga toxin B-subunit. *Mol Cancer Ther* 2016;15:1008–17.
- Martini M, Gnann A, Scheikl D, et al. The candidate tumor suppressor SASH1 interacts with the Actin cytoskeleton and stimulates cell-matrix adhesion. *Int J Biochem Cell Biol* 2011;43:1630–40.
- Kardinal C, Konkol B, Schulz A, et al. Cell-penetrating SH3 domain blocker peptides inhibit proliferation of primary blast cells from CML patients. *FASEB J* 2000;14:1529–38.
- Posern G, Zheng J, Knudsen BS, et al. Development of highly selective SH3 binding peptides for Crk and CRKL which disrupt Crk-complexes with DOCK180, SoS and C3G. *Oncogene* 1998;16:1903–12.
- Ran FA, Hsu PD, Wright J, et al. Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* 2013;8:2281–308.
- Rimkus C, Martini M, Friederichs J, et al. Prognostic significance of downregulated expression of the candidate tumour suppressor gene SASH1 in colon cancer. *Br J Cancer* 2006;95:1419–23.
- Nitsche U, Rosenberg R, Balmert A, et al. Integrative marker analysis allows risk assessment for metastasis in stage II colon cancer. *Ann Surg* 2012;256:763–71. discussion 71.
- Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal* 2013;6:pl1.
- Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov* 2012;2:401–4.
- Parekh S, Ziegenhain C, Vieth B, et al. The impact of amplification on differential expression analyses by RNA-seq. *Sci Rep* 2016;6:25533.
- Macosko EZ, Basu A, Satija R, et al. Highly parallel genome-wide expression profiling of individual cells using Nanoliter droplets. *Cell* 2015;161:1202–14.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
- Kuleshov MV, Jones MR, Rouillard AD, et al. Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 2016;44:W90–7.
- Wang H, Wang HS, Zhou BH, et al. Epithelial-mesenchymal transition (EMT) induced by TNF-alpha requires AKT/GSK-3beta-mediated stabilization of snail in colorectal cancer. *PLoS One* 2013;8:e56664.
- Toutant M, Costa A, Studler JM, et al. Alternative splicing controls the mechanisms of FAK autophosphorylation. *Mol Cell Biol* 2002;22:7731–43.
- Okada M. Regulation of the SRC family kinases by Csk. *Int J Biol Sci* 2012;8:1385–97.
- Sakkab D, Lewitzky M, Posern G, et al. 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* 2000;275:10772–8.
- Bos JL, de Bruyn K, Enserink J, et al. The role of Rap1 in integrin-mediated cell adhesion. *Biochem Soc Trans* 2003;31:83–6.
- McSherry EA, Brennan K, Hudson L, et al. Breast cancer cell migration is regulated through junctional adhesion molecule-A-mediated activation of Rap1 GTPase. *Breast Cancer Res* 2011;13:R31.
- Lilja J, Zacharchenko T, Georgiadou M, et al. SHANK proteins limit integrin activation by directly interacting with Rap1 and R-Ras. *Nat Cell Biol* 2017;19:292–305.
- Zhang YL, Wang RC, Cheng K, et al. Roles of Rap1 signaling in tumor cell migration and invasion. *Cancer Biol Med* 2017;14:90–9.
- Fischer KR, Durrans A, Lee S, et al. Epithelial-to-mesenchymal transition is not required for lung

- metastasis but contributes to chemoresistance. *Nature* 2015;527:472–6.
35. Zheng X, Carstens JL, Kim J, et al. Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* 2015;527:525–30.
 36. Chaffer CL, San Juan BP, Lim E, et al. EMT, cell plasticity and metastasis. *Cancer Metastasis Rev* 2016;35:645–54.
 37. Aiello NM, Brabletz T, Kang Y, et al. Upholding a role for EMT in pancreatic cancer metastasis. *Nature* 2017;547:E7–8.
 38. Ye X, Brabletz T, Kang Y, et al. Upholding a role for EMT in breast cancer metastasis. *Nature* 2017;547:E1–3.
 39. Jankowski W, Saleh T, Pai MT, et al. Domain organization differences explain Bcr-Abl's preference for CrkL over CrkII. *Nat Chem Biol* 2012;8:590–6.
 40. Cheung HW, Du J, Boehm JS, et al. Amplification of CRKL induces transformation and epidermal growth factor receptor inhibitor resistance in human non-small cell lung cancers. *Cancer Discov* 2011;1:608–25.
 41. Junttila MR, Westermarck J. Mechanisms of MYC stabilization in human malignancies. *Cell Cycle* 2008;7:592–6.
 42. Senechal K, Heaney C, Druker B, et al. Structural requirements for function of the Crkl adapter protein in fibroblasts and hematopoietic cells. *Mol Cell Biol* 1998;18:5082–90.
 43. Schaller MD, Hildebrand JD, Shannon JD, et al. Autophosphorylation of the focal adhesion kinase, pp125FAK, directs SH2-dependent binding of pp60src. *Mol Cell Biol* 1994;14:1680–8.
 44. Schaller MD, Parsons JT. pp125FAK-dependent tyrosine phosphorylation of paxillin creates a high-affinity binding site for Crk. *Mol Cell Biol* 1995;15:2635–45.
 45. Calalb MB, Polte TR, Hanks SK. Tyrosine phosphorylation of focal adhesion kinase at sites in the catalytic domain regulates kinase activity: a role for Src family kinases. *Mol Cell Biol* 1995;15:954–63.
 46. Sakakibara A. Novel function of chat in controlling cell adhesion via Cas-Crk-C3G-pathway-mediated Rap1 activation. *J Cell Sci* 2002;115:4915–24.
 47. Carmona G, Gottig S, Orlandi A, et al. Role of the small GTPase Rap1 for integrin activity regulation in endothelial cells and angiogenesis. *Blood* 2009;113:488–97.
 48. Watanabe T, Tsuda M, Makino Y, et al. 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* 2009;19:638–50.
 49. Kumar S, Lu B, Davra V, et al. Crk tyrosine phosphorylation regulates PDGF-BB-inducible Src activation and breast Tumorigenicity and metastasis. *Mol Cancer Res* 2018;16:173–83.
 50. Sanchez-Tillo E, Lazaro A, Torrent R, et al. ZEB1 represses E-cadherin and induces an EMT by recruiting the SWI/SNF chromatin-remodeling protein BRG1. *Oncogene* 2010;29:3490–500.