The MACC1 Dimerization and its Function in Tumor Metastasis

Inaugural-Dissertation to obtain the academic degree

Doctor rerum naturalium (Dr. rer. nat)

submitted to the Department of Biology, Chemistry, Pharmacy of Freie Universität Berlin

by

Malti Dumbani

Berlin 2023

This project was conducted from January 2019 until August 2023 in the research group “Translational Oncology of Solid Tumors” under the supervision of Prof. Ulrike Stein at the Experimental and Clinical Research Center, Charité-Universitätsmedizin Berlin and Max-Delbrück-Center for Molecular Medicine in the Helmholtz Association.

1st Reviewer: Prof. Dr. Ulrike Stein

2nd Reviewer: Prof. Dr. Gerhard Wolber

Thesis defense on: 26.01.2024

Herewith I certify that I have prepared and written my thesis independently and that I have not used any sources and aids other than those indicated by me.

I also declare that I have not applied for an examination procedure at any other institution and that I have not submitted the dissertation in this or any other form to any other faculty as a dissertation.

Table of Contents

Table of Figures

1 Summary

Colorectal cancer, a highly heterogeneous cancer, continues to be a leading cause of mortality worldwide. While the 5-year survival rates for patients with Stage I and II are high, there has been little or no improvement of survival for patients with metastases. To make matters worse, 20% of the patients already present with metastasis at the time of diagnosis. Therefore, early detection of patients who are at high risk of developing metastases using biomarkers is key to improving patient survival.

Metastasis-associated in colon cancer 1 (MACC1) is one such biomarker that has been directly linked to metastasis development, reduced survival, and worse overall outcomes. In addition to identifying high-risk patients, MACC1 is biologically linked to tumor and metastasis development. Specifically, the MACC1 structure contains diverse domains and several tyrosine sites capable of versatile interactions. Therefore, the aim of the first part of the project was to study the role of tyrosine sites close to the N-terminus of MACC1. Employing computational tyrosine phosphorylation prediction tools, site Tyr379 and SRC kinase as one of the promising kinases responsible for its phosphorylation were identified. Preliminary examination reveals an association between MACC1 and SRC.

Despite extensive evidence describing the functional diversity of MACC1, little is known about the structural features and self-association property of MACC1. To address this gap in knowledge, the goal of the second part of this project was to systematically evaluate the structural properties of MACC1 and the self-association capability of MACC1. Using AlphaFold2, the structures of MACC1 and MACC1 dimer were revealed. Val212, Ileu214, and Cys216 present in the ZU5 domain of MACC1 were found to be critical for dimerization. The knowledge gained from the AI prediction was transferred to set up a bioluminescence resonance energy transfer (BRET) assay to analyze MACC1 dimerization and the effect of mutation on dimerization. In addition to validating the presence of MACC1 dimer in living cells, the BRET assay confirmed reduced MACC1 self-association when the above residues were mutated. Ultimately, the impact of these mutations on MACC1 signaling and metastasis properties was verified using an *in vitro* metastasis assay.

In summary, these results shed new light on the MACC1 structural characteristics particularly the presence of MACC1 homodimer, and reveal the residues important for dimerization, thus providing a framework for future development of intervention strategies.

2 Zusammenfassung

Das kolorektale Karzinom, ein sehr heterogenes Karzinom, ist weiterhin eine der führenden Todesursachen weltweit. Während die 5-Jahre-Überlebensrate bei Patienten in Stufe I oder II hoch ist, hat sich diese bei Patienten mit Metastasen kaum verbessert. Allerdings werden bereits bei 20% der Patienten Metastasen zum Diagnosezeitpunkt detektiert. Deswegen stellt die frühzeitige Identifikation von Patienten mit hohem Risiko zur Entwicklung von Metastasen mit Hilfe von Biomarkern einen Schlüssel zur Verbesserung der Überlebensraten dar.

Metastasis-associated in colon cancer 1 (MACC1) ist ein solcher Biomarker, der in direkter Verbindung mit der Entwicklung von Metastasen und geringeren Überlebenschancen steht. Neben der Rolle als Identifikationsfaktor von Hoch-Risiko-Patienten hat MACC1 auch eine biologische Rolle in der Entwicklung von Tumoren und Metastasen. Hierbei spielen insbesondere die diversen Strukturdomänen und verschiedenen Tyrosinstellen von MACC1 eine Rolle, die unterschiedliche Interaktionen möglich machen. Deswegen war es Ziel des ersten Projektteils, die Funktion N-terminaler Tyrosinstellen von MACC1 zu untersuchen. Mit Hilfe verschiedener Modellierungs-Software zur Vorhersage von Tyrosin-Phosphorylierungen wurden das Tyr379 und die SRC Kinase als vielversprechender Kandidat für die Phosphorylierung dieses Tyrosins identifiziert. Vorläufige Untersuchungen haben bereits eine starke Assoziation zwischen MACC1 und SRC gezeigt.

Trotz vieler Evidenzen zur funktionellen Diversität von MACC1, gibt es bisher nur wenige Daten zu den strukturellen Eigenschaften, einschließlich der selbst-assoziierenden Fähigkeiten von MACC1. Um diese Lücke zu schließen, zielte der zweite Projektteil darauf ab diese strukturellen Eigenschaften und Selbst-Assoziationsfähigkeit systematisch zu untersuchen. Mit Hilfe von AlphaFold2 wurden die Strukturen von MACC1 und des MACC1-Dimers bestimmt. Hierbei wurden Val212, Ileu214 und Cys216 innerhalb der ZU5-Domäne von MACC1 als elementar für die Dimerisierung identifiziert. Basierend auf dem Wissen der KI-Vorhersage wurde ein bioluminescence resonance energy transfer (BRET) Assay durchgeführt, um die Dimerisierung von MACC1 und den Einfluss von Mutagenese dieser elementaren Interaktionsstellen zu analysieren. Durch den BRET Assay konnte somit nicht nur die Existenz von MACC1-Dimeren in vitro gezeigt werden, sondern auch die reduzierte Selbst-Assoziationsfähigkeit nach Mutagenese der oben genannten Aminosäuren. Letztendlich wurde zudem der Einfluss dieser Mutationen auf MACC1 Signalkaskaden und die metastasierenden Eigenschaften verifiziert.

Zusammenfassend werfen diese Ergebnisse ein neues Licht auf die strukturellen Charakteristika von MACC1, insbesondere durch die Entdeckung des MACC1 Homodimers. Dabei wurden außerdem die für die Dimerisierung entscheidenden Aminosäuren identifiziert, worauf zukünftige Entwicklungen von Interventionsstrategien basieren können.

3 Introduction

3.1 Metastasis: Present and Future

Metastasis, the final stage of cancer development, is a process in which cancer cells disseminate from primary tumors and seed in distant organs [1-3]. Mounting evidence clearly demonstrates that the higher mortality rate is directly linked with the development of metastasis in major cancers such as prostate [4], colorectal [5, 6], breast [7], and lung cancer [8]. Compared with primary tumors, metastasis is a dynamic process that comprises highly heterogeneous assemblies of cells with different phenotypes and genetic signatures [9, 10]. The dissemination of cancer cells is a well-orchestrated process involving local invasion into the circulatory system, proceeded by evasion of the immune system, and ultimately development of micro-metastatic colonies at a distant site [1-3, 11].

One important phenomenon that enables the dissemination of cancer cells is the epithelial-mesenchymal transition (EMT). EMT confers immotile epithelial cells with properties required to migrate and invade like mesenchymal cells [11, 12]. Normally, EMT is observed during wound healing and embryogenesis. However, in cancer development, EMT is aberrantly activated, enabling cancer cells to disseminate [1, 11, 13, 14]. This multi-step process is regulated by a group of transcription factors known as EMT-activating transcription factors (EMT-TFs) which downregulate the genes involved in cellular adhesion, and polarity and upregulate the genes involved in cell motility (Figure 1). The four core EMT-TFs include Snail, Slug, Twist-related protein 1 (Twist1) and zinc-finger E-box-binding homeobox (ZEB) [1, 11, 13, 15, 16]. The expressions of EMT-TFs are often correlated with a high risk of metastasis and poor disease outcome in various cancers [11, 13]. For example, Vimentin and E-cadherin are extensively studied as prognostic biomarkers for cancer aggressiveness in the clinic for breast, thyroid, and colorectal cancer (CRC) among others [17, 18]. Moreover, EMT-TFs also activate other hallmarks of cancer such as stemness, survival, and changes in cellular metabolism making them valuable in diagnosis and treatment [1, 19].

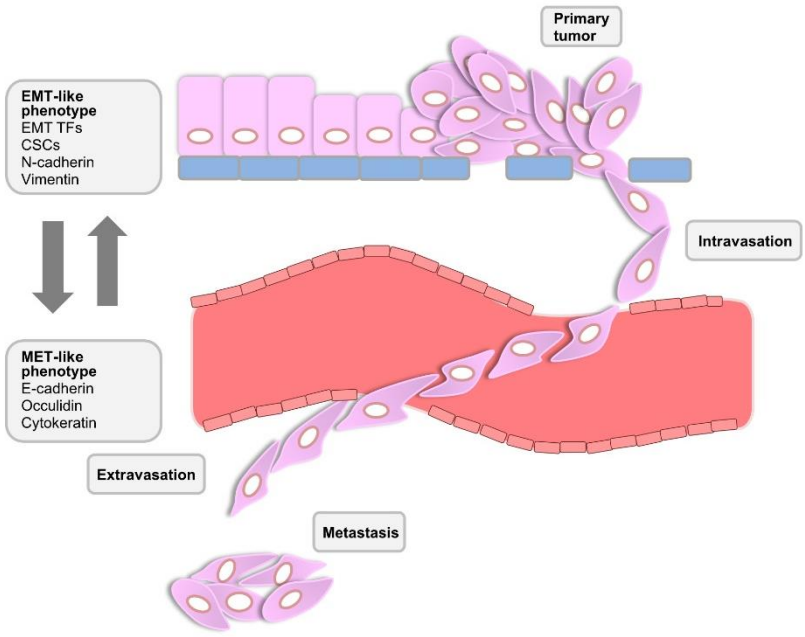

Figure 1: Schematic representation of the metastasis cascade

The dissemination of cells from the primary tumor site to a distant tissue comprises several steps. Through the activation of the EMT-like phenotype, firstly the tumor cells invade the surrounding tissue and enter the circulatory system (intravasation). In the circulatory system, tumor cells need to overcome sheer stress and immune destruction. Only a fraction of tumor cells that enter the circulation survive and can infiltrate the distant sites (extravasation). Notably after extravasation, the phenotype of the tumor cells is reversed to a MET-like phenotype to facilitate colonization and further metastasis formation [3, 11]. Adapted from: 1) Mittal V. (2018). Epithelial Mesenchymal Transition in Tumor Metastasis. 2) Fares et al. (2020) Molecular principles of metastasis: a hallmark of cancer revisited [3, 11, 20]. CSC: Cancer stem cell; EMT: Epithelial mesenchymal transition; MET: Mesenchymal epithelial transition; TF: Transcription factors.

Contrary to conventional assumption that cells undergoing EMT possess either distinct epithelial or mesenchymal properties, recent studies have identified that this transition is rather a gradual process [12, 21]. This is termed as hybrid/partial EMT where cells represent a range of epithelial-mesenchymal transition states and exhibit both epithelial and mesenchymal traits [12, 13]. As an example, hybrid EMT cells exhibit loss of cell polarity, and increase cell motility in combination with cell adhesion [12, 21]. Multiple factors such as the expression of transcription factors, tumor micro-environment, and growth factors contribute to hybrid EMT progression [12, 13, 20, 21]. It is noteworthy that cells exhibiting hybrid EMT traits pose a higher threat and poor clinical outcomes compared to cells that have completed EMT [12, 20].

Moreover, it remains debatable whether metastases arise from cells at later stages of tumorigenesis or cells that are disseminated early in primary tumor development [3, 9]. The linear progression model suggests that metastases are formed by a subpopulation of primary cells that disseminate at later stages of tumorigenesis [1, 9]. Several studies evaluating primary and secondary tissue have found genetic similarities between the two thus favoring the linear progression model. However, given the clonal diversity within a primary tumor, it is difficult to clearly relate the two tumors as clearly pointed out by several reports [1, 22]. On the other hand, the parallel progression model hypothesizes that cells disseminate early in the primary tumor development and attain additional favorable mutations that enable them to survive even in presence of chemotherapy (Figure 2) [23-25]. Then again, given the complexity and low efficacy of metastatic colonization this is unlikely [1, 9, 26]. An extension of the parallel model is a branched model which suggests that metastases develop from divergent cell populations of the primary tumor and evolve independently based on the cues from the new microenvironment [9, 27].

Advancements in sequencing technologies in combination with modern bioinformatic tools have made it possible to gain detailed insight into the dynamic transition of cells leading to metastasis [9, 28]. Primarily, next-generation sequencing of diverse tumors has highlighted the existence of inter- and intra-tumor heterogeneity among various cancer entities [9]. These studies combined with knowledge from mouse models not only provide valuable insights into the metastasis-promoting pathways but also determine the evolution of the clonal mutation from the primary tumor to the secondary tumor [29]. However, large-scale solid tumor studies struggled to identify the key driver mutation resulting in metastasis [9, 30-33]. The importance of epigenetic regulation and genes involved in chromatin remodeling and DNA methylation were suspected to be integral drivers of metastasis [34].

Other studies focusing on single tumor entities show that potential metastasis drivers could vary depending on the site of the primary tumor, for example in the case of metastatic prostate cancer mutations in tumor suppressor genes such as APC, TP53, RB1, PTEN, BRCA2 were commonly detected [35]. In addition to finding new subpopulations at secondary sites, these studies also found some similarities between the primary and secondary sites, thus, supporting the notion of metastasis initiation from primary tumor subpopulations [1, 9]. As pointed out by other researchers, the metastatic clone evolution detected in these cell populations could be due to chemotherapy or drug-resistance [27, 28, 36]. Although these studies offered valuable information about the mutational landscape and gene alterations in malignant progression, they lacked consensus due to tumor heterogeneity at the bulk level [9]. Therefore, novel methods such as

single-cell sequencing that characterize individual cells at different timepoints of metastasis progression are promising [9]. Like bulk sequencing, single-cell sequencing results have also displayed a lack of consensus in determining metastasis evolution and mechanism of dissemination [37, 38]. Therefore, hinting towards underlying biological complexity among various tumor entities [23, 27, 28, 39]. Another reason for contradicting results may be due to the limited number of paired primary-metastatic samples [9].

To surpass these challenges and to examine the evolutionary trajectory prospectively, scientists have developed various tracers that can be introduced stably into cells. Some examples of lineage-tracing approaches are optical barcoding based on fluorescence-based labeling models such as Brainbow and genetic barcoding established based on CRISPR-Cas9 are gaining interest [40-42]. However, it is important to note that these techniques are limited to preclinical systems, and thus the biological relevance of these findings in clinical settings is still debatable [9].

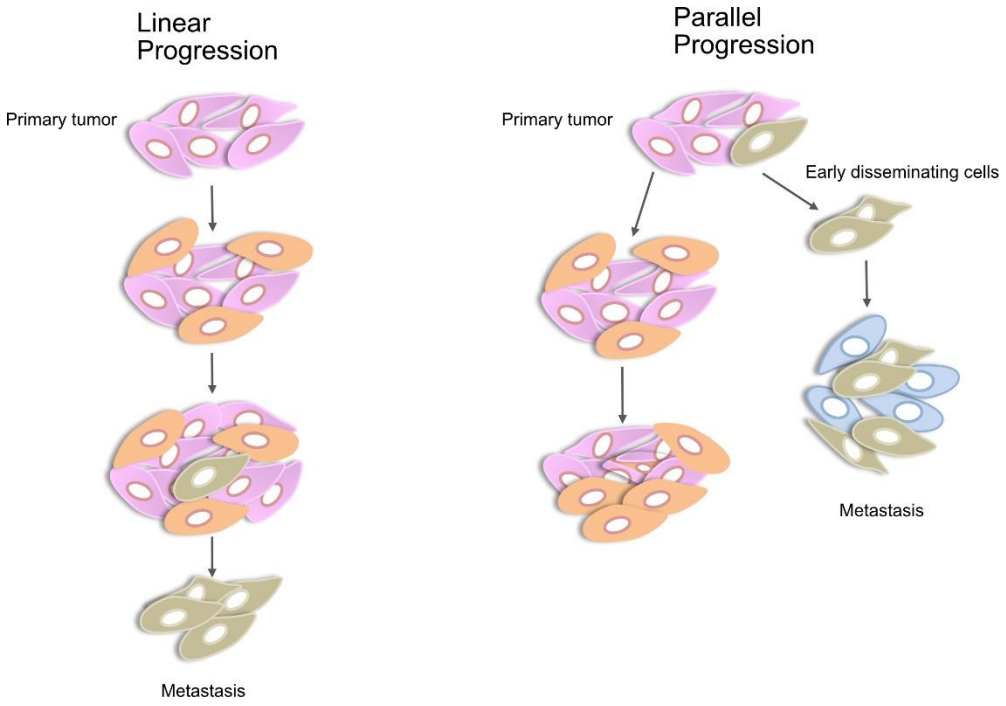

Figure 2: Two models of metastasis evolution

The two well-known models of metastasis progression are linear progression and parallel progression. While the linear progression model assumes that metastasis arises from cells at a later stage of the primary tumor, the parallel progression suggests that metastasis arises from cells that leave the primary tumor at an earlier stage and evolve separately from the primary tumor. Adapted from Gui et al. (2022). Evolution of metastasis: new tools and insights [9].

3.2 Colorectal Cancer: Journey from Benign Polyps to Colorectal Cancer

Each year around 1.9 million new cases of colorectal cancer (CRC) are reported worldwide making CRC the third most common cancer after breast and lung cancer [43, 44]. Metastasis is the main cause of mortality in CRC, accounting for 0.9 million mortalities, usually due to diagnosis at an advanced stage with metastasis presentation [10, 43, 45]. Therefore, making CRC the second leading cause of cancer-related deaths worldwide [44]. Although the incidence of CRC is highest in developed countries, higher mortalities are observed in lower- and middle-income countries in parts of Asia and Latin America due to limited resources for screening and detection [10, 43].

Despite the therapeutic advancements, the 5-year survival rate for patients with metastatic cancer is still below 20% [45, 46]. One major reason is the inherently complex nature and high inter-tumor heterogeneity [47]. Secondly, most cases have already developed metastases present with aggressive tumors at the time of diagnosis [16, 45, 48]. Another important reason is the failure of therapy, as most of the standard therapies for eliminating metastasis are identical to ones used for the treatment of primary tumors [10, 25, 49]. In such cases, the cells that have colonized in secondary sites have already evolved and developed adaptive programs to survive in the presence of therapy thus leading to inefficacy [25, 29, 50]. Finally, some clinics still utilize anatomical and histological details related to tumor size, number of lymph nodes, differentiation, and metastasis to predict prognosis and potential treatment strategies [39].

Most CRC cases are sporadic and not related to inherited mutations [51]. These cases commonly follow the classical adenoma-carcinoma sequence, wherein loss of tumor suppressor gene adenomatous polyposis coli (APC) is the first hit [47, 51] followed by the mutation in rat sarcoma viral oncogene homolog (RAS), tumor protein 53 (TP53), Sma-and mad-related protein 4 (SMAD4) among other genes [52]. 80% of sporadic colon cancer cases display loss of the APC gene [47] as shown in Figure 3. In the development and progression of CRC, APC functions as a gatekeeper and is involved in the regulation of intestinal stem cells through WNT signaling [53, 54].

In the serrated pathway, mutations in the BRAF gene, a prominent oncogene that encodes one of the serine/threonine kinases (STK) of the MAPK kinase signaling cascade, are observed. BRAF mutation is observed in 10-15% of the cases [39, 55]. As it is an early event in serrated CRC pathogenesis, it is associated with an aggressive phenotype and higher rates of peritoneal metastases [51, 56]. To make matters worse, these cases are commonly detected in advanced

stages and exhibit resistance to standard and targeted chemotherapy. BRAFV600E mutations result in constitutive active monomer kinase [57]. Interestingly, BRAF is mutually exclusive with KRAS and NRAS mutation [52, 57]. Other BRAF mutations include dimers, which are either dependent or independent of RAS. It is important to note that even within the BRAF mutation subgroups a high level of heterogeneity is observed due to activation of diverse signaling cascades and molecular profiles. Therefore, rational combinatorial therapy targeting diverse key pathways is required [57, 58].

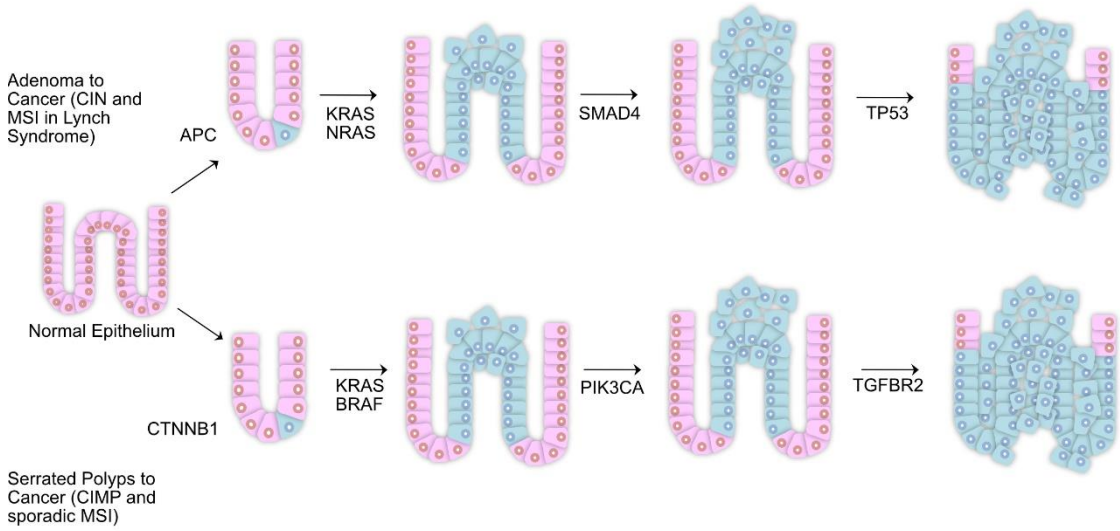

Figure 3: Major pathways of CRC progression

Most cases (50-60%) belong to the 'classic' or traditional pathway initiated by loss of the APC gene followed by mutations in KRAS or NRAS and mutations in tumor suppressor genes such as SMAD4 and TP53. The second pathway known as serrated polyps (bottom) contains activating mutations in the BRAF oncogene followed by mutation in PI3KCA and TGFBR2. Adapted from Walther et al. (2009) Genetic prognostic and predictive markers in CRC [54]. APC: Adenomatous polyposis coli; BRAF: v-RAF murine sarcoma viral oncogene homolog B; CIMP: CpG island methylator phenotype; CIN: Chromosomal instability; CTNNB1: Cantenin-β1; KRAS: Kirsten rat sarcoma viral oncogene homolog; MSI: Microsatellite instability; PI3KCA: Phosphatidylinositol-4,5-biphosphate 3-kinase catalytic subunit; SMAD4: SMA and MAD homolog 4; TGFBR2: Transforming growth factor-β receptor 2.

The third pathway is the alternative pathway and is associated with an activating KRAS mutation. 40% of CRC cases exhibit KRAS mutations in G12, G13 or Q61 residues [59]. Many studies have associated mutation in KRAS with worse prognosis and resistance to targeted receptor tyrosine kinase (RTK) drugs due to the presence of receptor cross talks and feedback loops [39, 59]. Drugs targeting downstream proteins of MAPK pathway in combination with CDK4/6 inhibitors have shown some promise in the clinical setting [59]. However, further research targeting main signaling hubs involved in multiple signaling pathways did not show much efficacy. Some research groups have attempted to design small-molecule inhibitors, but the absence of well-defined binding pockets and low affinity compared to abundantly present cellular GTP has resulted in limited success [59, 60]. Nevertheless, some progress has been made in targeting the KRAS G12 mutant variant [59, 61].

Together, three genomic pathways are attributed to CRC development. First, the chromosomal instability (CIN) pathway often associated with TP53 gene mutation is mainly observed in early stages [47, 57]. Second, the microsatellite instability pathway attributed to defective DNA mismatch repair (MMR) accounts for 15% of colon cancer cases [47, 51]. Microsatellite, a region of repetitive DNA sequences, is most prone to MMR and results in a microsatellite instability (MSI) phenotype. Thus, MSI is used as an indirect read-out to detect MMR deficiency [39, 47, 57]. Finally, the CpG island methylator phenotype (CIMP) pathway. The most prevalent cause of MMR is the inactivation of the MLH1 gene due to CpG island promoter methylation along with methylation of other promoters in the genome caused due to the BRAF mutation gene [47, 51]. Studies also point out the importance of CIMP and MSI in combination with BRAF is required for malignant transformation [47, 57].

To represent cell pathways and microenvironments and address the major inconsistencies, a unified consensus molecular subtypes (CMS) was developed [59, 62]. The first subtype, CMS1 is characterized by a high MSI/CIMP load, a high number of immune cells, and a high prevalence of BRAFV600E mutation [62]. The second subtype CMS2 also known as canonical displays high expression of WNT and MYC target genes and high levels of CIN [47, 62]. CMS3 is also referred to as a metabolic subtype due to the high activation of metabolic pathways having similar epithelial characteristics as CSM2, and a majority of cases exhibit KRAS mutation [47, 62]. The last subtype CMS4 represents an aggressive mesenchymal phenotype, which is associated with poor prognosis due to high activation of EMT and TGFβ pathways [47, 62]. CMS4 is further subdivided based on infiltration of immune cells and tumor microenvironment [47].

3.3 Current Challenges and Biomarkers in CRC

Biomarkers are specific biomolecular indicators produced by the tumor or the host in response to the tumor cells [63]. These molecules can be either DNA or RNA or proteins that are deregulated and result in tumor formation [63]. Additionally, biomolecules produced as a product of aggressive tumor progression can be used [5, 63, 64]. These molecules are shed in body fluids or tissues and are normally not detected in healthy individuals [65]. In addition to facilitating early detection (diagnosis), biomarkers offer other valuable insights such as the risk prediction (prognosis), biological characteristics of the tumor, appropriate therapeutic options, the effectiveness of the therapeutic intervention, and metastasis [5, 63-65].

Although the events leading to CRC development and progression have been known for decades, little progress has been made in the development of prognostic and predictive biomarkers for metastasis in CRC [5, 64]. The main reason for this is the high genomic, transcriptomic, proteomic, and metabolic heterogeneity found within CRC subtypes [39, 59]. Secondly, CRC patients are usually treated with drug combinations or drugs that act on multiple targets thus making it difficult to predict therapy response [5, 64]. Therefore, specific biomarkers that can stratify patients at high risk of metastasis formation and predict therapy response are highly valuable [5, 65].

Recent studies emphasize the investigation of biomarkers corresponding to a molecular pattern that represents tumor microenvironment and signaling pathways rather than single biomarkers denoting mutation to improve sensitivity and specificity [47, 66]. Currently, alterations in genetic and epigenetic factors such as mutation in the KRAS gene and MSI status govern the choice of treatment in CRC patients with metastasis (Figure 4) [5, 64]. Recent guidelines recommend testing the KRAS status before commencing anti-EGFR therapy as it may lead to unnecessary exposure to highly cytotoxic compounds in cases of KRAS mutated tumors [5, 45, 64].

Emerging evidence also points at the significance of HER2 (ERRB2) status in determining anti-EGFR efficacy [5]. It is also important to consider that constitutive activation of MAPK cascade could also mutate downstream effectors of EGFR leading to limited anti-EGFR efficiency [5, 64]. Interestingly, BRAF-activating mutations are mutually exclusive to RAS mutations, however, they are associated with a high number of MSI [5]. Studies show that the combination of anti-EGFR with MEK inhibitors and BRAF inhibitors is superior to standard anti-EGFR therapy [5]. Additionally, extended testing to evaluate ALK, ROS, and NTRK fusion are also beginning to emerge [5, 64]. Currently, carcinoembryonic antigen (CEA) is used in clinical settings to predict relapses in CRC cases. However, it suffers from low specificity and sensitivity [64, 67].

Further, an important element responsible for tumor heterogeneity and relapses in CRC is the tumor microenvironment (TME) and tumor stroma [39, 66]. Therefore, studies analyzing components of TME such as cancer-associated fibroblast markers as predictive biomarkers are gaining interest [5, 64, 66]. Another upcoming area in biomarker development is using non-invasive methods such as liquid biopsy to determine the circulating tumor cells, exosomes, and circulating tumor DNA (ctDNA) [5, 64, 68]. The analysis of ctDNA not only helps determine relapses and resistance to therapy but also to assess tumor mutation burden. However, detection can be challenging due to the limited number of ctDNA in the earlier stages [5, 64].

In the upcoming chapters, a novel metastasis biomarker discovered by our group and various protein prediction tools and methods to examine protein-protein interaction (PPI) are discussed.

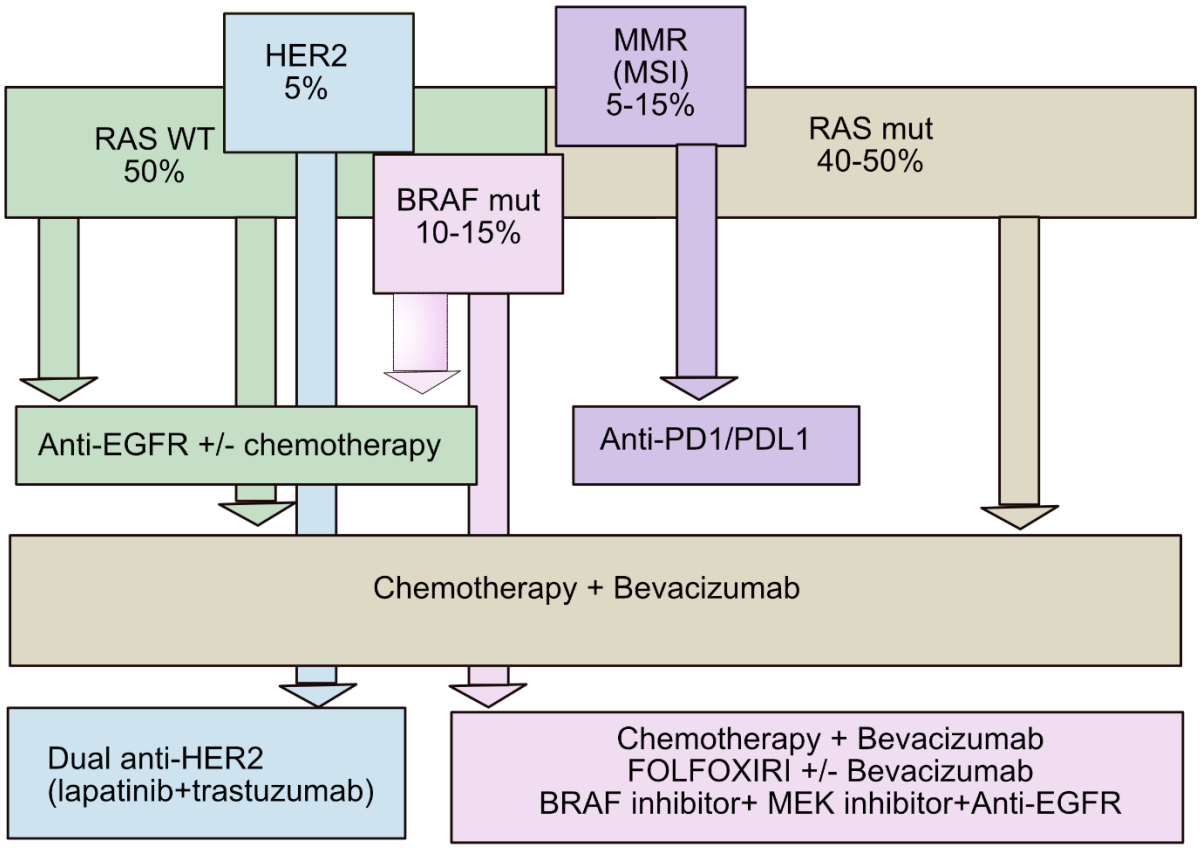

Figure 4: Summary of CRC treatment strategies based on the mutation profile

The first-line treatment for CRC mainly consists of FOLFIRI and FOLFOX combined with anti-EGFR or Bevacizumab based on the mutation profile of the tumor. RAS WT: Wild-type rat sarcoma viral oncogene homolog; Ras mut: Mutated rat sarcoma viral oncogene homolog; HER2: Human epidermal growth factor receptor 2; MSI: Microsatellite instability; BRAF mut: Mutated v-Raf murine sarcoma viral oncogene homolog B; EGFR: Epidermal growth factor receptor; PD1: Programmed cell death protein 1; PDL1: Programmed death-ligand 1; MEK: MAPK/ERK kinases; FOLFIRI: Folinic acid, Fluorouracil, irinotecan; FOLFOXIRI: Folinic acid, Fluorouracil, oxaliplatin, irinotecan. Adapted from Punt et al. (2017) From tumor heterogeneity to advances in precision treatment of CRC [39].

3.4 MACC1: A Promising Biomarker

Despite recent advances in sequencing technologies, biomarkers that can detect metastasis at the earlier stage are not yet well-established [5, 69]. As with many other cancers, the higher rate of mortality in CRC is due to metastasis rather than primary tumor [45, 70, 71]. Hence, biomarkers that can stratify patients at higher risk of developing metastasis are urgently required [5, 48, 64, 69, 70, 72]. In 2009, our group identified a novel gene known as Metastasis-associated in colon cancer 1 (MACC1) that was differentially expressed in colon cancer tissue, metastases, and normal colon tissue [48]. More than a decade of extensive research has shown that MACC1 is a prognostic biomarker in more than 20 solid tumor entities [48, 72, 73].

The 5-year survival for individuals with low MACC1-expressing tumors is 5.3 times higher than those with high MACC1 expression [48, 72, 74]. Strikingly, MACC1 expression in the primary tumor can independently stratify individuals at the risk of developing metastases even before initiation of metastases as observed in Figure 5 [48, 72]. Recent studies evaluating the expression of MACC1 in various CRC clusters confirm the association of MACC1 expression with metachronous metastases development [75]. A study by Ilm et al. (2015) established that high MACC1 expression along with KRAS G13 mutation compared to KRAS WT is linked with poor overall survival in CRC patients [74]. In addition to being a diagnostic and prognostic biomarker, MACC1 can predict the treatment outcome for standard chemotherapeutic agents such as 5-FU and platinum-based drugs [76-78].

Interestingly, the role of MACC1 is not limited to diagnosis but extends to enhancing biological traits required for metastasis such as migration, proliferation, invasion and suppressing apoptosis [48, 72, 73, 79]. Previous literature has pointed out that the activation of MACC1 and c-Met is interlinked and MACC1 transcriptionally regulates the expression of c-Met [48, 80]. Binding with HGF stimulates c-Met and triggers activation and recruitment of downstream proteins GRB2, SHC1, GAB1, SHP2, and STATs forming a signaling complex around GAB1 [81-83]. This signaling complex in turn activates the MAPK cascade and pivotal downstream effector – ERK [84, 85]. Interestingly, active ERK regulates MACC1 transcription through transcription factors AP1 and SP1 thus forming a positive feedback loop MACC1/c-Met/ERK to trigger sustained activation [73, 86, 87].

Notably, downregulation or knockout of MACC1 results in reduced migration, invasion, and higher apoptosis and therapy response compared to control as observed *in vitro* and *in vivo* models [48, 77, 78, 88-91]. An extensive high-throughput screening comprising more than 30,000 compounds

revealed that small molecules (statins and rottlerin) transcriptionally) transcriptional inhibit MACC1 expression and reduce cell invasion and cell migration [92]. Specifically, statin treatment in mouse models showed downregulation of MACC1 expression, and reduced tumor growth and metastasis formation [88, 92]. In line with these findings, emerging evidence from large trans-Atlantic cohorts proves that long-term statin intake is directly associated with the lower incidence of cancer and better cancer survival [93]. This could be correlated to the downregulation of MACC1 expression, thus implying an important role of MACC1 as a predictive biomarker for statin treatment [82, 93].

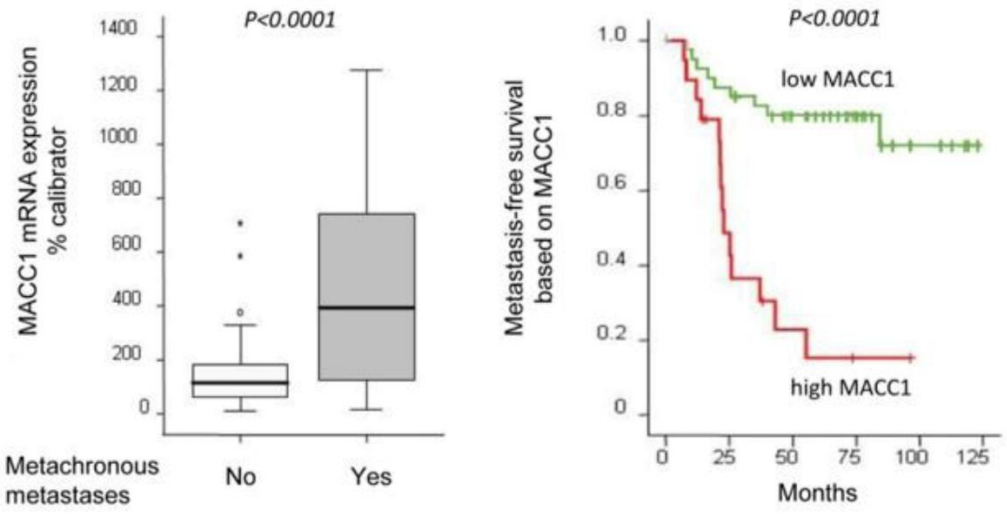

Figure 5: High MACC1 expression is linked with lower metastasis-free survival

MACC1 expression in primary not-yet-metastasized tumors determines the risk of metachronous metastases and metastasis-free survival in colon cancer. Source: Stein et al. (2019) MACC1, a newly identified key regulator of HGF-MET signaling, predicts colon cancer metastasis [48].

Moreover, co-expression of MACC1 and c-Met is observed in normal embryonic and tumor tissue [48, 72, 73, 94]. Remarkably, experiments with the ligand of c-Met, HGF have proved the shuttling of MACC1 from cytoplasm to nucleus on c-Met activation [48, 95]. MACC1 expression is tightly regulated in normal tissue and found mainly in tumor tissue indicating its relevance in cancer development and progression [48, 73]. Mechanistically through c-Met, MACC1 also activates AKT signaling and subsequently Wnt/β-catenin signaling via GSK3β inhibition thus inducing genes required for EMT and stemness such as TWIST1/2 SNAIL1/2 and MMPs [48, 73, 88, 91, 96, 97]. The silencing of MACC1 expression has been shown to attenuate AKT signaling and cause

G0/G1 arrest resulting in growth suppression [73, 98]. Additionally, silencing MACC1 also exaggerates apoptosis in cancer cells by reducing the MCL1 expression via Stat1/3 signaling [90]. Likewise, reports focusing on circadian rhythm have demonstrated that MACC1 in combination with core-clock genes regulates metastatic phenotype, and knock-out of MACC1 has an impact on expression of core-clock genes ARNTL and NR1D1 [99].

Increased expression of MACC1 is strongly associated with the transition from a benign to an aggressive malignant state [48, 79]. Importantly, immunohistochemical analyses reveal that the presence of MACC1 at the invasive front, and in tumor buds is related to higher metastatic risk and poor survival [100]. Moreover, enhanced expression of MACC1 in an APC-mutated mouse model significantly promotes tumor progression and expression of stem cell genes such as Nanog and Oct4 [101]. Consistently, crocin, which is the key compound present in saffron reduces the metastatic ability of MACC1-overexpressing cells *in vitro* via attenuation of cancer stem cell marker DCLK1 [102]. In a recent study, investigating the secretome of MACC1, Kortüm et al. (2022) showed that S100A4, a prominent metastasis biomarker, is upregulated in the MACC1 secretome and essential for MACC1-induced cell migration [88]. In agreement with this hypothesis, combination treatment with MACC1 transcriptional inhibitors (statins) and S100A4 (niclosamide) abrogated cell migration *in vitro* and metastasis *in vivo* [88, 92, 93].

Unlike other metastatic proteins, the structural composition of MACC1 is highly distinct and diverse [72, 73, 94]. Present abundantly in the cytoplasm of cancer cells, 852 amino acids forming MACC1 protein present the following features: N-terminal clathrin box followed by two motifs (NPF and DPF) [48, 73, 79, 94, 103, 104]. Next, a prominent feature of MACC1 is the ZU5 domain, usually found in proteins involved in cytoskeletal rearrangement such as Ankyrin B [73, 94, 105]. MACC1 also contains proline-rich motifs capable of binding to proteins with SH3 domain [79]. Surprisingly, a SH3 domain which is the complement of the proline-rich motif is found next to the proline-rich motif in the MACC1 protein [79]. Mounting evidence suggests that the SH3 domain is essential for MACC1 metastatic capabilities [48, 79, 104]. The composition of the domain at the C-terminus of MACC1 is also highly peculiar consisting of double death domains arranged in a tandem manner close to the C-terminal tail [73, 94, 103]. The closest homolog of MACC1 is SH3BP4, a protein that is also highly upregulated in CRC and involved in receptor recycling and regulation of WNT signaling [48, 95, 106, 107]. SH3BP4 has similar domains as MACC1 but a different position of the SH3 domain [104]. The SH3 domain located at the N-terminus of SH3BP4 is known to regulate mTOR signaling via Rag GTPases [108]. In a major study by Imbastari et al.

(2019) investigating the link between MACC1 and clathrin-mediated endocytosis concluded that MACC1 enhances the rate of receptor recycling leading to a stronger activation loop [104].

The sequence alignment data indicates that tyrosine sites Y379, Y598, Y695, Y768, Y789, and Y793 which are capable of phosphorylation are identical in different species [79]. In another study focusing on the tyrosine phosphorylation of MACC1, Kobelt et al. (2021) found that MACC1 is phosphorylated at sites Y673, Y695, and Y793 by MEK1 [103]. Consistently, treatment with MEK1 inhibitors interfered with MACC1's ability to induce cell migration, invasion *in vitro*, and metastasis *in vivo* [103].

Analyses by Koelzer et al. (2015) revealed that MACC1 expression is positively correlated with tumor grade, tumor size, and the number of metastases [100]. Another detailed analysis of MACC1 in the COAD-READ TCGA cohort illustrated that tumors with MACC1 somatic copy number alteration (SCNA) and increased mRNA expression are linked with poor outcomes and lower survival [75]. Surprisingly, the expression of MACC1 depends on the localization of the tumor and varies between tumors from proximal to distal colon (highest) [75].

Overall, these findings highlight the prominent role of MACC1 as a biomarker and metastasis protein. Further, these studies outline the ability of MACC1 to act as a transcription regulator and participate in diverse PPI through its peculiar domain arrangement [48, 73, 79, 94]. All this evidence combined with the biological impact of MACC1 on versatile signaling cascades makes MACC1 an attractive target for future therapeutic strategies to hinder metastasis [72, 73].

3.5 Role of Receptor Tyrosine Kinase in CRC Progression

The majority of kinases found in humans belong to the RTK subfamily and are fundamental to cancer development and progression [109-111]. In cancer development, RTK activity is highly deregulated due to gene amplification, mutations, chromosomal rearrangement, and/or autocrine activation loops [109, 111]. Under normal conditions, RTKs are activated by specific ligand binding followed by trans autophosphorylation and conformational changes that relieve the auto-inhibition state and expose the active binding site [111]. Commonly, ligand binding induces receptor dimerization, however, in some RTK receptor subfamilies dimers may be present either in the active or inactive state even in the absence of activating ligand [109-112].

Once activated, RTKs can bind to various downstream signaling proteins containing Src-homology-2 (SH2) or phosphotyrosine-binding (PTB) domains directly or through docking proteins such as Gab1 [81, 83, 109]. Mounting evidence shows that the conserved residues in the DFG motif present in the ATP pocket are often mutated in cancer and are responsible for sustained cell growth [109]. Another mode of abnormal RTK activation is chromosomal rearrangement which leads to the formation of chimeric fusion oncoproteins [109].

Based on the active state and binding site, tyrosine kinase inhibitors are divided into type I, II, allosteric, substrate-directed, and covalent inhibitors [50]. While the type I and type II inhibitors bind at the ATP binding pocket in the active state and inactive state, respectively, allosteric inhibitors bind to sites different from the catalytic ATP binding site and hence do not compete with ATP for binding [113]. Consistently, allosteric inhibitors offer higher specificity compared to type I and II as they target specific kinases [50, 113]. On similar lines, the substrate-directed inhibitor also binds at a different site than ATP but competes with the substrate for binding [114]. Finally, covalent inhibitors, as the name suggests, form covalent bonds with the non-catalytic cysteine residues in the active site [50, 115].

Among RTKs, c-Met, a prominent member of the MET family, is often deregulated in various cancer entities [81, 116-118]. The reason why c-Met is so integral for tumor development and progression is its ability to co-operate with other receptors forming crosstalk and activating versatile downstream signaling cascades such as RAS/MAPK, SRC, PI3K/Akt among others [116]. A study demonstrates how mutations in c-Met can influence endocytic pathways leading to lower degradation and enhancing receptor recycling to the plasma membrane [117, 119]. RAS/MAPK is a critical pathway regulated by c-Met that consists of a peculiar phospho-relay by

three protein kinases belonging to the mitogen-activated protein kinase (MAPK) family to activate various cell functions (Figure 6) [109, 111].

ERK is an important downstream effector of the MAPK pathway that facilitates the conversion of upstream signals to cell-specific responses [87]. ERK can phosphorylate a wide spectrum of substrates and activate diverse gene expressions, thus resulting in diverse cell responses. Importantly, the timescale of activation plays a fundamental role in deciding a cell's fate [84, 120]. The signal duration can vary from short (in minutes) to prolonged (in hours) depending on the strength and duration of the stimuli, expression of immediate-early and immediate-late genes, and rate of receptor internalization among others [84, 121, 122]. Particularly, a detailed study has revealed that immediate-early gene product, *c-fos* acts as a sensor for interpreting ERK signal strength and duration [121, 122].

Prolonged activation results in translocation and accumulation of ERK in the nucleus to activate gene expression [84, 120]. While the dimeric state of ERK is associated with cytoplasm proteins, the ERK monomers are mainly found in the nucleus [123]. Thus, the oligomeric state of ERK not only regulates the protein interaction but also the cellular localization of ERK [84, 123, 124].

SRC, another downstream protein of RTK signaling, is a prominent member of SRC family kinases (SFKs) which is associated with the regulation of cell-cell junction and cytoskeletal rearrangement required for cell migration [125-127]. Notably, SRC can interact with diverse tyrosine kinases such as c-Met, EGFR, VEGF, and PDGF among others, and contribute to receptor crosstalk [128, 129]. On activation, the change in conformation of SRC triggers dephosphorylation of Y530 and autophosphorylation of Y416 residues [128, 130]. This in turn activates downstream signaling cascades involved in cell proliferation and migration [126, 128, 131, 132].

Although the detailed mechanism of SFK in CRC is not completely understood yet, numerous studies indicate that SRC is highly deregulated in advanced stages and associated with poor outcomes [118, 129, 133]. Particularly, the upregulation of SRC expression has been linked to aggressive metastases and drug resistance in CRC [134, 135]. The increased SRC expression in turn also enhances the expression of tyrosine kinase receptors such as c-Met and downstream proteins involved in cell invasion [118, 132]. Consistently, some studies report that the expression of CSK, a negative regulator of SRC, is downregulated in CRC and inversely correlated with the outcome [132, 136]. In addition to the expression of CSK, researchers highlight the membrane localization of CSK to be integral for regulating SRC activity [125, 136].

A major problem noted with ATP-competitive kinase inhibitors used as a single agent is the development of drug resistance [49, 128]. Recent investigation indicated that the kinase inhibitors could unexpectedly activate SRC resulting in phosphorylation of FAK and in turn activation of the FAK/GRB2/ERK cascade [128]. In addition to regulating cell motility, a study by Lopez et al. (2012) also describes the important role of SRC in preventing apoptosis via enhanced BIK degradation [137].

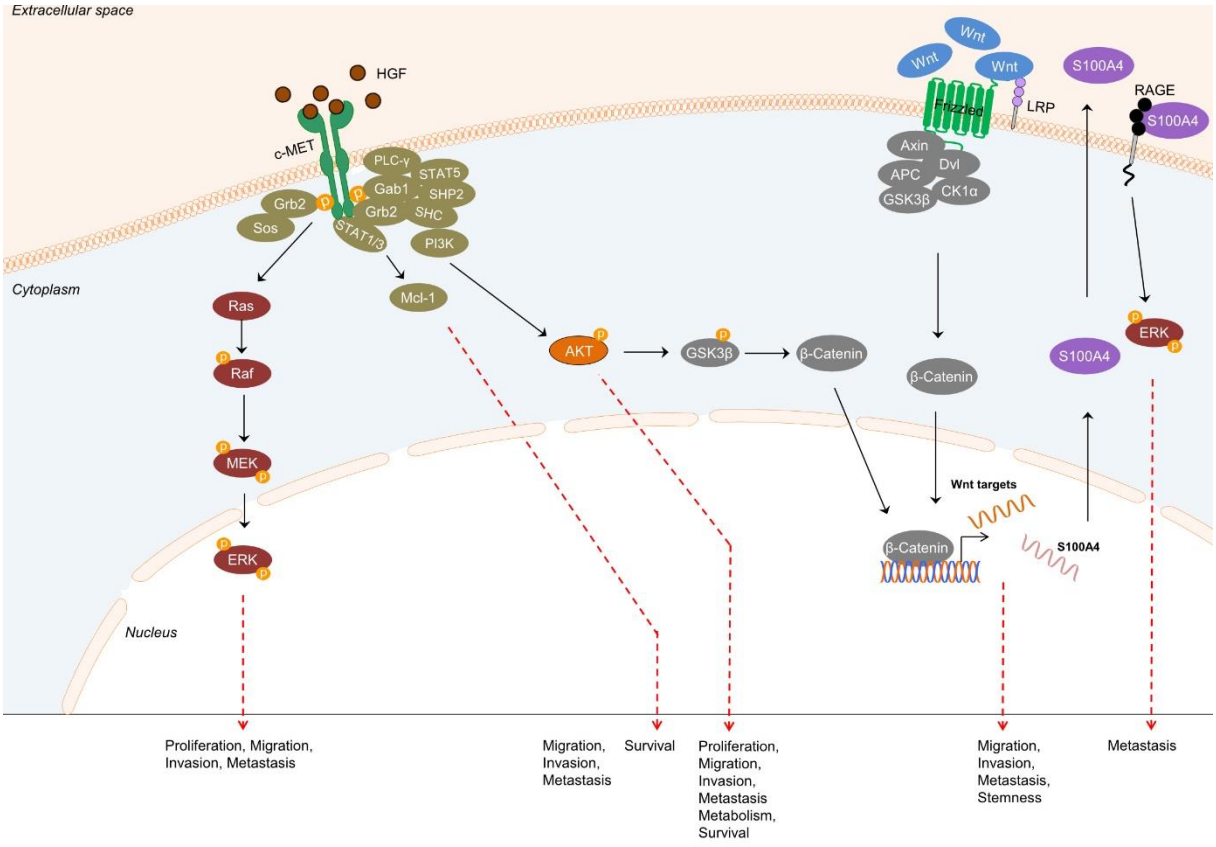

Figure 6: c-Met-HGF signaling landscape

Schematic representation of critical pathways linked with c-Met-HGF activation. Some of these pathways share common functional characteristics such as migration, invasion, proliferation, and survival. Created by Harikrishnan Radhakrishnan (AG Stein).

Given the important role of kinases in regulating practically every phase of cellular development, identification of phosphorylation sites and kinases involved in activation is essential to better understand disease progression, particularly cancer, and identify targets [50, 109, 110]. With the advancement of machine learning, numerous computational tools that can predict phosphorylation sites and kinases by recognizing patterns have greatly accelerated the

identification of novel phosphorylation sites [138-141]. Most of these techniques require only the amino acid sequence of the protein of interest as input and minimal computational resources and expertise [138-140]. As a result, they predict the likelihood of phosphorylation of each tyrosine/ serine/threonine residue in the sequence using numerical scores [138, 140, 141]. This rapid screen helps filter the promising sites that can be further investigated through site-directed mutation studies [139]. Netphos, one of the first tools developed to predict phosphorylation sites, utilizes complex artificial neural networks to recognize patterns formed by 9-33 residues neighboring the phosphorylation site [139, 141]. Group-based prediction system (GPS) is another highly accurate tool used to predict kinase-specific serine/threonine/tyrosine sites based on the sequence information [138]. It is an easy-to-use open-access tool that can predict up to 479 human protein kinases [138].

3.6 Innovative Tools for Protein Structure Determination

In addition to governing the physical characteristics of the protein, the structure of the protein is key to its cellular function [142]. Innumerable insights about the active conformation and degradation mechanism can be gained through the protein structure [143, 144]. Furthermore, understanding the characteristics and orientation of certain domains can facilitate the identification of novel drug targets and allosteric sites that might have remained unstudied [143, 145].

Modern techniques such as Cryogenic electron microscopy (Cyro-EM) can penetrate deep down and solve structures close to atomic resolution [146]. Of particular interest is its ability to detect various dynamic conformations in near-native environments and therefore reconstruct the transition between the states [147]. Cyro-EM has successfully solved protein structure complexes that were previously considered irresolvable. Nevertheless, the cost of setup, time, and expertise for sample preparation, image acquisition of smaller molecules, and analysis are still some of the main challenges [147, 148]. Consistently, other experimental techniques such as X-ray crystallography and NMR also face similar challenges [149]. This is clearly observed by the number of known structures compared to the number of known protein sequences [143].

In the meantime, great progress has been made in machine-learning-based structure prediction algorithms thus advancing the rate of protein structure prediction drastically [150]. Importantly, most of these prediction tools only require primary protein sequence as input to predict 3D-protein structure. Some tools develop structures based on evolutionary information, docking to predict

the 3D structure while others utilize complex neural networks and energy calculations [142, 144, 150]. The application of these methods paved a new path for the field of drug discovery and target identification [145, 151, 152]. For the first time, it was possible to conceive structures of prominent proteins that were considered difficult to crystallize [152, 153]. Altogether, computational protein prediction tools offer an unprecedented advantage and require limited resources [142, 143, 154]. However, given the overall complexity and high number of conformational states possible, it is challenging to determine accurate structures, especially in the case of large proteins with novel folds [143, 155, 156].

Currently, much focus has been switched towards the recent AlphaFold2 prediction tool, and many highly accurate structure predictions have been reported [152, 153, 157]. Numerous researchers have employed AlphaFold2 to predict the structure of proteins in viral research, the impact of oncogenic mutations, and drug development among others [145, 152, 157]. For the first time, predicted structures of more than 200 million proteins were freely available to researchers worldwide [151, 158]. As a result, providing valuable insights into complex biological protein structures such as nuclear pore complex, and facilitate future drug development [152, 153, 158].

AlphaFold2 algorithm can produce highly detailed 3D structure accurate to the side-chain level by employing an iterative process using templates of similar proteins [144]. Importantly, AlphaFold2 utilizes information from intermediate steps to refine the final model in an iterative manner thus making it more powerful and precise [144]. The result is a fine-tuned product obtained through repeated feedback between the different layers of the neural network [143, 144]. The query input sequence is screened against sequences from various databases to identify homologous sequences that act as building blocks for further refinement and structure prediction [143, 144]. Interestingly, AlphaFold can also compute novel structures that share no similarities with existing templates using coevolution patterns and recognizing local interaction patterns [143, 144, 151].

Accuracy is measured using robust readouts such as predicted template modeling score (pTM) and predicted local distance difference test (pLDDT) [144]. The pLDDT compares the local distance difference between atoms of the predicted structures to the reference structure or group of structures [144, 154]. This score ranges from 0-100, with 100 being the most certain [144]. Notably, Alphafold2 is trained on protein fragments to minimize the difference between the predicted and actual 3D structure [143, 144]. Nevertheless, AlphaFold's algorithms also have certain limitations especially when it comes to disordered regions [156]. It is important to consider that proteins are highly dynamic structures that frequently interact with other biomolecules under

different conditions to function [156]. However, AlphaFold cannot accurately accommodate these properties in calculation yet [143, 155]. Thus, caution in interpreting these models and combining them with further experimental techniques such as NMR or Cyro-EM is recommended [153, 159].

AlphaFold-Colab [154] is an easy-to-access Jupyter Notebook developed by DeepMind and Google to enable protein prediction even with unsophisticated hardware in minutes [154]. The reason AlphaFold-Colab functions so well and fast is the modified environmental database employed for building multiple sequence alignment (MSA) [144, 154]. It utilizes an alternative homology search method known as MMseq2 to quickly identify diverse templates [154]. Secondly, it is capable of early termination and thus conserving computation power [154].

Rarely do proteins function as individual units and several signaling proteins function as homooligomers or heteroligomers [160, 161], therefore various computation tools are developed to predict higher-order protein assembly using template-based or docking methods or a combination of both. GalaxyHomomer [162], is one such open-access web server that predicts homodimer based on protein sequence or monomer structure and utilizes structure symmetry to refine and remodel less accurate regions [162].

Recently, AlphaFold-Multimer, an extension of AlphaFold addresses this problem and facilitates reliable multiprotein complex prediction [151, 163]. AlphaFold-Multimer works in a similar way to AlphaFold but is trained on datasets containing multimers, and parameters such as model architecture are adapted for multimers [144, 151, 163]. Data suggests that AlphaFold-Multimer predicts homomeric interfaces with higher accuracy than heteromers due to the readily identification of co-evolutionary data from the MSA and structure symmetry [143, 151, 163]. Like AlphaFold, AlphaFold-Multimer also provides accuracy scores but for the interface and is termed as interface pTM or ipTM [144, 157, 163]. Considering the fast development of AI tools and the availability of vast protein structure datasets, the development of highly accurate protein structure and protein complex predictions representing various dynamic states is closer than ever imagined [142, 155].

3.7 Relevance of Dimerization

Self-association of proteins into smaller complexes forming dimers is an abundantly observed phenomenon in nature [161, 164, 165]. Over the past decade, numerous studies have shown that self-association offers several advantages such as increased stability, binding sites, and binding affinity, among others [160, 166, 167]. In addition, oligomerization between identical or nearly identical components enables proteins to form larger symmetrical structures without increasing the folding complexity and the genome size [165, 166].

Dimerization represents synergy and serves as a pivotal regulation mechanism enabling prompt binding and activation of downstream signaling under appropriate conditions [160, 166]. For instance, the prominent chemokine CXCL12 also exists in a monomer-dimer equilibrium. While the monomeric CXCL12 is known to interact with CXCR4 and regulate cell migration, the interactions were not observed in dimeric CXCL12. Thus, indicating a regulation mechanism mediated by the formation of higher-order structures [168].

Although it is speculated that most proteins exist in higher-order assemblies only a fraction of protein oligomers is identified [160, 161]. Nevertheless, with the advent of newer technologies focusing on resonance energy transfer, protein dimerization, and oligomerization can be readily detected in living cells in real-time [169]. In the next section, these techniques and the pros and cons associated with them are discussed briefly.

3.8 Techniques for Mapping Protein Interaction

Most proteins execute their function by participating in protein complexes rather than as a single unit [151, 165, 170]. Thus, PPIs are fundamental to all cellular processes in normal and disease states [171, 172]. The phenotype observed during disease progression is a product of different protein interactions within the cell [171, 173, 174]. The structural characteristics of proteins not only determine protein assembly, stability, and folding but also govern the choice of interaction partners and multiprotein complexes in which the protein participates [165, 174]. Protein interactions not only shape the signaling landscape of a cell but also dictate the localization of the proteins within the cell. Therefore, serving as attractive drug targets [151, 170, 172, 173, 175].

Protein interactions can be easily differentiated based on their binding interface size, binding strength, or duration of the interaction [175]. When considering binding strengths and duration of interaction, proteins can be subdivided into permanent or transient interactions [172, 175].

Permanent interactions are interactions in which the individual protein cannot exist without the protein complex [175-177]. On the other hand, transient interactions are observed often, as they function as switches to activate/deactivate diverse processes such as phosphorylation, enzymatic degradation, and endocytosis among others [170, 172, 173]. An example of transient interaction is the interaction of integrin with extracellular matrix (ECM) proteins to regulate cell motility [178]. In short, transient interactions play a key role in regulating various cell processes. Further, the transient interaction may be subdivided into weak or strong interactions [172]. As the name suggests, weak interaction can be easily formed or disrupted whereas strong interaction is regulated by specific signal or structural features [172]. While the dissociation constants of strong transient interactions range from nano to femtomolar, the dissociation constants for weak interaction lie in the micro-to-millimolar range [172, 175]. Therefore, weak interaction requires highly sensitive detection methods [172]. Protein chains can interact with identical protein chains forming homo-oligomers or with non-identical chains forming hetero-oligomers such as HER2, and TP53 among others [170, 172, 179, 180].

Due to ease of setup and simplicity, co-immunoprecipitation continues to be the most widely used technique for studying PPI [170]. Theoretically, the technique functions as fishing, where the protein of interest is precipitated with the antibody and functions as “bait” and the interacting proteins from the whole cell lysate are co-precipitated form the “prey” [181, 182]. In addition to its simplicity, the technique offers other advantages such as determining complexes in their native form and incorporating the post-translational modification observed in eukaryotic cells [170]. While useful in routine protein interaction studies, co-immunoprecipitation suffers from disadvantages such as the inability to differentiate between direct and indirect binders and low sensitivity [182].

Another classical method for detecting protein interaction is the yeast two-hybrid assay. It was developed more than 30 years ago and continues to be one of the common high-throughput techniques to identify binary interactions [183, 184]. This technique is based on the reconstitution of two domains of a transcription factor when two proteins interact, and this ultimately activates the transcription of a reporter gene. These reporter genes can be essential enzymes such as β-galactosidase or auxotrophic markers [175, 185]. Although informative, this technique has prominent limitations such as a high false positive rate, interactions that need to occur in the nucleus, and challenges to draw biologically relevant conclusions as the interaction are not detected in their native environment [170, 175, 185, 186].

Given the limitations associated with classical techniques, modern approaches such as bimolecular fluorescence complementation (BiFC), fluorescence resonance energy transfer

(FRET), and bioluminescence resonance energy transfer (BRET) are gaining popularity [187-190].

FRET is a technique in which non-radiative energy transfer occurs between two appropriately chosen fluorophores when they are less than 10 nm (100 Å) apart [188, 191]. The emission spectrum of one fluorophore (donor) should overlap with the excitation spectrum of another fluorophore (acceptor) to enable dipole-dipole interaction [175, 188]. FRET offers numerous applications including the detection of transient and stable interaction, multiprotein complex detection, and importantly detection of self-association of proteins in living cells [175, 188]. BRET is another protein interaction technique inspired by the natural phenomenon observed in sea creatures [192, 193]. It is based on the transfer of energy between bioluminescent donor and fluorescent acceptor when less than 10 nm apart [169, 175, 194]. When compared to FRET, BRET offers several benefits such as i) Reduced autofluorescence ii) No need for external light sources for donor excitation thus avoiding photobleaching and cross acceptor activation iii) High sensitivity making it possible to detect interaction close to endogenous level iv) Increase signal to noise ratio. Nevertheless, BRET also encompasses limitations such as low light intensity of luciferase on excitation compared to excitation of fluorescence protein in FRET [175, 192]. Thus, it requires a highly sensitive microscope and detectors to study PPI using microscopy at a single-cell level. However, it is not a major concern when studying PPI in cell populations using a conventional plate reader [175, 195].

With the advent of newer generations of stable and intense luciferase donors, for instance, NanoLuc (NLuc), detection even at lower levels is possible now [187]. Additionally, the high intensity of NLuc is highly advantageous when studying dynamic interactions at a single-cell level using live imaging [187, 196]. Precisely, the narrow spectrum enables better signal separation and a lower tendency of donor emitted light to be detected at the acceptor emission wavelength (bleed-through) when using appropriate acceptor pairs [175, 196]. A few important requirements for BRET include i) The distance between the speculated interacting protein should be less than 10 nm ii) The excited wavelength needed to excite the acceptor should match the emission wavelength of the donor iii) The position of the tag should facilitate maximum resonance transfer and minimum interference iv) Appropriate negative control, control for donor bleed-through, and positive control [175, 192].

4 Project Goal

Despite recent advancements, metastatic cancer still represents a terminal illness for large numbers of cancers [45, 197]. Low survival rates, development of resistance, and poor quality of life are some of the factors indicating the current gap in effective biomarkers and therapies [5, 29, 64, 72]. MACC1, a metastasis biomarker established in several cancer entities, represents an attractive target [48, 73, 198]. Most importantly, previous *in vitro* and *in vivo* studies have proved reduced metastasis and tumor development when MACC1 expression is reduced (RNAi) or completely abolished (MACC1-knockout) in various cancer types [78, 88, 91, 101]. Besides the distinct domains that form the MACC1 protein and important tyrosine residues close to the C-terminus [48, 73, 82, 94, 103, 104], very little is known about the MACC1 structure and its self-association capabilities. This interesting background and the need to understand the structural characteristics of MACC1 to develop rational therapies form the framework for my project. The two main goals of my project are as follows:

a) **Investigating the role of tyrosine sites located close to the N-terminus of MACC1:** This part of the project aimed to study the phosphorylation potential of tyrosine site close to the N-terminus of MACC1 using computational tools and determination of the impact of mutation of this tyrosine residue on MACC1 signaling.

b) **Systematically characterizing the MACC1 structure and its homodimer:** The second part of the study focused on understanding the structural properties of MACC1 and its ability to self-associate. This part comprised detailed analysis and validation of the predicted MACC1 homodimer, and effects of dimer interface mutation on MACC1 signaling and function.

5 Materials and Methods

5.1 Cell Culture

All CRC cell lines (HCT116, SW480, SW620) used in this study were purchased from the American Type Culture Collection (ATCC, Manassas, USA). Cells were cultivated in RPMI-1640 or DMEM (Gibco, Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS) (Bio & Sell, Feucht, Germany) (Table 1).

Human embryonic kidney cells (HEK293) cells were obtained as a gift from the group of AG Wanker, Neuroproteomics, MDC. The HEK293 cells were maintained in DMEM without pyruvate (Gibco, Thermo Fisher Scientific) supplemented with 10% FBS and 1x penicillin/streptomycin (Corning). For the bioluminescence resonance energy transfer (BRET) experiment, DMEM or RPMI (Gibco, Thermo Fisher Scientific) without phenol-red was used. All cell lines were passaged regularly (every 3-4 days) using 0.05% Trypsin-EDTA (Gibco, Thermo Fisher Scientific). Cells were maintained at 37°C in a humidified, 5% CO_2 ventilated incubator.

Table 1: Summary of the used cell lines

Cell line	Medium	Characteristics
HCT116 (CCL-247)	RPMI-1640, 10% FBS	Carcinoma
SW480 (CCL-228)	RPMI-1640, 10% FBS	Adenocarcinoma, Dukes' type B
SW620 (CCL-227)	DMEM, 10% FBS	Adenocarcinoma, Dukes' type C
HEK293	DMEM w/o pyruvate, 10% FBS	Epithelial morphology isolated from human embryo kidney

5.2 Virus Production and Generation of Stable Modified Cell Lines

To create stable cell lines overexpressing MACC1, a lentivirus system was applied. The lentiviral plasmid expressing open reading frame (ORF) of MACC1 tagged with GFP RC224774L2 (Origene, USA) was used. For virus production, HEK293T cells were seeded at a density of

7.5 x 10^6 in a T-75 flask. After 24 h, a transfection mix containing 10 µg transfer vector, 4 µg envelope plasmid, 6 µg packaging plasmid, and 60 µL TransIT-2020 transfection reagent (Mirusbio, USA) was prepared in the serum-free medium and incubated for 15 min. The cell medium was changed, and the transfection mix was added to the flask. After 48 h, the supernatant from HEK293T cells collected and filtered through a 0.45 µm filter was either used directly or stored at -80°C. For transduction, 3 x 10^5 cells were seeded per well in a 6-well plate and allowed to attach for 24 h. After 24 h incubation, the cell medium was replaced with virus-containing medium. GFP-expressing cells were sorted 72 h later using FACS.

5.3 Plasmid Construction

The mutant MACC1 plasmid containing triple mutation V212D, I214D, and C216D MACC1 (3xMut MACC1) was generated from MACC1-V5 by site-directed mutagenesis (SDM) using QuikChangeTM II XL site-directed mutagenesis kit (Agilent, USA) that includes PfuUltra High Fidelity DNA Polymerase. The primers for SDM were designed using the Agilent Quick Primer Design Platform. The primers were synthesized and HPLC purified by Biotez are listed in Table 2. The SDM was performed as per manufacturer's instruction with minor adaptations as follows: 200 ng MACC1-V5 plasmid was used as a template and 4 µL Quick Solution was utilized. XL10-Gold cells were transformed using 3 µL of the reaction mixture as per the manufacturer's guidelines. From the transformation mixture, 200 µL were plated on agar plates containing suitable antibiotics. Single clones were selected and expanded. Plasmid DNA was isolated using the plasmid miniprep kit PureYield™ (Promega, USA) according to manufacturers' instructions. The plasmid DNA from the clones was evaluated by restriction digestion with appropriate FastDigest™ restriction enzymes (Thermo Scientific, USA). The band pattern was observed using 1% Agarose gel stained with RedSafe™. The samples showing the expected band pattern were sent for DNA sequencing performed by LGC Genomics. The primer and cycling parameters used for mutation are listed in Table 2 and Table 3, respectively. Sequencing results were analyzed and selected clones were further expanded.

Table 2: Primers used for site-directed mutagenesis

Site of Mutation	Sequence 5' – 3'
MACC1 212V,214I,216C-D fwd	5'-ccctccttgatggtttactttgtcagcatcggtgtcctccgcaagttgtgtctgggc-3'
MACC1 212V,214I,216C-D rev	5'-gcccagacacaacttgcggaggacaccgatgctgacaaagtaaaccatcaaggaggg-3'
MACC1 I214P fwd	5'-atggtttactttgcaagctggggtgacctccgcaagttgt-3'
MACC1 I214P rev	5'-acaacttgcggaggtcaccccagcttgcaaagtaaaccat-3'

Table 3: Cycling Parameters

Step	Cycles	Temperature (°C)	Time
1	1	95°C	1 min
2	18	95°C	50 sec
		60°C	50 sec
		68°C	1 min/kb of plasmid length
3	1	68°C	7 min

To generate MACC1 fused to BRET reporters, the coding sequence of MACC1 was amplified using the primers containing the Gateway cassette elements namely attB1 and attB2 (Table 4). For N-terminal and C-terminal MACC1 fusions, cDNAs were amplified using primers listed in Table 4. The resulting PCR fragment containing the attB1 and attB2 sites was cleaned up using the Wizard® SV Gel and PCR Clean-Up System (Promega). The size of PCR products was evaluated by running the sample on 1% Agarose gel. In the next step, 60 fmol of the PCR product was mixed with 75 ng Gateway entry vector (received as a gift from AG Wanker), 1x TE-Buffer, and

BP clonase™ (Thermo Fischer Scientific). This mixture was incubated overnight at 25°C and then used for transformation. On the next day, the colonies were selected and expanded. The plasmid DNA was extracted as mentioned earlier and evaluated by restriction digestion. The entry clone containing MACC1 or MACC1 mutant (75 ng) was then shuttled into a destination vector (150 ng) containing mCitrine (mCit) or NanoLuc (NLuc) (kind gift from AG Wanker) using LR clonaseTM (Thermo Fischer Scientific). After colonies were selected and expanded, plasmid DNA containing MACC1 fused to mCit or NLuc was evaluated by restriction digestion and sequencing (LGC Genomics).

Table 4: Primer used for generation of MACC1 BRET vectors

MACC1	**Primer Sequence 5'- 3'**
n-term fwd	5'-GGGGacaagtttgtacaaaaaagcaggctTCatgctaatcactgaaagaaaacattttcg-3'
c-term fwd	5'GGGGacaagtttgtacaaaaaagcaggctTCGaaGgaGatAgaAccAtgGatgctaatcactg aaagaaaaca-ttttcg-3'
n-term rev	5'-GGGGaccactttgtacaagaaagctgggtCtacttcctcagaagtggagaatgc-3'
c-term rev	5'-GGGGaccactttgtacaagaaagctgggtCCTAtacttcctcagaagtggagaatgc-3'

5.4 BRET

To identify direct protein interactions, an innovative technique known as BRET was employed in this study, the protocol was adapted from a publication by Trepte et al. (2018) with minor adaptations as mentioned below [195]. HEK293 or HCT116 cells at a density of 4.5 x 10^4 per well were reverse transfected in a white 96-well plate (Greiner, 655983) using TransIT-2020 transfection reagent (Mirusbio, USA). The total amount of DNA in each well of 96-well was 200 ng. Plasmids encoding MACC1/SRC tagged with NLuc (donor) or MACC1 tagged with mCit (acceptor) were co-transfected at a ratio of 1:10 or 1:20 (10-15 ng of donor and 150-185 ng of acceptor) and empty pcDNA3.1 expression plasmid was added to make up the DNA amount to 200 ng in each well. Each interaction was tested in 3 technical replicates. After 72 h, the fluorescence in intact cells was measured (Ex/Em: 500/530 nm) using an M1000 pro plate reader

(Tecan, Switzerland). In the next step, a luciferase substrate coelenterazine-h to a final concentration of 5 µM (pjk, Germany) was added to each well and incubated for 15 min. The short-wavelength luminescence (370-480 nM), long-wavelength luminescence (520-570 nM) and total luminescence were measured after incubation. Additionally, two control interactions were included for each tested interaction pair. Control 1 consists of co-transfection of the empty NLuc donor vector with PA-mCit-MACC1 and control 2 consists of a combination of empty PA-mCit acceptor vector with NLuc-MACC1. For interaction between MACC1 and SRC, control 2 was made up of the empty PA-mCit acceptor vector and NLuc-SRC. The BRET ratio values of Control 1 and Control 2 were compared, and the higher control BRET ratio value was selected for correction. The threshold value of 0.01 for the corrected BRET ratio as determined by Trepte et al. [195] was adopted. The corrected BRET ratio was calculated as summarized by Trepte et al. [195].

5.5 Protein Extraction and Western Blot (WB)

Western blot was performed as a standard technique to analyze protein expression. For protein extraction, cells were seeded in 6-well or 12-well plates and allowed to attach for 24h. To study the effect of growth factors, cells were starved overnight before treatment. On day 3 if applicable, cells were treated with respective growth factors for respective time intervals. Immediately after treatment, cells were washed with cold PBS and then scraped off in ice-cold RIPA buffer (50 mM Tris, 150 mM NaCl, Igepal 1%, Sodium deoxycholate 0.5%, pH 7.5) supplemented with complete protease inhibitor tablets (Complete, Roche) and phosphatase inhibitor tablets (PhosStop, Roche). Cells were allowed to lyse for 30 min on ice with intermediate vortexing, and centrifuged for 30 min at 14,800 rpm at 4°C. The supernatant was transferred to a new tube, stored at -20°C or directly used for protein quantification.

Protein concentration was quantified using Bicinchoninic Acid Protein Assay Reagent (BCA, Thermo Scientific) according to manufacturer's instruction using dilutions of 2 mg/mL albumin standard solution (Thermo Scientific) for the standard curve. BCA reaction was incubated at 37°C for 30 min and absorption was measured using a SpectraFluor Plus Tecan plate reader at 560 nm. For WB, 25-30 µg of protein was mixed with 4x NuPAGE LDS sample buffer supplemented with 10% DTT and boiled at 98°C for 10 min. Prepared samples were either stored at -20°C or directly subjected to WB analysis.

To analyze the protein expression levels samples were separated by SDS-PAGE with 10% Tris-HCL separating gels. Protein electrophoresis was performed using Tris/glycine/SDS buffer as a running buffer at 100 V for 2 h. Spectra Multicolor Broad Range (Thermo Scientific) or Page Ruler (Thermo Scientific) were applied to determine the band size. The proteins were transferred on the PVDF membrane previously activated with methanol using the TransBlot® Turbo™ semidry transfer system (Bio-Rad) at 25V for 10 min. The membrane was blocked with 5% bovine serum albumin (BSA) in TBST buffer (10 mM Tris-HCl, 150 mM NaCl, 0.1% Tween 20, pH 7.5) for 1 h at room temperature. The membranes were then incubated with a primary antibody in 5% BSA in TBST solution overnight at 4°C (summarized in Table 5). The next day, the membranes were washed 3 times with TBST followed by incubation with horseradish peroxidase (HRP)-conjugated secondary antibody (Table 5, bottom two rows) diluted in TBST for 1 h at room temperature. After 1 h, membranes were washed with TBST and antibody-protein complexes were observed using WesternBright® ECL HRP substrate (Advansta, CA, USA) and subsequent exposure to Fuji medical X-ray film SuperRX® (Fujifilm, Japan). Western blot for alpha-vinculin served as a protein loading control. Protein expressions were quantified using ImageJ (version 1.53, NIH, USA). The signal intensity of the target protein was normalized to loading control and for phosphorylated protein intensity was normalized to the total protein intensity of the respective protein. For experiments that required reprobing, membranes were stripped using the mild stripping protocol by Abcam (Cambridge, UK), blocked, and incubated with respective primary and secondary antibodies as mentioned in the previous steps.

Table 5: Summary of Antibodies used for WB

Antibody	Dilution	Source	Product number
MACC1	1:3000	Rabbit, Polyclonal	HPA020081/Sigma-Aldrich
ERK	1:1000	Rabbit, Polyclonal	9102/Cell Signaling
pERK (T202/Y204)	1:1000	Rabbit, Polyclonal	9101/Cell Signaling
SRC	1:1000	Rabbit, Polyclonal	36D10/Cell Signaling
pSRC (Y416)	1:1000	Rabbit, Polyclonal	2101/Cell Signaling
Vinculin	1:1000	Mouse, Monoclonal	VLN01/Thermo Fisher Invitrogen
Anti-Rabbit HRP	1:20,000	Goat, Polyclonal	W4011/Promega
Anti-Mouse HRP	1:20,000	Goat, Polyclonal	31430/Thermo Fisher

5.6 Co-immunoprecipitation (Co-IP)

In order to investigate interactions of proteins, Co-IP experiments were performed using whole cell lysates. $6x10^6$ cells were seeded in a 10 cm dish and allowed to assimilate for 24 h. Then cells were starved overnight and treated for respective time intervals. Immediately after the treatment, cells were washed once with cold 1x PBS and scratched with 600 µL ice-cold IP lysis buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 0.1% NP40, 1 mM EDTA pH 8) supplemented with complete protease inhibitor cocktail and phosphatase inhibitors (Roche, Switzerland). Cells were allowed to lyse for 30 min with intermediate vortexing followed by centrifugation for 30 min at 14,800 rpm at 4°C. Supernatants were collected in new tubes and incubated with 2 µg respective target antibody (Table 6) overnight at 4°C on a rotational shaker. On the next day, the protein-antibody complex was incubated with Dynabeads® Protein G (Invitrogen) for 4 h at 4°C on a rotational shaker. The antigen-antibody-magnetic bead complex was washed 3 times with 200 µL cold PBS. After washing, the complex was suspended in PBS and transferred to a clean tube. The target antigen was eluted using 50 mM glycine (pH 2.8) as an elution buffer. The eluted complex was mixed with 4x NuPAGE™ LDS sample buffer (Invitrogen) and 10% DTT and boiled for 8 min at 95°C. Using a magnet, the beads were separated, and the supernatant was collected and either directly loaded on the gel or stored at -20°C for SDS-PAGE and Western blotting.

Table 6: Summary of Antibodies employed for Co-IP

Antibody	Concentration/ Dilution	Source	Product number
Normal Mouse IgG	2 µg	Mouse, Polyclonal	12-371/Millipore
Normal Rabbit IgG	2 µg	Rabbit, Polyclonal	12-370/Millipore
pY1000	2 µg	Rabbit, Multimab	8954/Cell Signaling
MACC1	1:3000	Rabbit, Polyclonal	HPA020081/Sigma-Aldrich
SRC	1:1000	Mouse, Monoclonal	2110/Cell Signaling
Anti-Rabbit HRP	1:20,000	Goat, Polyclonal	W4011/Promega
Anti-Mouse HRP	1:20,000	Goat, Polyclonal	31430/Thermo Fisher

5.7 RNA Extraction and qRT-PCR

To determine the mRNA or DNA expression of different genes qRT-PCR was employed. 3 x 10^5 cells were seeded in a 6-well plate and allowed to adapt for 1 day and then starve overnight. On day 3, cells were treated for 90 min and immediately washed with cold PBS. Total RNA was isolated using the Universal RNA Purification Kit (Roboklon, Germany) according to the manufacturer's instructions. In the final step, RNA was eluted with nuclease-free water and quantified using a NanoDrop™ 1000 Spectrophotometer (Peqlab, Germany). For reverse transcription (RT), 50 ng RNA was reverse transcribed using Biozym (Germany) cDNA synthesis kit according to the manufacturer's instructions at 30°C for 10 min, 50°C for 40 min, 99°C for 5 min with subsequent cooling at 4°C for 5 min. Complementary DNA (cDNA) was either stored at -20°C or directly used for quantitative real-time PCR. cDNA was diluted 1:2 or 1:5 and amplified using the Biozym Blue S'Green qPCR Mix. Gene-specific primers were synthesized and HPLC purified by Biotez (Berlin, Germany) and are listed in Table 7. Using LightCycler® 480 (Roche Diagnostics, Switzerland) cDNA was amplified in a qPCR under the following conditions: 95°C for 2 min followed by 45 cycles of 95°C for 7 sec, 60°C for 10 sec, and 72°C for 5 sec. Data was analyzed employing LightCycler® 480 Software release 1.5.0SP3 (Roche Diagnostics, Switzerland). Each 10 µL PCR reaction was performed in 96-well plates in duplicates. The mean values of the expressed genes were normalized to the respective mean GAPDH expression which was used as a housekeeping gene.

Table 7: Sequence of qRT-PCR primers

Primer name	Primer Sequence 5' - 3'
MACC1 fwd	TTC TTT TGA TTC CTC CGG TGA
MACC1 rev	ACT CTG ATG GGC ATG TGC TG
c-fos fwd	CTGGCGTTGTGAAGACCAT
c-fos rev	TCCCTTCGGATTCTCCTTTT
G6PDH	ATC GAC CAC TAC CTG GGC AA
G6PDH	TTC TGC ATC ACG TCC CGG A

5.8 Proliferation Assay

For identification of the impact of mutation on MACC1's functional capability, a proliferation assay was performed using the IncuCyte® Zoom system. $6x10^3$ cells/well were seeded in the 60 inner wells of a Sarsedt 96-well plate (flat bottom) and the peripheral wells were filled with PBS. The cells were incubated overnight at 37°C in a humidified incubator with 5% CO_2. After the cells were settled, the medium was replaced with fresh medium with or without HGF 20 ng/mL. Subsequently, the plate was transferred to an incubator with an attached IncuCyte® Zoom system. Using the IncuCyte® Zoom 2019B software, a scan was scheduled to capture 4 images/well every 2 h continually. Cell proliferation was monitored for a period of 5 days and then the raw data was exported for analysis.

5.9 Wound Healing Assay

The scratch assay, also known as wound healing assay, was performed to evaluate the migration ability of MACC1 WT and MACC1 mutant cells. $1x10^5$ cells/well were seeded in the 60 inner wells of the 96-well IncuCyte® ImageLock plate. Then the cells were incubated overnight at 37°C in a humidified incubator. This short period enabled the cells to settle and form a confluent layer. After the cells were settled, the cells were starved by replacing the medium with reduced serum medium (2% FBS) for around 4 h. Subsequently, the wound was inflicted using the IncuCyte® WoundMaker™. Directly after creating the wound, the medium was removed to eliminate the reattachment of any floating cells. In the next step, a fresh medium containing reduced serum with or without HGF treatment (20 ng/mL) was added gradually to the cells without disturbing the wound. The plate was then transferred to an incubator connected to the IncuCyte® Zoom system. Employing the IncuCyte® software 2019B, a scanning schedule scanning 2 images/well every 2 h was set up to monitor the wound closure. The cells were observed for up to 3 days after the creation of the wound. After 3 days, the raw data and images were exported for analysis.

5.10 Colony Formation

Colony formation evaluates the ability of single cells to form colonies. For this study, a 2D clonogenic assay without any matrix was employed to study the effect of mutation on the colonization ability of MACC1. MACC1 WT and MACC1 Mutant cells at a very low density of 400 cells/well were seeded in a 6-well plate. The plates were incubated for a week at 37°C in a

humidified incubator with 5% CO_2. After 7 days, the medium was removed, and the cells were washed with PBS. To visualize the colonies, the cells were incubated with a solution containing fixing and the staining agent for 20 min. For this purpose, a solution of 1% formaldehyde and 0.1% crystal violet was applied. Next, the staining solution was removed, and the plates were washed with water until the background was clear. The plates were photographed, and the colony area was analyzed using the colony area plugin for ImageJ (version 1.53) [199].

5.11 Immunofluorescence

For fluorescent imaging, $1.5x10^4$ cells were seeded on a glass coverslip (VWR, 13 mm diameter, 1.5 µm thickness) in a 24-well plate and allowed to assimilate for 1 day and then starved using serum-free medium overnight. Cells were then either left untreated or treated with HGF for respective time intervals as mentioned in the figures. Post-treatment cells were washed with PBS and then fixed with 4% Paraformaldehyde for 15 min at room temperature. Fixed cells were then quenched with 0.1 M Glycine for 20 min. Following three consecutive washing steps with PBS-Tween (phosphate-buffer saline, 0.2% tween), the cells were permeabilized with 0.2% Triton-x 100 for 2 min. After washing with PBS-Tween, cells were then incubated with a blocking solution comprising 5% BSA for 1 h at room temperature. In the next step, cells were incubated with primary antibodies diluted in 2.5% BSA overnight at 4°C (Table 8). On the following day, the cells were washed 5x with PBS-Tween and incubated with the respective secondary antibody for 1 h at room temperature. The antibodies used for immunofluorescence are summarized in Table 8. After washing 3x, cells were incubated with DAPI for 3 min, washed with PBS, mounted with Dako Fluorescent mounting media (Agilent, USA), and allowed to dry. Slides were imaged using a Leica Stellaris 8 confocal microscope (Leica Microsystems, Germany) equipped with Power HyDS1, HyDS2, and HyDX3 detectors. Images were captured at a pixel size of 60.13 nm x 60.13 nm using a 63x (NA 1.5) objective, a zoom factor of 3x, and step size of 0.25 µm. Images were saved in original (.lif) format and quantified using Imaris 8 with colocalization add-on (Bitplane, USA). Colocalization was quantified using the automatic threshold by Imaris.

Table 8: Antibodies used for immunofluorescence

Antibody	Dilution	Source	Company
Anti-MACC1	1:100	Mouse, Monoclonal	Sigma Aldrich
Anti-SRC	1:500	Rabbit, Polyclonal	Sigma Aldrich
Anti-Mouse Alexa Fluor® 488	1:200	Goat, Polyclonal	Thermo Fisher
Anti-Rabbit Alexa Fluor® 647	1:200	Donkey, Polyclonal	Thermo Fisher

5.12 Mass Spectrometry Analysis

The global protein profiling of HCT116/WT and HCT116/Mut cells was performed using the data-independent acquisition (DIA)-based Mass spectrometry (MS) strategy on an Exploris 480 mass spectrometer by the MDC proteomic facility (O.Popp/P. Mertins). In the next step, data was analyzed in the Spectronaut 17 due to the improved availability of downstream analysis tools. A minimum of 3 biological replicates were used for the global protein and phosphoprotein profiling. The proteome and phosphoproteome data matrix were filtered for ≥75% valid values per row and then imputations were applied to the log2-LFQ values. The phospho-site table was created using the Perseus plugin tool.

The gene enrichment analyses were performed using Metascape [200] and AmiGO2 [201-203]. The top 200 downregulated genes in the 3xMut MACC1 cells were selected for gene enrichment analyses. The expression was analyzed using the ANOVA test employing a 5% False Discovery Rate (FDR).

5.13 Prediction of MACC1 Phosphorylation Tyrosine Site

NetPhos 3.1, a prediction tool developed by Blom N, et al. (2014) was employed to identify the putative phosphorylation sites close to the N-terminus of MACC1 [140]. Additionally, a kinase prediction tool known as the Group-based prediction system (GPS 5.0) was used to predict the potential protein kinases responsible for phosphorylating the putative phosphorylation sites [138]. The FASTA sequence of MACC1 (Q6ZN28) from UniProt [204] was inserted as an input. The

phosphorylation tyrosine predictions were carried out with default settings. For kinase prediction, a high threshold cut-off was selected to improve selectivity and sensitivity of the kinase prediction.

5.14 Protein Structure and Dimer Prediction

The recent development in AI has advanced protein prediction to a new level. One such example is AlphFold2 developed by DeepMind [144]. AlphaFold2 predicts highly accurate protein structure even when templates are not available. For predicting the MACC1 structure, AlphaFold2 Colab [154], a simplified version of AlphaFold2 was utilized. The structure was predicted based on the FASTA sequence of MACC1 with accession number Q6ZN28 from UniProt [204] using default parameters. The output yielded 5 models which were ranked based on accuracy.

The MACC1 homo-oligomer state was predicted by using a freely accessible web server named GalaxyHomomer [162]. The amino acid sequence of MACC1 was inserted as an input without any template or oligomeric state. In addition, MACC1 dimer predictions were performed by our collaboration partner C.Buhlheller (FMP, Berlin) using AlphaFold-Multimer [163]. A cropped region of MACC1 was utilized to predict the MACC1 dimer. The cropped region consists of residues from Ser201 to Pro555 excluding the beginning disorder region. For C-terminal MACC1 dimer prediction, the amino acid sequence Ser551 to Val840 was used as an input. The predicted 3D structures of MACC1 were visualized using PyMol version 2. (Schrödinger, LLC, New York, USA).

5.15 Statistical Analysis

Statistical analyses were performed with GraphPad Prism version 6 (GraphPad Software, Massachusetts, USA) and Microsoft Excel (part of Microsoft 365 version 2307). Student t-test and one-way analysis of variance (ANOVA) were applied to compare two or more than two groups, respectively. p-values below 0.05 were considered statistically significant.

6 Results

6.1 MACC1 Residue Y379 is Integral for Constitutive Activation of MACC1

Extensive research over a decade has established MACC1 as a key player in the progression of metastasis in various types of cancer [72, 73, 79, 205-207]. High expression of MACC1 is associated with increased metastatic properties such as increased proliferation, migration, and invasion in cell models and increased tumor metastasis in the mouse model [101]. In patients, MACC1 expression levels not only help identify patients at high risk of developing metastasis but also treatment responses to standard chemotherapeutics [73, 76-78, 101].

Interestingly, MACC1 is a transcriptional regulator of a proto-oncogenic receptor tyrosine kinase, c-Met, and previous MS analyses have suggested interaction between MACC1 and downstream signaling components of the HGF/c-Met signaling axis [48, 73, 82]. Although the precise protein interactome of MACC1 involved in metastasis formation remains to be determined, it is evident that the unique domain structure composition of MACC1 and tyrosine sites are fundamental to its metastatic ability [72, 73, 79, 103]. This project focuses on the structural features, tyrosine sites of MACC1, and MACC1's ability to self-associate. To achieve this, various protein interaction, cell signaling, and metastasis assays were employed in CRC cell lines with modified MACC1 expression.

6.1.1. Identification of Tyrosine Phosphorylation sites of MACC1 *in silico*

Tyrosine phosphorylation is an important step in the regulation of tyrosine kinase signaling not only under physiological conditions but also in tumor progression [138, 139, 208]. Tyrosine phosphorylation sites act as platforms for protein interactions and thus link various signaling cascades [81, 83, 109]. To investigate the probability of phosphorylation of tyrosine sites of MACC1, a prediction tool known as NetPhos 3.1 [141] was used. NetPhos employs artificial neural networks to predict posttranslational modifications such as the phosphorylation of tyrosine in an amino acid sequence [140, 141]. A threshold score of 0.5 as determined by the server was applied for MACC1 tyrosine predictions. Only prediction scores above 0.8 were considered for further analysis. Putative tyrosine phosphorylation sites were distributed throughout the MACC1 sequence, however, more frequently observed close to the C-terminal double death domain. Out of 25 tyrosine residues, 6 tyrosine sites at 379, 598, 695, 768, 789, and 793 showed prediction scores above 0.8 and 15 other tyrosine sites showed prediction scores above 0.5. As the sites close to the C-terminus were analyzed in another study [82, 103], the role of individual mutation at Y379 close to the N-terminus was evaluated in this study.

In the next step, kinases responsible for phosphorylating Y379 were investigated using GPS 5.0 [138]. GPS stands for Group-based Prediction System and is a highly accurate tool for the prediction of kinase-specific phosphorylation sites [138]. Using the default setting and high cut-off threshold, the analysis predicted Abelson-related gene (ABL2), fibroblast growth factor receptor (FGFR), Janus kinase 3 (JAK3), and SRC kinase as some of the hits capable of phosphorylating Y379.

Extensive evidence links SRC to malignant tumors of the colon [133, 209]. Like MACC1, upregulation of SRC is also observed in the progression of CRC from benign polyps to carcinoma and metastasis [118, 125, 133]. A previous study by Zincke F. (2019) showed that SRC and FAK activation is increased in cells overexpressing MACC1 [82]. However, not much is known about the association between SRC and MACC1 in tumor development [82]. Therefore, the correlation between SRC and MACC1 was investigated in patient cohorts. The open-access Pan-cancer Analysis of Whole Genome (ICGC/TCGA) database [30] was evaluated using the cBioportal platform [210, 211]. Results analyzing the mRNA expression of 1210 cancer samples from various cancer types showed a positive correlation between MACC1 and SRC with a Spearman correlation coefficient of 0.36. Additional analysis employing the GEO Omnibus CRC patient cohort dataset, GDS4516 [212] revealed a similar trend with a Spearman coefficient of 0.42. Surprisingly, a negative correlation was observed between MACC1 and CSK, a negative regulator

of SRC. In line with the previous report examining the influence of MACC1 on FAK phosphorylation [82], a high correlation was also observed between MACC1 and other interaction partners of SRC such as FAK (PTK2) and RAC1 (Figure 7). Therefore, it suggests a high likelihood of association between MACC1 and SRC signaling in tumor progression.

To examine whether MACC1 and SRC interact in addition to co-expression, a classical PPI assay known as co-immunoprecipitation (Co-IP) was employed. SW620, a metastatic CRC cell line known to have high expression of MACC1 was used for the Co-IP experiment [88, 95, 103, 213]. Strikingly, MACC1 co-immunoprecipitated with the anti-SRC antibody indicating a positive interaction. The pull-down with an IgG antibody served as a negative control and displayed no unspecific binding. In addition, whole cell lysate was used as a positive control and revealed high intensity MACC1 bands as expected. The interaction was observed strongly in cells without HGF treatment indicating that the interaction could be ligand-independent.

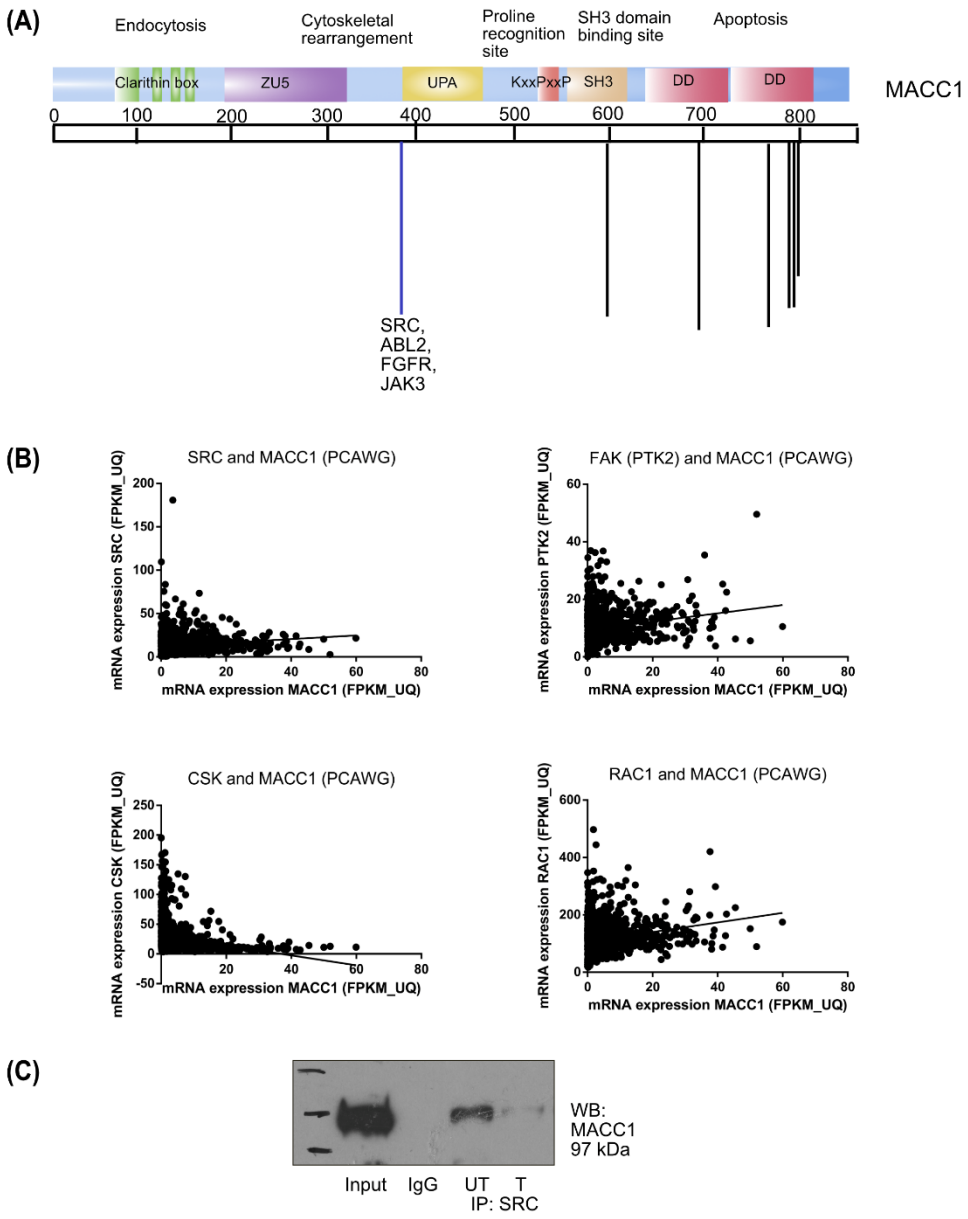

Figure 7: Association between MACC1 and SRC

(A) Schematic representation of MACC1 along with the predicted phosphorylated tyrosine sites and kinases. (B) Correlation of MACC1 with SRC, FAK (PTK2), the negative regulator of SRC (CSK), and RAC1 in tumor samples. (C) Co-immunoprecipitation (Co-IP) showing the interaction of MACC1 with SRC in untreated (UT) and treated (T) with HGF 20 ng/mL.

6.1.2. Impact of Single Mutation at Y379F on MACC1 Signaling

Based on these findings, tyrosine site Y379 was mutated to phenylalanine (Y379F), a non-phosphorylable form [82]. HCT116, a CRC cell line expressing moderate levels of MACC1, expressing GFP only (Control), MACC1-GFP (WT MACC1-GFP), and Y379F (TyrMut MACC1-GFP) which was developed for a former project was employed for further investigation [82]. ERK1/2, a downstream effector of the mitogen-activated protein kinase (MAPK) cascade, is central to tumor metastasis and acts as a central regulator in the conversion of an extracellular signal to cell-specific response depending on the stimuli and cell state [84, 122]. Thus, activation of ERK through phosphorylation of sites T202 and Y204 was investigated using WB.

To examine the impact of the tyrosine mutation on MACC1 signaling, Control (GFP only), WT (MACC1-GFP), and TyrMut (Y379F MACC1-GFP) were starved after seeding and treated with HGF 20 ng/mL for distinct time intervals. Cell lysates were collected after treatment and protein expression of downstream effector ERK was analyzed. Additionally, the mRNA expression of *MACC1* and *cfos* was analyzed to corroborate the effect of treatment on the gene level. Consistent with previous reports [82] WT MACC1 cells displayed ERK activation before growth factor treatment indicating a ligand-independent signal activation. The TyrMut cells showed reduced ERK activation compared to WT cells in the absence of growth factor (Figure 8). This is in line with the previous evidence investigating the double mutant (Y379+Y789F) [82].

As SRC was identified as one of the key kinases responsible for the phosphorylation of Y379F, the presence of a feedback mechanism leading to SRC activation post-treatment was hypothesized. Phosphorylation of SRC at residue Y416 served as a read-out for SRC activation. Although TyrMut cells showed slightly higher SRC phosphorylation without treatment compared to WT, the intensity of SRC phosphorylation was similar in both groups post-HGF treatment. In summary, these findings emphasize an important role of Y379 MACC1 residue in the activation of ERK in a ligand-independent manner.

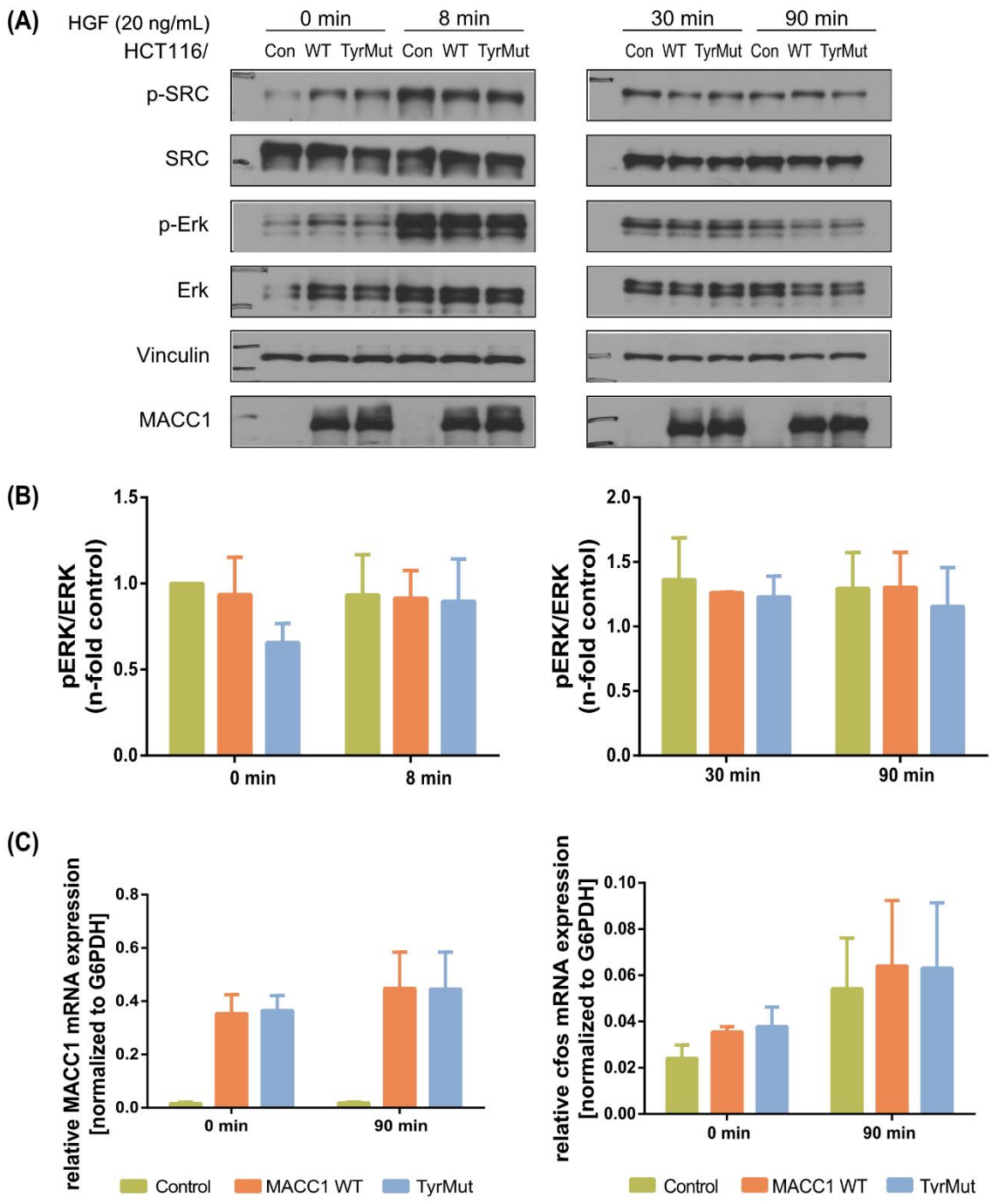

Figure 8: Characterization of MACC1 TyrMut

(A) HCT116 cells overexpressing Control (GFP only), WT (MACC1-GFP), and TyrMut (Y379F MACC1-GFP were treated with HGF 20 ng/mL for 0-, 8-, 30- and 90-minute intervals. Cell lysates were collected and analyzed for downstream effectors of the c-Met-HGF pathway such as ERK and SRC by WB. The effect of mutation on the activation of SRC was analyzed in at least three independent

biological replicates through phosphorylation of site Y416. (B) Results were quantified using spot densitometry in ImageJ (version 1.53, National Institutes of Health, USA) [214]. Although differences were observed in untreated conditions between the WT and TyrMut, no significant differences were observed on treatment between the two groups. (C) MACC1 mRNA expression in Control, MACC1 WT, and TyrMut normalized to G6PD with and without treatment verified the mRNA expression of *MACC1* and *c-fos*.

6.1.3. **SRC Colocalizes with MACC1 in Colorectal Cancer Cells**

SRC plays a key role in regulating cell-cell contacts and dysregulation of SRC has a direct impact on the actin-cytoskeleton dynamics [127, 215, 216]. Data from other groups suggest that the localization of SRC is critical for the regulation of cellular junctions and cell polarity. The active form of SRC is found at the periphery of the cell and in the cytoplasm while the inactive form is localized perinuclearly [217, 218].

On the other hand, MACC1 is expressed mainly in the cytoplasm and to some extent at the cell membrane [48, 104]. Interestingly, research from our group reveals that MACC1 is capable of shuttling from the cytoplasm to the nucleus on treatment with HGF [48, 95]. Moreover, the whole tissue section expression analysis showed that the expression of MACC1 at the invasive front is correlated with metastasis and reduced survival [100]. Using the metastatic CRC cell line SW620, which expresses a high level of MACC1, the probability of MACC1 colocalization with SRC was determined in this study.

Consistent with the previous observation [48, 95, 104] MACC1 was mainly found in the cytoplasm and SRC along the cell-cell junction. Confocal microscopy images show that MACC1 is abundantly present throughout the SW620 cells (Figure 9). Strikingly, SRC and MACC1 accumulate at the cell periphery without treatment. This was clear also in Pearson's co-efficient value of 0.5218 obtained for SRC and MACC1 in the absence of growth factor treatment.

Subsequently, the effect of treatment with HGF on the colocalization of SRC with MACC1 was investigated at two different time points explicitly 8 min (Figure 9) and 90 min representing a short and long duration of activation. Contrary to expectation, MACC1 and SRC correlation decreased upon HGF stimulation. This statement is based on Pearson's coefficient values of 0.3608 and 0.3129 attained for SW620 cells treated with HGF for 8 and 90 min, respectively.

Taken together, these findings highlight the association of MACC1 and SRC in metastatic CRC cell line SW620. Specifically, a higher degree of colocalization was observed in the untreated condition and decreased gradually on treatment with HGF.

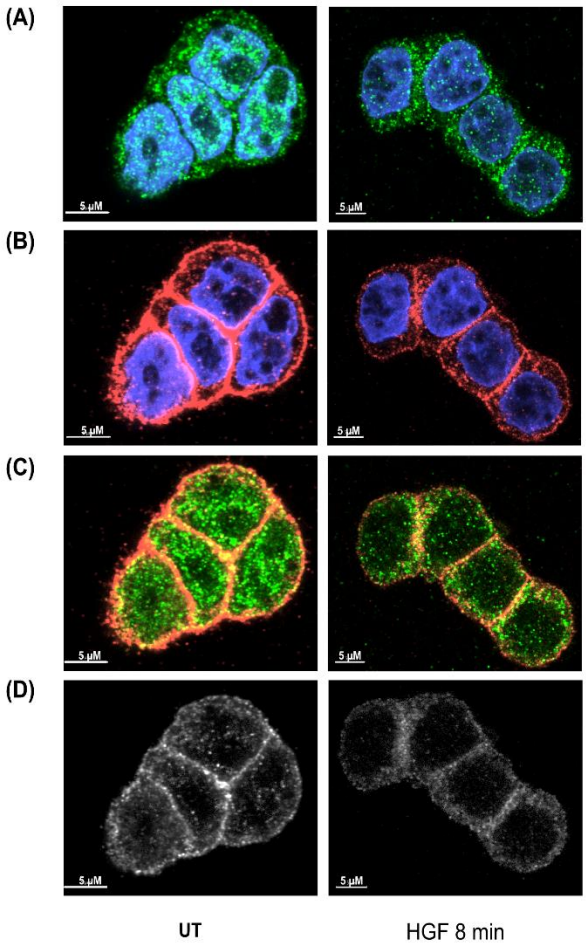

Figure 9: Cellular distribution of MACC1 and SRC in SW620 cells

SW620 ($1.5x10^4$) cells were seeded on coverslips and serum-starved overnight. On the following day, cells were either left untreated or treated with HGF for 8 min. Cells were prepared for immunofluorescence analysis to evaluate the colocalization of MACC1 with SRC. Single channels showing the distribution of MACC1 (A) and SRC (B) with DAPI (blue). (C) Colocalization of both MACC1 and SRC in untreated and treated cells. (D) The colocalization channel depicts the distribution pattern of MACC1 and SRC.

6.2 The Power of Two MACC1s: MACC1 Dimers

The first part of this project sought to determine the potential tyrosine phosphorylation sites of MACC1 and the association between SRC and MACC1. The most interesting finding was that ABL2, JAK, SRC, SYK, and RET are the key kinases involved in the phosphorylation of tyrosine sites of MACC1. Another important observation was the accumulation of MACC1 and SRC at the cell periphery in CRC cells.

Although prior studies have noted the importance of MACC1 structure [73, 79, 82, 94, 103, 104], a closer in-depth characterization, and likelihood of MACC1 self-association have not been addressed. With this in mind, the next part of the study was designed to understand in detail the structural features of MACC1 and explore possibilities of MACC1 oligomerization in living cells.

The upcoming part of this project sets out to investigate the self-association ability of MACC1 using novel techniques such as BRET and explore MACC1 residues involved in dimerization through innovative prediction tools. Here, the important goal was to identify the self-association of MACC1 and its influence on metastasis progression (Figure 10).

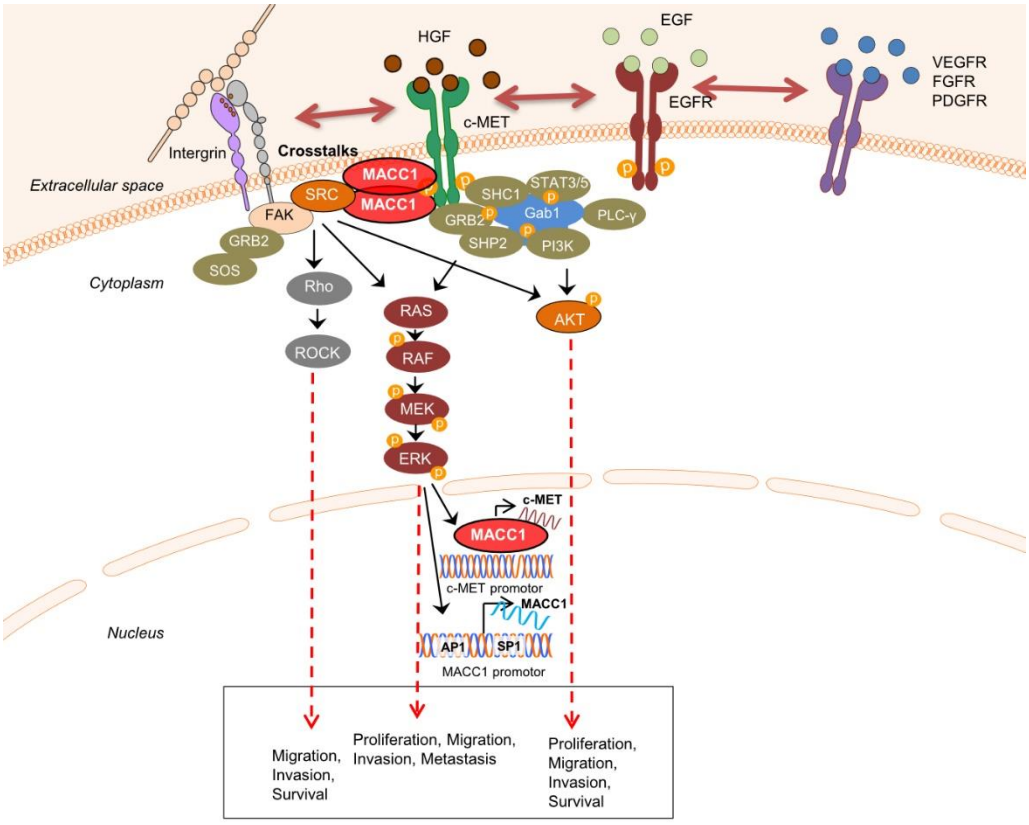

Figure 10: Investigating MACC1 dimerization and its role in tumor metastasis

6.2.1. Colabfold: An AlphaFold Platform Predicts the 3D structure of MACC1

Despite numerous studies highlighting the role of MACC1 in metastasis progression, little is known about the 3D structure of MACC1 [72, 73, 82, 94, 95, 103]. AlphaFold2 [144], a highly accurate protein prediction tool based on novel neural networks was utilized for MACC1 structure prediction. AlphaFold2 employs available information from similar sequences and protein structures in the same living organism and refines it through repeated cycles to predict a representative structural model. The information from similar sequences, similar structures, and a predicted representative structure is combined and refined again through repeated cycles to predict a highly accurate 3D structure model [143, 144].

Colabfold [144, 154] is an open-source platform based on AlphaFold2 but with a modified environment and smaller database. Thus, making it possible to run simulations on systems with limited computing power and without coding expertise [154]. For this study, the MACC1 sequence consisting of 852 amino acids stored under UniProt [204] accession *Q6ZN8* was inserted as an input. Predictions were made using the default settings and with no template. As an output, five models were generated with three recycles each.

In the next step, the five models were ranked by AlphaFold based on accuracy and reliability of the prediction. The sequence coverage graph indicated the highest sequence identity for residues between 200 and 600. The score was lowest for the first 200 N-terminal residues and at the end with minor drops observed around 400 and 550 residues.

The accuracy of the prediction was determined by AlphaFold using two parameters. First, the predicted local distance difference test (pLDTT) [144, 219], which ranges from 0-100, values above 70 indicate good prediction. The values for disordered regions usually lie below 50 and need to be interpreted cautiously [144]. Second, the prediction aligned error (PAE) which is represented as a 2D plot, denotes the position error if the predicted and true structure were aligned on the X and Y axis, respectively. Additionally, the predicted template modeling score (pTM) derived from the PAE ranging from 0-1 is assigned to each model, with 1 denoting the best prediction [144].

Overall, the pLDTT values were similar between the five models illustrated by the superimposed curve. Values above 70 were observed for most of the sequence indicating good MACC1 backbone prediction. The pLDDT and pTM score for the best model was 70.5 and 0.468, respectively.

The PAE plot for MACC1 model 1 displays well-defined domains. The upper left quadrant corresponds to the N-terminus and the lower right quadrant to the C-terminus of MACC1. As illustrated in Figure 11, the values for the PAE plot range from 0 (blue) for the least probability of error to 30 (red) representing the highest probability of error in Angstrom units [144]. The plot for MACC1 shows the presence of two major well-defined regions with additional smaller domains present within these regions. The starting region until residue 200 is denoted in red representing a high level of uncertainty in prediction.

Overall, the confidence scores such as the pTM and pLDDT values for the five predicted models were high, ranging from 0.468 to 0.432 and 70.5-68.3, respectively. The structural features of the ZU5 domain and double-death domain were all well-defined and provided a detailed insight into MACC1 structural characteristics. The subsequent chapter inspects the probability of MACC1 oligomerization and the number of MACC1 units involved.

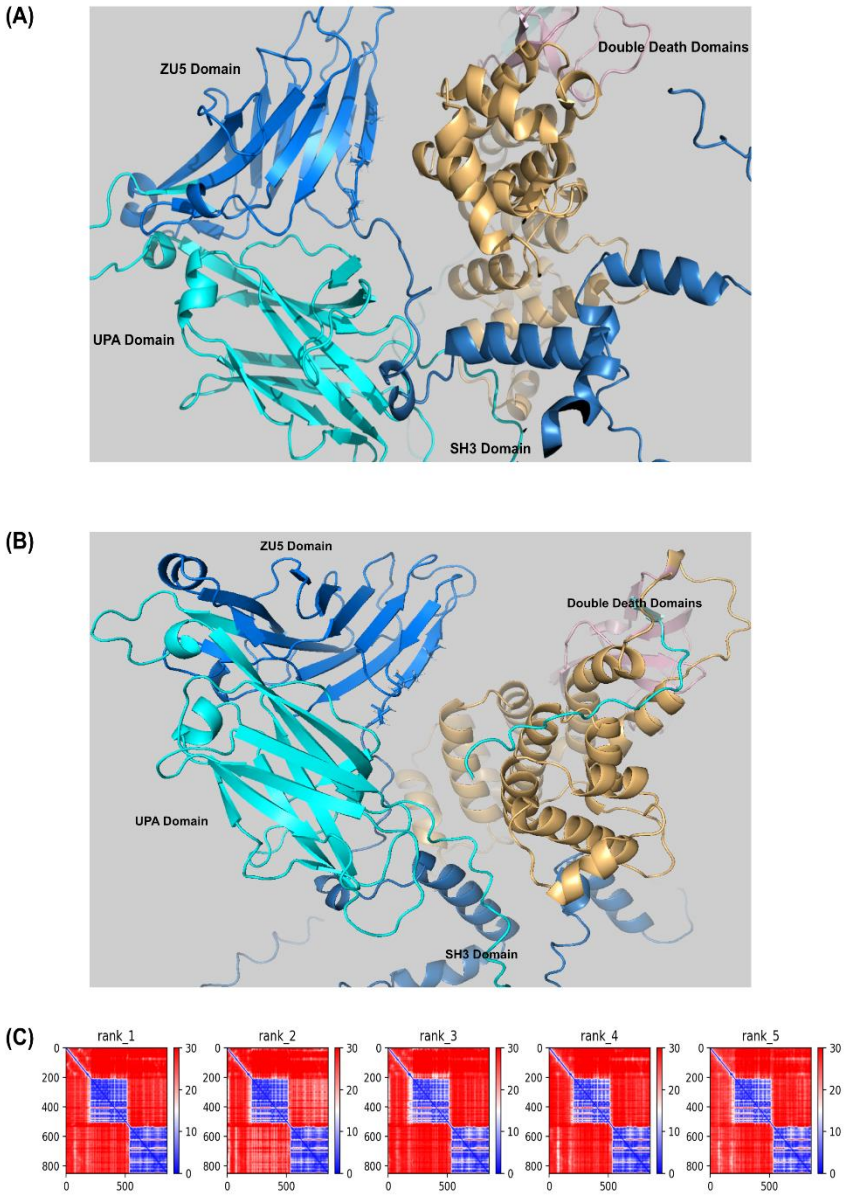

Figure 11: Predicted MACC1 structure

(A) Overview of the predicted model of MACC1 depicting the four important domains of MACC1 from N-to C-terminus (counterclockwise), namely the ZU5 domain followed by UPA, SH3, and a double death domain. (B) Representation of the predicted model from above exhibited the domain organization between the N- and C-terminus domains. (C) Predicted aligned error (PAE) score for the predicted MACC1 model as retrieved from Colab [154]. The calculated error ranges from blue (lowest predicted error) to red (highest predicted error) depicted in angstrom units [154].

6.2.2. Higher Probability of MACC1 Forming Dimers

Protein homo-oligomerization is a common phenomenon observed in numerous protein families including metastasis proteins [180, 220, 221]. In addition to improving stability, the self-association of proteins plays an important role in regulating various cellular processes and providing additional binding sites [160, 161, 166]. Interestingly, self-association can also contribute to prime proteins for their functional characteristics [222].

To investigate whether MACC1 is capable of oligomerization, an open-access homo-oligomer prediction tool known as GalaxyHomomer [162] was employed. For the prediction, an amino-acid sequence of MACC1 from UniProt [204] was inserted as an input without providing the specific oligomeric state. The GalaxyHomomer server utilizes information from similar sequences to predict the structure, perform *ab initio* docking and predicts the oligomer state of the protein [162]. Structural features generated with low level of confidence are refined and relaxed further [162, 223]. The output consisted of 5 models, ranked based on the *ab initio* docking score.

The region between 620-628 amino acids was refined for all 5 models. Out of the 5 models generated, the first three high-ranking models predicted MACC1 as a dimer whereas the fourth model predicted it as a pentamer, and the last model was a hexamer. From the prediction data, it was evident that there is a higher tendency of MACC1 to form dimers than higher-order structures. The *ab initio* docking score for the 5 models ranged from 1766 to 1079 with scores for dimer considerably higher than that for pentamer or hexamer. The interface area for the highest ranked dimer model was 5601 A° (Table 9).

In summary, the prediction tool suggested a higher likelihood of MACC1 to exist as a dimer as opposed to higher-order oligomers such as pentamer or hexamer. The following section explores in detail the likelihood of MACC1 to form a homodimer.

Table 9: Prediction of MACC1 homooligomer units using GalaxyHomomer [162]

Model no.	Number of subunits	Interface area	Docking score
1	2-mer	5601.9	1766.73
2	2-mer	1985.6	1654.45
3	2-mer	2607.5	1651.66
4	5-mer	157789.4	1072.06
5	6-mer	189772.4	1079.09

The Fasta-sequence of MACC1 from UniProt [204] was used as input. The interface area and *ab initio* docking score for the 5 predicted MACC1 models are displayed in the Table showing the highest docking scores for the MACC1 dimer structure.

6.2.3. Characterization of MACC1 Dimer using AlphaFold-Multimer

Self-association of proteins to form higher-order oligomers is a cornerstone of several biological pathways [160, 161, 166, 222]. Although, studies estimate the majority of proteins to exist in oligomer states rather than monomeric units, only a fraction of proteins existing as dimers or in higher-order assembly have been identified [160]. AlphaFold-Multimer [151, 163], an extension of AlphaFold, with features adapted to handle multiple chains helps to predict highly accurate homomeric protein complexes [163].

To gain unbiased insight into the ability of MACC1 to dimerize and the putative interface involved in dimerization, three different scenarios were postulated. In the first scenario, the dimerization of N-terminus MACC1 (ZU5 domain) with N-terminus (ZU5) of another MACC1 molecule was hypothesized (N-domain:N-domain). To simulate this model, a region of MACC1 N-terminus starting from amino acid Ser201 to Pro555 was utilized. 25 models were generated and ranked based on accuracy. A readout comparable to predicted TM (pTM) [144, 163] described in the previous section was used for determining the accuracy score between residues at the interfaces. This score was denoted as ipTM [163]. The pTM score of 25 models generated for N-domain:N-domain dimerization ranged from 0.86 to 0.51. On the other hand, remarkable differences were observed in the ipTM. While the ipTM value was between 0.8-0.47 for the 9 top ranked models, the ipTM value was below 0.24 for the rest of the models. As expected, the contact map demonstrating the distances between N-domain:N-domain molecular residues were distinct and well-defined. Moreover, the predicted aligned error (PAE) was also considerably smaller, indicating a reliable 3D prediction of the complex.

The prediction shows that the hydrophobic residues Val212, Ile214 and Cys216 present at the beginning of the ZU5 domain of one MACC1 chain interact with residues on the second MACC1 chain. The interaction between the ZU5-ZU5 domain mainly comprise the first β-strand of the ZU5 domain (Val212-Val218) (Figure 12). In depth analysis reveals that the MACC1 homodimer interface contains a hydrophobic core formed by residues Val212, Ile214 and Cys216. The speculated residues from the two chains are around 4.9 to 4.5 A° apart. Interestingly, all the top ranked 15 models show involvement of these specific residues. This implies a pivotal role of these residues at the dimer interface.

In the second scenario, the involvement of the C-terminus in MACC1 oligomerization (C-domain:C-domain) was questioned. To test this case, a similar procedure as before was followed. A cropped C-terminus region of MACC1 from Ser551 to Val840 was used as an input.

As an output, 25 models ranked by accuracy were generated. The ipTM scores for these models were notably lower than for N-terminus prediction, with 0.29 being the highest value. In agreement with ipTM values, the pTM scores were also considerably lower with values ranging from 0.57-0.48 suggesting lower prediction confidence. The results of the contact map matched the scores and were not as distinct as observed in the previous case.

In the final case, the prospect of interdomain dimerization comprising of C-terminus interaction of one MACC1 chain with N-terminus of another MACC1 chain (C-domain:N-domain) was investigated with similar settings. Strikingly, all the 25 models created showed a drastically low ipTM value below 0.2. As suspected, the pTM scores were also minimal and no clear pattern was observed in the contact maps for this interaction pair.

Among all the modeled dimer combinations, the N domain:N domain (ZU5-ZU5) dimer was the best hit with prominently higher ipTM, pTM scores and lowest PAE values. Thus, the N-terminus dimerization was selected for in-depth characterization.

A careful examination of the dimer structure verifies that the dimerization of MACC1 is largely a result of interactions between the residues of the first β-strand of the two MACC1 proteins. This region is part of the ZU5 domain and corresponds to the non-polar residues Val212, Ile214, and polar residue Cys216. Remarkably, all the models generated pointed towards the involvement of these residues. Hence, these residues were identified as good candidates for further examination and mutation studies.

Evidence suggests that it is challenging to disrupt β-strand interactions. Precisely, these complexes are shallower and mainly involve complex interaction patterns [224, 225]. Therefore, the exchange of postulated residues forming the dimer interface with negative repulsive charges such as aspartic acid was recommended by our collaboration partners.

In summary, the results in this section identify the region of the dimer interface and residues pivotal for MACC1 dimerization, therefore, forming the foundation for the upcoming experiments. In the subsequent section, the importance of these residues is analyzed in-depth.

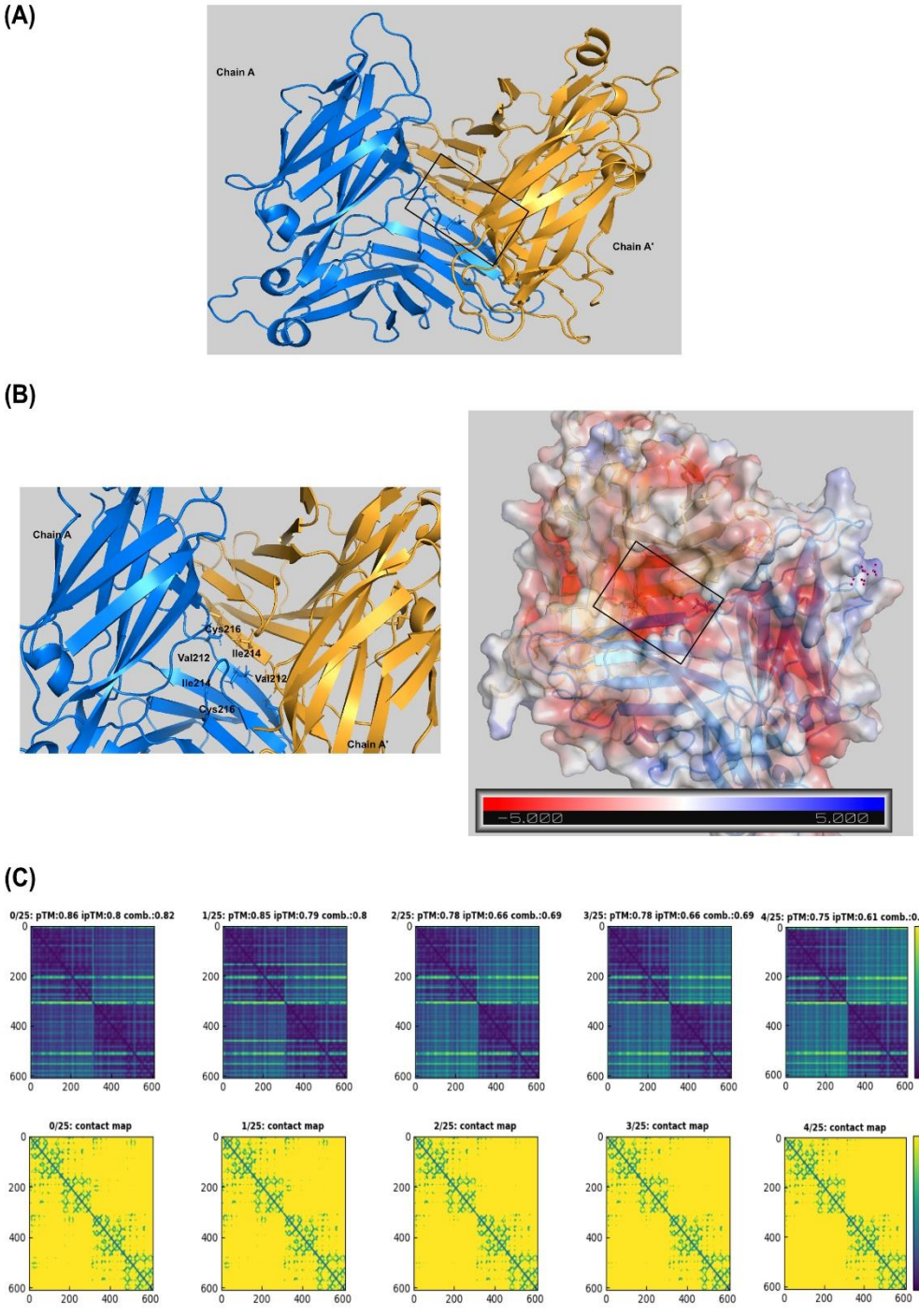

Figure 12: Structure of the MACC1 dimer

(A) The structure of MACC1 dimer prediction using AlphaFold-Multimer [163] shows the interaction of two ZU5 domains resulting in a MACC1 dimer. (B) A thorough analysis of the MACC1 dimer structure revealed the involvement of key residues namely Val212, Ile214, and Cys216 in the dimer formation. These residues are present close to the N-terminus of the MACC1 ZU5 domain and interact with the complimentary residues of the second chain thus forming a deep hydrophobic pocket. (C) Accuracy

score matrix and contact maps based on the results from AlphaFold-Multimer [163] for the top 5 predicted structures from 25 predicted models. The accuracy score matrix depicts the predicted aligned score ranging from a high level of confidence (blue) to the lowest level of confidence (yellow). The contact maps illustrate the intermolecular contact points for the top 5 models.

6.2.4. Systematic Validation of MACC1 Expression in BRET Vectors

PPIs are fundamental for all cell processes from beginning of the cell cycle to cell death [173, 192]. Noninvasive techniques such as FRET or BRET that can study protein interactions in a living cell are indispensable [187, 188, 192]. Intervening with MACC1's expression and cell signaling complex to hinder metastasis has been a long-term goal of our group [82, 88, 92, 103]. In this study, BRET, a technique based on transfer of resonance energy between two fluorophores when in proximity (max 10 nm) was employed to confirm MACC1 dimerization and potential target site important for MACC1 dimerization [169, 192]. The BRET fluorescence vector expressing mCit (GFP), and luminescence reporter vector expressing NLuc were received as a kind gift from the group of E. Wanker (MDC, Neuroproteomics) [175, 195, 196].

The cDNA fragment encoding WT MACC1 was shuttled into plasmids containing the BRET reporter using Gateway® Cloning as described in the methods section. MACC1 was expressed with luciferase donor (NLuc) at the N-terminus resulting in NLuc-MACC1 referred to as MACC1(N') and MACC1 along with BRET acceptor (mCit) at the N-terminal producing PA-mCit-MACC1 referred to as MACC1(N). Additionally, WT MACC1 at the C-terminal was tagged with luciferase donor giving rise to MACC1-NLuc denoted as MACC1(C'). The correct insertion of MACC1 was evaluated using restriction digestion and Sanger sequencing (LGC).

The controls for BRET experiments were received as a kind gift from the group of E.Wanker. The positive control comprised the acceptor and donor linked to each other (PA-mCit-NLuc), negative control comprised empty BRET reporter vector without the protein of interest, PA-NLuc control to compensate for the donor luminescence bleed-through and an empty vector pcDNA3.1 control were included in the assay setup [175, 195].

Interestingly, BRET is also suitable for studying the impact of point mutations and domain deletions on the interaction strength. Thus, making it valuable to identify and validate key contact points between the proteins [175, 195, 226]. A similar BRET assay setup as described earlier in HEK293 and HCT116 cells was selected to assess the influence of mutation of the putative residues on MACC1 dimerization. For this purpose, first the three putative residues Val212, Ile214 and Cys216 were mutated using SDM kit. The mutated MACC1 was then inserted in BRET fluorophore containing plasmids. In the final step, plasmids encoding the mutated MACC1 proteins along with BRET reporter were validated using restriction digestion and Sanger sequencing (LGC).

Protein expression of MACC1-tagged BRET hybrid proteins was verified using WB. A model system of HEK293 cells that endogenously express low MACC1 was utilized to set up the assay conditions. Additionally, as MACC1 is a well-known biomarker in CRC, a CRC cell line was also included for testing. HCT116, a human colorectal carcinoma cell line, which shows moderate endogenous expression of MACC1 was selected for this purpose. Conservation analysis of MACC1 ZU5 residues forming the hydrophobic pocket demonstrated that the residues Val212 and Ile214 are dispensable across species. In comparison, the postulated Cys216 is highly conserved among various species and may be integral for dimerization.

HEK293 and HCT116 cells were transfected with BRET vectors expressing WT MACC1, 3xMut MACC1 and empty vector. After 72 h, cells were lysed and the protein expression of MACC1 was investigated. Previously established MACC1-V5 vector was used as a positive control and non-transfected cells served as negative control. The expression of MACC1 was evaluated. In addition, β-actin expression was analyzed as a loading control for protein. The molecular weight of PA-mCit tag (~40 kDa) and cmyc-NLuc tag (~20 kDa) were added to the molecular weight of MACC1 respectively. As expected, all plasmids expressed MACC1 in both cell lines as observed in Figure 13. There was no visible difference in the band intensity between the WT MACC1 and 3xMut MACC1, suggesting that the mutation of the ZU5 domain residues does not interfere with MACC1 expression. As the excitation and emission wavelength of mCit tag is similar as GFP, the fluorescence of WT MACC1 tagged with mCit and 3xMut MACC1 was investigated using fluorescence microscopy.

From these findings, two important conclusions can be drawn - Firstly, the dimer interface contains a highly conserved residue, Cys216, which could be integral for dimerization. Secondly, the selected HEK293 and CRC cell line HCT116 express MACC1 BRET vectors. Importantly, protein expression of MACC1 was not affected by the mutation. In the subsequent sections, these vectors are employed to study the interaction between the two MACC1 proteins resulting in MACC1 dimer.

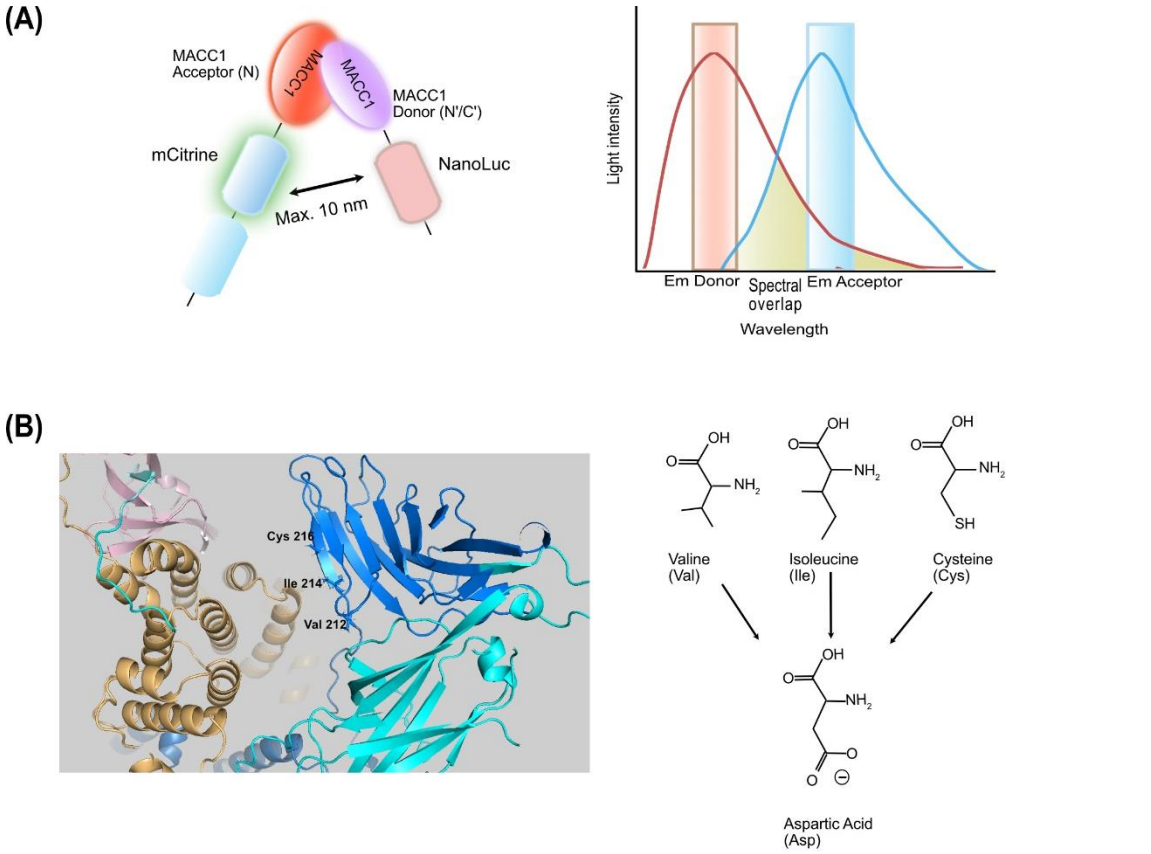

Figure 13: Generation and characterization of MACC1 dimer hindering mutant

(A) Schematic representation of the interaction between MACC1 fused with BRET acceptor (mCit) at N-terminus and MACC1 fused with BRET donor (NLuc) at the N- or C-terminus (N'/C'). Graph showing the importance of spectral overlap between the emission of BRET donor and excitation of BRET acceptor (B) MACC1 structure depicting the residues involved in dimerization and mutation scheme adopted to inhibit MACC1 dimerization. (C) Comparison of MACC1 sequences depicting conserved Cys216 across species. (D) Protein analysis confirming the expression of MACC1 WT and 3xMut vector in HEK293 and HCT116 cells, respectively.

6.2.5. BRET Establishes the Presence of MACC1 Dimers in Living Cells

BRET continues to be a powerful method to detect direct associations between proteins forming homo-oligomer complexes in living cells [227, 228]. In addition to circumventing problems associated with FRET, the development of the next generation of high-intensity BRET luminescence donors such as NLuc significantly improve the sensitivity and applicability of the technique [187, 196].

To investigate whether MACC1 forms a homodimer, HEK293 cells were transiently co-transfected with two vectors: The first vector comprising MACC1 expressed with BRET acceptor (mCit) at the N-terminus (MACC1(N)) and the second vector containing MACC1 expressed with the BRET donor (NLuc) either at the N-terminus (MACC1(N')) or at the C-terminus (MACC1 (C')). Additionally, respective negative controls for each interaction pair were included. Using the 96-well transfection protocol, HEK293 cells were reverse co-transfected with MACC1(N)/ MACC1(N') or MACC1(N)/MACC1(C'). A fusion construct of donor and acceptor in a single vector received as a gift from AG Wanker (MDC) was employed as a positive control for the study. On Day 3, the fluorescence and luminescence intensities were measured. The net BRET signal/corrected BRET ratio was quantified as the ratio of the energy emitted by the acceptor to the energy emitted by the donor. This was then subtracted by the background luminescence. The cut-off value for positive interaction ≥0.01 was employed as determined by previous reports [175, 195].

Both the tested combinations, MACC1(N)/MACC1(N') and MACC1(N)/MACC1(C'), resulted in values above the threshold indicating a positive interaction. However, the value obtained by MACC1(N)/MACC1(N') was much higher compared to the interaction between MACC1(N)/ MACC1(C') as observed in Figure 14. The mean BRET signal obtained for the N-terminus interaction after subtraction of background luminescence was 0.06313 ± 0.0066 SEM for three independent experiments. On the other hand, the corrected BRET ratio for C-terminus interaction was 0.0181 ± 0.0053 SEM. The comparison of the N-domain:N-domain with N-domain:C-domain using an unpaired t-test resulted in a p-value of 0.0061, $p < 0.05$.

In order to understand the effect of various stimuli on the dimerization, two conditions were selected. In the first condition, the effect of HGF stimulation was evaluated along with untreated control. The values for the investigated MACC1(N)/MACC1(N') interaction for the untreated and treated were similar. No noticeable difference was observed with or without treatment with HGF.

In addition to positive stimuli, it was important to understand the effect of inhibitory signals on MACC1 oligomerization. To test this question, the impact of a SRC inhibitor (Bosutinib) on MACC1 dimerization was investigated using a similar assay setup. The HEK293 cells were co-transfected with MACC1 vectors and further systematically tested with SRC inhibitor Bosutinib 0.1 µM or DMSO (negative control). The average BRET values from three independent experiments for DMSO and Bosutinib were 0.049 and 0.037, respectively. As expected, the BRET signal obtained for DMSO (control) was not significantly different than the untreated state indicating no interference of DMSO with the dimerization. However, only a small reduction in the BRET ratio was noticed in cells treated with the SRC inhibitor Bosutinib (Figure 14). In line with these findings, the one-way ANOVA statistical analysis also showed no remarkable differences between the groups. These results hint that SRC inhibitors might not influence the MACC1 dimerization. Nevertheless, more detailed investigations with various SRC inhibitors over different time intervals are required to examine the link between SRC inhibition and MACC1 dimerization. A closer look at the results from the two conditions reveals that MACC1 dimers are present in living cells without treatment.

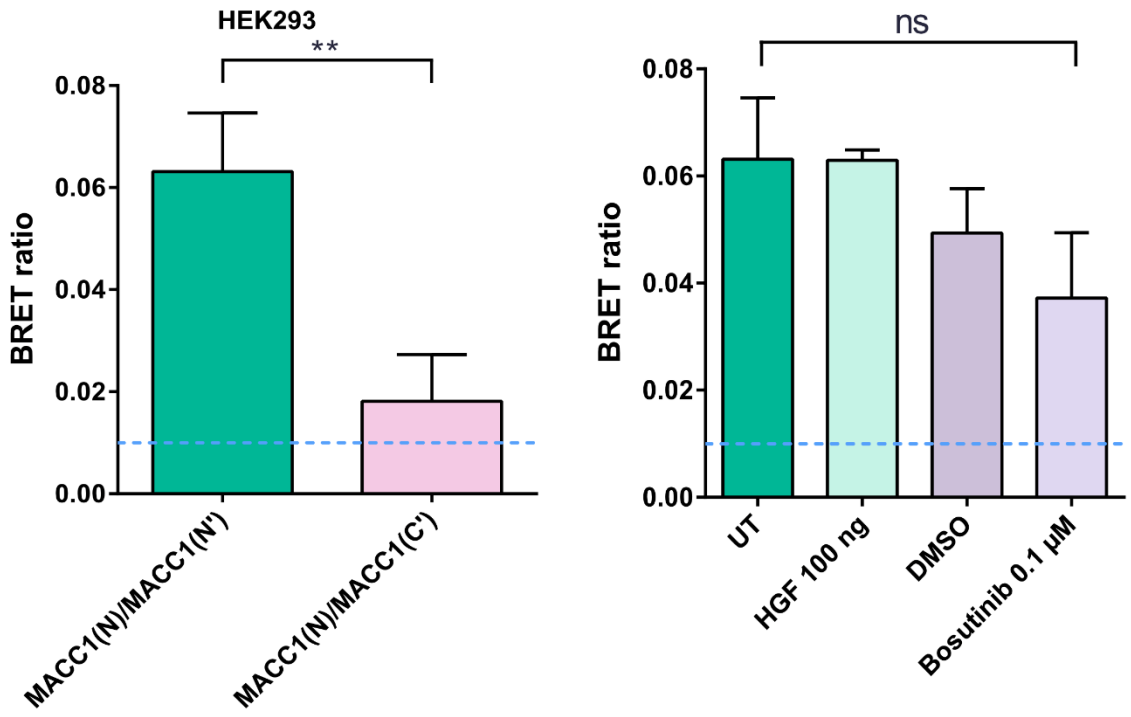

Figure 14: BRET detects MACC1 dimer in living cells

HEK293 cells were co-transfected with MACC1 fused with the BRET acceptor mCit [MACC1 (N)] and MACC1 fused with the BRET donor NLuc either at the N-terminus or C-terminus [MACC1 (N') or

MACC1 (C')]. The cut-off value of ≥0.01 for a positive interaction as established by Trepte P. et al [195]. was applied. The BRET ratio for interaction between two N-termini was 2-fold higher than the interaction between N- and C-terminus. The treatment with HGF or SRC inhibitor (Bosutinib) showed no significant changes on the MACC1 dimerization.

In the next step, the possibility of a direct interaction between MACC1 and SRC was investigated. The gateway entry vector containing the cDNA fragment of SRC was shuttled into an expression vector containing the BRET donor at the C-terminal. The reverse co-transfection protocol described previously was utilized to co-transfect MACC1(N) and SRC(C') vectors in HEK293 cells. The corrected BRET signal value for the interaction was below the threshold of 0.01 (data not shown) suggesting either no direct interaction or greater distance between the two protein tags [175, 195]. One possible reason for the lower BRET value for interaction between SRC and MACC1 could also be the orientation of the BRET tags [226]. Another likely reason that could interfere with the resonance energy transfer is the size of the two proteins, especially considering MACC1 exists as a dimer.

So far, this section focused on identifying the MACC1 dimerization and the influence of various external factors on the MACC1 dimerization and association of SRC with MACC1. In the following experiments, the influence of mutation of residues in ZU5 domain is studied.

The results from BRET experiments using MACC1 WT vectors demonstrated that the most efficient resonance energy transfer occurred between the two N-terminus-tagged MACC1 protein. Additionally, the most favorable condition for energy transfer between MACC1 fusion proteins persisted to be the untreated condition without any treatment. Therefore, this setup was taken ahead for further experiments with the mutant MACC1 in HEK293 and CRC cells HCT116. It was suspected that the mutation would dissociate or reduce the self-association of MACC1. To test this hypothesis, the vectors containing mutated MACC1 residues Val212, Ileu214, and Cys216 MACC1 (3xMut) were developed and shuttled into BRET expression vectors namely with the BRET acceptor (mCit) at the N-terminus forming 3x Mut(N) or with BRET donor (NLuc) at the N-terminus resulting in 3xMut(N').

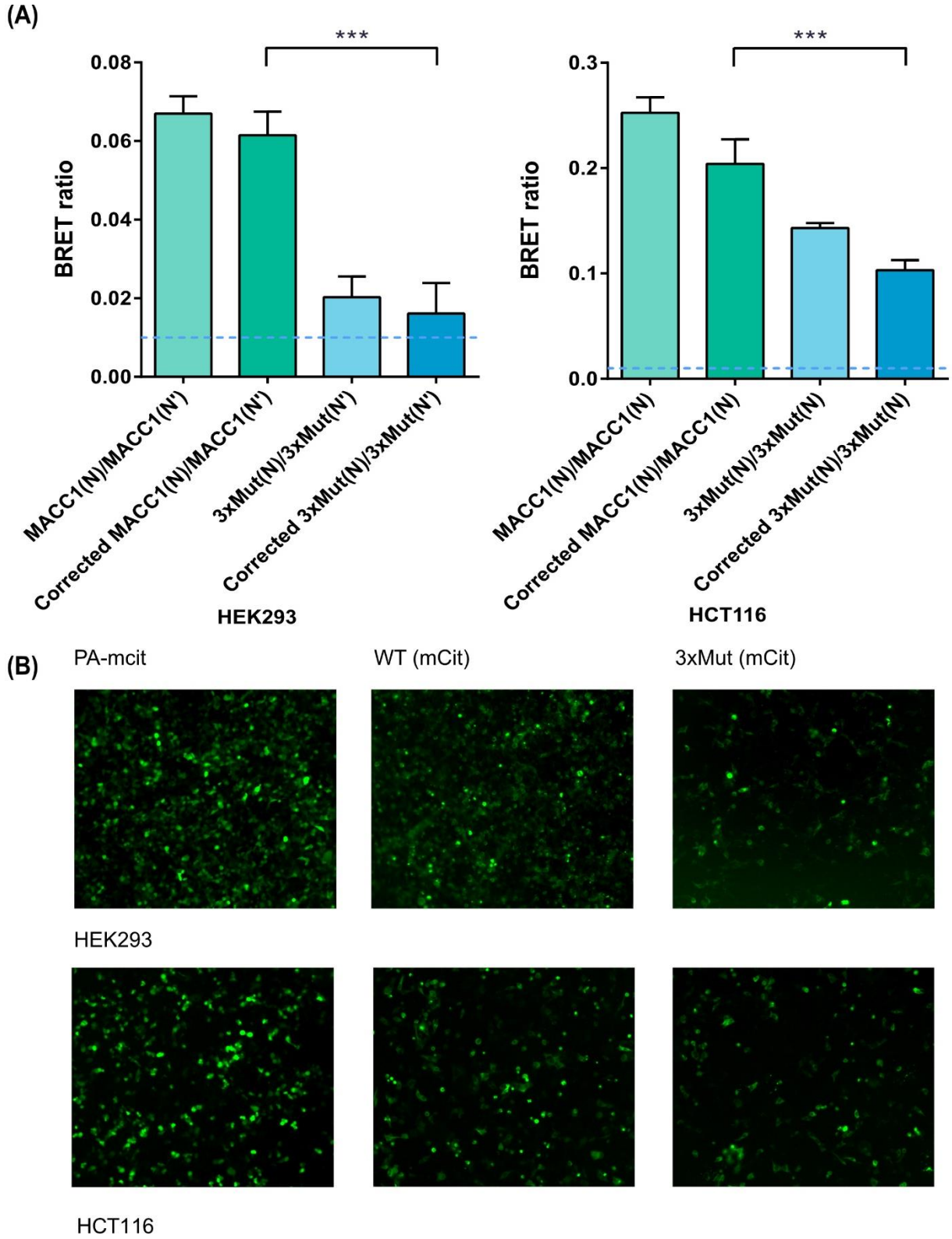

Figure 15: Studying the impact of mutation on MACC1 dimerization

(A) HEK293 and HCT116 cells were co-transfected either with MACC1 WT (MACC1) or MACC1 ZU5 mutant (3xMut) fused with BRET reporters. BRET measurements revealed a significant decrease in

the BRET ratio when the residues Val212, Ile214, and Cys216 were mutated to aspartic acid (Asp) indicating an obstruction in dimer formation in both cell lines. For the correction of BRET ratio, values of two corresponding negative controls were compared and the higher value was subtracted to obtain the corrected BRET (cBRET) ratio. Values above the threshold of 0.01 [195] were classified as positive. (B) Expression of MACC1 fused to BRET acceptor (mCit) which has similar characteristics as GFP was evaluated using a fluorescence microscope 72 h post-transfection in HEK293 and HCT116 cells.

To assess the effect of selected mutations on dimerization via BRET, 3xMut(N) and 3xMut(N') vectors were reverse co-transfected in HEK293 and HCT116 cells for BRET analysis and incubated for 72 h. For all the interactions tested, the cBRET values were above the defined threshold of 0.01. However, a remarkable reduction in BRET signal was observed in cells co-transfected with 3xMut MACC1 compared to MACC1 WT cells in both the cell line models as observed in Figure 15. In HEK293 cells, the mean cBRET ratio for MACC1 WT was 0.0615, and for 3xMut was 0.0161 obtained from three independent experiments. A remarkable mean difference of 0.0454 was observed between the two cell lines. In addition, statistical analysis using one-way ANOVA analysis resulted in a p-value of 0.0001, $p < 0.05$ indicating a significant difference between the WT and 3xMut MACC1 cells.

In CRC cells (HCT116) a similar trend was observed. Interestingly, values detected in HCT116 were much higher than those observed in HEK293 cells for both the WT and 3xMut cells. This is clearly demonstrated by the mean cBRET ratio values for MACC1 WT 0.2041 and 3xMut 0.1031. One possible explanation could be the role of MACC1 in CRC development. Although both values were above the threshold, strikingly the cBRET ratio for 3xMut decreased by 50% compared to the MACC1 WT group in HCT116 cells.

Together this section has addressed three key aspects of MACC1 self-association:

(1) MACC1 self-associates in living cells forming dimers

(2) MACC1 dimers are not easily affected by external stimuli, and

(3) The residues in the ZU5 domain of MACC1 are involved in the MACC1 dimerization.

In the subsequent sections, the effects of dimer-hindering mutations on MACC1 signaling and metastasis development are investigated in detail.

6.2.6. Effect of Mutations in the MACC1 ZU5 domain on MACC1 Signaling

In the previous section, findings established that MACC1 exists as dimeric proteins and mutation of the residues present at the dimer interface hinders self-association, but the consequences of the mutation on MACC1 function need to be investigated. Before proceeding to examine the functional consequences, it was necessary to study the impact of these mutations on MACC1 signaling.

The activation of downstream proteins in the HGF/c-Met signaling cascade was evaluated to gain insight into the potential effect of the dimer hindering mutation on the signaling cascade. As past studies have revealed that MACC1 overexpression causes higher ERK phosphorylation [48, 82, 103, 104], preliminary experiments were performed to determine if mutation of the suspected residues affected ERK activation.

Using growth factor treatment with HGF, the downstream activation of ERK was investigated at different time points in HCT116/V5 (Cont), HCT116/WT MACC1-V5 (MACC1 WT) and HCT116/Mutated MACC1-V5 (3xMut MACC1). Previous results from Zincke F. (2019) suggest that MACC1 overexpressing cells also have higher activation of SRC [82]. Therefore, the effect of dimer-hindering mutation was examined not only on ERK phosphorylation but also on the activation of SRC. For this purpose, serum-starved cells were treated with HGF 20 ng/mL for 0, 8, 30, and 90 min and cell lysates were analyzed using WB.

Phosphorylation of ERK was detected at Thr202 and Tyr204. Consistent with the previous findings, WT MACC1 induced activation of ERK without treatment [82, 103]. Interestingly, the mutant cell line (3xMut) also displayed ERK activation at all the investigated time points. The phosphorylation levels of WT MACC1 and 3xMut MACC1 were similar as observed in Figure 16. The highest level of ERK activation was observed at 8 min in WT MACC1 and 3xMut MACC1 group. Concomitantly, the level of ERK activation in WT MACC1 and 3xMut MACC1 were similar in untreated and 90 min post-treatment.

Activation of SRC on the other hand showed differences between the WT MACC1 and 3xMut MACC1. As suspected, the phosphorylation of SRC was much higher in WT MACC1 overexpressing cells compared to the control and mutant cell line post-treatment. This was highly evident at 8 min after HGF stimulation (mean difference = 1.331). Interestingly, the level of SRC phosphorylation was similar between WT MACC1 and 3xMut MACC1 in untreated conditions. The phosphorylation of SRC in the 3xMut MACC1 was less compared to WT MACC1 at longer

time intervals after 30 min and 90 min HGF treatment. This indicates that mutation of MACC1 residues 212, 214, and 216 interferes with sustained activation of SRC kinase in CRC cells.

In summary, this section demonstrated that the activation of ERK in the 3xMut was similar between the WT MACC1 and 3xMut MACC1. However, the activation of SRC varied between the groups with 3xMut MACC1 showing lower signal intensity and a longer time to reach peak activation compared to WT MACC1. These findings once more confirm the assumption that the residues in the ZU5 domain are critical for MACC1 signal transfer. In the upcoming section, the effects of the dimer hindering mutation on MACC1 function are evaluated by employing *in vitro* metastasis functional assays.

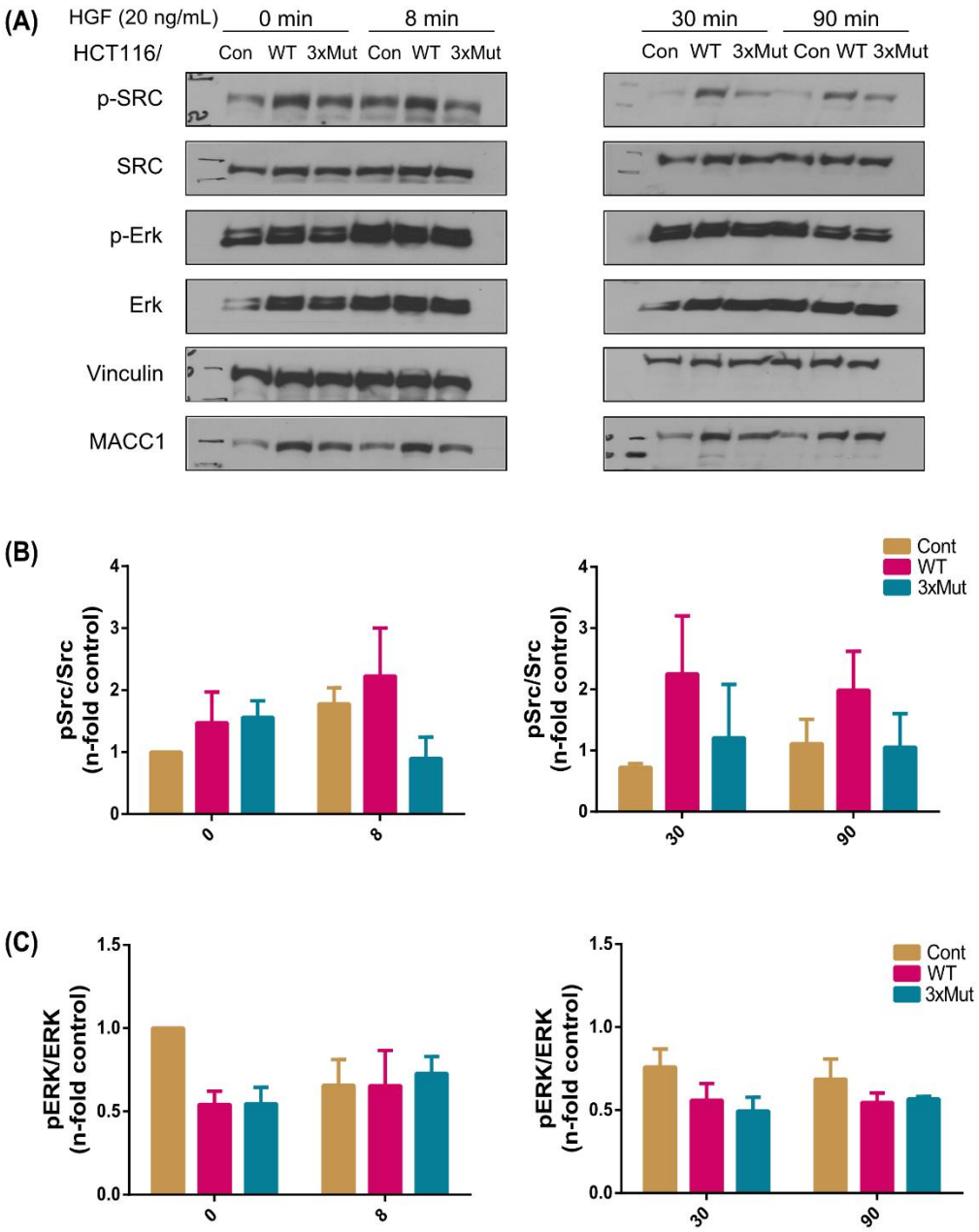

Figure 16: Mutation at the dimer interface affects the MACC1 signaling

HCT116 cells stably expressing empty vector (Cont), MACC1 (WT), and MACC1 vector containing mutations at Val212, Ile214, and Cys216 (3xMut) were seeded and subsequently, serum-starved overnight. On Day 3, cells were treated with HGF 20 ng/mL for 0, 8, 30, and 90 min, respectively. (A) Activation of downstream targets of RTK such as SRC and ERK through phosphorylation was evaluated from at least three independent biological replicates. Treatment with HGF enhanced the activation of downstream targets, especially SRC. While the WT and 3xMut cells exhibited similar pSRC levels in the absence of treatment, the difference in pSRC between WT and 3xMut cells was observed on treatment. (B) The WB results were analyzed by spot densitometry using ImageJ (version

1.53). The expression of total and pSRC was first normalized with the loading control (vinculin) and in the next step with pSRC was normalized with total SRC. Values for control cells (Cont) were set to 1 thus resulting in relative expression values for each time point. A similar analysis was carried out for ERK (C) as well; however, no difference was observed. The WT cells showed higher phosphorylation of SRC in all the treated conditions with a peak at 8 min. On the other hand, 3xMut cells exhibited a reduced pSRC compared to WT cells. In addition to exhibiting early response, WT cells also displayed sustained activation of pSRC until 90 min.

6.2.7. MS reveals Differences in Protein expression in 3xMut MACC1 cells compared to WT MACC1 Cells

Global protein and phosphoprotein analysis were carried out to determine the differential expressed proteins in WT and 3xMut. As the data-independent acquisition (DIA)-based mass spectrometry (MS) strategy offers higher selectivity, sensitivity, and proteome coverage compared to classical MS approaches [90], it was chosen for profiling the differential expressed proteome of WT and 3xMut cells. The global protein profiling was performed on the Exploris 480 mass spectrometer by the MDC proteomic facility (O. Popp/P. Mertins).

The proteins and phosphoproteins downregulated in the MACC1 3xMut compared to WT MACC1 were determined. To understand the impact of these proteins on the biological function of MACC1, a gene enrichment analysis was carried out using Metascape [200] and AmiGO2 [201-203]. The analysis revealed the enrichment of three main processes (1) Actin filament-based process (Figure 18) (2) Intracellular signal transduction Figure 19 (3) Regulation of the actin filament-based process. The intracellular signal transduction process contained the highest number of proteins found in the MACC1 phosphoprotein and proteome dataset. Other clusters in which the downregulated genes were enriched included cell-cell junction organization, apoptotic cleavage of cellular proteins, EGFR signaling, VEGFA-VEGFR2 signaling, and Ras protein signal transduction.

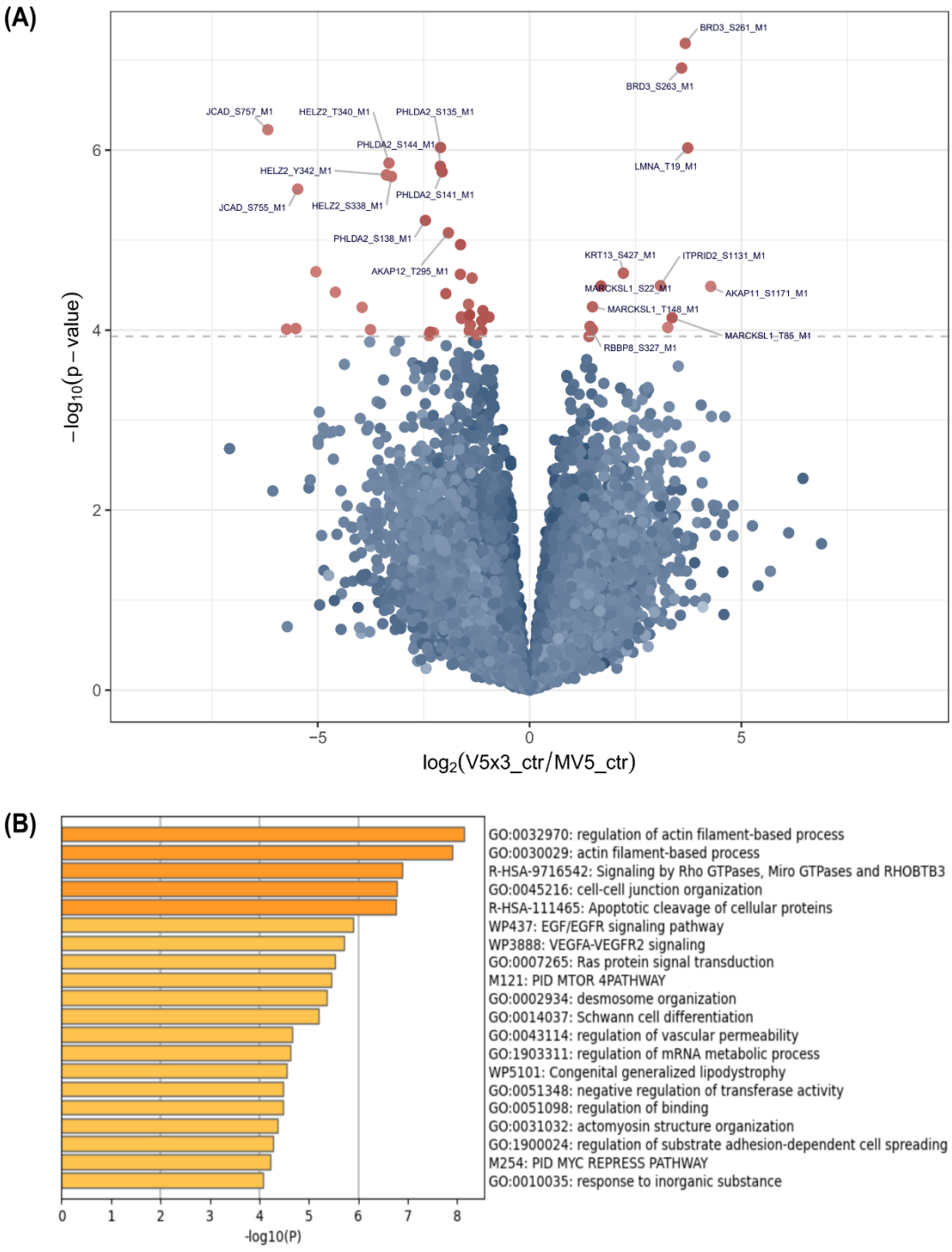

Figure 17: Differentially expressed phosphoprotein in 3xMut compared with WT MACC1 sample

(A) Volcano plot illustrating the effect of 3xMut on the phosphoprotein expression. Significant genes are highlighted in red. (B) Bar graph depicting the enriched biological process as retrieved from Metascape [200].

The commonly downregulated genes in 3xMut cells were epidermal growth factor receptor pathway substrate 8 (EPS8), FH1/FH2 domain-containing protein 1 (FHOD1), Plakophilin-2 (PKP2), Mothers against decapentaplegic homolog 3 (SMAD3), Tight junction protein ZO-1 (TJP1) and TRAF2 and NCK-interacting protein kinase (TNIK). Interestingly, proteins involved in filament networks such as Plectin (PLEC), LIM domain and actin-binding protein 1 (LIMA1) responsible for actin cytoskeleton regulation and dynamics [229-231] were also found to be downregulated in 3xMut MACC1 cells.

Another key cytoskeleton protein known as SRC substrate cortactin (CTTN) linked to metastasis and cell migration [232, 233] was also reduced in 3xMut cells. BRAF, a prominent oncogene responsible for tumor progression [6, 56, 57], was also downregulated in cells containing dimer-hindering mutations. Further, the heatmap of significantly downregulated phosphoproteins in 3xMut MACC1 showed differently regulated BRAF phosphorylation sites S363-S365, and T373 which are involved in actin filament-based process. Further, BRAF phosphorylation sites S394, T395, T396, S399, and T401 which are involved in the regulation of actin filament-based process were also differently regulated in 3xMut when compared to WT MACC1.

In summary, the global protein and phosphoprotein analysis showed downregulation of proteins involved in the actin filament-based process, intracellular signaling, and regulation of actin filament-based process in 3xMut MACC1 cells compared to WT MACC1 indicating an important role of the ZU5 residues of MACC1 in regulating these processes.

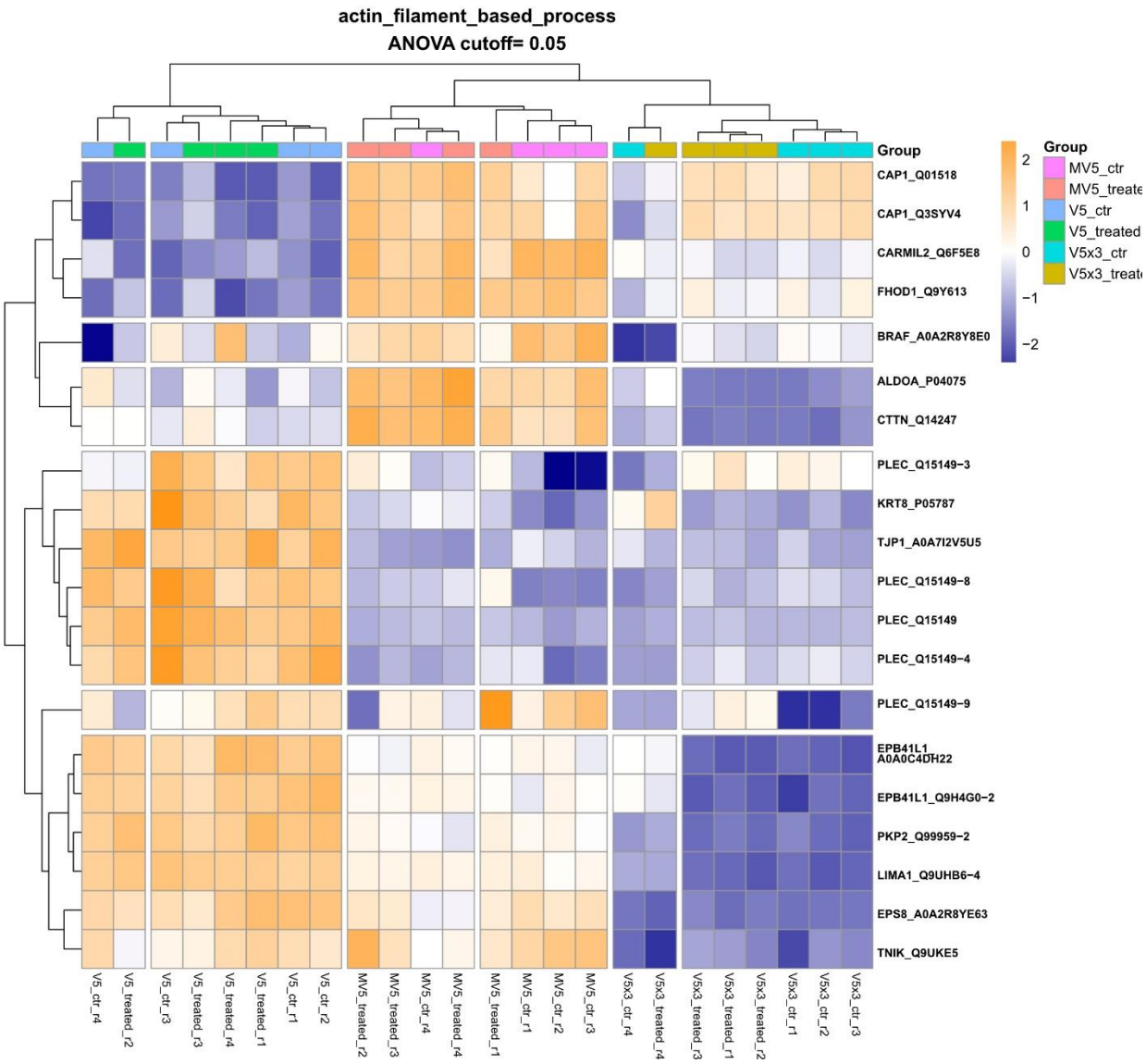

Figure 18: Heatmap illustrating the differential expressed proteins in WT and 3xMut MACC1 involved in the actin filament-based process

The heatmap of the differentially regulated proteins in the 3xMut MACC1 sample compared to the WT MACC1 shows differences in the expression of proteins such as PLEC and BRAF among others which are commonly enriched in the actin filament-based process. BRAF: v-RAF murine sarcoma viral oncogene homolog B; PLEC: Plectin.

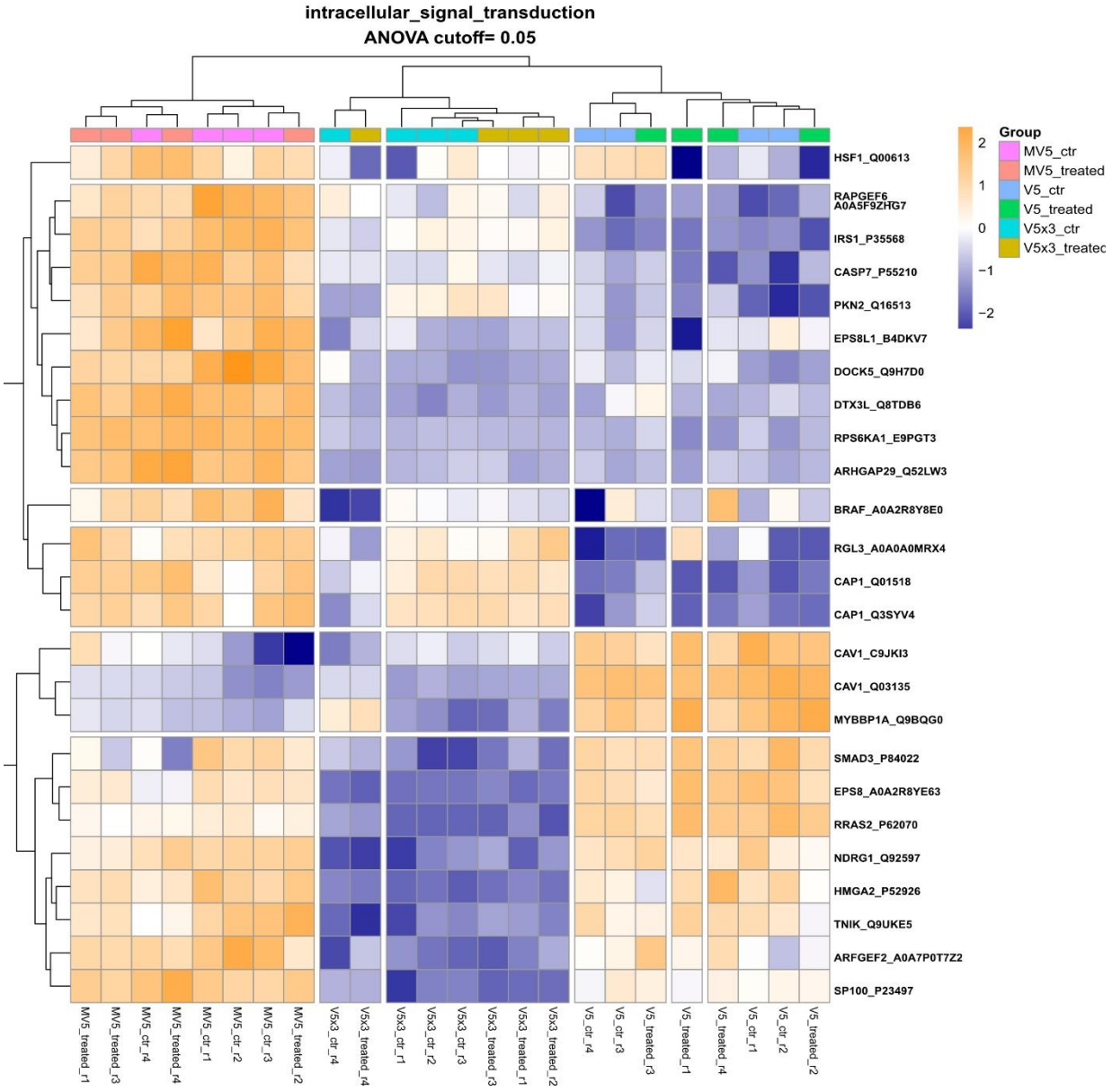

Figure 19: The differential expressed proteins involved in the intracellular signaling process

The heatmap of the differential regulated proteins in the 3xMut MACC1 sample compared to the WT MACC1 shows differences in the expression of proteins such as TNIK and SMAD3 among others which are commonly involved in intracellular signal transduction. SMAD3: Mothers against decapentaplegic homolog 3; TNIK: TRAF2 and NCK-interacting protein kinase.

6.2.8. MACC1 cells with Dimer Hindering Mutation show Reduced Proliferation

The above section analyzes the impact of the mutation on downstream signaling proteins, in this section the effect of diminished SRC activation on cellular functions is examined. Past studies clearly demonstrate a pivotal role of SRC signaling in proliferation, invasion, and migration, thus contributing to the development of metastasis in CRC [125, 129, 133]. In addition to SRC activation, downstream effectors of MAPK signaling are also highly activated in metastasis progression [116, 234]. For instance, enhanced ERK activation promotes the entry of cells to the S phase of the cell cycle in the presence of growth factors [84]. Removal of critical growth factors or serum from the medium has also been shown to block the G1/S phase transition and result in cell-cycle arrest [235].

In this study, the effect of MACC1 mutation on proliferation was evaluated using a live-cell imaging system (IncuCyte®). HCT116/WT and HCT116/3xMut MACC1 cells overexpressing WT MACC1 and expressing mutated MACC1, respectively were seeded in a 96-well plate and allowed to attach. On the next day, the cells were either left untreated or treated with HGF 20 ng/mL. The cells were observed over a period of 5 days with images captured every 2 h for each well. The image capture was stopped after 130 h and the data were analyzed. In general, WT MACC1 cells showed enhanced proliferation compared to 3xMut MACC1 (Figure 20). The mean value for WT MACC1 was 48.35 ± 2.877 SEM and 3xMut MACC1 was 33.90 ± 2.453 SEM. Further statistical analysis employing an unpaired t-test revealed a p value of 0.0002, $p < 0.05$ for the 66 time points. It was interesting that the difference between the means of WT MACC1 and 3xMut MACC1 was smaller in the HGF treated group compared to the absence of HGF (12.90 ± 3.943 SEM vs. 14.45 ± 2.877 SEM). These findings suggest that the ZU5 domain residues of MACC1 are integral for MACC1 function in CRC. In the following section, therefore, other metastasis characteristics of cancer cells such as migration, and colony formation are investigated using the corresponding *in vitro* assays.

The colony formation assay was used to study the ability of WT MACC1 and 3xMut MACC1 on distant organ seeding. As only a few cells survive after EMT, thus the ability of cells to form colonies from single cells is an essential step post-extravasation [1, 82]. To test this in a 2D model, V5 only (Cont), WT MACC1 and 3xMut MACC1 cells were seeded at a very low density of about 200 cells/mL. The cells were monitored for 7 days and then the cells were stained with crystal violet in formaldehyde solution.

Colonies were observed in all wells after 7 days. A representative example and average colony area of three independent experiments is shown in Figure 20. All values were normalized to the V5-only vector (Cont). An increase in colony formation was observed in MACC1 overexpressing cells compared to V5-only vector [82]. Consistent with the previous observation, WT MACC1 showed significantly higher colonies than the control V5 empty vector in this study. However, no significant differences were observed between the WT MACC1 and 3xMut MACC1 when analyzed using the ImageJ plugin ColonyArea [199]. It is noteworthy that the size and shape of the colonies were distinct in WT MACC1 compared to 3xMut MACC1 cells.

Altogether, the results demonstrate that the ZU5 residues of MACC1 play a critical role in cell proliferation in CRC cells but do not significantly affect MACC1-mediated colony formation. Nonetheless, there are morphological differences in the colonies formed by WT MACC1 and 3xMut MACC1.

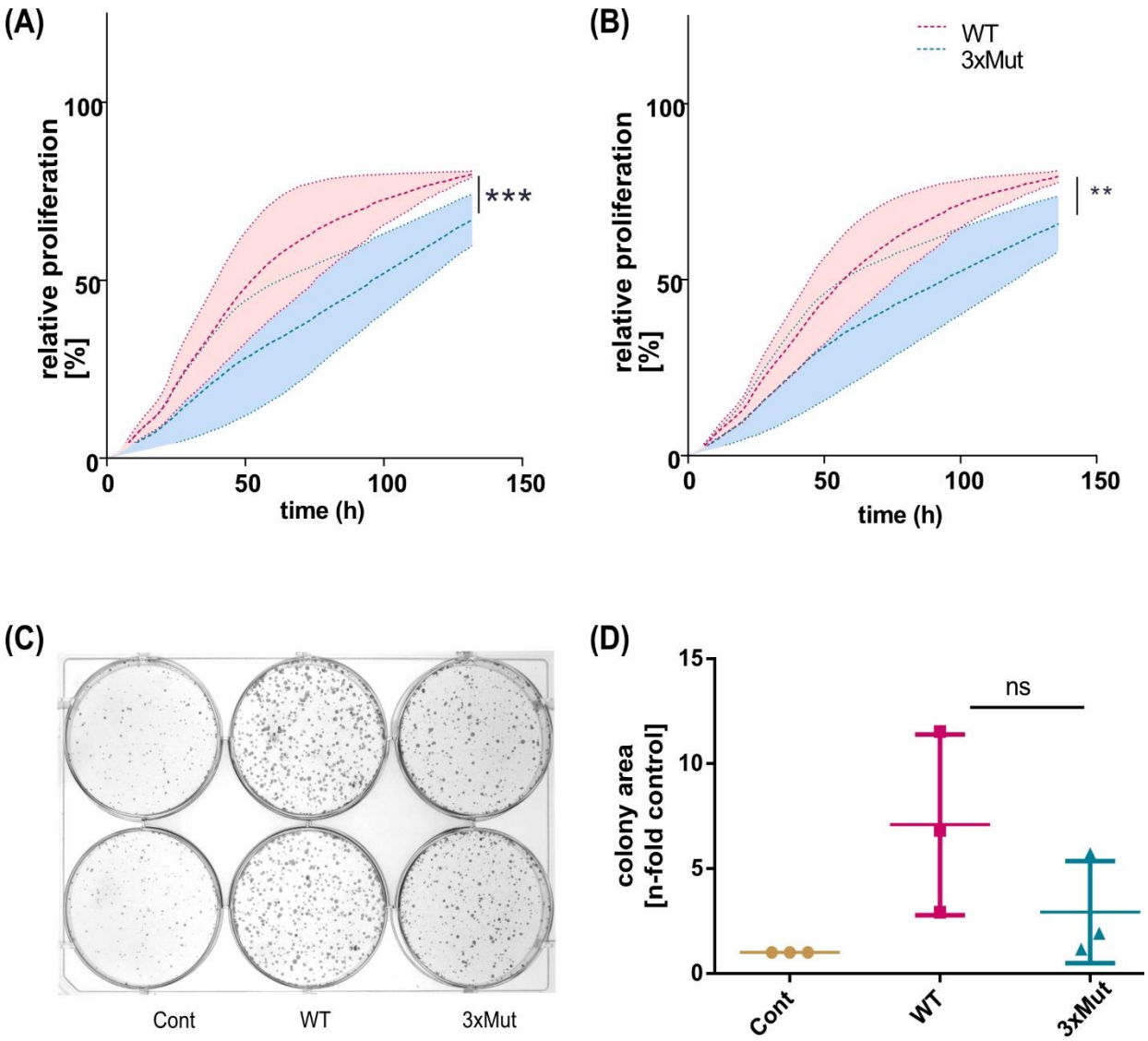

Figure 20: Mutations at the dimer interface reduce MACC1-mediated proliferation

Using the IncuCyte® ZOOM system, the impact of mutation on MACC1-mediated proliferation was analyzed. HCT116 cells expressing MACC1 (WT) and mutated MACC1 (3xMut) were seeded in a 96-well plate and imaged every 2 h for a period of 130 h. Data is shown as mean ± SEM for three independent experiments. (A) The relative proliferation (%) of MACC1 WT cells (red) and 3xMut cells (blue) without any treatment displays a higher proliferation in WT cells compared to 3xMut cells. (B) The right-side graph represents the relative proliferation in the presence of treatment with HGF 20 ng/mL, as observed treatment with HGF improves proliferation in 3xMut cells but does not rescue it completely. (C) The clonogenic assay was used to study the effect of dimer-hindering mutation on MACC1's ability to form colonies. The representative image of the colonies from one single experiment displaying colonies formed by Control (Cont), MACC1 (WT), and MACC1 mutant (3xMut) cells. A difference is observed in the morphology and size of the colonies between the WT and 3xMut cells (D) Results are represented as a means of three independent experiments normalized to values from the control group.

6.2.9. Integral Role of ZU5 domain residues of MACC1 in Cell Migration

Considering the diminished activation of SRC in the mutant MACC1 cells, the effect of the mutation on the migration ability of MACC1 was investigated in the CRC cell line. The wound healing assay, also known as scratch assay, is a conventional two-dimensional *in vitro* method to evaluate the effect of mutations or treatment on the migratory ability of the cells [236]. For the scratch assay, cells are cultivated to form a confluent monolayer and then scratched mechanically to create a gap (wound). The images are captured at regular intervals to monitor the closure of the wound specifically until the cell-cell contacts are established again [103, 236].

To investigate whether the mutation of residues Val212, Ile214, and Cys216 influences the migratory ability of MACC1, wound healing assays were performed in the presence and absence of HGF. Confluent monolayers of WT MACC1 and 3xMut MACC1 cells were cultivated, then cells were starved for 6 h and scratched. After scratching, the medium was replaced with medium containing reduced serum in the presence or absence of HGF. Migratory cells were observed in both conditions, untreated and HGF-treated. During the first 24 hours following the scratch, cells close to the wound edge in WT MACC1 already displayed migratory characteristics. In contrast, cells in 3xMut MACC1 still retained cell-cell contacts. Overall, 3xMut MACC1 cells showed a reduced migration compared to WT MACC1 when evaluated over a period of 108 h. On Day 4, the wound gap was significantly higher in 3xMut MACC1 compared to WT MACC1 cells (Figure 21). Statistical comparison using Mann-Whitney test results between the WT MACC1 and 3xMut MACC1 cells resulted in a p-value of $p < 0.0001$. The complete wound closure did not occur in 3xMut MACC1 cells in the untreated condition. These findings indicate an integral role of MACC1 ZU5 domain residues in cell migration.

From the images, it was apparent that in the presence of HGF higher wound confluence was observed in both groups. Statistical analysis comparing the WT MACC1 with 3xMut MACC1 reveals a p-value of $p < 0.0001$ using the Mann-Whitney test. In addition to being a ligand for the c-Met receptor, HGF is also known to induce branching morphogenesis called cell scattering [48, 237]. Cell scattering is a result of actin cytoskeletal rearrangement and is the driving force for cell motility [48, 237, 238]. Treatment with HGF promoted a more pronounced migratory phenotype in cells expressing WT MACC1 but not in 3xMut MACC1.

Taken together, these data demonstrate that residues in the ZU5 domain are crucial for MACC1 migratory characteristics. Furthermore, the ability to respond to growth factor treatment such as HGF is diminished when these residues are mutated. In summary, the results obtained from

various metastasis assays under different conditions highlight the importance of residues Val212, Ile214, and Cys216 in the function of MACC1.

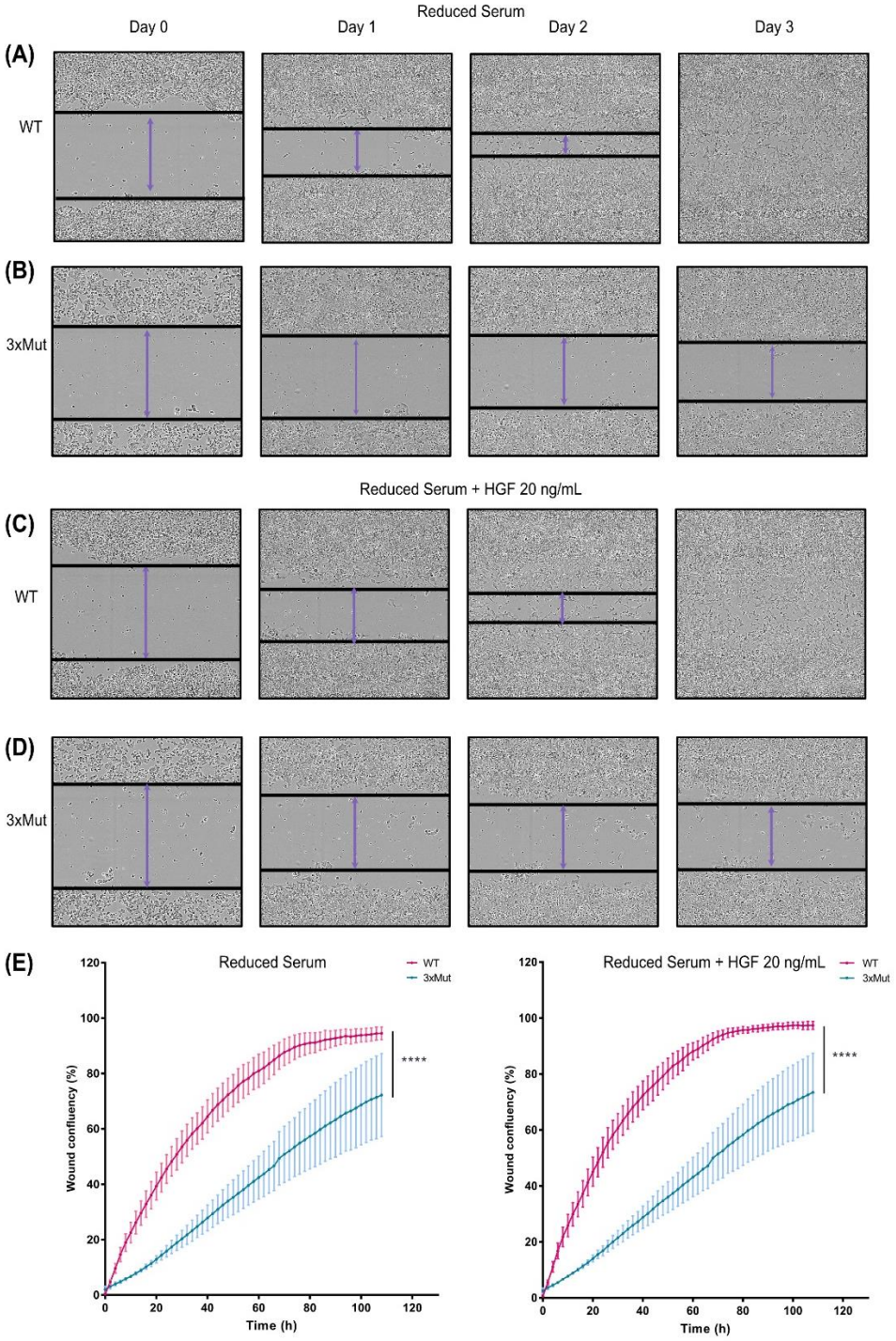

Figure 21: Dimer hindering mutations restrict MACC1-induced migration

HCT116 cells expressing MACC1 (WT) and MACC1 mutated at Val212, Ile214, and Cys216 (3xMut) were seeded in the IncuCyte® ImageLock plate ($1x10^5$ cells per well). After the cells were settled and

starved, a scratch was created using WoundMaker, and wound closure was observed until Day 3. The WT cells (A) displayed migratory properties soon after the wound formation in comparison to 3xMut cells (B). Complete wound closure was observed in WT cells irrespective of treatment (C), whereas a clear gap was still observed at the end of Day 3 in 3xMut cells (D). (E) Data was analyzed using the Mann-Whitney test WT vs. 3xMut $p < 0.0001$.

7 Discussion

7.1 MACC1: More than a Metastasis Biomarker

CRC, one of the most preventable cancers when detected early, continues to be the second leading cause of cancer deaths worldwide [44]. At present, the number of CRC cases is rising not only in Western countries but globally [10, 55, 64, 239]. While the survival rates are above 90% for early-stage localized disease, the value drops to 14% for patients with later-stage metastasis [240]. Unfortunately, around 20% of cases already present distant metastasis at the time of diagnosis [6]. Even with the advancements in the field of precision medicine, treatment options for metastatic CRC are still limited [58, 71]. Most of the known targets to date have either been reported as undruggable or developed resistance gradually [58]. Therefore, early detection will not only substantially improve the quality of life but also significantly reduce mortality and economic burden [64, 68].

MACC1, a well-established prognostic and predictive biomarker, has been shown to have a causal relationship with tumor progression and metastasis formation in CRC [48, 80]. Most importantly, the poor prognosis associated with high MACC1 expression has been consistently observed across tumor entities in independent studies [73, 89, 207, 241]. A recent study by Kobelt et al. (2021) identifies a hierarchical mode of phosphorylation of tyrosine sites at the C-terminus of MACC1 by MEK1 which is suggested to be essential for MACC1 function [103]. However, the tyrosine site close to the N-terminus and the kinase responsible for phosphorylating this tyrosine site is only partially addressed [82].

In this part of the study employing computational tools, the tyrosine site Y379 located close to the N-terminus of MACC1 and SRC kinase capable of phosphorylating this site is investigated. Further, the spatial localization of MACC1 and SRC is detected in the metastatic CRC cell model and the impact of the mutation on MACC1 signaling is investigated. However, additional work is still required to identify the precise mechanism of regulation between SRC and MACC1 to translate these findings to the clinic. With the advancements in AI-driven predictions, more innovative solutions could be utilized to address this knowledge gap.

Protein phosphorylation is an important reversible post-translational modification which in combination with phosphatases is responsible for regulating diverse cellular processes [208]. Using NetPhos 3.1 [141], a phosphorylation tyrosine prediction tool, 6 tyrosine sites of MACC1 capable of phosphorylation were identified. Among these 6 identified sites, the sites close to the C-terminus and a combination of the tyrosine site close to the N-terminus with the C-terminus tyrosine site were already analyzed in another study [82, 103]. Therefore, for this study, the impact of individual mutation on the Y379 close to the N-terminus was undertaken. To investigate the tyrosine residues kinases capable of phosphorylating the Y379 MACC1 site, the prediction tool GPS 5.0 which is based on the amino acid substitution matrix was used [138]. GPS identified the ABL2, JAK3, FGFR, and SRC as putative kinases capable of phosphorylation of the Y379 site of MACC1.

A previous report by Kobelt et al. (2021) examines the MACC1 kinome using MS and identifies MEK1 as the kinase within the MACC1 kinome able to phosphorylate MACC1 [103]. In comparison, in this report, two computational tools were employed to dissect the kinase responsible for MACC1 phosphorylation. In addition to finding MEK kinases, the computational screen identified additionally seven other tyrosine kinases, that are capable of phosphorylating MACC1. Previous work by Zincke F. (2019) demonstrates a stronger activation of SRC in MACC1 overexpressing CRC cells. This in turn facilitates the FAK/SRC signaling resulting in metastasis phenotype [82]. Based on the findings from this study, SRC was selected for further analysis.

The present data show a positive correlation between MACC1 and SRC in TCGA [30] and GEO Omnibus [212] cancer patient cohort samples. The finding that MACC1 and SRC correlate positively in various cancer entities raises the hypothesis that MACC1 may be involved in the SRC signaling landscape. Consistently, a positive correlation has also been found between MACC1 and FAK. This is in line with the previous study indicating a worse prognosis for patients with high MACC1 and high FAK expression in hepatocellular carcinoma [205]. Interestingly, the patient cohorts investigated in this study also show an inverse correlation between MACC1 and CSK which is the negative regulator of SRC. Together, this suggests an important link between MACC1 and SRC in tumor progression.

Studies have reported higher expression levels of SRC in CRC metastases cells compared to poorly metastatic CRC cells [125]. Interestingly, the upregulation of SRC is accompanied by higher expressions of RTK such as EGFR, HER2, and c-Met [135]. The downregulation of c-Met has a direct impact on SRC activity [118]. Initial reports on MACC1 established that c-Met is transcriptionally regulated by MACC1 [48]. Through the c-Met/Akt axis, MACC1 induces

migration, proliferation, and vessel development in gastric cancer [48, 73, 86]. Potentially, MACC1 regulates SRC activity through c-Met expression and induces cell migration.

Based on this evidence, the interaction of MACC1 with SRC using Co-IP assay in SW620 cells was examined. A positive interaction between MACC1 and SRC confirmed that MACC1 was involved in SRC kinase signaling in metastatic cancer cells. The most probable mode of interaction might occur between the SH2 domain of SRC and the pY-site on MACC1. This mode of binding to MACC1 has been observed previously. For instance, a report examining the phospho-interactome of MACC1 by Küster and Zincke (unpublished) demonstrates that SHP2, PLCG1, and GRB2 interact with phosphorylated Y379 peptide in an *in vitro* pull-down. The author shows that Y379 residue of MACC1 contains potential binding motifs capable of binding to SH2 domain of PLCG1 or GRB2 [82]. Besides the pY-sites, MACC1 also contains proline-rich motifs. Taking this into account, another possibility could be an interaction between the SH3 domain of SRC with the proline-rich motif of MACC1. A study shows that SH3 domain of SRC interacts with cytoskeleton proteins such as paxillin and may be responsible for subcellular localization of proteins to the cellular membrane [242].

Despite its wide applicability, Co-IP is incapable of characterizing direct binders and indirect interaction partners and subcellular localization of the interaction [182]. It is also possible that MACC1 might bind to SRC indirectly or as a part of a multi-protein complex instead of direct binary interaction. Thus, a detailed analysis of the binding mode and stoichiometry of the interaction between MACC1 and SRC is necessary to better understand the interaction model and how one regulates the other. Another critical limitation of Co-IP is that some transient interactions may be lost due to repeated washing [181]. Taking this into consideration, it is possible that some interactions that are part of a MACC1-SRC co-complex may be missed. Therefore, employing other sensitive methods as described in the later sections provides additional benefits to identify transient direct interactions that might link MACC1 to SRC.

The Co-IP results showed that MACC1 interacted with SRC in metastatic colon cancer cells - SW620. The presence or absence of growth factor treatment showed differences in interaction of SRC and MACC1. As phosphorylation events are highly regulated and dynamic [138], it would be beneficial to include other timepoints for further characterization. Nevertheless, phosphorylation without treatment suggests a pre-activation state even in absence of growth factor treatment. This is consistent with the past studies and data from this project examining MAPK pathway at various time intervals [82, 103, 104]. A ligand-independent activated state is also frequently observed in the RTK protein family. Ligand-independent activation has shown to be the cornerstone of tumor

progression and responsible for various tumor attributes ranging from early intense activation loop to drug resistance [109].

Unlike the common oncogenes such as BRAF and KRAS that are consistently mutated in CRC progression [51, 54], MACC1 activation by mutation is a seldom event (Figure 22). Therefore, suggesting that the activation of MACC1 is most often caused due to overexpression. It is important to note that MACC1 is also expressed in normal tissue, however, the MACC1 expression is very low and tightly regulated thus preventing oncogenic transformation [48, 72, 79]. Therefore, suggesting a mechanism of regulation that is disrupted and leads to oncogenic transformation. A study by Kim et al. (2018) indicates that MACC1 expression is regulated by deleted in breast cancer (DBC1) protein in association with β-catenin in colonosphere cells [97]. Previous studies have shown that mutation of SRC is also a rare event. It is possible that MACC1 and SRC are regulated via similar mechanisms leading to overexpression of both proteins in tumor tissue. Further studies, investigating MACC1 regulators and how they are dysregulated in metastasis, would offer valuable insights for development of future therapeutic strategies [136].

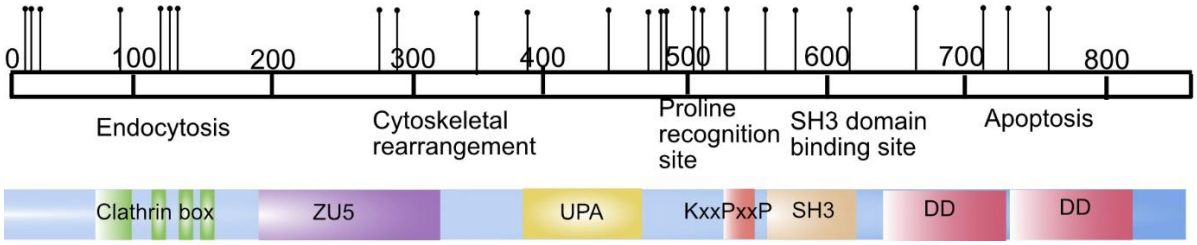

Figure 22: Schematic representation of the mutation landscape of MACC1

Overview of the MACC1 mutation as retrieved from cBioPortal. The lollipop graph shows the sites of missense and truncating mutation in MACC1.

The initial screen indicated an integral association between MACC1 and SRC. To test the impact of the Y379 mutation on MACC1 signaling downstream intermediates of the MAPK signaling pathway were analyzed in CRC HCT116/WT MACC1 and HCT116/Y379F mutant MACC1 cells. The MAPK signaling cascade and in turn the ERK activation are highly regulated by various factors such as different phosphatases, negative feedback loops, signal duration, and intensity and scaffold proteins among others. Considering the past literature and the influence of MACC1 on the c-Met signaling [48, 80, 82, 243], treatment with c-Met ligand, HGF, was applied for different time intervals and phosphorylation levels of SRC and ERK were evaluated.

In agreement with the previous reports, a higher activation of ERK was observed in untreated MACC1-GFP cells compared to Control cells. This indicates a pre-active state in MACC1 overexpressing cells even in the absence of ligand. This is consistent with previous reports from our group [82, 103]. This ligand-independent activation has been observed in RTKs and other oncogenes. Contrary to previous reports demonstrating a higher activation in MACC-GFP compared to control cells, in this study, only a minimal hyperactivation of ERK and SRC was observed in MACC1 overexpressing cells compared to control cells [103]. A possible explanation could be the loss of sensitivity to growth factor treatment over a period or due to batch changes in HGF or serum content.

With the help of imaging studies, colocalization between MACC1 and SRC was detected in the metastasis cell line SW620. Conventionally, under serum-starved condition, SRC is located perinuclearly [217] but in this experiment with the metastatic cell line SW620, SRC was already found at the plasma membrane. On the other hand, on treatment with HGF, SRC relocated from the periphery to the cytoplasm. The finding is unexpected, and this pattern of localization could be attributed to the presence of SFK in different pools as established by Veracini et al. (2005) in response to PDGF [244]. The authors suggest that the presence of distinct pools facilitates SFK to regulate diverse events like mitogenesis through STAT3 phosphorylation, *c-myc* expression, and cytoskeletal rearrangement associated with dorsal ruffle formation [244]. Thus, the presence of SFK in different pools facilitates specific cellular responses such as mitogenesis or actin assembly. Another unrelated study by Imbastari et al. (2019) examining MACC1's cellular localization, reveals that MACC1 is capable of changing localization depending on the stimulus [104]. This raises the possibility that MACC1 might exist in several distinct pools to activate specific signaling cascades. Activation of the c-Met/HGF pathway renders MACC1 translocation to the nucleus and localization of SRC into the cytoplasmic compartment. Considering that MACC1 is present at the periphery of the cell, it is possible that MACC1 is involved in the regulation of actin-cytoskeleton in addition to RTK signaling. In contrast to other studies, only the total SRC expression was investigated in this colocalization experiment. Evaluating the spatial distribution of MACC1 with the phosphorylated form of SRC (active) would provide additional information to reconstruct a precise picture of MACC1's involvement in SRC signaling.

Donepudi and colleagues show that SRC moves together with EGFR through the endocytic pathway resulting in prolonged receptor activation even in the absence of a ligand [245]. In the most recent model of MACC1's contribution to receptor endocytosis, our group reveals that MACC1 overexpression increases the rate of EGFR recycling and results in a stronger activation

loop [104]. The author indicates an integral role of MACC1 in regulating the balance between EGFR degradation and recycling. In accordance with this data, it is tempting to speculate a mechanism of synergism between MACC1 and SRC to regulate EGFR recycling and stronger activation [82, 104].

7.2 Is MACC1 Dimerization the Missing Piece of the Puzzle?

Overexpression of MACC1 has been linked to an increase in tumor malignancy [72, 73, 76, 78, 89, 96, 198, 206, 207, 246]. Although MACC1 has been largely implicated in metastasis progression, the structure of MACC1 is partially understood [48, 73, 94, 103]. Here, using novel protein prediction platforms the structural properties of MACC1 are closely reviewed [144, 163].

The established model for MACC1 function is based on MACC1 existing as a monomer. Until the present work, however, neither the dimeric state nor the possibility of functional regulation through dimerization has been questioned. The purpose of this study is to evaluate whether MACC1 dimerizes and the role of the dimeric state in tumor metastasis. In contrast to earlier studies that focused mainly on evaluating MACC1 expression and inhibition strategies [88, 93], this study addresses the gaps in the understanding of MACC1 structural features and how they contribute to MACC1 function.

This project utilizes different approaches to identify the possibility of MACC1 dimerization and the region involved in dimerization. Two key approaches help to address this question and form the foundation of this study. While the first approach is classical and based on resonance energy transfer to study PPI in living cells [169, 192, 195], the second one is a novel prediction tool that employs AI to gain insight into the multimeric protein complex [144, 163].

Employing the combination of signaling and metastasis assays, this study evaluates the link between sites involved in MACC1 dimerization and MACC1 function. In summary, this project exploits the possibility of MACC1 oligomerization, residues important for MACC1 dimerization, the role of MACC1 dimerization in the regulation of metastasis signaling, and finally consequences of MACC1 dimer hindering mutation on metastasis.

7.2.1. Highly Accurate AlphaFold Predicts MACC1 Structure

Numerous studies confirm the direct correlation of MACC1 with metastasis in various cancer entities [73, 78, 90, 96, 198, 206, 207, 246]. However, a complete understanding of the structural characteristics of MACC1 is still missing. As a first step, a highly accurate protein prediction tool known as AlphaFold2 [144] was employed to understand the structural composition of MACC1. Colab, a Jupyter Notebook developed based on AlphaFold2 with modified environment, was utilized to predict the structure of MACC1 [144, 154].

The global error prediction for the MACC1 structure shows that the individual domain of MACC1 is distinct and nicely folded with high values of confidence. However, some level of uncertainty is observed for interdomain regions and N-terminal regions. In addition, the overlapping pIDDT curve and nearly identical differences in residue distance of the five models indicate that the predictions are consistent. As anticipated, the N-terminus of MACC1 exhibits an unstructured region of up to 160 residues suggesting the presence of an intrinsically disordered region. Extensive evidence suggests an integral role of intrinsically disordered regions in regulating diverse cellular functions especially in cancers [247].

The ZU5 and UPA domains of MACC1 expand around 550 residues and displays a β-sandwich formed by the antiparallel sheets of the β-strand. Despite lower protein sequence homology [94], the ZU5 domain of MACC1 exhibits similar structural features to other ZU5 proteins. Consistently, the ZU5 domain and UPA domain of MACC1 also interact through the loop regions, and only loose interactions are observed between UPA and DD like other ZU5 domain proteins [248]. However, one striking difference in MACC1 structure is the presence of tandem DD in the MACC1 structure next to each other. In line with the previous observations, MACC1 structure shows the double death domain located close to the C-terminus [79, 94, 95, 103]. Death domains as the name suggests are domains critical for regulating apoptosis and inflammatory pathways [249]. The key function of proteins containing DD is promotion of apoptosis through activation of caspases and pro-inflammatory factors. DD commonly oligomerizes leading to activation of diverse signaling cascades [249-251]. However, in MACC1 two double domains are observed at the C-terminus. It is possible that the presence of double domains may facilitate easy formation of higher-order oligomers.

As expected, the folds of the MACC1 ZU5 domain are structurally similar to the β-strands of other ZU5 domain-containing proteins. In comparison with Ankyrin, the most extensively studied ZU5 domain-containing protein [105, 248], MACC1 contains only one ZU5 domain and lacks the

binding site for β-spectrin. Studies have identified differences between the two ZU5 domains present in the Ankyrin-B, where only the N-terminal ZU5 is capable of binding to β-spectrin and interacting with the UPA domain [105, 248]. It is likely that the ZU5 domain of MACC1 shares more characteristics with the non-spectrin binding ZU5 domain. It remains to be determined why MACC1 does not contain two ZU5 domains arranged next to each other as observed in Ankyrin B and PIDD and whether dimerization of the two MACC1 ZU5 domains compensates for the lack of a second ZU5 domain.

It is interesting to note that not all ZU5-containing proteins have a tandem of ZU5 domains next to each other. As an example, UNC5b, a protein involved in apoptosis and angiogenesis, contains similar domain arrangement as MACC1. The crystal structure of UNC5b reveals that the DD domain of UNC5b along with the ZU5 and UPA domain results in a closed auto-inhibited conformation that is released on activation [252]. Based on the structure composition, it is perhaps possible that the ZU5, UPA, and DD domains of MACC1 may form a similar auto-inhibited conformation. However, no evidence of signal inhibition in MACC1 overexpressing cells in untreated conditions has been reported till date. On the contrary, higher MAPK signal activation and activation of ERK have been observed in MACC1 overexpressing cells [82, 93, 103]. Therefore, the possibility of an auto-inhibited MACC1 state seems unlikely. Nevertheless, evaluating MACC1 signaling in an alternate cell model system could be one solution to address this question.

Another outstanding feature in the MACC1 structure is the presence of a SH3 domain. SH3BP4, the closest homolog of MACC1, is another example of protein that contains an SH3 domain along with the ZU5 domain (Figure 23). Precisely, the ZU5 domain of SH3BP4 is involved in the inhibition of Wnt/β-catenin signaling by preventing the nuclear accumulation of β-catenin and resulting in tumor suppression [106-108]. Notably, like MACC1, no experimental structure of SH3BP4 is available till date. Considering the similarities in the protein sequence between the two proteins, it is tempting to speculate that a similar condition could be applicable for the structure determination of both proteins. For now, comparing the predicted AlphaFold structures of the two proteins shows striking similarities in the domain arrangements and overall topology. This is in line with the previous report that highlights the similarities in domain composition and function [95]. It is important to bear in mind that both are predicted structures generated using the same computational tool. Therefore, these comparisons need to be interpreted with caution.

Although AlphaFold2 is the most accurate protein prediction tool till date, it suffers from certain drawbacks such as limited capability to predict disordered regions, inadequate post-translational

modifications such as phosphorylation, and inability to provide insights into the dynamic conformation yet [157]. Further work focusing on crystallization of MACC1 would provide valuable insights into the structural features involved in the activation and regulation of MACC1.

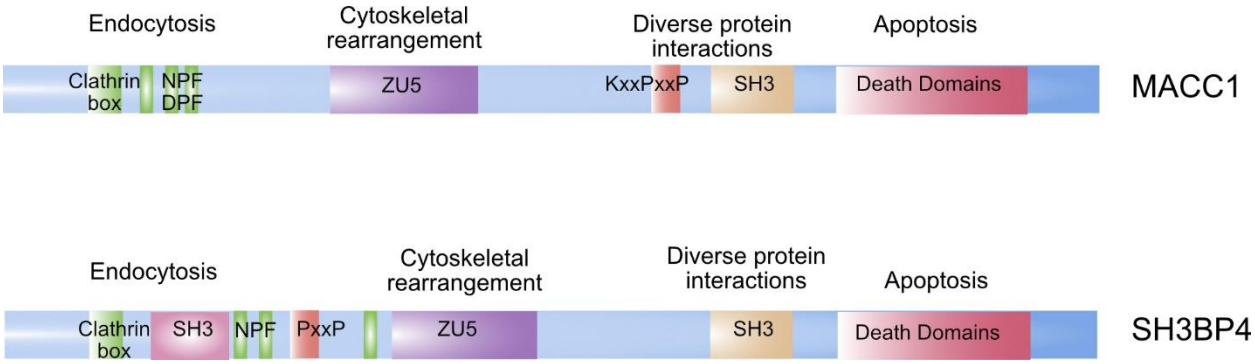

Figure 23: Domain structure of MACC1 and SH3BP4

Schematic representation of the domain structures of MACC1 and its closest homolog SH3BP4. In addition to sharing more than 40% protein homology, MACC1 and SH3BP4 also share similar functional characteristics such as endocytosis [95, 104]. At the N-terminus, both show the presence of endocytosis elements such as Clathrin box, NPF, DPF motifs followed by the ZU5, UPA, SH3, and double Death domains. One major difference is the presence of two SH3 domains in SH3P4 compared to one SH3 domain in MACC1.

7.2.2. Unraveling the MACC1 Dimers using AlphaFold-Multimer

MACC1 is a key player in tumorigenesis, but despite intense clinical implications none of the studies have ever questioned MACC1's ability to oligomerize [72, 73, 79]. In this project, GalaxyHomomer [162, 223], a template-based protein oligomer prediction web server, and AlphaFold-Multimer [163], a novel prediction system for protein complexes, was utilized to analyze the ability of MACC1 to form a higher-order assembly. Results reveal that MACC1 has a high tendency to form homodimers. With the help of collaboration partners, a detailed evaluation of the MACC1 dimer structure using AlphaFold-Multimer [144, 163] uncovered the MACC1 dimer interface along with the residues involved in dimer assembly.

The examination of the predicted dimer structure illustrates the involvement of the first β-sheet of the ZU5 domain. Precisely, the interaction between the residues Val212, Ileu214, and Cys216 of the one MACC1 chain with the other MACC1 chain serves as the framework for dimerization. The AlphaFold accuracy score for prediction ranged from 0-1 with 1 representing the highest prediction accuracy [144, 163]. The homodimer of MACC1 involving the N-term ZU5 domain showed a maximum score of 0.815. Other combinations such as C-domain:C-domain, N-domain:C-domain showed strikingly lower prediction scores of 0.343 and 0.259, respectively. Further inspection of the corresponding dimer structure in the C-terminus scenarios did not show a distinct dimer interface as clearly observed in N-domain:N-domain. This indicates a higher likelihood of dimerization mediated by the N-terminal ZU5 domains of MACC1. Nonetheless, it is important to point out that only cropped regions of MACC1 were used for prediction. It is probable that the global structure of MACC1 favors another structural conformation.

Despite the high accuracy of AlphaFold2, it suffers from limitations when identifying indirect and transient interactions [151]. In a recent report, Ruff and Pappu advised to proceed with caution when interpreting predicted structures of intrinsically disordered protein regions [156]. Taking this into consideration, it is crucial to bear in mind the contributions of N-terminal MACC1 disordered regions in dimer formation and its impact on dimer stabilization. A rational approach would be to analyze dimerization of full-length MACC1 through another independent state-of-the-art computational tool.

It is worth mentioning that the 24 other models generated for N-domain:N-domain MACC1 dimer were also strikingly similar to the top-ranked model. These results suggest that the residues of the first β-sheet may be indeed integral for dimerization. On the other hand, the scores were highly variable among the predicted models in the other two scenarios with C-terminus. Concisely, the

predicted aligned error matrix between residues of the neighboring molecules was small for N-domain:N-domain but large for C-domain:C-domain and N-domain:C-domain. In line with the above observations, the contact maps displaying the intermolecular contact points at the interface quadrants also displayed a distinct contact pattern for N-domain:N-domain MACC1 dimerization structure compared to other scenarios.

A remarkable study by Tosoni et al. (2005) demonstrates that SH3BP4, the closest homolog of MACC1, is present as dimers and eluted from size columns at higher molecular weight than expected [107]. The authors suggest that SH3BP4 could act as a scaffold protein involved in the multiprotein complex as a dimer or multimer [107]. Most importantly, the cysteine residue involved in MACC1 dimerization is not only conserved among species in MACC1 but also across species in SH3BP4 [94]. Although speculative, this observation suggests that perhaps MACC1 and SH3BP4 could share a similar mechanism for dimerization that facilitates their function as scaffold/adaptor protein.

Numerous proteins related to tumor progression exist as homodimers or higher-order oligomers and can increase the stability, activity, or DNA binding affinity of the protein [160, 168, 179, 229, 253, 254]. For instance, a report by Khanal et al. (2018) shows that YAP2L and TAZ, crucial components of the Hippo pathway, exist as homodimers in HEK293 cells and *in vivo* models through forming disulfide bonds between the cysteine residues [255]. Extensive literature indicates that the presence of structure symmetry observed in homodimers facilitates crystallization up to 1.5-fold compared to monomeric proteins [256]. Previous attempts to produce and crystallize MACC1 protein have resulted in limited success. A study by Joseph et al. (2013) shows that substituting β-strand interface residues with proline disrupts the H-bonding of a dimer and results in a functional monomer [257]. The knowledge that the first β-strand of the ZU5 domain in MACC1 is involved in dimerization could be exploited to attain the appropriate conditions for crystallization of MACC1.

Notably, the postulated mode of the MACC1 dimer assembly is consistent with its biological function, i.e., the active sites such as the SH3 domain and double Death domains are accessible for protein interaction [48, 90, 95]. With the MACC1 N-termini engaged in dimer assembly, the active sites are available for activation and participation in the signaling cascade. Past studies have shown that the SH3 domain is essential for MACC1 function *in vitro* and *in vivo*. Particularly, without the SH3 domain, MACC1 is neither able to translocate into the nucleus nor regulate the transcription of c-Met. Concomitantly, the metastasis phenotype of MACC1 is also lost in the absence of the SH3 domain *in vivo* [48, 95]. This domain is known to play a critical role in

establishing an association between signaling cascades through adaptor proteins [258, 259]. These interactions not only form the foundation of kinase activation but also help in the ring assembly of dynamin in endocytosis, thus rendering these proteins indispensable for many cellular functions [258, 259].

Several studies have highlighted the key role of SH3 domains and their potential as anti-cancer targets. It is interesting to note that the SH3 domain of SH3BP4 is also vital for its function. As demonstrated by Kim et al. (2012) the SH3 domain of SH3BP4 is indispensable for the interaction of SH3BP4 and Rag GTPases and consequently inhibition of mTORC1 signaling [108]. An additional study shows that the SH3 domain of SH3BP4 plays a critical role in transferrin internalization [107]. On similar lines, an independent study by Imbastari et al. (2019) reveals that the SH3 domain of MACC1 is important for MACC1's distribution at the plasma membrane and interaction with endocytosis proteins clathrin heavy chain 1 and dynamin 2 [104]. Given that MACC1 is implicated in endocytosis and leads to an increased rate of EGFR recycling, it is tempting to speculate that the oligomeric state of MACC1 might be important for the regulation of RTK signaling and endocytic trafficking. Some authors have speculated about the possibility of a dimer-monomer equilibrium in cells [260]. It is likely that MACC1 might also co-exist as a dimer, monomer, or other oligomer in living cells and have different functions based on the disease state. Gaining additional evidence regarding the different oligomer states of MACC1 in living cells will open a new area of regulation of MACC1 function based on structure assembly.

Another example is the functional unit of Ras, a prominent oncogenic protein, which is also present in a dimer form. Detailed *in vitro* and *in vivo* experiments reveal that targeting the RAS dimer interface hinders signal transduction and tumor growth [261]. Considering this notion, it is possible that targeting the MACC1 dimerization could influence MACC1's phenotype *in vivo*.

Recent studies have demonstrated that high levels of S100A4, a prominent metastasis-associated protein, are found in the MACC1 secretome [88]. Evidence shows that S100A4 is essential for MACC1-mediated cell migration. Importantly, S100A4 and MACC1 are positively correlated in CRC cohorts and the elevated expression of both genes is associated with poor outcome [88]. Interestingly, S100A4 is also present as homodimers within the cells. Crystal structure and NMR studies establish the interaction of S100A4 dimers with myosin IIA to regulate cell migration [220, 253, 262]. Interestingly, the MACC1 interactome also shows the presence of myosin proteins (data unpublished). It is probable that similar stoichiometry could be applicable to MACC1's interaction with myosin to mediate cell migration.

In some cases, dimerization provides an alternative binding site or increased binding affinity for protein interaction [166]. Most of the MACC1 biological studies until date are interpreted based on MACC1 present in living cells as a monomeric protein. However, MACC1 is shown to interact with several proteins containing similar interaction sites at the same time [82]. The possibility of a spatial temporal regulation is hypothesized yet the dynamics of such a system are still not clear. The findings that MACC1 exists as a dimer could offer an alternative explanation for this phenomenon suggesting the presence of identical sites available on two MACC1 units. Thus, interactions that were previously considered mutually exclusive may exist together. MACC1 dimerization could enhance stability and increase the number of binding sites without increasing the overall genome size as observed in other proteins [166].

In this study, all three putative residues namely Val212, Ile214 and Cys216 present at the MACC1 dimer interface were mutated. Curiously, not all the three residues involved in dimerization are conserved among species. Sequence analysis reveals that only residue Cys216 is conserved (Figure 13). It is likely that the inhibition of Cys216 alone may reduce the self-association of MACC1. Developing small-molecule inhibitors to intervene with MACC1 function has been a challenging undertaking [92], the search for small-molecule inhibitors targeting constitutive present MACC1 dimers within living cells could be even more difficult. Interestingly, the dimer structure predicted by AlphaFold-Multimer reveals the presence of a hydrophobic pocket formed by Val212, Ile214 and Cys216 on dimerization (Figure 12). Thus, exploiting this information from the predicted model could expedite the future development of small-molecule inhibitors capable of binding to this hydrophobic pocket and interfering with MACC1 function.

To sum up, employing AlphaFold-Multimer, the dimerization of MACC1 was predicted and the residues Val212, Ileu214, and Cys216 involved in dimerization were identified. Based on the MACC1 dimer structural analysis, all three residues, Cys216 in particular, may be important in the formation of MACC1 dimers. The hydrophobic pocket formed by the MACC1 dimer interface residues serves as the potential target site for future drug design and screening.

7.2.3. BRET: A Valuable Tool for Detecting MACC1 Homodimers

Co-immunoprecipitation is a standard technique to identify PPI. However, the major disadvantage is that direct (binary) interactions cannot be separated from complex interactions [170, 181]. To overcome this limitation, an innovative technique known as BRET was employed to study MACC1 dimerization. BRET is a valuable technique that can detect transient and stable binary protein interactions with high levels of sensitivity and specificity [169, 192, 195].

Using BRET fluorophores (mCit and NLuc) tagged to MACC1, homodimerization of MACC1 was examined in living cells. Previous studies state that the position of the BRET reporter tag is crucial for identification of a positive interaction [195, 226]. As anticipated, the highest BRET ratio was observed when the fluorophores were placed at the N-terminus. Moreover, reports suggest that the closer the BRET reporter tag is to the binding region the higher the BRET ratio value [195]. Thus, these findings indicate that the interaction site could be close to the N-terminus. Another explanation for the high BRET ratio values could be that the N-terminus fusion construct allowed for an efficient transfer of resonance energy [226] between the two MACC1 units compared to the N-terminus to C-terminus fusion construct. Therefore, this suggests that the N-terminus conformation or orientation is more stable or preferable compared to interaction with the C-terminus. This is in agreement with the prediction scores observed using the AlphaFold-Multimer [163].

While the development of mCit tagged to N-terminus MACC1 was successfully achieved, several attempts to develop the mCit tagged to C-terminus MACC1 yielded no success. Thus, only mCit tagged to N-terminus of MACC1 was used as a BRET acceptor for this study. With this setup, the acceptor mCit tag was always positioned at the N-terminus whereas two different positioned vectors for the donor NLuc tag were applied namely the N-terminal NLuc and C-terminal NLuc. It is important to mention that as the mCit tagged to C-terminus was not included in this study, therefore the interaction between two C-termini of MACC1 could not be examined using BRET in living cells.

A closer look at the results of MACC1 dimerization under different conditions reveals that the MACC1 dimer is present even in the absence of growth factors. It is interesting to note that other proteins involved in tumor progression such as STAT1 and STAT3 also exist in stable dimeric state prior to activation [221]. In this study, the presence of MACC1 dimer prior to stimulation could explain the activation of ERK observed in untreated conditions in MACC1 overexpressing cells in previous studies from our group [82, 103]. Although no significant differences in BRET

values were observed between the untreated and treated cells, these findings do not completely exclude the effect of growth factor stimulation on MACC1 dimer regulation. Further work, which examines different concentrations of HGF at various time points, needs to be performed to clearly understand the dynamics of MACC1 dimer.

Contrary to expectation, no interaction was observed between MACC1 and SRC. This is contrary to the findings from Co-IP where an interaction was observed between MACC1 and SRC. These results can be explained in part by the fact that Co-IP does not distinguish between direct and indirect partners, and in part by the proximity restraint of tags in the BRET assay. It is noteworthy to state here that the absence of a BRET signal does not completely exclude interaction but rather asserts that the proteins were not found in the preferable orientation in the proximity to facilitate dipole-dipole coupling. Based on this data, it can be assumed that SRC could be a potential MACC1 binder but may not be a specific interaction partner.

In the next step, residues involved in dimerization namely Val212, Ileu214, and Cys216 were mutated. These mutated MACC1 constructs were then shuttled into plasmids containing BRET fluorophores. The WB SDS-PAGE analysis (Figure 13) confirmed that all the fusion proteins were well-expressed, and the mutations did not affect the protein expression of MACC1. In addition to detecting weak transient interactions, BRET enables us to study the interactions in their native environment [169]. This makes it superior to the traditional techniques that require cell lysis. This improved method can easily be adapted to high-throughput screens, systematically mapping vast multi-protein complexes. Though the construction of the tagged vectors can be challenging, the assay workflow once established can be easily adapted to study the interactions under various conditions [169, 195].

The BRET results show a 2.5-fold reduction in BRET ratio in 3xMut MACC1 cells compared to WT MACC1 (Figure 15). A possible explanation could be the change in the distance between the NLuc (BRET donor) and mCit (BRET acceptor) moieties caused due to repulsion between the mutated residues. Another feasible explanation could be a change in the conformation or opening of the MACC1 dimer structure due to mutation. It is important to note that the mutations diminish the interaction between the MACC1 monomers but did not completely abolish the association. Additionally, the BRET system used in this study is only capable of analyzing the interaction between two proteins, thus higher-order oligomers of MACC1 could not be evaluated. It is important to mention that the likelihood of a higher-order MACC1 oligomeric form is possible and should not be completely dismissed based on these findings.

In this study, only BRET was used to analyze the MACC1 dimerization in part because of its high sensitivity but also due to the lack of large amounts of purified MACC1 protein. Hence, additional investigations such as size exclusion chromatography, blue native gel electrophoresis, and analytical ultracentrifugation are warranted to verify the intriguing findings made in this project. These studies will undoubtedly improve our understanding of MACC1 dimerization and additionally address the possibility of higher-order MACC1 oligomer assembly.

As expected, similar results were obtained in HCT116 cells and a reduction of 50% was observed between the WT MACC1 and 3xMut MACC1 cells. An important advantage of BRET is that the interactions are analyzed in their native environment. In contrast to HEK293 cells, the BRET signal ratio observed in HCT116 was much higher than HEK293 cells (0.2 vs. 0.06). This difference can be explained in part by the important role of MACC1 in CRC signaling, in the HCT116 system as shown in previous studies [82, 88, 103]. Collectively, the BRET results of 3xMut MACC1 expressing cells strengthen the proposed binding mode for MACC1 dimer, where the hydrophobic residues Val212, Ile214, and Cys216 form the dimer interface.

As stated in other studies, the main drawback of transient co-transfection is that the expression varies between cell lines [226, 263]. As clearly observed in this study, the BRET ratio obtained in HCT116 cells is significantly higher than the BRET ratio in HEK293 cells (Figure 15). Additionally, exogenous protein expression suffers from inherent drawbacks such as expression variability as observed also in the case of WT MACC1-mCit and 3xMut-mCit. The plate reader measurements recorded may not consider the low fluorescence intensity generated by some cells. Therefore, it is possible that signals from some subpopulations may be neglected.

It is important to note that the tagged MACC1 protein was expressed exogenously in cell lines expressing low or moderate amounts of MACC1. The non-native expression of MACC1 may trigger the formation of false interactions that may not exist endogenously. Further, the transient overexpression might also interfere with the posttranslational modifications and modify the distribution of MACC1 within various subcellular compartments. A good alternative would be the establishment of stable cell lines [263] that express the MACC1 hybrid protein or using methods that employ endogenously tagged proteins. Newer technologies that produce endogenously tagged proteins will not only surpass these issues but also aid in monitoring this complex at different stages of cell development including when MACC1 cells enter senescence or EMT and post-therapy MACC1 cells. Therefore, the development of stable cell lines will be highly valuable to study the overall impact of MACC1 in tumor development (Figure 24).

Nonetheless, the findings presented above are particularly interesting for several reasons: First, BRET confirmed the existence of MACC1 as a dimer not only in HEK293 cells but also in CRC cells, HCT116. Secondly, the assay provided preliminary insights into the impact of various conditions on the MACC1 dimerization. Another important implication of knowing the oligomeric state of MACC1 is that it enables us to correctly assess the sites available for protein interactions.

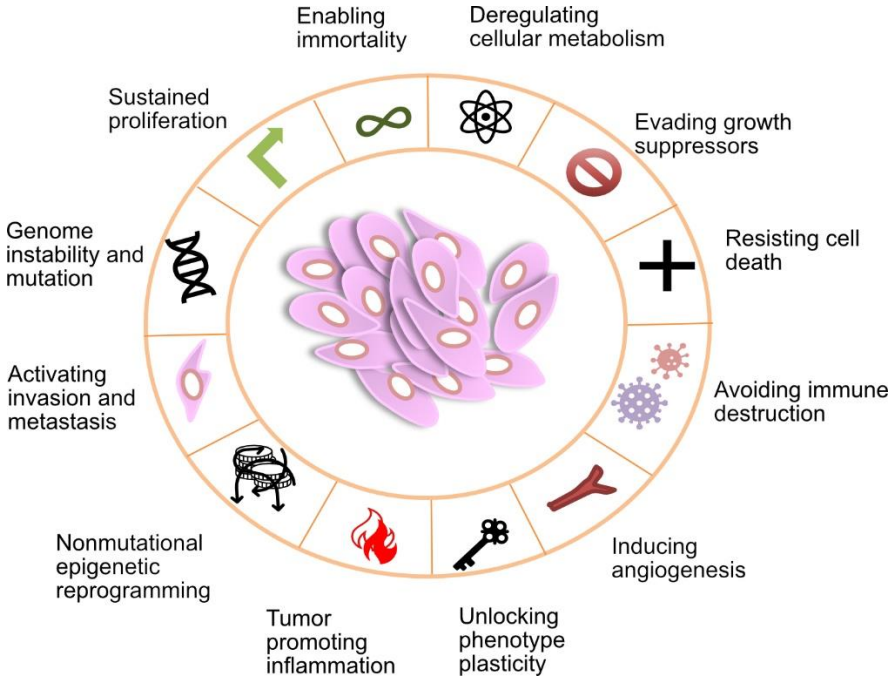

Figure 24: Overview of the role of MACC1 in tumorigenesis

MACC1 expression impacts critical cancer hallmarks through critical pathways involved in invasion, cell death resistance, sustained proliferation, angiogenesis, and stemness.

7.2.4. Impact of Dimer Hindering mutations on MACC1 Signaling

Throughout tumor progression, higher activation of RTKs leading to sustained activation of downstream signaling pathways such as the MAPK among others is commonly observed [85, 109]. SRC, another integral protein involved in cell motility, is also associated with higher activation of the c-Met signaling cascade [125]. Previous literature indicates that MACC1 is involved in regulating the SRC activation and MAPK pathway post-HGF treatment [48, 82, 103].

It was hypothesized that mutation of residues in the ZU5 domain of MACC1 would also have an impact on the MAPK signaling. Contrary to expectation, no differences were observed in the activation of downstream effector ERK between WT MACC1 and 3xMut MACC1 cells. On the other hand, considerable differences were observed in the activation of SRC in WT MACC1 and 3xMut MACC1 cells. It was clear from these results that mutation of residues in the ZU5 domain not only influenced the structural features of MACC1 but also the activation of SRC-mediated pathways. Higher activation of SRC was observed in WT MACC1 cells treated with HGF compared to 3xMut MACC1 cells. Most certainly, SRC activation is one among several pathways that are highly activated in tumor progression. This study addresses the impact of MACC1 mutation on SRC as it is one of the integral proteins involved in cell migration and drug resistance [128, 132, 133, 135]. Furthermore, this difference was clear when treating the cells with HGF for longer intervals. This could suggest that the dimerization of MACC1 could facilitate stability and prolonged signal activation.

Comparing the activation of ERK and SRC in MACC1 overexpressing cells shows that the two proteins are activated independently of each other. This finding corroborates a previous study in colorectal carcinoma that noted the importance of c-Met in SRC activity. The report shows that the downregulation of c-Met reduces SRC activity but has a minimal effect on ERK1/2 and Akt activation [118]. The present findings in WT MACC1 and 3xMut MACC1 cells seem to be consistent with this report suggesting an independent regulation of SRC and ERK activation. Furthermore, the authors reveal that SRC inhibitors reduce the production of Vascular Endothelial Growth Factor (VEGF) resulting in lower basal proliferation, anchorage-independent growth, and migration which could be rescued by HGF treatment [118]. In this study, however, only the activation of downstream intermediates of the c-Met was evaluated. Additionally, preliminary investigations evaluating the activation of the MAPK pathway post-VEGF and post-PDGF treatment were performed (data not shown), but more detailed experiments need to be undertaken to evaluate the role of MACC1 in the expression and signaling of VEGF. Another study by Veracini et al. (2005) proves that the SFK signaling is mainly independent of the MAPK

pathway and is capable of phosphorylating substrates such as SHC, ABL, and STAT3 to induce *c-myc* expression and cell-cycle progression [244]. It is likely that a similar mechanism occurs in MACC1 overexpressing CRC cells in which SRC and MAPK are regulated independently.

Interestingly, MACC1 is regulated post-transcriptionally by tumor suppressor miR, particularly, miR-218. In addition to inhibiting MACC1 protein expression, miR-218 impairs MACC1-mediated cell migration and invasion [213]. In another independent study, a feedback loop between SRC and miR-218 was reported to regulate receptor protein tyrosine phosphatase expression [264]. It is possible that a similar mechanism could exist between SRC and miR-218 to regulate MACC1 function. However, a different study by Shi et al. (2017) investigating the role of miRNAs in EMT-associated properties such as migration and invasion detected Slug and ZEB2 as direct targets of miR-218 in lung cancer [265]. Although the above findings suggest the probable role of SRC in MACC1 regulation through miR-218, the possibility of MACC1 regulation via Slug and ZEB2 cannot be completely dismissed. In an unrelated study by Huang et al. (2015) in gastric cancer, the author reports that miRNA-338 directly targets ZEB2 and MACC1. Precisely through MACC1, miR-338 can decrease the activation of the c-Met /Akt axis leading to EMT suppression [266].

It is reasonable to expect that the MACC1 mutant may also have an impact on the proteins downstream of SRC. Correlation studies in patient cohorts indicate a positive correlation of MACC1 not only with SRC but also with FAK. Additionally, the ZU5 domain also contains phosphorylation site Y379 predicted to bind with SRC (Figure 7). This may suggest the role of MACC1 dimerization in the regulation of migration induced by interaction with SRC and in turn facilitating a positive feedback loop.

In a detailed study, Ahmed et al. (2015) demonstrates that GRB2, an adaptor protein downstream of RTK, exists in a constitutive equilibrium between the monomeric and dimeric states [254, 260]. The authors state that the tyrosine phosphorylation Y160 on GRB2 regulates the dimer dissociation and initiation of the MAPK signaling. The authors describe this monomer-dimer equilibrium as a switch that regulates tumor progression where phosphorylation on Y160 could interfere with the dimerization through repulsive charge or steric clashes [260]. Interestingly, another independent study demonstrates that SRC is one of the kinases that phosphorylate Y160 on GRB2 and regulate its activity [267]. Taking this into consideration, it is possible that an equilibrium might also exist between the MACC1 monomeric and dimeric state to regulate MACC1 function and this could be regulated via tyrosine phosphorylation as observed with GRB2. Consistently, a report from different cancer types shows that the active monomeric GRB2 is linked with malignant cancer [260]. It is possible that MACC1 functions through a similar mechanism

where the phosphorylation of Y379 in the ZU5 domain of MACC1 could be a switch that regulates the dimeric state and in turn the MAPK signaling. In this context, it would be interesting to examine the impact of Y379 mutation on the dimer formation. In addition to tyrosine phosphorylation, other mechanisms could regulate the association and dissociation of MACC1 dimer such as the expression level of MACC1 or localization of MACC1 as observed in other dimer-forming proteins [124, 255, 261].

Global proteomic and phosphoprotein profiling revealed significant differences between WT MACC1 and 3xMut MACC1. As expected, and in accordance with the wound healing and cell proliferation assay, there were differences between the WT and 3xMut cells in the actin filament-based process and intracellular signal transduction. Among the commonly downregulated proteins, PLEC, a cytoskeletal protein involved in the filament organization, was also found to be downregulated in 3xMut MACC1 cells. Recently, a study by Yuan et al. (2022) indicates that PLEC along with other hypoxia-related genes confers drug resistance in CRC [205].

Considering all these findings, it can be inferred that the residues Val212, Ile214, and Cys216 are important for MACC1-mediated SRC activation and may regulate processes involved in cell migration and intracellular signal transduction. The results suggest the functional signaling unit of MACC1 may be a dimer instead of monomer as assumed till date.

7.2.5. Functional Characterization of MACC1 Dimer Hindering Mutation

Previous studies have established that cells overexpressing MACC1 demonstrate a metastasis phenotype with higher proliferation and migration characteristics in addition to hyperactivated MAPK signaling [48, 82, 103]. The initial WB screen indicated the impact of dimer interface mutation on MACC1 signaling and its influence on SRC activation. Previous literature highlights the role of SRC along with integrins and focal adhesion proteins in the regulation of cytoskeletal reorganization and cell migration [268, 269]. Studies in different tumor entities support this observation and further intervening with SRC signaling has been demonstrated to reduce metastasis properties such as migration, invasion, and proliferation [129].

To assess the biological relevance of the mutations studied in this project, various metastasis functional assays were performed. Firstly, the ability of 3xMut MACC1 cells to colonize was evaluated using 2D colony formation. The 3xMut showed only a slight reduction in the colonization ability of MACC1, and no significant differences were observed between the WT MACC1 and 3xMut MACC1 cells. Nevertheless, mutation of the residues in the ZU5 domain of MACC1 resulted in altered cellular morphology. Precisely, the WT MACC1 cells formed bigger round colonies compared to 3xMut MACC1 cells. This finding was unexpected and suggests that perhaps dimerization may not affect all functional characteristics of MACC1, particularly in conditions without any treatment. In future investigations, the use of reduced serum or inhibitors could be a solution to evaluate the consequences of mutation on colony formation under stress.

Although no differences in ERK activation were observed in the previous signaling assay, the proliferation assay displayed significant differences between the WT MACC1 and 3xMut MACC1 cells. A possible explanation might be that the proliferation in MACC1 overexpressing cells is partly regulated by mechanisms other than ERK. Another important factor may be that the proliferation assay is analyzed continually over a long period. These findings thus attribute the long-term effect of the treatment on the function of the cells compared to WB analysis that investigated specific intervals only up to 90 min. In the wound healing assay, 3xMut cells showed significantly lower wound confluence compared to WT MACC1 cells. A further aspect of these assays was to evaluate if treatment with HGF could rescue the reduced metastasis ability due to the mutation. Contrary to expectations, the proliferation and migration ability of 3xMut MACC1 cells were not rescued by HGF treatment, indicating a strong influence of mutation on the metastasis phenotype.

A study by Hohmann et al. (2023) reports that MACC1 mainly influences collective cell migration rather than single-cell migration in a proliferation-dependent manner [270, 271]. Therefore, for this project, the wound healing assay (also known as scratch assay) which analyzes collective migration was employed. Some of the previous migration studies on MACC1 were performed with the same setup as the proliferation assay and as an end-point measurement [243, 272]. However, the drawback of this approach is that the proliferation and migration ability of cells cannot be differentiated. An alternative to completely abrogate proliferation is using mitomycin. In this study, a reduced serum medium and live-cell imaging system with continuous analyses (IncuCyte®) was employed to address this issue. In addition to providing continuous data regarding collective migration, this setup potentially decoupled MACC1-induced proliferation from migration. An additional inherent drawback of wound healing assay is that migration is observed without applying chemotactic gradients. Further examinations using Transwell Boyden Chamber migration would complement the current findings and examine the influence of chemoattractant and repellent on 3xMut MACC1 cells.

Mounting evidence indicates that CRC cells exhibit variable properties and cytoskeletal rearrangement in 2D and 3D models [273]. The metastasis assays used for this study investigated cell motility, proliferation, and colonization in 2D systems without any matrix coating. ECM, another important component of cell adhesion and motility, plays a critical role in metastasis. Studies show that the composition and stiffness of the coating matrices can have a significant impact on cell invasion and migration [241, 274]. Therefore, additional experiments employing 3D systems such as spheroids and various matrix compositions could verify metastasis phenotype in a more native environment and further validate these findings.

Since the discovery of MACC1, functional characteristics of MACC1 such as migration, proliferation, and invasion have been tested in more than 20 different cancer entities [73]. A study by Sueta et al. (2015) reveals that MACC1 expression and its influence on HGF/c-Met has different consequences in breast cancer compared to CRC [243]. The current study only examined the role of MACC1 dimers in CRC cells. As the role of MACC1 is established in 20 different tumor types, it is likely that MACC1 dimer would show similar characteristics in other types of cancer. Nevertheless, future research in other cancer entities would be beneficial to establish the dependence of MACC1 function on MACC1 dimer in the progression of other cancers such as breast, pancreatic, gastric, and prostate among others. Additionally, these results will verify the impact of mutation of residues in the ZU5 domain of MACC1 at the cellular level *in vitro* in different cell models. All together these findings laid down the first evidence for MACC1

dimerization, however further *in vitro and in vivo* research needs to be conducted to validate this hypothesis.

To sum up, in wound healing and cell proliferation assay, the 3xMut MACC1 cells displayed reduced cell proliferation and migration which is consistent with the lower SRC phosphorylation observed in the signaling assay. Even more striking was that the reduction of proliferation in the mutant cells was observed even in the presence of HGF. These findings imply that the self-association of MACC1 is integral for its activation of cell migratory pathways and once more emphasizes the importance of the three residues in the ZU5 domain for MACC1 dimerization. Future structure-based inhibitor design experiments are needed to understand the precise picture of MACC1 dimerization and whether MACC1 dimer-specific inhibitors display anti-tumorigenic and anti-metastasis properties.

8 Concluding Remarks and Future Perspectives

Metastasis continues to be a major hurdle in the effective management of tumors across various cancer types [9]. MACC1 is a promising biomarker that has been associated with metachronous metastasis and low survival in numerous cancer entities [48, 72, 73]. In this study, the first evidence of MACC1 existing as a homodimer is provided using AlphaFold2 [144, 154, 163] and BRET [192, 195]. Importantly, the residues critical for MACC1 dimerization are identified and the impact of mutation of the postulated residues on MACC1 signaling and function is presented.

Several independent studies establish the role of MACC1 in metastasis development and demonstrate tumor inhibition when the expression of MACC1 is altered [78, 91, 92, 103, 198, 206]. Together with the notion of targeting the transcription of MACC1, inhibition of MACC1 protein interactions, especially by targeting the MACC1 dimerization could offer a promising new strategy. Other studies including this one (Dumbani et al., in preparation) provide insights into the structural features of MACC1 and potential novel MACC1 targeting sites. The predicted structure of a MACC1 dimer could serve as a starting point for designing and understanding the binding mode of novel MACC1 inhibitors.

The understanding of MACC1 dimer has just started and the findings presented here provide a framework for future development of strategies to intervene with MACC1-mediated metastasis and to better understand the versatility of MACC1 as a metastasis protein.

9 List of Abbreviations

ABL	Abelson murine leukemia viral oncogene homolog 1
AI	Artificial intelligence
ANOVA	Analysis of variance
APC	Adenomatous polyposis coli
BCL2	B-cell lymphoma 2
BiFC	Bimolecular fluorescence complementation
BRAF	v-RAF murine sarcoma viral oncogene homolog B
BRET	Bioluminescence resonance energy transfer
CDK4/6	Cyclin-dependent kinase 4/6
CEA	Carcinoembryonic antigen
CIMP	CpG island methylator phenotype
CIN	Chromosomal instability
CMS	Consensus molecular subtypes
CRC	Colorectal cancer
CREB	cAMP response element-binding protein
CSK	C-terminal Src kinase
ctDNA	Circulating tumor DNA
CTNNB1	Catenin-β1
CTTN	SRC substrate cortactin
Cyro-EM	Cryogenic electron microscopy
DCLK1	Doublecortin like kinase 1
ECM	Extracellular matrix
EGFR	Epidermal growth factor receptor
EMT	Epithelial-mesenchymal transition
EMT-TFs	EMT-activating transcription factors
EPS8	Epidermal growth factor receptor pathway substrate 8
ERK	Extracellular signal–regulated kinase
FAK	Focal adhesion kinase
FHOD1	FH1/FH2 domain-containing protein 1
FOLFIRI	Folinic acid, fluorouracil, irinotecan
FOLFOXIRI	Folinic acid, fluorouracil, oxaliplatin, irinotecan
FRET	Fluorescence resonance energy transfer
GAB1	GRB2-associated binder 1
GPCR	G protein-coupled receptor
GPS	Group-based prediction system
GRB2	Growth-factor-receptor-bound protein 2
GSK3β	Glycogen synthase kinase 3 beta
HER2	Human epidermal growth factor receptor 2

HGF/SF	Hepatocyte growth factor/scatter factor
HGFR	Hepatocyte growth factor receptor
HRP	Horseradish peroxidase
KRAS	Kirsten rat sarcoma viral oncogene homolog
LIMA1	LIM domain and actin-binding protein 1
MACC1	Metastasis-associated in colon cancer 1
MCL-1	Myeloid leukemia cell differentiation protein
MEK	MAPK/ERK kinases
MMP	Matrix metalloprotease
MMR	DNA mismatch repair
MS	Mass spectrometry
MSAs	Multiple sequence alignments
MSI	Microsatellite instability
NF-kB	nuclear factor kappa-light-chain-enhancer of activated B-cells
PD1	Programmed cell death protein 1
PDGF	Platelet-Derived Growth Factor
PDGFR	Platelet-derived growth factor receptor
PI3KCA	Phosphatidylinositol-4,5-biphosphate 3-kinase catalytic subunit
PKA	Protein kinase A
PKP2	Plakophilin-2
PLCG	Phospholipase C gamma
pLDDT	Predicted local-distance difference test
PLEC	Plectin
PPI	Protein-Protein Interaction
PTB	Phosphortyrosine-binding domain
pY	Phosphorylated tyrosine
qRT-PCR	Quantitative real-time polymerase chain reaction
RAS	Rat sarcoma viral oncogene homolog
RTK	Receptor Tyrosine Kinase subfamily
S100A4	S100 calcium-binding protein A4
SCNA	Somatic copy number alteration
SDM	Site-directed mutagenesis
SDS-PAGE	Sodium dodecyl sulphate - polyacrylamide gel electrophoresis
SFK	SRC family kinase
SH2	Src-homology-2 domain
SH3BP4	SH3 domain-binding protein 4
SHC1	SH2 domain containing transforming protein
SHP2	SH2 domain-containing protein tyrosine phosphatase 2
SMAD3	Mothers against decapentaplegic homolog 3
SMAD4	Mothers against decapentaplegic homolog 4
SRC	Non-receptor tyrosine kinase SRC
STAT	Signal transducer and activator of transcription
STK	Serine/threonine kinases

TCGA	The Cancer Genome Atlas
TGFBR3	Transforming growth factor-β receptor 2
TJP1	Tight junction protein ZO-1
TM-score	Global superposition metric template modelling score
TNIK	TRAF2 and NCK-interacting protein kinase
TP53	Tumor protein 53
Twist1	Twist-related protein 1
VEGF	Vascular endothelial growth factor
WB	Western blot
WNT	Wingless/integrated
ZEB	Zinc-finger E-box-binding homeobox

10 References

1. Lambert, A.W., D.R. Pattabiraman, and R.A. Weinberg, *Emerging Biological Principles of Metastasis.* Cell, 2017. **168**(4): p. 670-691.
2. Welch, D.R. and D.R. Hurst, *Defining the Hallmarks of Metastasis.* Cancer Res, 2019. **79**(12): p. 3011-3027.
3. Fares, J., et al., *Molecular principles of metastasis: a hallmark of cancer revisited.* Signal Transduct Target Ther, 2020. **5**(1): p. 28.
4. Damodaran, S., C.E. Kyriakopoulos, and D.F. Jarrard, *Newly Diagnosed Metastatic Prostate Cancer: Has the Paradigm Changed?* Urol Clin North Am, 2017. **44**(4): p. 611-621.
5. Koncina, E., et al., *Prognostic and Predictive Molecular Biomarkers for Colorectal Cancer: Updates and Challenges.* Cancers (Basel), 2020. **12**(2).
6. Rumpold, H., et al., *Prediction of mortality in metastatic colorectal cancer in a real-life population: a multicenter explorative analysis.* BMC Cancer, 2020. **20**(1): p. 1149.
7. Wang, R., et al., *The Clinicopathological features and survival outcomes of patients with different metastatic sites in stage IV breast cancer.* BMC Cancer, 2019. **19**(1): p. 1091.
8. Campos-Balea, B., et al., *Prognostic factors for survival in patients with metastatic lung adenocarcinoma: An analysis of the SEER database.* Thorac Cancer, 2020. **11**(11): p. 3357-3364.
9. Gui, P. and T.G. Bivona, *Evolution of metastasis: new tools and insights.* Trends Cancer, 2022. **8**(2): p. 98-109.
10. Chan, D.K.H. and S.J.A. Buczacki, *Tumour heterogeneity and evolutionary dynamics in colorectal cancer.* Oncogenesis, 2021. **10**(7): p. 53.
11. Mittal, V., *Epithelial Mesenchymal Transition in Tumor Metastasis.* Annu Rev Pathol, 2018. **13**: p. 395-412.
12. Sinha, D., et al., *Emerging Concepts of Hybrid Epithelial-to-Mesenchymal Transition in Cancer Progression.* Biomolecules, 2020. **10**(11).
13. Stemmler, M.P., et al., *Non-redundant functions of EMT transcription factors.* Nat Cell Biol, 2019. **21**(1): p. 102-112.
14. Goossens, S., et al., *EMT transcription factors in cancer development re-evaluated: Beyond EMT and MET.* Biochim Biophys Acta Rev Cancer, 2017. **1868**(2): p. 584-591.
15. Ansieau, S., G. Collin, and L. Hill, *EMT or EMT-Promoting Transcription Factors, Where to Focus the Light?* Front Oncol, 2014. **4**: p. 353.
16. Eide, P.W., et al., *Metastatic heterogeneity of the consensus molecular subtypes of colorectal cancer.* NPJ Genom Med, 2021. **6**(1): p. 59.
17. Burandt, E., et al., *E-Cadherin expression in human tumors: a tissue microarray study on 10,851 tumors.* Biomark Res, 2021. **9**(1): p. 44.
18. Toiyama, Y., et al., *Increased expression of Slug and Vimentin as novel predictive biomarkers for lymph node metastasis and poor prognosis in colorectal cancer.* Carcinogenesis, 2013. **34**(11): p. 2548-57.
19. Sciacovelli, M. and C. Frezza, *Metabolic reprogramming and epithelial-to-mesenchymal transition in cancer.* FEBS J, 2017. **284**(19): p. 3132-3144.
20. Seo, J., et al., *The role of epithelial-mesenchymal transition-regulating transcription factors in anti-cancer drug resistance.* Arch Pharm Res, 2021. **44**(3): p. 281-292.
21. Pastushenko, I. and C. Blanpain, *EMT Transition States during Tumor Progression and Metastasis.* Trends in Cell Biology, 2019. **29**(3): p. 212-226.
22. Haffner, M.C., et al., *Tracking the clonal origin of lethal prostate cancer.* J Clin Invest, 2013. **123**(11): p. 4918-22.
23. Yates, L.R., et al., *Genomic Evolution of Breast Cancer Metastasis and Relapse.* Cancer, Cell, 2017(1878-3686 (Electronic)).

24. Klein, C.A., *Parallel progression of primary tumours and metastases.* Nature Reviews Cancer, 2009. **9**(4): p. 302-312.
25. Venkatesan, S., et al., *Treatment-Induced Mutagenesis and Selective Pressures Sculpt Cancer Evolution.* Cold Spring Harb Perspect Med, 2017. **7**(8).
26. Gundem, G., et al., *The evolutionary history of lethal metastatic prostate cancer.* Nature, 2015. **520**(7547): p. 353-357.
27. Gerlinger, M., et al., *Intratumor heterogeneity and branched evolution revealed by multiregion sequencing.* N. Engl J Med, 2012 Mar 8(1533-4406 (Electronic)).
28. Hinohara, K. and K. Polyak, *Intratumoral Heterogeneity: More Than Just Mutations.* Trends in Cell Biology, 2019. **29**(7): p. 569-579.
29. Venkatesan, S. and C. Swanton, *Tumor Evolutionary Principles: How Intratumor Heterogeneity Influences Cancer Treatment and Outcome.* American Society of Clinical Oncology Educational Book, 2016(36): p. e141-e149.
30. Consortium, I.T.P.-C.A.o.W.G., *Pan-cancer analysis of whole genomes.* Nature, 2020. **578**(7793): p. 82-93.
31. Priestley, P., et al., *Pan-cancer whole-genome analyses of metastatic solid tumours.* Nature, 2019. **575**(7781): p. 210-216.
32. Bailey, M.H., et al., *Comprehensive Characterization of Cancer Driver Genes and Mutations.* Cell,, 2018(1097-4172 (Electronic)).
33. Zehir, A., et al., *Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients.* Nature Medicine, 2017. **23**(6): p. 703-713.
34. Ganesh, K. and J. Massagué, *Targeting metastatic cancer.* Nature Medicine, 2021. **27**(1): p. 34-44.
35. Armenia, J., et al., *The long tail of oncogenic drivers in prostate cancer.* Nature Genetics, 2018. **50**(5): p. 645-651.
36. Yachida, S., et al., *Distant metastasis occurs late during the genetic evolution of pancreatic cancer.* Nature,, 2010(1476-4687 (Electronic)).
37. Rizzetto, S., et al., *Impact of sequencing depth and read length on single cell RNA sequencing data of T cells.* Scientific Reports, 2017. **7**(1): p. 12781.
38. Lähnemann, D., et al., *Eleven grand challenges in single-cell data science.* Genome Biology, 2020. **21**(1): p. 31.
39. Punt, C.J., M. Koopman, and L. Vermeulen, *From tumour heterogeneity to advances in precision treatment of colorectal cancer.* Nat Rev Clin Oncol, 2017. **14**(4): p. 235-246.
40. Livet, J., et al., *Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system.* Nature,, 2007(1476-4687 (Electronic)).
41. McKenna, A., et al., *Whole-organism lineage tracing by combinatorial and cumulative genome editing.* Science,, 2016(1095-9203 (Electronic)).
42. Gilbert, L.A., et al., *Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation.* Cell,, 2014(1097-4172 (Electronic)).
43. Xi, Y. and P. Xu, *Global colorectal cancer burden in 2020 and projections to 2040.* Transl Oncol, 2021. **14**(10): p. 101174.
44. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.* CA Cancer J Clin, 2021. **71**(3): p. 209-249.
45. Morris, V.K., et al., *Treatment of Metastatic Colorectal Cancer: ASCO Guideline.* Journal of Clinical Oncology, 2023. **41**(3): p. 678-700.
46. Sawicki, T., et al., *A Review of Colorectal Cancer in Terms of Epidemiology, Risk Factors, Development, Symptoms and Diagnosis.* Cancers (Basel), 2021. **13**(9).
47. Buikhuisen, J.Y., A. Torang, and J.P. Medema, *Exploring and modelling colon cancer inter-tumour heterogeneity: opportunities and challenges.* Oncogenesis, 2020. **9**(7): p. 66.

48. Stein, U., et al., *MACC1, a newly identified key regulator of HGF-MET signaling, predicts colon cancer metastasis.* Nat Med, 2009. **15**(1): p. 59-67.
49. Piawah, S. and A.P. Venook, *Targeted therapy for colorectal cancer metastases: A review of current methods of molecularly targeted therapy and the use of tumor biomarkers in the treatment of metastatic colorectal cancer.* Cancer, 2019. **125**(23): p. 4139-4147.
50. Metibemu, D.S., et al., *Exploring receptor tyrosine kinases-inhibitors in Cancer treatments.* Egyptian Journal of Medical Human Genetics, 2019. **20**(1).
51. Yamagishi, H., et al., *Molecular pathogenesis of sporadic colorectal cancers.* Chin J Cancer, 2016. **35**: p. 4.
52. Kuipers, E.J., et al., *Colorectal cancer.* Nat Rev Dis Primers, 2015. **1**: p. 15065.
53. Smit, W.L., et al., *Driver mutations of the adenoma-carcinoma sequence govern the intestinal epithelial global translational capacity.* Proc Natl Acad Sci U S A, 2020. **117**(41): p. 25560-25570.
54. Walther, A., et al., *Genetic prognostic and predictive markers in colorectal cancer.* Nat Rev Cancer, 2009. **9**(7): p. 489-99.
55. Brenner, H. and C. Chen, *The colorectal cancer epidemic: challenges and opportunities for primary, secondary and tertiary prevention.* Br J Cancer, 2018. **119**(7): p. 785-792.
56. Yamane, L., et al., *Serrated pathway in colorectal carcinogenesis.* World J Gastroenterol, 2014. **20**(10): p. 2634-40.
57. Fanelli, G.N., et al., *The heterogeneous clinical and pathological landscapes of metastatic Braf-mutated colorectal cancer.* Cancer Cell Int, 2020. **20**: p. 30.
58. Di Nicolantonio, F., et al., *Precision oncology in metastatic colorectal cancer - from biology to medicine.* Nat Rev Clin Oncol, 2021. **18**(8): p. 506-525.
59. Dienstmann, R., et al., *Precision Therapy in RAS Mutant Colorectal Cancer.* Gastroenterology, 2020. **158**(4): p. 806-811.
60. Ryan, M.B. and R.B. Corcoran, *Therapeutic strategies to target RAS-mutant cancers.* Nature Reviews Clinical Oncology, 2018. **15**(11): p. 709-720.
61. Liu, J., R. Kang, and D. Tang, *The KRAS-G12C inhibitor: activity and resistance.* Cancer Gene Therapy, 2022. **29**(7): p. 875-878.
62. Guinney, J., et al., *The consensus molecular subtypes of colorectal cancer.* Nat Med, 2015. **21**(11): p. 1350-6.
63. Tainsky, M.A., *Genomic and proteomic biomarkers for cancer: a multitude of opportunities.* Biochim Biophys Acta, 2009. **1796**(2): p. 176-93.
64. Puccini, A., A. Seeber, and M.D. Berger, *Biomarkers in Metastatic Colorectal Cancer: Status Quo and Future Perspective.* Cancers (Basel), 2022. **14**(19).
65. Henry, N.L. and D.F. Hayes, *Cancer biomarkers.* Mol, Oncol, 2012(1878-0261 (Electronic)).
66. Calu, V., et al., *Key biomarkers within the colorectal cancer related inflammatory microenvironment.* Sci Rep, 2021. **11**(1): p. 7940.
67. Chao, M. and P. Gibbs, *Caution is required before recommending routine carcinoembryonic antigen and imaging follow-up for patients with early-stage colon cancer.* J Clin Oncol, 2009. **27**(36): p. e279-80; author reply e281.
68. Mo, S., et al., *Early detection and prognosis prediction for colorectal cancer by circulating tumour DNA methylation haplotypes: A multicentre cohort study.* EClinicalMedicine, 2023. **55**: p. 101717.
69. Jelski, W. and B. Mroczko, *Biochemical Markers of Colorectal Cancer - Present and Future.* Cancer Manag Res, 2020. **12**: p. 4789-4797.
70. Dmello, R.S., S.Q. To, and A.L. Chand, *Therapeutic Targeting of the Tumour Microenvironment in Metastatic Colorectal Cancer.* Int J Mol Sci, 2021. **22**(4).

71. Malki, A., et al., *Molecular Mechanisms of Colon Cancer Progression and Metastasis: Recent Insights and Advancements.* Int J Mol Sci, 2020. **22**(1).
72. Stein, U., *MACC1 - a novel target for solid cancers.* Expert Opin Ther Targets, 2013. **17**(9): p. 1039-52.
73. Radhakrishnan, H., et al., *MACC1-the first decade of a key metastasis molecule from gene discovery to clinical translation.* Cancer Metastasis Rev, 2018. **37**(4): p. 805-820.
74. Ilm, K., et al., *High MACC1 expression in combination with mutated KRAS G13 indicates poor survival of colorectal cancer patients.* Mol Cancer, 2015. **14**: p. 38.
75. Vuaroqueaux, V., et al., *Elevated MACC1 Expression in Colorectal Cancer Is Driven by Chromosomal Instability and Is Associated with Molecular Subtype and Worse Patient Survival.* Cancers (Basel), 2022. **14**(7).
76. Duan, J., et al., *MACC1 decreases the chemosensitivity of gastric cancer cells to oxaliplatin by regulating FASN expression.* Oncol Rep, 2017. **37**(5): p. 2583-2592.
77. Wang, C., et al., *MACC1 mediates chemotherapy sensitivity of 5-FU and cisplatin via regulating MCT1 expression in gastric cancer.* Biochem Biophys Res Commun, 2017. **485**(3): p. 665-671.
78. Zhang, R., et al., *Knockdown of MACC1 expression increases cisplatin sensitivity in cisplatin-resistant epithelial ovarian cancer cells.* Oncol Rep, 2016. **35**(4): p. 2466-72.
79. Stein, U., M. Dahlmann, and W. Walther, *MACC1 - more than metastasis? Facts and predictions about a novel gene.* J Mol Med (Berl), 2010. **88**(1): p. 11-8.
80. Boardman, L.A., *Overexpression of MACC1 leads to downstream activation of HGF/MET and potentiates metastasis and recurrence of colorectal cancer.* Genome Med, 2009. **1**(4): p. 36.
81. Boccaccio, C. and P.M. Comoglio, *Invasive growth: a MET-driven genetic programme for cancer and stem cells.* Nat Rev Cancer, 2006. **6**(8): p. 637-45.
82. Zincke, F., *Biomarker based therapies in high risk cancer patients - MACC1 as molecular target*, in *Natural Sciences*. 2019, Humboldt University Berlin: Berlin.
83. Gu, H. and B.G. Neel, *The "Gab" in signal transduction.* Trends Cell Biol, 2003. **13**(3): p. 122-30.
84. Lavoie, H., J. Gagnon, and M. Therrien, *ERK signalling: a master regulator of cell behaviour, life and fate.* Nat Rev Mol Cell Biol, 2020. **21**(10): p. 607-632.
85. Katz, M., I. Amit, and Y. Yarden, *Regulation of MAPKs by growth factors and receptor tyrosine kinases.* Biochim Biophys Acta, 2007. **1773**(8): p. 1161-76.
86. Juneja, M., et al., *Promoter identification and transcriptional regulation of the metastasis gene MACC1 in colorectal cancer.* Mol Oncol, 2013. **7**(5): p. 929-43.
87. Zhang, W. and H.T. Liu, *MAPK signal pathways in the regulation of cell proliferation in mammalian cells.* Cell Research, 2002. **12**(1): p. 9-18.
88. Kortum, B., et al., *Combinatorial treatment with statins and niclosamide prevents CRC dissemination by unhinging the MACC1-beta-catenin-S100A4 axis of metastasis.* Oncogene, 2022. **41**(39): p. 4446-4458.
89. Qian, L.Q., et al., *Downregulation of MACC1 inhibits the viability, invasion and migration and induces apoptosis in esophageal carcinoma cells through the phosphatase and tensin homolog/phosphoinositide 3-kinase/protein kinase B signaling pathway.* Oncol Lett, 2017. **14**(4): p. 4897-4905.
90. Radhakrishnan, H., et al., *MACC1 regulates Fas mediated apoptosis through STAT1/3 - Mcl-1 signaling in solid cancers.* Cancer Lett, 2017. **403**: p. 231-245.
91. Meng, F., et al., *MACC1 down-regulation inhibits proliferation and tumourigenicity of nasopharyngeal carcinoma cells through Akt/beta-catenin signaling pathway.* PLoS One, 2013. **8**(4): p. e60821.
92. Juneja, M., et al., *Statin and rottlerin small-molecule inhibitors restrict colon cancer progression and metastasis via MACC1.* PLoS Biol, 2017. **15**(6): p. e2000784.

93. Gohlke, B.O., et al., *Real-world evidence for preventive effects of statins on cancer incidence: A trans-Atlantic analysis.* Clin Transl Med, 2022. **12**(2): p. e726.
94. Kokoszynska, K., et al., *Unexpected domain composition of MACC1 links MET signaling and apoptosis.* Acta Biochim Pol, 2009. **56**(2): p. 317-23.
95. Imbastari, F., et al., *MACC1 regulates clathrin-mediated endocytosis and receptor recycling of transferrin receptor and EGFR in colorectal cancer.* Cell Mol Life Sci, 2021. **78**(7): p. 3525-3542.
96. Zhang, X., et al., *MACC1 promotes pancreatic cancer metastasis by interacting with the EMT regulator SNAI1.* Cell Death Dis, 2022. **13**(11): p. 923.
97. Kim, H.J., et al., *DBC1 regulates Wnt/beta-catenin-mediated expression of MACC1, a key regulator of cancer progression, in colon cancer.* Cell Death Dis, 2018. **9**(8): p. 831.
98. Zhang, K., et al., *MACC1 is involved in the regulation of proliferation, colony formation, invasion ability, cell cycle distribution, apoptosis and tumorigenicity by altering Akt signaling pathway in human osteosarcoma.* Tumour, Biol, 2014(1423-0380 (Electronic)).
99. Basti, A., et al., *Core-Clock Genes Regulate Proliferation and Invasion via a Reciprocal Interplay with MACC1 in Colorectal Cancer Cells.* Cancers (Basel), 2022. **14**(14).
100. Koelzer, V.H., et al., *Heterogeneity analysis of Metastasis Associated in Colon Cancer 1 (MACC1) for survival prognosis of colorectal cancer patients: a retrospective cohort study.* BMC Cancer, 2015. **15**: p. 160.
101. Lemos, C., et al., *MACC1 Induces Tumor Progression in Transgenic Mice and Colorectal Cancer Patients via Increased Pluripotency Markers Nanog and Oct4.* Clin Cancer Res, 2016. **22**(11): p. 2812-24.
102. Güllü, N., et al., *Saffron Crudes and Compounds Restrict MACC1-Dependent Cell Proliferation and Migration of Colorectal Cancer Cells. LID - 10.3390/cells9081829 [doi] LID - 1829.* Cells,, 2020(2073-4409 (Electronic)).
103. Kobelt, D., et al., *The newly identified MEK1 tyrosine phosphorylation target MACC1 is druggable by approved MEK1 inhibitors to restrict colorectal cancer metastasis.* Oncogene, 2021. **40**(34): p. 5286-5301.
104. Imbastari, F., *Impact of MACC1 in Cargo Specific Clathrin-Mediated Endocytosis* in *Natural Sciences*. 2019, Humboldt University Berlin.
105. Yasunaga, M., J.J. Ipsaro, and A. Mondragon, *Structurally similar but functionally diverse ZU5 domains in human erythrocyte ankyrin.* J Mol Biol, 2012. **417**(4): p. 336-50.
106. Antas, P., et al., *SH3BP4 Regulates Intestinal Stem Cells and Tumorigenesis by Modulating beta-Catenin Nuclear Localization.* Cell Rep, 2019. **26**(9): p. 2266-2273 e4.
107. Tosoni, D., et al., *TTP specifically regulates the internalization of the transferrin receptor.* Cell, 2005. **123**(5): p. 875-88.
108. Kim, Y.M., et al., *SH3BP4 is a negative regulator of amino acid-Rag GTPase-mTORC1 signaling.* Mol Cell, 2012. **46**(6): p. 833-46.
109. Du, Z. and C.M. Lovly, *Mechanisms of receptor tyrosine kinase activation in cancer.* Mol Cancer, 2018. **17**(1): p. 58.
110. Regad, T., *Targeting RTK Signaling Pathways in Cancer.* Cancers (Basel), 2015. **7**(3): p. 1758-84.
111. Lemmon, M.A. and J. Schlessinger, *Cell signaling by receptor tyrosine kinases.* Cell, 2010. **141**(7): p. 1117-34.
112. Paul, M.K., & Mukhopadhyay, A. K. , *Tyrosine kinase - Role and significance in Cancer.* International journal of medical sciences, , 2004: p. 101–115.
113. Wu, P., M.H. Clausen, and T.E. Nielsen, *Allosteric small-molecule kinase inhibitors.* Pharmacology & Therapeutics, 2015. **156**: p. 59-68.
114. Gumireddy, K., et al., *A non-ATP-competitive inhibitor of BCR-ABL overrides imatinib resistance.* Proceedings of the National Academy of Sciences, 2005. **102**(6): p. 1992-1997.

115. Zhao, Z., et al., *Determining Cysteines Available for Covalent Inhibition Across the Human Kinome.* Journal of Medicinal Chemistry, 2017. **60**(7): p. 2879-2889.
116. Zhang, Y., et al., *Function of the c-Met receptor tyrosine kinase in carcinogenesis and associated therapeutic opportunities.* Mol Cancer, 2018. **17**(1): p. 45.
117. Joffre, C., et al., *A direct role for Met endocytosis in tumorigenesis.* Nat Cell Biol, 2011. **13**(7): p. 827-37.
118. Herynk, M.H., Zhang, J., Parikh, N. U., & Gallick, G. E. , *Activation of Src by c-Met overexpression mediates metastatic properties of colorectal carcinoma cells.* Journal of experimental therapeutics & oncology, 2007. **6(3)**: p. 205–217.
119. Clague, M.J., *Met receptor: a moving target.* Sci Signal, 2011. **4**(190): p. pe40.
120. Guo, Y.J., et al., *ERK/MAPK signalling pathway and tumorigenesis (Review).* Exp Ther Med, 2020. **19**(3): p. 1997-2007.
121. Uhlitz, F., et al., *An immediate-late gene expression module decodes ERK signal duration.* Mol Syst Biol, 2017. **13**(5): p. 928.
122. Murphy, L.O., et al., *Molecular interpretation of ERK signal duration by immediate early gene products.* Nat Cell Biol, 2002. **4**(8): p. 556-64.
123. Casar, B., A. Pinto, and P. Crespo, *Essential Role of ERK Dimers in the Activation of Cytoplasmic but Not Nuclear Substrates by ERK-Scaffold Complexes.* Molecular Cell, 2008. **31**(5): p. 708-721.
124. Casar, B., A. Pinto, and P. Crespo, *ERK dimers and scaffold proteins: unexpected partners for a forgotten (cytoplasmic) task.* Cell Cycle, 2009. **8**(7): p. 1007-13.
125. Jin, W., *Regulation of Src Family Kinases during Colorectal Cancer Development and Its Clinical Implications.* Cancers (Basel), 2020. **12**(5).
126. Kim, L.C., L. Song, and E.B. Haura, *Src kinases as therapeutic targets for cancer.* Nat Rev Clin Oncol, 2009. **6**(10): p. 587-95.
127. Ortiz, M.A., et al., *Src family kinases, adaptor proteins and the actin cytoskeleton in epithelial-to-mesenchymal transition.* Cell Commun Signal, 2021. **19**(1): p. 67.
128. Higuchi, M., et al., *Paradoxical activation of c-Src as a drug-resistant mechanism.* Cell Rep, 2021. **34**(12): p. 108876.
129. Martellucci, S., et al., *Src Family Kinases as Therapeutic Targets in Advanced Solid Tumors: What We Have Learned so Far.* Cancers (Basel), 2020. **12**(6).
130. Harvey, R., et al., *pp60c-src variants containing lesions that affect phosphorylation at tyrosines 416 and 527.* Mol Cell, Biol, 1989(0270-7306 (Print)).
131. Kopetz, S., et al., *Src activity is modulated by oxaliplatin and correlates with outcomes after hepatectomy for metastatic colorectal cancer.* Bmc Cancer, 2014. **14**.
132. Roskoski, R., *Src kinase regulation by phosphorylation and dephosphorylation.* Biochemical and Biophysical Research Communications, 2005. **331**(1): p. 1-14.
133. Chen, J., et al., *The role of Src in colon cancer and its therapeutic implications.* Clin Colorectal Cancer, 2014. **13**(1): p. 5-13.
134. Emaduddin, M., et al., *Cell growth, global phosphotyrosine elevation, and c-Met phosphorylation through Src family kinases in colorectal cancer cells.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(7): p. 2358-2362.
135. Mao, W., et al., *Activation of c-Src by receptor tyrosine kinases in human colon cancer cells with high metastatic potential.* Oncogene, 1997. **15**(25): p. 3083-3090.
136. Sirvent, A., et al., *Src family tyrosine kinases-driven colon cancer cell invasion is induced by Csk membrane delocalization.* Oncogene, 2010. **29**(9): p. 1303-15.
137. Lopez, J., et al., *Src tyrosine kinase inhibits apoptosis through the Erk1/2- dependent degradation of the death accelerator Bik.* Cell Death Differ, 2012. **19**(9): p. 1459-69.

138. Wang, C., et al., *GPS 5.0: An Update on the Prediction of Kinase-specific Phosphorylation Sites in Proteins.* Genomics Proteomics Bioinformatics, 2020. **18**(1): p. 72-80.
139. Trost, B. and A. Kusalik, *Computational prediction of eukaryotic phosphorylation sites.* Bioinformatics, 2011. **27**(21): p. 2927-35.
140. Blom, N., et al., *Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence.* Proteomics, 2004. **4**(6): p. 1633-49.
141. Blom, N., Gammeltoft, S., & Brunak, S. , *Sequence and structure-based prediction of eukaryotic protein phosphorylation sites.* Journal of molecular biology, 1999(294(5)): p. 1351–1362.
142. Kuhlman, B. and P. Bradley, *Advances in protein structure prediction and design.* Nat Rev Mol Cell Biol, 2019. **20**(11): p. 681-697.
143. Bertoline, L.M.F., et al., *Before and after AlphaFold2: An overview of protein structure prediction.* Front Bioinform, 2023. **3**: p. 1120370.
144. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
145. Borkakoti, N. and J.M. Thornton, *AlphaFold2 protein structure prediction: Implications for drug discovery.* Curr Opin Struct, Biol, 2023(1879-033X (Electronic)).
146. Dubochet, J., et al., *Cryo-electron microscopy of vitrified specimens.* Quarterly Reviews of Biophysics, 1988. **21**(2): p. 129-228.
147. Murata, K. and M. Wolf, *Cryo-electron microscopy for structural analysis of dynamic biological macromolecules.* Biochim Biophys Acta Gen Subj, 2018. **1862**(2): p. 324-334.
148. Renaud, J.P., et al., *Cryo-EM in drug discovery: achievements, limitations and prospects.* Nat Rev Drug Discov, 2018. **17**(7): p. 471-492.
149. Samant, S. *The Rise of Cryo-EM Among Structural Characterization Methods: NMR, X-Ray Crystallography*. Reviewing NMR, X-Ray crystallography and electron microscopy 2020 [cited 2023.
150. AlQuraishi, M., *Machine learning in protein structure prediction.* Curr Opin Chem Biol, 2021. **65**: p. 1-8.
151. Burke, D.F., et al., *Towards a structurally resolved human protein interaction network.* Nat Struct Mol Biol, 2023. **30**(2): p. 216-225.
152. Gutnik, D.A.-O.X., et al., *Using AlphaFold Predictions in Viral Research.* Curr Issues Mol, Biol, 2023(1467-3045 (Electronic)).
153. Fontana, P., et al., *Structure of cytoplasmic ring of nuclear pore complex by integrative cryo-EM and AlphaFold.* Science, 2023. **376**(6598): p. eabm9326.
154. Mirdita, M., et al., *ColabFold: making protein folding accessible to all.* Nat Methods, 2022. **19**(6): p. 679-682.
155. D, L. *Why AlphaFold won't revolutionise drug discovery*. OPINION 2022.
156. Ruff, K.M. and R.V. Pappu, *AlphaFold and Implications for Intrinsically Disordered Proteins.* J Mol Biol, 2021. **433**(20): p. 167208.
157. Olsvik, H.L. and T. Johansen, *AlphaFold-multimer predicts ATG8 protein binding motifs crucial for autophagy research.* PLoS Biol, 2023. **21**(2): p. e3002002.
158. Callaway, E. *The entire protein universe': AI predicts shape of nearly every known protein*. The world this week News in focus 2022; 608, 15-16:[
159. Stephens, D.C., et al., *In the Age of Machine Learning Cryo-EM Research is Still Necessary: A Path toward Precision Medicine.* Advanced Biology, 2023. **7**(8): p. 2300122.
160. Danielli, L., et al., *Quantifying the distribution of protein oligomerization degree reflects cellular information capacity.* Sci Rep, 2020. **10**(1): p. 17689.

161. Hashimoto, K. and A.R. Panchenko, *Mechanisms of protein oligomerization, the critical role of insertions and deletions in maintaining different oligomeric states.* Proc Natl Acad Sci U S A, 2010. **107**(47): p. 20352-7.
162. Baek, M., et al., *GalaxyHomomer: a web server for protein homo-oligomer structure prediction from a monomer sequence or structure.* Nucleic Acids Res, 2017. **45**(W1): p. W320-W324.
163. Evans, R., et al., *Protein complex prediction with AlphaFold-Multimer.* bioRxiv 2022. **2021.10.04.463034**.
164. Dang, D.T., *Molecular Approaches to Protein Dimerization: Opportunities for Supramolecular Chemistry.* Front Chem, 2022. **10**: p. 829312.
165. Ispolatov, I., et al., *Binding properties and evolution of homodimers in protein-protein interaction networks.* Nucleic Acids, Res, 2005(1362-4962 (Electronic)).
166. Marianayagam, N.J., M. Sunde, and J.M. Matthews, *The power of two: protein dimerization in biology.* Trends Biochem Sci, 2004. **29**(11): p. 618-25.
167. Gaber, A.A.-O. and M.A.-O. Pavšič, *Modeling and Structure Determination of Homo-Oligomeric Proteins: An Overview of Challenges and Current Approaches. LID - 10.3390/ijms22169081 [doi] LID - 9081.* Int, J. Mol Sci, 2021(1422-0067 (Electronic)).
168. Drury, L.J., et al., *Monomeric and dimeric CXCL12 inhibit metastasis through distinct CXCR4 interactions and signaling pathways.* Proc Natl Acad Sci U S A, 2011. **108**(43): p. 17655-60.
169. Pfleger, K.D., R.M. Seeber, and K.A. Eidne, *Bioluminescence resonance energy transfer (BRET) for the real-time detection of protein-protein interactions.* Nat Protoc, 2006. **1**(1): p. 337-45.
170. Rao, V.S., et al., *Protein-protein interaction detection: methods and analysis.* Int J Proteomics, 2014. **2014**: p. 147648.
171. Ivanov, A.A., F.R. Khuri, and H. Fu, *Targeting protein-protein interactions as an anticancer strategy.* Trends Pharmacol Sci, 2013. **34**(7): p. 393-400.
172. Perkins, J.R., et al., *Transient protein-protein interactions: structural, functional, and network properties.* Structure, 2010. **18**(10): p. 1233-43.
173. Farooq, Q.U.A., et al., *Protein-protein interactions: Methods, databases, and applications in virus-host study.* World J Virol, 2021. **10**(6): p. 288-300.
174. Ryan, D.P. and J.M. Matthews, *Protein–protein interactions in human disease.* Current Opinion in Structural Biology, 2005. **15**(4): p. 441-446.
175. Trepte, P., *Establishment and Systematic Benchmarking of Innovative Methods for Quantitative Interactome Mapping in Mammalian Cells*, in *Biology, Chemistry and Pharmacy*. 2016, Freie Universität Berlin: Berlin.
176. Nooren, I.M.A. and J.M. Thornton, *Diversity of protein–protein interactions.* The EMBO Journal, 2003. **22**(14): p. 3486-3492.
177. Jones, S. and J.M. Thornton, *Principles of protein-protein interactions.* Proc Natl Acad Sci, U. S. A., 1996(0027-8424 (Print)).
178. Coppolino MG, D.S., *Ligand-specific, transient interaction between integrins and calreticulin during cell adhesion to extracellular matrix proteins is dependent upon phosphorylation/dephosphorylation events.* Biochem J. , 1999 **May 15;340 (Pt 1)(Pt 1)**: p. 41-50.
179. Tarunina, M. and J.R. Jenkins, *Human p53 binds DNA as a protein homodimer but monomeric variants retain full transcription transactivation activity.* Oncogene, 1993. **8(11)**(0950-9232 (Print)): p. 3165–3173.
180. Rubin, I. and Y. Yarden, *The basic biology of HER2.* Ann Oncol, 2001. **12 Suppl 1**: p. S3-8.

181. Burckhardt, C.J., J.D. Minna, and G. Danuser, *Co-immunoprecipitation and semi-quantitative immunoblotting for the analysis of protein-protein interactions.* STAR Protocols, 2021. **2**(3).
182. Shi, G., *Immunoprecipitation (IP) Principles and Factors to Consider to Obtain Quality Results*, in *BenchSci*. 2018.
183. Vidal, M. and S. Fields, *The yeast two-hybrid assay: still finding connections after 25 years.* Nat, Methods, 2014(1548-7105 (Electronic)).
184. Fields, S. and O. Song, *A novel genetic system to detect protein-protein interactions.* Nature,, 1989(0028-0836 (Print)).
185. Geldermalsen, M.v. *An Overview of Yeast Two-Hybrid (Y2H) Screening*. Protein Expression and Analysis 2015.
186. Bruckner, A., et al., *Yeast two-hybrid, a powerful tool for systems biology.* Int J Mol Sci, 2009. **10**(6): p. 2763-2788.
187. Machleidt, T., et al., *NanoBRET--A Novel BRET Platform for the Analysis of Protein-Protein Interactions.* ACS Chem Biol, 2015. **10**(8): p. 1797-804.
188. Vogel, S.S., S.V. Thaler C Fau - Koushik, and S.V. Koushik, *Fanciful FRET.* Sci, Stke, 2006(1525-8882 (Electronic)).
189. Morell, M., S. Ventura, and F.X. Avilés, *Protein complementation assays: Approaches for the in vivo analysis of protein interactions.* FEBS Letters, 2009. **583**(11): p. 1684-1691.
190. Johnsson, N. and A. Varshavsky, *Split ubiquitin as a sensor of protein interactions in vivo.* Proc Natl Acad Sci, U. S. A., 1994(0027-8424 (Print)).
191. Förster, T., *Zwischenmolekulare Energiewanderung und Fluoreszenz.* Annalen der Physik, 1948. **437**(1-2): p. 55-75.
192. Pfleger, K.D. and K.A. Eidne, *Illuminating insights into protein-protein interactions using bioluminescence resonance energy transfer (BRET).* Nat Methods, 2006. **3**(3): p. 165-74.
193. Shimomura O Fau - Johnson, F.H., Y. Johnson Fh Fau - Saiga, and Y. Saiga, *Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea.* J. Cell Comp Physiol, 1962(0095-9898 (Print)).
194. Heim, R. and R.Y. Tsien, *Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer.* Current Biology, 1996. **6**(2): p. 178-182.
195. Trepte, P., et al., *LuTHy: a double-readout bioluminescence-based two-hybrid technology for quantitative mapping of protein-protein interactions in mammalian cells.* Mol Syst Biol, 2018. **14**(7): p. e8071.
196. England, C.G., E.B. Ehlerding, and W. Cai, *NanoLuc: A Small Luciferase Is Brightening Up the Field of Bioluminescence.* Bioconjug Chem, 2016. **27**(5): p. 1175-1187.
197. Esposito, M., S. Ganesan, and Y. Kang, *Emerging strategies for treating metastasis.* Nat Cancer, 2021. **2**(3): p. 258-270.
198. Hagemann, C., et al., *Impact of MACC1 on human malignant glioma progression and patients' unfavorable prognosis.* Neuro Oncol, 2013. **15**(12): p. 1696-709.
199. Guzmán, C., et al., *ColonyArea: An ImageJ Plugin to Automatically Quantify Colony Formation in Clonogenic Assays.* PLOS ONE, 2014. **9**(3): p. e92444.
200. Zhou, Y., et al., *Metascape provides a biologist-oriented resource for the analysis of systems-level datasets.* Nat, Commun, 2019(2041-1723 (Electronic)).
201. Aleksander Sa Auid- Orcid: --- Fau - Balhoff, J., et al., *The Gene Ontology knowledgebase in 2023. LID - 10.1093/genetics/iyad031 [doi] LID - iyad031.* Genetics, 2023(1943-2631 (Electronic)).
202. Carbon, S., et al., *AmiGO: online access to ontology and annotation data.* Bioinformatics, 2009(1367-4811 (Electronic)).

203. Ashburner, M., et al., *Gene ontology: tool for the unification of biology. The Gene Ontology Consortium.* Nat, Genet, 2000(1061-4036 (Print)).
204. Consortium, T.U., *UniProt: the Universal Protein Knowledgebase in 2023.* Nucleic Acids Research, 2022. **51**(D1): p. D523-D531.
205. Gao, S., et al., *Role of overexpression of MACC1 and/or FAK in predicting prognosis of hepatocellular carcinoma after liver transplantation.* Int J Med Sci, 2014. **11**(3): p. 268-75.
206. Wang, L., et al., *MACC-1 Promotes Endothelium-Dependent Angiogenesis in Gastric Cancer by Activating TWIST1/VEGF-A Signal Pathway.* PLoS One, 2016. **11**(6): p. e0157137.
207. Dong, G., et al., *MACC1 and HGF are associated with survival in patients with gastric cancer.* Oncol Lett, 2018. **15**(3): p. 3207-3213.
208. Singh, V., et al., *Phosphorylation: Implications in Cancer.* Protein J, 2017. **36**(1): p. 1-6.
209. Cartwright, C.A., Meisler, A. I., & Eckhart, W., *Activation of the pp60c-src protein kinase is an early event in colonic carcinogenesis* Proceedings of the National Academy of Sciences of the United States of America, 1990. **87(2)**: p. 558–562.
210. Gao, J., et al., *Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.* Sci Signal, 2013. **6**(269): p. pl1.
211. Cerami, E., et al., *The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.* Cancer Discov, 2012. **2**(5): p. 401-4.
212. Tsukamoto, S., et al., *Clinical significance of osteoprotegerin expression in human colorectal cancer.* Clin Cancer Res, 2011. **17**(8): p. 2444-50.
213. Ilm, K., Fuchs, S., Mudduluru, G., & Stein, U. , *MACC1 is post-transcriptionally regulated by miR-218 in colorectal cancer.* Oncotarget, 2016. **7**: p. 53443–53458.
214. Schneider, C.A., W.S. Rasband, and K.W. Eliceiri, *NIH Image to ImageJ: 25 years of image analysis.* Nature Methods, 2012. **9**(7): p. 671-675.
215. Tehrani, S., et al., *Src phosphorylation of cortactin enhances actin assembly.* Proc Natl Acad Sci, U. S. A., 2007(0027-8424 (Print)).
216. Fincham, V.J. and M.C. Frame, *The catalytic activity of Src is dispensable for translocation to focal adhesions but controls the turnover of these structures during cell motility.* Embo, J., 1998(0261-4189 (Print)).
217. Reinecke, J.B., et al., *Regulation of Src trafficking and activation by the endocytic regulatory proteins MICAL-L1 and EHD1.* J Cell Sci, 2014. **127**(Pt 8): p. 1684-98.
218. Sandilands, E., et al., *RhoB and actin polymerization coordinate Src activation with endosome-mediated delivery to the membrane.* Dev Cell, 2004. **7**(6): p. 855-69.
219. Mariani, V., et al., *lDDT: a local superposition-free score for comparing protein structures and models using distance difference tests.* Bioinformatics, 2013. **29**(21): p. 2722-8.
220. Kim, E.J. and D.M. Helfman, *Characterization of the metastasis-associated protein, S100A4. Roles of calcium binding and dimerization in cellular localization and interaction with myosin.* J Biol Chem, 2003. **278**(32): p. 30063-73.
221. Braunstein, J., et al., *STATs dimerize in the absence of phosphorylation.* J Biol Chem, 2003. **278**(36): p. 34133-40.
222. Garton, M., et al., *Interplay of self-association and conformational flexibility in regulating protein function.* Philos Trans R Soc Lond B Biol Sci, 2018. **373**(1749).
223. Ko, J., et al., *GalaxyWEB server for protein structure prediction and refinement.* Nucleic Acids Res, 2012. **40**(Web Server issue): p. W294-7.
224. Hegedus, Z., et al., *Identification of beta-strand mediated protein-protein interaction inhibitors using ligand-directed fragment ligation.* Chem Sci, 2021. **12**(6): p. 2286-2293.
225. Watkins, A.M. and P.S. Arora, *Anatomy of beta-strands at protein-protein interfaces.* ACS Chem Biol, 2014. **9**(8): p. 1747-54.

226. Kostova, S., *Functional characterization of a novel p97:ASPL protein complex*, in *Biology, Chemistry and Pharmacy*. 2020, Freie Universität Berlin: Berlin.
227. Angers, S., et al., *Detection of beta 2-adrenergic receptor dimerization in living cells using bioluminescence resonance energy transfer (BRET).* Proc Natl Acad Sci, U. S. A., 2000(0027-8424 (Print)).
228. Percherancier, Y., et al., *Bioluminescence Resonance Energy Transfer Reveals Ligand-induced Conformational Changes in CXCR4 Homo- and Heterodimers*.* Journal of Biological Chemistry, 2005. **280**(11): p. 9895-9903.
229. Fontao, L., et al., *The interaction of plectin with actin: evidence for cross-linking of actin filaments by dimerization of the actin-binding domain of plectin.* J. Cell Sci, 2001(0021-9533 (Print)).
230. Wiche, G., *Role of plectin in cytoskeleton organization and dynamics.* J. Cell Sci, 1998(0021-9533 (Print)).
231. Wang, X., et al., *Characterization of LIMA1 and its emerging roles and potential therapeutic prospects in cancers.* Frontiers in Oncology, 2023. **13**.
232. Moon, S.-J., et al., *CTTN Overexpression Confers Cancer Stem Cell-like Properties and Trastuzumab Resistance via DKK-1/WNT Signaling in HER2 Positive Breast Cancer.* Cancers, 2023. **15**(4): p. 1168.
233. Luo, M.L., et al., *Amplification and overexpression of CTTN (EMS1) contribute to the metastasis of esophageal squamous cell carcinoma by promoting cell migration and anoikis resistance.* Cancer, Res, 2006(0008-5472 (Print)).
234. Koveitypour, Z., et al., *Signaling pathways involved in colorectal cancer progression.* Cell Biosci, 2019. **9**: p. 97.
235. Shin, J.S., et al., *Serum starvation induces G1 arrest through suppression of Skp2-CDK2 and CDK4 in SK-OV-3 cells.* Int, J. Oncol, 2008(1019-6439 (Print)).
236. Kobelt, D., W. Walther, and U.S. Stein, *Real-Time Cell Migration Monitoring to Analyze Drug Synergism in the Scratch Assay Using the IncuCyte System.* Methods Mol, Biol, 2021(1940-6029 (Electronic)).
237. Stella, M.C. and P.M. Comoglio, *HGF: a multifunctional growth factor controlling cell scattering.* Int, J. Biochem Cell Biol, 1999(1357-2725 (Print)).
238. Grotegut, S., et al., *Hepatocyte growth factor induces cell scattering through MAPK/Egr-1-mediated upregulation of Snail.* Embo, J., 2006(0261-4189 (Print)).
239. Xie, Y.H., Y.X. Chen, and J.Y. Fang, *Comprehensive review of targeted therapy for colorectal cancer.* Signal Transduct Target Ther, 2020. **5**(1): p. 22.
240. Wang, J., et al., *Metastatic patterns and survival outcomes in patients with stage IV colon cancer: A population-based analysis.* Cancer Med, 2020. **9**(1): p. 361-373.
241. Hohmann, T., et al., *MACC1 driven alterations in cellular biomechanics facilitate cell motility in glioblastoma.* Cell Commun Signal, 2020. **18**(1): p. 85.
242. Weng, Z., et al., *Detection of Src homology 3-binding proteins, including paxillin, in normal and v-Src-transformed Balb/c 3T3 cells.* Journal of Biological Chemistry, 1993. **268**(20): p. 14956-14963.
243. Sueta, A., et al., *Differential role of MACC1 expression and its regulation of the HGF/c-Met pathway between breast and colorectal cancer.* Int J Oncol, 2015. **46**(5): p. 2143-2153.
244. Veracini, L., Franco, M., Boureux, A., Simon, V., Roche, S., & Benistant, C. , *Two functionally distinct pools of Src kinases for PDGF receptor signalling.* Biochemical Society transactions, 2005. **33(Pt 6)**: p. 1313–1315.
245. Donepudi, M. and M.D. Resh, *c-Src trafficking and co-localization with the EGF receptor promotes EGF ligand-independent EGF receptor activation and signaling.* Cellular Signalling, 2008. **20**(7): p. 1359-1367.

246. Chen, S., et al., *The role of metastasis-associated in colon cancer 1 (MACC1) in endometrial carcinoma tumorigenesis and progression.* Mol Carcinog, 2017. **56**(4): p. 1361-1371.
247. Iakoucheva, L.M., et al., *Intrinsic disorder in cell-signaling and cancer-associated proteins.* J Mol Biol, 2002. **323**(3): p. 573-84.
248. Wang, C., et al., *Structure of the ZU5-ZU5-UPA-DD tandem of ankyrin-B reveals interaction surfaces necessary for ankyrin function.* Proc Natl Acad Sci U S A, 2012. **109**(13): p. 4822-7.
249. Park, H.H., et al., *The death domain superfamily in intracellular signaling of apoptosis and inflammation.* Annu Rev Immunol, 2007. **25**: p. 561-86.
250. Ferrao, R. and H. Wu, *Helical assembly in the death domain (DD) superfamily.* Curr Opin Struct, Biol, 2012(1879-033X (Electronic)).
251. Park, H.H., et al., *Death domain assembly mechanism revealed by crystal structure of the oligomeric PIDDosome core complex.* Cell,, 2007(0092-8674 (Print)).
252. Wang, R., et al., *Autoinhibition of UNC5b revealed by the cytoplasmic domain structure of the receptor.* Mol Cell, 2009. **33**(6): p. 692-703.
253. Kiss, B., et al., *Crystal structure of the S100A4-nonmuscle myosin IIA tail fragment complex reveals an asymmetric target binding mechanism.* Proc Natl Acad Sci U S A, 2012. **109**(16): p. 6048-53.
254. Maignan, S., et al., *Crystal Structure of the Mammalian Grb2 Adaptor.* Science, 1995. **268**(5208): p. 291-293.
255. Khanal, P., Z. Jia, and X. Yang, *Cysteine residues are essential for dimerization of Hippo pathway components YAP2L and TAZ.* Sci Rep, 2018. **8**(1): p. 3485.
256. Banatao, D.R., Cascio, D., Crowley, C. S., Fleissner, M. R., Tienson, H. L., & Yeates, T. O, *An approach to crystallizing proteins by synthetic symmetrization.* Proceedings of the National Academy of Sciences of the United States of America 2006. **103(44), 16230–16235**.
257. Joseph, P.R., et al., *Proline substitution of dimer interface beta-strand residues as a strategy for the design of functional monomeric proteins.* Biophys J, 2013. **105**(6): p. 1491-501.
258. Kurochkina, N. and U. Guha, *SH3 domains: modules of protein-protein interactions.* Biophys Rev, 2013. **5**(1): p. 29-39.
259. E., S.T., *SH2 and SH3 domains: potential targets for anti-cancer drug design. .* Journal of pharmacological and toxicological methods, 1995. **34(3)**: p. 125–132.
260. Ahmed, Z., et al., *Grb2 monomer-dimer equilibrium determines normal versus oncogenic function.* Nat Commun, 2015. **6**: p. 7354.
261. Rudack, T., et al., *The Ras dimer structure.* Chem Sci, 2021. **12**(23): p. 8178-8189.
262. Ramagopal, U.A., et al., *Structure of the S100A4/myosin-IIA complex.* BMC Struct Biol, 2013. **13**: p. 31.
263. Lan, T.H., et al., *BRET evidence that beta2 adrenergic receptors do not oligomerize in cells.* Sci Rep, 2015. **5**: p. 10166.
264. Lai, X., et al., *Regulation of RPTPα–c-Src signalling pathway by miR-218.* The FEBS Journal, 2015. **282**(14): p. 2722-2734.
265. Shi, Z.M., et al., *Downregulation of miR-218 contributes to epithelial-mesenchymal transition and tumor metastasis in lung cancer by targeting Slug/ZEB2 signaling.* Oncogene, 2017. **36**(18): p. 2577-2588.
266. Huang, N., et al., *MiR-338-3p inhibits epithelial-mesenchymal transition in gastric cancer cells by targeting ZEB2 and MACC1/Met/Akt signaling.* Oncotarget, 2015. **6**(17): p. 15222-34.

267. Jones, D.A. and C.W. Benjamin, *Phosphorylation of Growth Factor Receptor Binding Protein-2 by pp60c-srcTyrosine Kinase.* Archives of Biochemistry and Biophysics, 1997. **337**(2): p. 143-148.
268. Bolós, V., et al., *The dual kinase complex FAK-Src as a promising therapeutic target in cancer.* Onco Targets, Ther, 2010(1178-6930 (Electronic)).
269. Mitra, S.K. and D.D. Schlaepfer, *Integrin-regulated FAK-Src signaling in normal and cancer cells.* Curr Opin Cell, Biol, 2006(0955-0674 (Print)).
270. Hohmann, T., et al., *MACC1-Induced Collective Migration Is Promoted by Proliferation Rather Than Single Cell Biomechanics.* Cancers (Basel), 2022. **14**(12).
271. Hohmann, T., U. Hohmann, and F. Dehghani, *MACC1-induced migration in tumors: Current state and perspective.* Front Oncol, 2023. **13**: p. 1165676.
272. Zhang, C., et al., *MicroRNA-338-3p suppresses cell proliferation, migration and invasion in human malignant melanoma by targeting MACC1.* Exp Ther Med, 2019. **18**(2): p. 997-1004.
273. Fraley, S.I., et al., *A distinctive role for focal adhesion proteins in three-dimensional cell motility.* Nature Cell Biology, 2010. **12**(6): p. 598-604.
274. Kai, F., A.P. Drain, and V.M. Weaver, *The Extracellular Matrix Modulates the Metastatic Journey.* Dev Cell, 2019. **49**(3): p. 332-346.

11 List of Relevant Publications

- Kobelt D, Dahlmann M, **Dumbani M,** et al. Small Ones to Fight a Big Problem-Intervention of Cancer Metastasis by Small Molecules. Cancers (Basel). 2020;12(6):1454.
- Basti A, Malhan D, **Dumbani M,** et al. Core-Clock Genes Regulate Proliferation and Invasion via a Reciprocal Interplay with MACC1 in Colorectal Cancer Cells. Cancers (Basel). 2022 Jul 16;14(14):3458.

12 Acknowledgements

It was an incredible journey filled with moments of excitement, setbacks, and lots of patience. First and foremost, I would like to thank Prof. Ulrike Stein and Prof. Dr. Wolfgang for believing in me and supporting me throughout this journey. This encouraged me to develop my critical thinking and pushed me to think outside the box. My sincere gratitude to my PhD committee members, Prof. Dr. Gerhard Wolber and Prof. Dr. Oliver Daumke for their guidance and valuable discussions that helped shape this work.

I am extremely grateful to my research supervisors, Dr. Fabian Zincke and Dr. Benedikt Kortüm. It was a pleasure working together and learning from you both. Fabian, you showed me the importance of consistency and inspired me to be a better scientist. Thank you for your patience in the lab and while proofreading my thesis. Benedikt, thank you for guiding me and encouraging me to look for innovative ways to solve problems. I was truly fortunate to have two great mentors.

My deepest gratitude to our collaboration partner Dr. Simona Kostova, AG E. Wanker. Thanks to your help and dedication we were able to move this project forward. I admired your positive and helpful attitude. You were a great role model and inspired me to be a better researcher. It was a privilege to have met a kind person like you.

I would like to acknowledge the help of Alexandra Forrai in setting up the BRET experiments. I am highly fortunate to have mentored such a talented and hardworking young scientist.

Special thanks to Dennis, Danai, Mathias, Nan, Nazli, Paul, Sebastian, Tam, and Yan for the positive energy and delightful discussions. I sincerely thank Janice, Pia, and Rita for their support in technical and administrative matters.

I would like to express my deepest gratitude to my family and friends who encouraged me to keep going. Thank you!