Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences
(Dr. rer. nat.)

presented by

Diplom-Biologin Seda Ballikaya
born in Üsküdar, Turkey
Oral examination:
28.07.2014

Activin Receptor Type 2 A (ACVR2A)-dependent Proteomic and Glycomic Alterations in a Microsatellite Unstable (MSI) Colorectal Cancer Cell Line Model System

Referees: Prof. Dr. rer. nat. Dieter Kübler

Prof. Dr. rer. nat. Jürgen Kopitz

Aileme…

Meiner Familie…

Ya Xızır,
bana değiştiremeyeceğim şeyleri kabul etmek için sabır,
değiştirebileceğim şeyleri değiştirmek için cesaret,
ve ikisini ayırt etmek için akıl ver

Gib mir die Ruhe hinzunehmen, was nicht zu ändern ist,
den Mut zu ändern, was zu ändern ist,
und die Weisheit, das eine vom anderen zu unterscheiden

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

Completing this dissertation has been a long journey that has encountered numerous ups and downs, moments of joy and aspiration as well as uncertainty, but never hopelessness or boredom. Therefore, I would like to start by saying thank you to all the people who enabled and supported my project during the last years.

First and foremost, my utmost thankfulness goes to my supervisors Dr. Johannes Gebert and Prof. Dr. Jürgen Kopitz. Their enthusiasm, optimism and patience together with their experience and scientific guidance have proved out to be of invaluable importance to me. I would like to express my sincere gratitude to them for providing me with many helpful suggestions and constant encouragement during the entire time of this work.

I would like to gratefully acknowledge Prof. Dr. Magnus von Knebel Doeberitz giving me the opportunity to work on this interesting project and accepting me as a part of his team. In addition, I would like to thank all the members of the Department of Applied Tumor Biology for creating a nice atmosphere in the group. I would like to thank Marcel Karl and Sigrun Himmelsbach for their great technical assistance and helping me to accomplish the studies. Special thanks go to Jennifer Lee, Gergana Bozukova, Anna Schulz, Madeleine Sauer and Elena Prigge for not only being colleagues, but also friends. Thank you for the pleasant atmosphere and always helping me.

I am very grateful to Dr. Martina Schnölzer for her interest in my project and kind cooperation as well as the helpful feedback.

I would like to thank the members of my thesis committee Prof. Dr. Dieter Kübler and Prof. Dr. Bernd Bukau for helpful discussions during the annual meetings, their input during progress reports and for their helpful suggestions on how to proceed with my project. I also want to thank Prof. Dr. Buselmaier and Prof Dr. Frings for taking their time to be a part of mine examining committee.

Many thanks to Sevil and Nil for the funny chats as well as for critical proof-reading.

Thank you to all my dear friends in Germany and Turkey, you have been the source of inspiration, laughter and empathetic listening that empowered me to get through difficult days of my personal and professional life. Especially I want to name my dearest friend, my cousin Betül for walking together on the same path and her continuous help. Thank you and I wish you the best of bests!

Finally, I would like to dedicate this thesis to my family, who has never stopped believing in me. My heart-felt gratitude goes to my sisters Belinda and Selina and my brother Berhat. Thank you for your support, prayers and encouragement. Lastly, my profound gratitude goes to my parents, who were and still are the motor driving my education. They supported me whenever possible and were a source of motivation for me during the entire period of my thesis.

SUMMARY

DNA mismatch repair-deficient colorectal tumors exhibit a high-frequency of microsatellite instability (MSI-H) and accumulate somatic frameshift mutations in genes harboring repetitive DNA sequences. Biallelic frameshift mutations in the A8 coding repeat of the activin receptor type 2 A (*ACVR2A* or *ACVR2*) gene occur at high frequency in these tumors thereby abrogating normal receptor and signaling function. Nevertheless, it is not clear if protein deficiency is modifying the glycosylation pattern as well as the whole proteomic constellation of cells, since it has been shown that proteomic and glycomic alterations have emerging significance in cancer cells. Here, we compensated the loss of function by reconstitution of ACVR2 into a MSI-H colon cancer cell line and analyzed its impact on the protein pattern of these cells. As a model system we used the MSI colorectal cancer cell line HCT116-AWE that enables doxycycline-inducible expression of target genes. Applying retroviral genomic targeting and recombination-mediated cassette exchange (RMCE) technology we have generated stable clones that allow dox-regulated expression of a single copy ACVR2 transgene. Fragment analysis, determination of the transgene transcript level, dox-inducible expression of wildtype ACVR2 protein and functional analysis by ligand-stimulated activation of signal transduction and expression of specific target genes confirmed successful gene reconstitution. Upon induction of receptor expression, glyco-gene chip expression analysis of ACVR2-induced versus -uninduced cells revealed ACVR2-dependent upregulation of the glycosyltransferase LNFG, an important regulator of Notch signaling. Moreover, metabolic labeling experiments showed a significant decrease in fucose incorporation and a modest increase in mannosamine uptake, indicating significant glycan alterations of newly synthesized proteins due to ACVR2 re-expression. By applying a Click-it chemistry approach and subsequent mass spectrometry analysis a list of proteins, differentially expressed between ACVR2-deficient and -proficient cells, was identified.

These results suggest that ACVR2 signaling can affect glycomic and proteomic modifications that might cause alterations in many cellular processes like growth suppression, cell death, cell adhesion and communication properties or invasion and metastasis of these MSI tumor cells. This study is of major relevance because it may

define and provide a novel source of MSI tumor-specific carbohydrate epitope changes in cell surface levels, thereby initiating analysis of glycan function related to MSI. At the same time new tumor markers could be established, which are not only monitors for diagnosis or therapy, but also represent the biological characteristics of cancer cells.

ZUSAMMENFASSUNG

Mikrosatelliten-Instabilität (MSI) tritt in etwa 15% aller Darmtumoren auf und beruht auf dem Funktionsverlust des zellulären DNS Mismatch-Reparatursystems (MMR). Das MMR-System ist für die Korrektur von Basenfehlpaarungen und kleinen Insertionen / Deletionen, die während der DNS-Replikation auftreten, von essentieller Bedeutung. MSI führt zur Anhäufung von somatischen Frameshift-Mutationen, insbesondere in Genen mit repetitiven DNS Sequenzen. In ~85% aller MSI-Darmtumoren kommt es zu Frameshift-Mutationen des Gens, das für den Activin Rezeptor Typ 2 A (ACVR2A oder ACVR2) kodiert. Hierbei handelt es sich um eine kleine Deletion einer Polyadenin-Mikrosatellitensequenz (A8) in Exon 10 dieses Gens, die beide Allele betrifft und damit zur Inaktivierung der normalen Signal-transduktionsfunktion des ACVR2 Rezeptors führt. Gegenwärtig ist noch unklar, wie sich der Funktionsverlust von ACVR2 auf das Proteom und auf postranslationale Proteinmodifikationen, insbesondere das Glykosylierungsmuster, auswirkt. Es gibt bereits eine Vielzahl von Hinweisen, dass neben Veränderungen im Tumor-Proteom auch spezifische Veränderungen im Tumor-Glykom auftreten können. Um den Einfluss des ACVR2-Funktionsverlusts sowohl auf das Glykom als auch auf das Proteom an einem geeigneten MSI-Tumorzelllinien Modell untersuchen zu können, kompensierten wir den ACVR2-Funktionsverlust in einer MSI-Darmkrebszelllinie durch die Rekonstitution des ACVR2-Gens. Dazu verwendeten wir die MSI-Darmkrebszelllinie HCT116-AWE, die eine Doxycyclin-induzierbare Expression von Zielgenen ermöglicht. Durch die Anwendung eines retroviralen genomischen Targetings und der Rekombinations-vermittelten Kassettenaustausch Technologie (RCME) stellten wir stabile Zellklone her, bei denen die erfolgreiche Rekonstitution des ACVR2-Gens auf DNA-, RNA- und Protein-Ebene sowie durch eine funktionelle

Analyse der Signaltransduktion bestätigt werden konnte. Eine RNA-Chip-Expressions-Analyse von Genen, die an der Glykosylierung beteiligt sind, zeigte außerdem eine ACVR2-abhängige Hochregulation der Glykosyltransferase LNFG. Dies belegt erstmals einen direkten Zusammenhang zwischen veränderter ACVR2-Signaltransduktion und der Expression einer Glykosyltransferase. Darüber hinaus ist LFNG ein wichtiger Regulator der Notch-Signaltransduktion. Radioaktive Markierungsexperimente zeigten eine signifikante Abnahme der Protein-Fucosylierung sowie eine geringe Zunahme in der Protein-Sialylierung. Damit konnten erstmals signifikante ACVR2-abhängige Änderungen in den Oligosaccharidstrukturen neu synthetisierter Glykoproteine nachgewiesen werden. Zur Identifizierung der davon betroffenen Glykoproteine als auch von ACVR2-abhängigen Veränderungen im Proteom der Zellen wurde die sogenannte Click-it Technologie zur Isolierung neu-synthetisierter Proteine mit massenspektrometrischer Identifizierung kombiniert. So konnte eine Liste der durch den ACVR2-Verlust in MSI-Kolonkarzinomzellen auftretenden Veränderungen im Proteom als auch der Proteine mit veränderter Glykosylierung erstellt werden. Diese Listen stellen die Grundlage für weiterführende Untersuchungen dar, mit dem Ziel, neue Tumormarker, die für die Diagnostik und Therapie Verwendung finden können, zu etablieren. Darüber hinaus eröffnen die beobachteten Veränderungen neue Möglichkeiten für detaillierte Untersuchungen zum Einfluss des ACVR2-Verlustes auf Tumor-relevante zelluläre Prozesse.

INDEX OF ABBREVIATIONS

ACVR2(A)	Activin receptor type 2 A
Act A	Activin A
AIM2	Absent in melanoma 2
APC	Adenomatous polyposis coli
BMP	Bone morphogenetic protein
BSA	Bovine serum albumin
CIN	Chromosomal instability
CRC	Colorectal cancer
DCC	Deleted in colorectal cancer
DMEM	Dulbecco's modified eagle's medium
DMSO	Dimethyl sulfoxide
dNTP	Deoxynucleotide triphosphate
dox	Doxycycline
DTT	1',4'-Dithiothreitol
ECL	Enhanced chemiluminescence
E. coli	Escherichia coli
EDTA	Ethylendiaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
FACS	Fluorescence activated cell sorting
FAP	Familial adenomatous polyposis
FBS	Fetal bovine serum
FSC	Forward scatter
Gan	Ganciclovir
GFP	Green fluorescent protein
HNPCC	Hereditary nonpolyposis colorectal cancer
HPLC	High pressure liquid chromatography
HRP	Horseradish peroxidise
HygTK	Hygromycin B phosphotransferase-thymidine kinase
IC	Immunocytology
IP	Immunoprecipitation
IRES	Internal ribosomal entry site

LOH	Loss of heterozygosity
LTR	Long terminal repeat
mAb	Monoclonal antibody
MLH	MutL homolog 1
MMR	Mismatch repair
MOI	Multiplicity of infection
MSH	MutS protein homolog
MSI	Microsatellite instability
MSS	Microsatellite stability
MW	Molecular weight
$OD_{n\ nm}$	Optical density at a wavelength of n nanometer
ON	Overnight
pAb	Polyclonal antibody
PAGE	Polyacrylamide-Gelelectrophorese
PBS	Phosphate buffer
PCR	Polymerase chain reaction
PE	Phycoerythrin
PMS	Postmeiotic segregation increased
qRT	Quantitative real time
RIPA	Radioimmunoprecipitation assay
RMCE	Recombinase-mediated cassette exchange
RNA	Ribonucleic acid
rpm	Rotations per minute
RPMI	Roswell Park Memorial Institute medium
RT	Room temperature
rtTA	reverse transcriptional transactivator gene
SA	Streptavidin
SDS	Sodium dodecyl sulfate
TBS	Tris buffered saline
tet	Tetracycline
TGF-ß	Transforming growth factor beta
TGFBR2	Transforming growth factor beta receptor 2
Tris	Trishydroxymethylaminomethane
Triton X-100	$Octylphenolpoly(ethyleneglycolether)_x$

1. INTRODUCTION

1.1 Colorectal Cancer

1.1.1 Epidemiology and Prevalence

Cancer is a leading cause of death worldwide, accounting for 7.6 million deaths (Globocan 2008, World Health Organization (WHO)). Colorectal cancer (CRC includes cancerous growth in the colon, rectum and appendix. According to the World Health organization (WHO), it is the third most common form of cancer and the second leading cause of cancer-related death in the western world. The incidence of colorectal cancer in the western world is relatively high and steadily increasing [1]. In the general population, the life time risk of developing sporadic colorectal cancer is 5-6% [2]. Population screening for sporadic colorectal cancer, mostly based on faecal blood tests, is planned or implemented in many European countries. [3-5]. Optimizing population screening with e.g. colonoscopy, CT colonography, or DNA-based stool tests is currently under investigation [6-8]. Colorectal carcinomas arise both sporadically at a median age of 67 years and hereditarily at a median age of 42 years [9]. Approximately 80% of colorectal cancers are sporadic tumors, 15% occur on a familial basis without a known underlying genetic event and 5% of all colorectal cancers develop as part of rare hereditary syndromes [10, 11]. These hereditary syndromes include hereditary non-polyposis colorectal cancer (HNPCC or Lynch syndrome), familial adenomatous polyposis syndrome (FAP) and even more rare, hamartomatous polyposis syndromes including Peutz-Jeghers, Juvenile Polyposis and Cowden Syndrome [10, 12]. Patients with FAP are almost certain to develop colorectal cancer before the age of 40 and patients with HNPCC have a 70-80% life-time risk of colorectal cancer and intensive screening or preventive surgery is indicated.

1.1.2 Progression of Colorectal Cancer

Colorectal adenoma, a benign neoplasm, remains the one reliable biomarker of risk for colorectal cancer. Adenoma patients have markedly higher risk of developing colorectal cancer and removing adenomatous polyps reduces the risk for future

colorectal cancer [13-15]. The colonic epithelium contains about 10^7 crypts that are the main morphologic units of the colorectal mucosa (Figure 1.1). Colonic epithelial cells originate from stem cells at the base of the crypt and migrate upward towards the surface epithelium layer. The cells differentiate and mature during their migration to the surface where they are replaced by a new generation of cells about every 3 – 6 days.

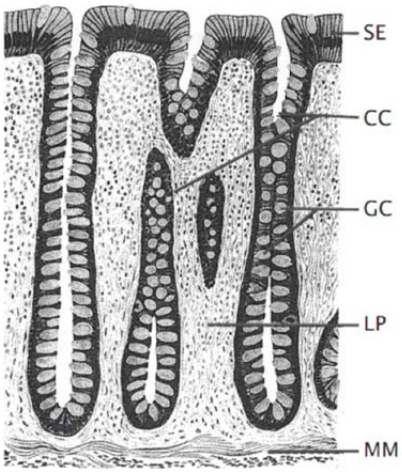


Figure 1.1: Morphology of normal colon tissue [16]. Labels show surface epithelium (SE), colon crypts (CC), goblet cells (GC), lamina propria (LP), and muscularis mucosa (MM).

The progression from an adenoma to a cancer passes through a series of defined histological stages referred to as the adenoma-carcinoma sequence (Figure 1.2) [17-19]. Colorectal cancer arises in epithelial cells lining the interior of the large intestine [20]. The pathologic transformation of normal colonic epithelium to benign tumors and finally invasive tumors requires several years and multiple genetic alterations. A genetic pathway of carcinogenesis is a process in which one particular type of genomic instability predominates, causing tumors to progress through characteristic histopathological stages with similar genetic alterations [21]. There are two main categories of genomic instability in colorectal cancer. The most common one is chromosomal instability (CIN), characterized by accumulation of numerical or structural chromosomal abnormalities. The second type is microsatellite instability (MSI), which is a consequence of impaired recognition and repair of mismatched bases in the daughter strand of the DNA during DNA replication. Either pathway is sufficient to drive colorectal carcinogenesis.

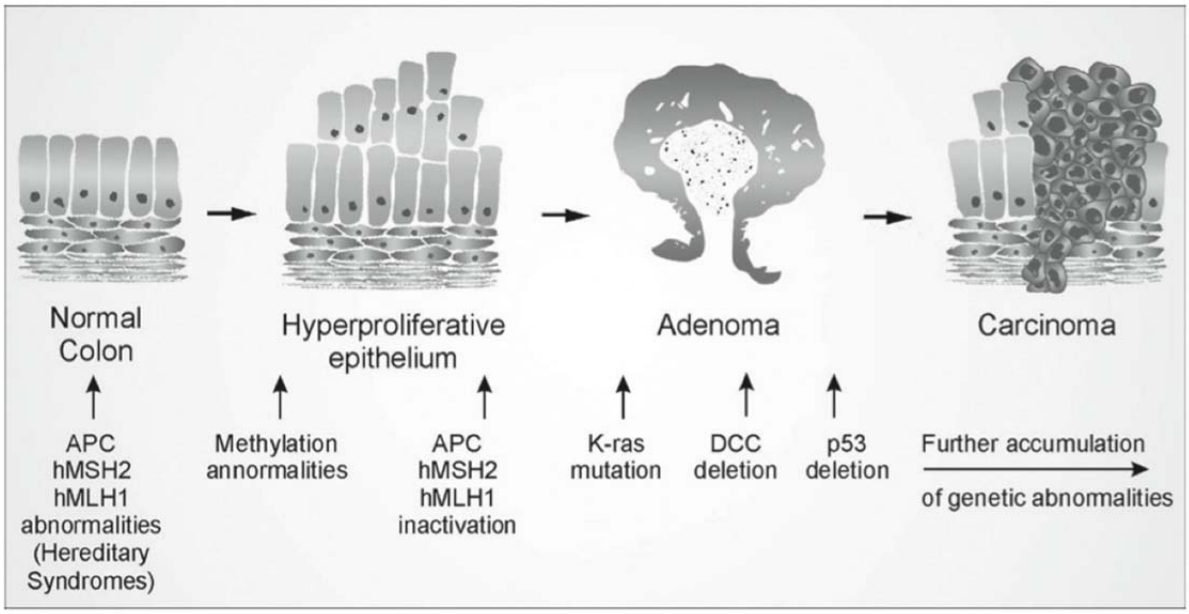


Figure 1.2: The adenoma-carcinoma sequence [22]. Colorectal cancers arise from areas of hyperproliferative epithelium through early, intermediate and finally late adenomas. APC indicates adenomatous polyposis coli, hMSH2 human mutS homolog 2, hMLH1 human mutL homolog 1, K-ras kirsten rat sarcoma viral oncogene homolog, DCC deleted in colorectal carcinoma and p53 indicates the p53 tumor suppressor.

1.1.3 Pathogenesis of MSI and CIN

Approximately 85% of colorectal cancers develop via the traditional chromosomal instability pathway [23]. Mostly one oncogene (K-ras) and three tumor suppressor genes (APC, DCC and p53) are sequentially genetically altered. Whereas the oncogene K-ras only requires a genetic event in one allele, the tumor suppressor genes require genetic events in both alleles [12] according to Knudson's two-hit hypothesis [24, 25]. The event that triggers the adenoma-carcinoma sequence and thereby leads to the development of malignant cancers is the activation of the Wnt signaling pathway in consequence of mutations in the adenomatous polyposis coli (APC) tumor suppressor gene [26, 27]. According to a current model, wild-type APC binds nuclear β-catenin and exports it to the cytoplasm [28], where it is phosphorylated by a complex of various proteins. Phosphorylated β-catenin becomes ubiquitinated and therefore targeted for degradation by the proteasome [29, 30]. Mutations in APC prevent degradation of β-catenin and lead to its accumulation in the nucleus [12]. Nuclear β-catenin functions in association with the HMG (high mobility group) box protein T cell factor (TCF)-4 as a transcriptional coactivator and thereby enables the expression of genes controlled by promoters with TCF4 binding

sites [31], for example, c-MYC [32] and Cyclin D2 [33]. These in turn activate cell proliferation. The subsequent genetic alteration occurring in the adenoma-carcinoma sequence affects the oncogene K-ras. The K-ras gene encodes a membrane-localized G-protein involved in signal transduction critical for normal proliferation. Mutations lead to constitutively activated Ras protein, which is stimulating cell proliferation [34]. Proceeding in the adenoma-carcinoma sequence, alterations in components of the TGF-ß (transforming growth factor beta) signaling pathway like SMAD2, SMAD4 and DCC (deleted in colorectal cancer) are affected. Through inactivation of these tumor suppressor genes cells become resistant to TGF-β-mediated growth suppression [35]. The tumor suppressor gene p53 was termed as “guardian of the genome” [36], as p53 mediates growth arrest or apoptosis as response to various cellular stresses, like DNA-damage or oncogenic activation. Therefore, inactivation of p53 allows the survival of aberrant cells. p53 is esteemed, for example, to be responsible for the transition from adenoma to carcinoma [34].

Although the majority of CRCs shows CIN, about 15% are attributable to defects in the DNA mismatch repair (MMR) system [37, 38]. It seems, however, that patients with MMR-deficient adenomas have greater risk of progressing into invasive cancer than those having CIN adenomas [21]. MMR-deficient CRCs may develop sporadically or in the context of hereditary non-polyposis colorectal cancer [39, 40]. The main feature of the MMR pathway is the interruption of the normal review and repair of DNA after replication. The MMR system is composed of at least 7 proteins: hMLH1, hMLH3, hMSH2, hMSH3, hMSH6, hPMS1 and hPMS2, which are associated with specific partners to form functional heterodimers [41]. DNA MMR deficiency induces a high number of mutational events characterized by small alterations at the nucleotide level like missense, nonsense and frameshift mutations [42]. These mutations occur mainly at short repetitive DNA sequences, termed microsatellites, because these structures are particularly prone to DNA polymerase slippage during DNA replication. The resulting phenotype is termed microsatellite instability (MSI). Five microsatellites (BAT25, BAT26, D5S346, D2S123, and D17S250) were identified as the most common sites of such mutations. Cancers that have mutations in two or more of these microsatellites are defined as high microsatellite instability (MSI-H) cancers; tumors with only one of these sites mutated are regarded as low microsatellite instability (MSI-L) cancers. Microsatellite stable (MSS) neoplasms do not have mutations in any of the five microsatellites. While CIN

cancer is characterized by large chromosomal alterations like allelic losses, chromosomal amplifications and translocations, as well as loss of heterozygosity (LOH), tumors with MSI show, after the initial loss of the MMR function, mutations in specific target genes [43, 44].

1.2 Microsatellite Instability

Depending on the length of the repeated unit, microsatellites are classified as mono-, di-, tri-, tetra-, penta- and hexanucleotide repeats. Microsatellites represent the most variable types of DNA sequences in the genome and are estimated to account for about 3% of the human genome. They are non-randomly distributed throughout the human genome [45], thus they can be distributed in non-coding regions (intragenic or intergenic) or in coding regions. Intragenic regions like promoters, 3`-untranslated regions and introns can be important regulators of gene expression, while intergenic regions could have functions in chromatin organization and recombination [46]. Because of their high variability, microsatellites in intergenic and non-coding regions are frequently used as molecular markers in forensics, paternity testing and linkage mapping. The maintenance of microsatellite stability is controlled by the MMR system, which is able to correct base substitutions and mismatches as well as insertion / deletion mutations, which can occur during DNA replication due to a polymerase slippage [46]. Figure 1.3 shows the role of the MMR system in maintaining the length of microsatellite sequences. This MMR system is comprised of two major heterodimeric protein complexes, MutS and MutL: initiation of the DNA repair is performed by recognition and binding to the mismatch mainly by MutSα (MSH2 and MSH6) or MutSß (MSH2 and MSH3), depending on the length and type of the mismatch. After recruitment of the key heterodimer MutLα (MLH1 and PMS2) the endonuclease activity of PMS2 introduces single-strand breaks proximal to the mismatch, followed by additional proteins that excise the mispaired DNA section, synthesize the missing nucleotides and finally ligate the corrected DNA.

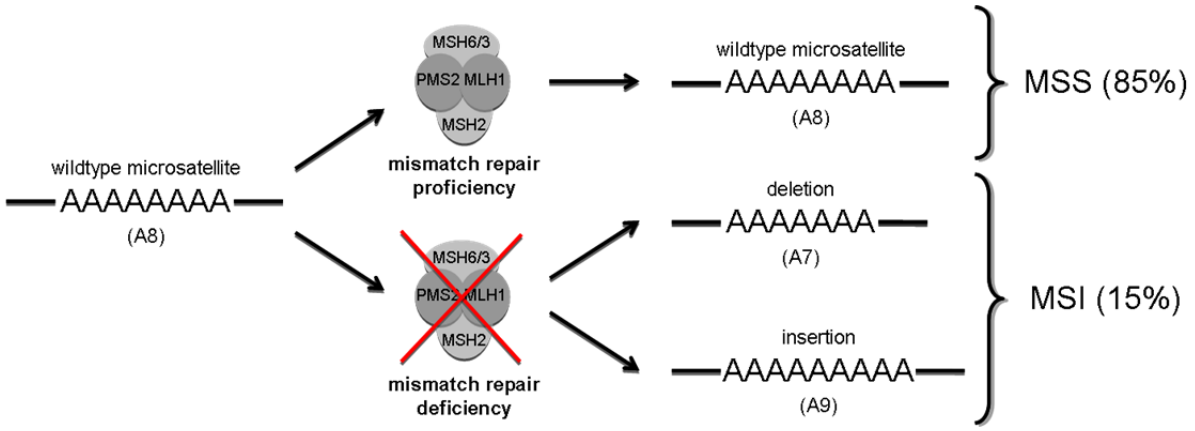


Figure 1.3: Molecular mechanisms underlying microsatellite instability (MSI) (modified according to [47]). For illustration, a microsatellite of eight adenine nucleotides (A8) is represented. Functional mismatch repair system recognizes slippages and repairs them. Accumulations of deletion (A7) or insertion (A9) mutations arise in 15% of all CRC cases during DNA replication due to a non-functional mismatch repair system.

In Lynch syndrome, germline mutations in one of four MMR genes (MLH1, MSH2, MSH6 and PMS2) are responsible for dysfunction of the DNA MMR system. However, in sporadic MSI tumors epigenetic silencing of MHL1 by promoter methylation is responsible for loss of DNA MMR function and consequently development of colorectal cancer [42, 48]. In addition to MLH1 promoter methylation, these sporadic MSI-H tumors are associated with BRAF mutations [49]. BRAF is a serine / threonine kinase involved in the MAPK signaling pathway [50]. The occurrence of MLH1 promoter hypermethylation and BRAF mutations distinguish HNPCC from sporadic MSI-H colon cancers. Heterozygous germline mutations in Lynch syndrome patients most frequently affect the MLH1 (50%) and MSH2 (40%) genes [51] and to a lesser extent PMS2 and MSH6 (10%). In addition to the germline mutated MMR gene, a second somatic mutation in the remaining wildtype allele is required to cause MMR deficiency in tumor cells of Lynch syndrome patients.

Mutations in microsatellites located in coding regions of expressed genes might lead to a loss of protein function or the translation of truncated proteins due to frameshift mutations. Frameshift-derived truncated proteins may harbor neopeptide tails that represent potentially antigenic epitopes capable to induce cellular or humoral immune responses [52, 53]. It is generally assumed that coding microsatellite frameshift mutations in some of these genes provide a growth advantage to MMR-deficient cells and hence drive MSI tumorigenesis [54]. Coding microsatellites

consisting one type of nucleotide (like the A8 repeat shown in Figure 1.3) are known as mononucleotide repeats (cMNRs). Woerner et al. has established a database of human cMNR mutations (www.seltarbase.org) and proposed a statistical model that allows the prediction of genes, that when mutated might provide a growth advantage to affected cells [54]. The statistical model is based on a sigmoid regression analysis aiming at the identification of genes involved in MSI carcinogenesis by their mutation frequency in cMNRs. A given cMNR length correlated with the average mutation rate, revealing genes with increased or decreased mutation frequency. Among the most frequently mutated genes frequently affected by cMNR mutations in MSI tumors are two members of the TGF-ß superfamily, *TGFßR2* (transforming growth factor beta receptor type 2) and *ACVR2A* (activin receptor type 2 A) [54-56]. In this thesis, the ACVR2A (or shortly ACVR2) is of particular interest, because it is a key mediator of the activin A-mediated signaling pathway that regulates normal growth and differentiation of colon epithelium. More than 80% of MSI colorectal tumors are known to exhibit biallelic cMNR frameshift mutations in the *ACVR2* gene that inactivate this pathway [57].

1.3. Activin Receptor Type 2

The initial discovery of activin was for its ability to regulate follicle stimulating hormone production and is among the first identified members of TGF-ß superfamily [58]. Activin exists in three isoforms that are homo- and heterodimers and it regulates cell differentiation, proliferation, and apoptosis in many epithelial and mesenchymal cells [59]. A family of transmembrane kinases, present on the cell surface, act as receptors for these TGF-ß superfamily ligands. These receptors fall into two distinct subfamilies known as type 1 and type 2 receptors that are distinguished by the level of sequence homology of their kinase domains and by other structural and functional features. Type 1 and type 2 receptors act cooperatively to bind ligand and transduce signals. Type 2 receptors bind ligand on their own, whereas the type 1 receptors bind ligand only when co-expressed with the corresponding type 2 receptors [60, 61]. Genetic and biochemical evidence indicates that co-expression of type 1 and type 2 receptors in the same cell is required for signaling [62, 63]. Both activin receptor type 1 (ACVR1) and activin receptor type 2 (ACVR2) are transmembrane proteins with ligand-binding activity in the extracellular domain and serine / threonine kinases in

their cytoplasmic domain. Potential ligands for activin receptors type 2 include activin A, activin B, and inhibin A [64], however they also work as receptor for a subset of bone morphogenetic proteins (BMP) [65, 66]. There are two subtypes of ACVR2 receptors, designated ACVR2A and ACVR2B. ACVR2A, also often termed ACVR2, is located on chromosome 2 and harbors 2 coding polyadenine tracts (A8). ACVR2B, located on chromosome 3, is 69% identical to ACVR2A and has similar binding characteristics to ACVR2A but does not contain any polyadenine tract. Control of expression of both receptors appears to occur in a tissue and gene-specific manner during human development [67].

During activin signaling, ligand induced activation of ACVR2 allows the type 2 receptor to phosphorylate serine and threonine residues in the GS (glycerine- and serine-rich) -domain of the type 1 receptor, thus, inducing its kinase activity (Figure 1.4) [68]. Once phosphorylated, it will in turn transduce the signal to the downstream signaling molecules, the Smad proteins. Thus, activated type 1 receptor leads to the phosphorylation of receptor specific Smads (Smad2 / 3) which form a complex with the common Smad4 protein and subsequently this complex translocates to the nucleus where it regulates together with different transcription factors the expression of specific target genes like *SMAD7* [69], *SERPINE* [70] and *C-MYC* [71].

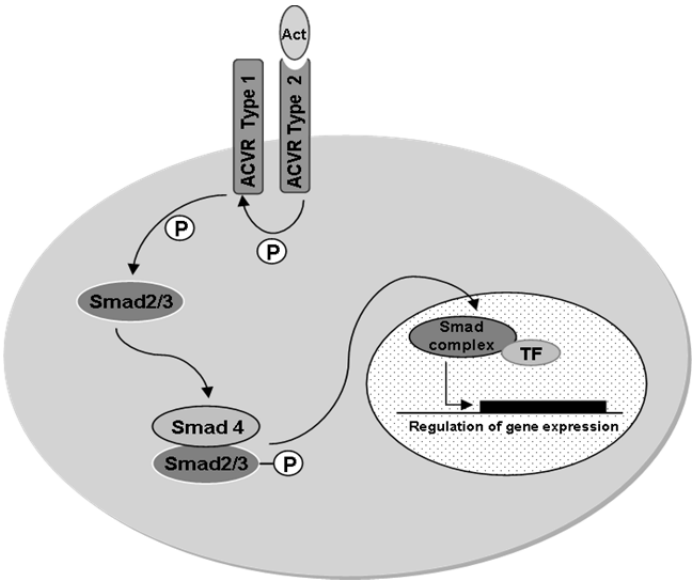


Figure 1.4: Overview of activin receptor mediated signaling transduction pathway (modified according to [68]). Act indicates activin, ACVR activin receptor and TF transcription factor.

ACVR2 contains polyadenine tracts at both exons 3 and 10 but only the A8 tract in exon 10 is mutated in 85% of MSI-H colorectal cancers [72, 73]. The biallelic frameshift mutation causes ACVR2 protein loss, and is associated with histologically poor grade tumors and significantly larger volume tumors [73, 74]. Restoration of ACVR2 in colon cancer cells causes growth suppression [57]. Both ACVR2 and TGFBR2 mutations often occur simultaneously in MSI cancers [73] and in several MSI-H CRC cell lines [75]. However, both receptors are less commonly mutated in MSS colon cancers, which tend to have a worse prognosis than MSI-H colon cancers [9] and both pathways may function independently. Indeed, recent findings demonstrate signaling induced by TGF-ß is leading to growth suppression, while activin signaling is targeting apoptosis [76]. Further, growth suppression and apoptosis by both ligands are dependent on Smad4. However, activin down-regulates p21 protein in a Smad4-independent fashion leading to increased ubiquitination and proteasomal degradation, while p21 is up-regulated by TGF-ß in a Smad4-dependent fashion to affect growth arrest. Activin-induced growth suppression and cell death are dependent on p21, while activin-induced migration is counteracted by p21 [76].

Unlike TGFBR2, in MSS tumors loss of ACVR2 arises through a combination of LOH (loss of heterozygosity) at ACVR2 and distinct ACVR2 promoter methylation, but not genetic mutation [77]. In colon cancer cell lines, mechanisms for ACVR2 loss also segregate according to microsatellite status, with MSI-H cell lines showing ACVR2 polyadenine tract mutation and MSS colon cancer cells demonstrating promoter hypermethylation [77]. Thus, disruption of activin signaling occurs in MSI and MSS colon cancers by distinct mechanisms, revealing activin signaling as an important target in the two most common genomic subtypes of colon cancer.

1.4 Glycosylation

1.4.1 Structure and Synthesis of Glycans

Considering the central role of activin signaling in the regulation of cellular behavior also significant biochemical alterations, including changes in protein expression as well as post-translational protein modifications, like glycosylation, are expectable. Posttranslational modification of proteins is an important biological feature, which expands the functional space of the proteome manifold [78]. This expansion is

believed to be a highly efficient way of evolving more complex and robust organisms. A single posttranslational modification is able to change the biochemical, functional or structural properties of many gene products at once, virtually doubling the genome. For example chemical groups like phosphates are added or removed from proteins to modulate their function [79]. Lipids are transferred to proteins to anchor them in membranes [80]. The polypeptide ubiquitin is attached to proteins to mark them for degradation [81]. The human genome contains just twice the number of genes comparing to fly or worm genomes. However, the increased complexity of the human organism is mainly the result of the posttranscriptional and posttranslational modification and alternative splicing. Typically, the genomes of more complex organisms contain more genes for the expression of various components of the posttranslational modification machineries [78].

The most frequent protein modification, glycosylation, has a huge variety of roles since it is estimated that more than 50% of all proteins are glycosylated [82]. Glycosylation is the enzymatic attachment of carbohydrates to proteins or lipids by glycosyltransferases and is found in all kingdoms of life, including some viruses [83]. Thus, glycosylation is a highly conserved and important protein modification from single cell to multicellular organisms.

The high level of complexity of the oligosaccharide chains attached to proteins is achieved by the combination of nine different monosaccharides: glucose (Glc), galactose (Gal), mannose (Man), N-acetylglucosamine (GlcNAc), N-acetylgalactosamine (GalNAc), fucose (Fuc), sialic acid (Sia), xylose (Xyl) and glucuronic acid (GlcA) [84] (Figure 1.5). Further modifications of these monosaccharides such as epimerization of GlcA to iduronic acid (IdoA) and sulfation of Gal, GlcNAc, GalNAc, GlcA and IdoA can occur after the incorporation of the monosaccharides into a glycan chain [85, 86]. Depending on the carbohydrate-peptide linkage, different types of protein glycoconjugates are formed like proteoglycans and glycophosphatidylinositol (GPI)-anchored-, O-GlcNAc-, C-mannosylation-, O-linked- and N-linked- glycoproteins [87].

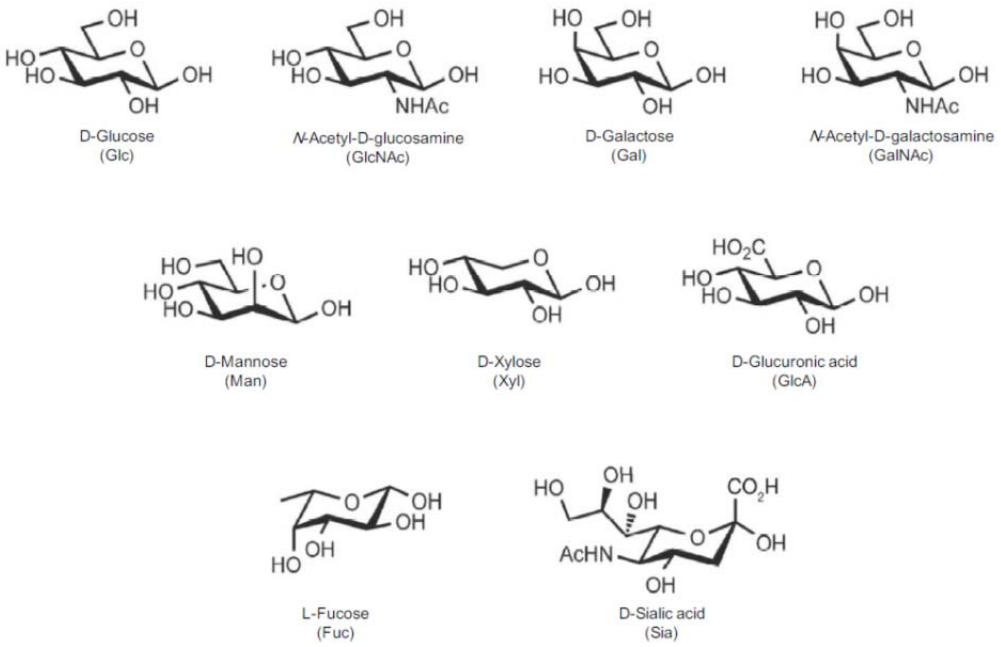


Figure 1.5: Chemical structure of the nine carbohydrates used in mammalian glycoconjugates (adapted from [88]).

Glycan synthesis is initiated by attachment of a monosaccharide to different chemical groups of amino acids or lipids. The type of chemical group modified, on a specific peptide or lipid substrate, by a specific monosaccharide provides a first level of diversity and complexity to the process of glycosylation. A second level of complexity derives from a diversity of chemical linkages between monosaccharides of carbohydrate chains. Further complexity arises from the length of the carbohydrate chain and the presence or absence of branches in the chain. A final level of complexity arises from heterogeneity at sites of glycosylation, where specific sites in a protein or lipid can be modified by different glycan structures or no glycan at all depending on the cell type or the stage of development. The level of diversity is such that even identical substrates produced in a single cell can acquire different glycan compositions. This high degree of heterogeneity and complexity implies a broad range of functions for glycans and requires a high level of control of glycosylation reactions [88].

Glycan biosynthesis mainly takes place at the endoplasmic reticulum (ER) membrane and in the Golgi apparatus. The exception is cytoplasmic / nuclear glycosylation. Key components of glycosylation reactions are the glycosyl-transferases, which catalyze the formation of a glycosidic bond. Most Golgi

glycosyltransferases are type II transmembrane proteins with a short cytosolic tail and a large luminal catalytic domain, whereas many ER glycosyltransferases are proteins spanning the transmembrane several times [89]. With few exceptions, glycosyltransferases are specific for acceptor substrate, donor substrate, and linkage. Therefore, they can be named by the donor substrate that they utilize and the carbohydrate linkage they produce.

1.4.2 Glycosylation Changes in Cancer

The addition of complex carbohydrate groups to biomolecules creates several populations of biological entities that are rich in information and diverse in function. Glycosylation plays a central role in many biological processes, including protein folding, oligomerization and stability, cell signaling, cell-cell communication, cell growth, immune response, host-pathogen interaction and inflammation [90]. Links between specific glycosylation states and diseases were first postulated more than 40 years ago [91]. Evidence has linked certain inflammatory, immune and neurodegenerative diseases, several cancers and other pathological states with alterations of glycan structures. Aberrant patterns of N- and O-linked oligosaccharide modification in cell surface proteins and in circulating plasma proteins are found in almost all human cancers and are correlated with the progression and metastatic potential of the disease [92, 93]. However, aberrant glycosylation is not a random outcome due to disordered biology in cancer cells. Numerous studies have shown that aberrant glycosylation could be both a result of oncogenic transformation as well as a crucial event for survival of cancer cells and induction of metastasis [94-96]. The expression level of GlcNAc transferase V (GlcNAcT-V) enzyme and thus the level of ß(1,6) branching of N-glycans has been shown to increase during carcinogenesis [97]. A further example is the increased expression of N-acetylglucosaminyl-transferase V (MGAT5) shown in a number of tumors, which is leading to increased glycan branching on proteins and thereby to increased tumor growth and metastasis [98, 99]. A classic cancer associated change in glycosylation is increased sialylation often manifested as specific increases in α(2,6)-linked sialic acids attached to outer N-acetyllactosamine units [100]. Increased outer-chain polyfucosylation and sialyl Lewis^x production is another important feature of many cancer cell types [101]. In fact, some glycan epitopes, shown in Figure 1.6, are of well-established value for detecting and monitoring the growth status of tumors. Sialyl Lewis^x, sialyl Lewis^a, and

sialyl T_n are well established tumor antigens. Sialyl T_n is a glycan epitope over expressed in ovarian cancer. Overexpression of sialyl Lewis^a-related glycans in lung and breast cancers has long been shown [102]. Several other cancer biomarkers are mucin type proteins: a class of heavily O-glycosylated proteins produced by epithelial tissues and secreted onto mucosal surfaces. The glycoepitope CA19-9 (sialyl Lewis^a) is a prognostic marker for pancreatic cancer. Also the high-molecular weight carbohydrate antigen CA15-3 also known as mucin 16 or MUC16 is aberrantly expressed in more than 90% of breast carcinomas and appears to strongly correlate with invasiveness of breast adenocarcinomas [103]. Using monoclonal antibody based assays, the level of these glycan biomarkers in the serum are measured in order to monitor the amount of tumor remaining in patients after surgery or chemotherapy.

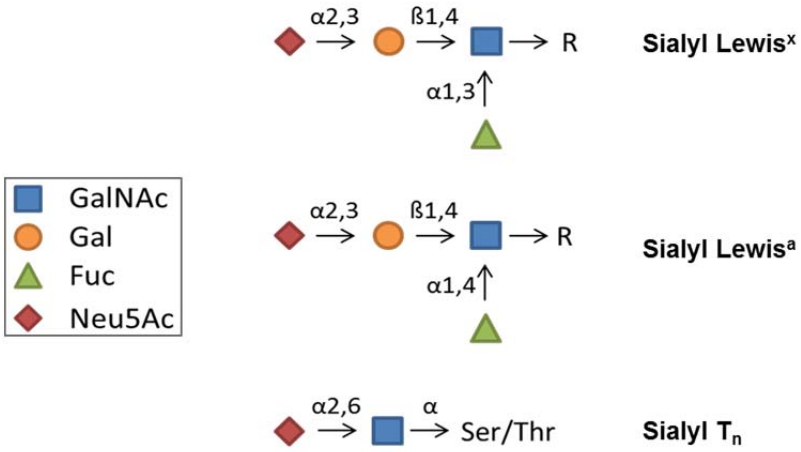


Figure 1.6: Schematic overview of selected glycan determinants (adapted from [104]). Sialyl Lewis^x, sialyl Lewis^a tumor antigen carried on O-glycans as well as on N-glycans. Sialyl Tn, commonly seen in carcinoma mucins, are the results of incomplete glycosylation in the O-linked pathway.

1.4.3 Glycosylation Changes in CRC

Altered protein glycosylation is a characteristic feature of human colon cancer. The colon is particularly rich in glycoproteins, such as mucins that are either secreted or membrane bound. In the normal colon, mucins are highly O-linked glycosylated proteins produced by goblet cells and play an important protective role in the intestine. In CRC there is a dysregulation in mucin gene expression. MUC1 is an independent prognostic factor found at higher levels in metastatic CRC, while MUC5,

which is normally not switched on in the colon, becomes expressed in CRC. In the tumors of the colon, changes in the glycans on mucins have been described, the most notable being an increase in truncated and negatively charged glycans as a result of incomplete glycosylation reactions. It is important to note that when cancer cells are treated with molecules that inhibit the attachment of O-linked glycans, mucin glycosylation is largely abrogated and cancer cell attachment to endothelial cells is much reduced. Thus, CRC uses glycan-endothelial cell interactions as an important step in the metastatic process. Furthermore MUC1 overexpression has been shown to be associated with poor prognosis in MSS tumors [105, 106]. In comparison to normal colonic mucins, cancerous human colonic mucins have reduced oligosaccharide length and total carbohydrate content [107, 108]. These changes have often been studied using lectins. Lectins are multimeric proteins, which normally have a single carbohydrate binding site per domain. Using the lectin from the pea, Arachis hypogea (PNA), the tumor-associated glycoprotein of colon polyps detected in tissue and serum is the Thomsen-Friedenreich “T-antigen” (Gal(β-1,3)GalNAc-Ser/Thr) [109]. T, T_n (GalNAc(α1)-Ser/Thr) and sialyl-T_n (Neu5Ac(α2-6)GalNAc-Ser/Thr) antigens are colon cancer associated antigens. T_n and sialyl-T_n may be useful markers of poorly differentiated adenocarcinomas and mucinous carcinomas [110]. T_n and T antigens were found to be sialylated in the majority of cancer metastatic colon cells [111].

Moreover, highly metastatic human colon carcinomas express more poly-N-acetyllactosaminyl side chains in N-glycans of lysosomal membrane glycoproteins, which are more sialylated but less fucosylated, and correlate with the increased expression of sialyl Lewis^x structures, than cells with low metastatic potential [112]. Increased N-glycan branching detected by the usage of lectins is associated with colorectal tumor recurrence, patient survival and the presence of lymph node metastases [113]. Additionally, applying lectin glycol-array and lectin blotting revealed that elevated N-glycan sialylation in complement C3, histidine-rich glycoprotein and kininogen-1 occurs in the colorectal adenoma to carcinoma sequence [114]. Galectins-1 and -3, animal lectins that are specific for N-glycan associated galactose, are up-regulated in human colon tumor and increased Galectin-3 expression correlates with colorectal tumor progression, liver metastasis, poor patient survival and tumor aggressiveness [115-117].

However most of these previous studies on glycosylation alterations did not account for the different genetic characteristics of CRC's, e.g. MSI versus MSS tumors. In particular, there are only a few anecdotal reports on MSI-specific glycosylation changes. For example, gene expression of GALNT5, that catalyzes the formation of the T-antigens sugar structure, was found to be up-regulated by microarray analysis on MSI-H compared to MSS colorectal carcinomas [118]. The expression of two other glycosyltransferases was also found to be up-regulated in a MSI-specific manner. B4GALT1 that is known to transfer galactose to the branched oligosaccharide chains of N-glycans [119] and SIAT4B that is responsible for the α2,3-sialylation of O-glycans [120]. Further, coding microsatellites of several genes involved in the glycosylation machinery are frequently mutated, including lectin mannose-binding 1 (LMAN1), xylosyltransferase 2 (XYLT2) and OGT [56, 121]. Preliminary evidence indicates that the inactivation of some MSI targets, ACVR2, TGFBR2 and AIM2, may influence protein glycosylation pattern of CRC cells, which has been shown by transient reconstitution experiments [122].
In conclusion, MSI-specific glycosylation changes are a rather unexplored field and thus, more work needs to be carried out to obtain a better understanding in this field. Particularly, despite the obvious important functions of glycan structures only few reports are available on the regulation of glycosylation by cellular signaling [123, 124].

1.5 Challenges in Analyzing Proteomic Alterations in Cancer

Besides tumor-associated glycosylation changes human tumors acquire a large number of genetic and epigenetic alterations that arise during progression from preneoplastic lesions to metastatic disease. However, the diversity of these alterations reflects the intratumoral heterogeneity and represents the genomic landscape of tumors. Among a high background of irrelevant passenger alterations, only a limited number of genetic alterations are considered to be driving events that confer a selective advantage to tumor cells. Major signaling pathways affected by such driver mutations include the TGF-ß, BMP, Activin, Wnt and Notch pathways thereby abrogating normal regulation of key cellular processes like cell fate, cell survival and genome maintenance. Both, tumor-relevant driver mutations in a major signaling receptor as well as tumor-irrelevant passenger mutations can also cause

changes at the proteomic level. Passenger mutation-associated proteomic patterns are propagated randomly and do not represent generic tumor-associated changes [125]. Therefore, focusing on proteome alterations associated with single driver mutations is necessary to identify specific changes that underlie tumor development. However, such analyses encounter two major limitations at different levels.

At the molecular level, the genetic heterogeneity of tumors especially those of the microsatellite unstable and mutator phenotype poses a significant problem in determining mutation-specific effects. Two principal strategies to detect cellular consequences of a single mutation have been applied. First, targeted gene knock out in target gene proficient cell lines by homologous recombination, adeno-associated viral delivery or zinc finger nucleases has been used successfully applied [126-128]. However, these approaches are often limited by their low efficiency, are laborious and time-consuming and bear the potential for confounding off-target effects. Second, transfer of the target gene into deficient cell lines by gene insertion or gene targeting methods has been extensively applied. However, insertion methods are often affected by random insertion, variable number of integrated gene copies per cell and inconsistent integration sites eventually resulting in unpredictable expression patterns [129]. On the other hand, many non-integrating vectors, like adenoviral DNA, are not often replicated during cell division thereby limiting their use in basic research.

At the protein level, sample complexity is a major limiting factor. In addition to prefractionation methods, metabolic labeling is a versatile tool to focus on proteomic changes induced by gene activation. Since activation of tumor suppressor pathways directly regulates target gene expression, analysis of tumor suppressor dependent alterations of newly synthesized proteins by metabolic labeling is a reasonable approach to restrict the proteomic complexity. Conventional methods for metabolic labeling usually rely on amino acids containing either radioactive or stable isotopes. Although radioactive labeling enables extremely sensitive detection methods, its use for proteomic analysis is limited due to the need of special handling precautions and contamination of the analytical instrumentation. Stable isotopic labeling, in particular the SILAC (stable isotope labeling by amino acids in cell culture) methodology, is currently the preferred method for most metabolic labeling approaches in proteomic analyses especially for cell lines [130]. However, when applying the SILAC technology, mass spectrometric detection of labeled peptides has to be conducted in

the presence of high abundant irrelevant unlabeled peptides, thereby hampering the detection of labeled low-abundance peptides. A relatively new method, termed Click-it labeling, that enables labeling of nascent proteins comparable to a radioactive compound can overcome this problem since upon incorporation of the labeled compound also a handle for specific extraction of the labeled protein is worked in. The click reaction makes use of a methionine derivative that is functionalized with an azide (L-Azidohomoalanine; AHA). During protein synthesis AHA is incorporated into proteins which can then be tagged with a biotin alkyne (PEG4 carboxamide-Propargyl Biotin) resulting in the specific biotinylation of the metabolically labeled proteins [131]. The final step is extraction of the labeled proteins by streptavidin beads. This new method provides a fast, sensitive, nontoxic and non-radioactive alternative to the traditional radioactive technique and enables the detection of newly synthesized proteins as well as post-translational protein modifications. Irrespective of the utilized technique the advances in the fields of proteomics are hoped to help understanding the molecular protein complexity, especially of the disease process and thus enable the development of tools to use in treatment as well as in detection and prevention of diseases.

1.6 Aims of the Present Study

Microsatellite unstable (MSI) tumors exhibit a variety of histoclinicopathological features including mucinous histology and improved prognosis when compared to their microsatellite stable (MSS) counterparts. Biallelic mutational inactivation of some MSI target genes is generally believed to drive MSI tumorigenesis. Focusing on the pathogenetic mechanisms of these MSI tumors, a potential correlation between functional inactivation of specific MSI target genes and cell surface glycosylation pattern has been recently uncovered [122]. The present study aims to explore MSI target gene-dependent alterations of cell glycosylation and also glycomic signatures in MSI colorectal cancer cells. Thereby one of the most frequently mutated MSI target genes, the activin receptor type 2 (*ACVR2*), is of special interest in order to obtain a better understanding of ACVR2-dependent MSI tumorigenesis. The biological effects of activin in colon cancer have not been previously appreciated and recently came to light when the ACVR2 receptor was discovered to be mutated in the majority of MSI-H colon cancers [73, 132]. Although ACVR2 signaling is

involved in many major cellular processes, including cell differentiation, proliferation and apoptosis, the role of this pathway with respect to the proteomic constellation, thereby mainly focusing on the glycosylation of proteins, is largely unexplored.

Therefore, the first specific aim was to generate a MSI colorectal cancer cell line model system enabling a stable and doxycycline inducible ACVR2 expression. After generating this model cell line the signaling ability of reconstituted ACVR2 needs to be confirmed. Thus, the effects of activin-mediated signaling on the glycosylation profile could be examined by using different approaches. First fluorescence activated cell sorting (FACS) analysis using a panel of different plant lectins was applied to determine ACVR2-dependent changes of cell surface glycoproteins. At the same time, the transcript level of genes involved in glycosylation should be obtained by performing a Glyco-Gene Chip analysis in the presence or absence of ACVR2 expression. Since these approaches determine the steady state levels of the proteins, a further goal was to analyze also newly synthesized proteins. For this purpose, radioactive labeling experiments were performed in order to evaluate the incorporation of ^{3}H-labeled saccharides, L-fucose and ManNAc. Furthermore, a relatively new method, termed Click-it labeling, which enables labeling of nascent proteins and at the same time an easy extraction of the labeled compounds, was established.

The results of these diverse experiments should provide the basis for the identification and characterization of candidate proteins and possibly deliver insights into the mechanism of ACVR2-dependent proteomic and glycomic alterations (Figure 1.7). Taken together, this study should give a better understanding of MSI tumor-specific protein profile in general and particularly of the MSI target gene ACVR2-dependent protein changes.

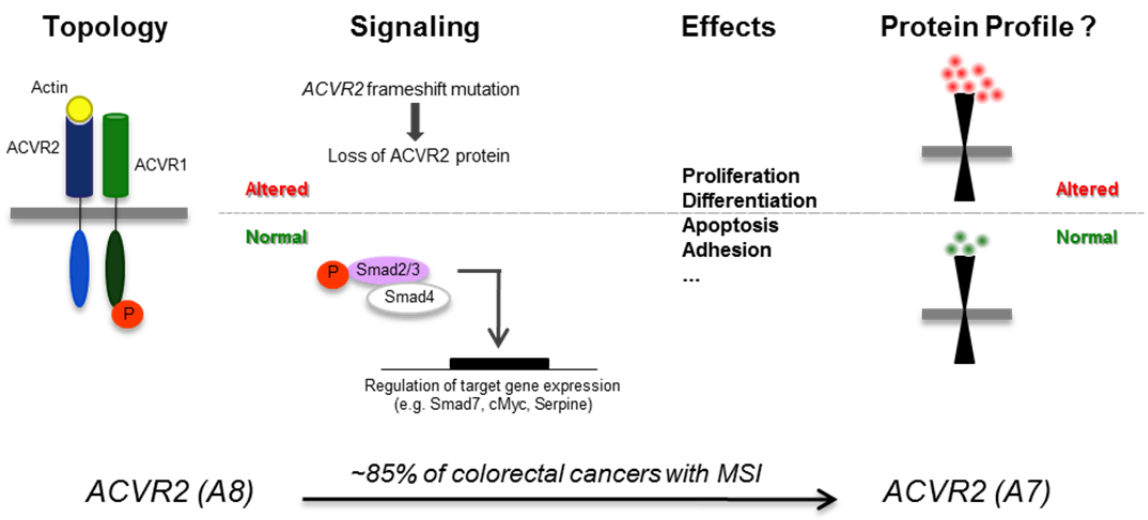


Figure 1.7: Schematic overview of the hypothesis. Upon ligand binding, activin-mediated signaling is initiated and Smad proteins become phosphorylated inducing the transcription of target genes involved in many cellular processes like proliferation, differentiation, apoptosis and adhesion. Since in over 85% of MSI tumors the microsatellite of the *ACVR2* gene is mutated leading to inactivation of its protein, the question is if the absence of ACVR2 may have effects on the protein profile of these cells.

2. MATERIAL

2.1 Instruments

Device	Supplier
Agarosegel Chamber Sub Cell GT	Biorad (München, Germany)
ÄktaPurifier FPLC System	GE Healthcare (München, Germany)
Analytical scale (BP 210 D)	Sartorius (Göttingen, Germany)
Camera (Electrophoresis Docu System 120)	Kodak (Stuttgart, Germany)
Camera (Biorad Gel Doc 2000)	Biorad (München, Germany)
Centrifuge (5810R)	Eppendorf (Hamburg, Germany)
Centrifuge (Biofuge 13)	Heraeus Holding (Hanau, Germany)
Centrifuge (Microcentrifuge 1-14)	Sigma Laborzentrifugen (Osterode, Germany)
Centrifuge (Sigma 3MK)	Sigma Laborzentrifugen (Osterode, Germany)
Centrifuge (Varifuge 3.0R)	Heraeus Holding (Hanau, Germany)
Centrifuge Heraeus (Modell T 110 L)	Heraeus Holding (Hanau, Germany)
Digital pH Meter pH 525	WTW (Weilheim, Germany)
Electrophoresis chamber (Sub Cell GT)	Biorad (München, Germany)
Electroporator for cells (Nucleofector 1)	Lonza Biosystems (Basel, Switzerland)
ELISA-Reader (GENios)	GENios Tecan (Crailsheim, Germany)
FACS (FACSCalibur)	Becton Dickinson (Franklin Lakes, USA)
Genetic analyzer (ABI 3100)	Life Technologies (Darmstadt, Germany)
Incubator (BD6220)	Fisher Scientific (Loughborough, UK)
Liquid Scintillation Counter (TRI-CARB 2900)	Perkin Elmer (Boston, USA)
LTQ Orbitrap XL mass spectrometer	Thermo Scientific (Bremen, Germany)
Luminometer (Lumat LB9507)	Berthold (Bad Wildbad, Germany)
Magnetic Separation Rack (6-Tube)	New England Biolabs (Frankfurt, Germany)
Microscope (DMBRE)	Leica (Bensheim, Germany)
Microscope (Leica DMIL)	Leica (Bensheim, Germany)
Microscope CK 40	Olympus Holding (Hamburg, Germany)
Microwave	Siemens (Munich, Germany)
nanoAcquity UPLC system	Waters GmbH (Eschborn, Germany)
NuPAGE X-Cell Sure Lock	Life Technologies (Darmstadt, Germany)
NuPAGE ZOOM IPG Runner Cassette	Life Technologies (Darmstadt, Germany)

NuPAGE Protein Electrophoresis System	Life Technologies (Darmstadt, Germany)
PCR system (GeneAmp 22400)	Perkin Elmer (Waltham, USA)
PCR System (Robo Cycler Gradient96)	Stratagene (Böblingen, Germany)
PH meter (Calimatic 761)	Knick (Berlin, Germany)
Photometer (Ultrospec 3300)	Amersham Pharmacia (Cambridge, UK)
Pipet aid (Pipetman)	Gilson (Limburg-Offheim, Germany)
Pipetboy plus	TecNoMara (Fernwald, Germany)
Pipettes (2 µl; 10 µl; 20 µl; 200 µl; 1000µl)	Gilson (Limburg-Offheim, Germany)
Power supply (Consort E835)	Peqlab (Erlangen, Germany)
Power supply (Power Pac 300)	Bio-Rad (München, Germany)
Rotating mixer RM5	NeoLab (Heidelberg, Germany)
Shaker (Certomat H)	Sartorius (Göttingen, Germany)
Sonopuls-Bandelin (ultrasonic homogenizer)	Bandelin electronic (Berlin, Germany)
StepOnePlus™ Real-Time PCR Systems	ABI, (Darmstadt, Germany)
Thermomixer (5436)	Eppendorf (Hamburg, Germany)
Ultracentrifuge (TLA-100.2 rotor)	Beckmann Coulter (Krefeld, Germany)
Ultra-Low Temperatur Freezer MDF-U53V	Sanyo (München, Germany)
UV Transilluminator	Konrad B. Laborgeräte (Wiesloch, Germany)
Vortex (MS1 Minishaker)	IKA (Staufen, Germany)
Waterbath (Grant SUB6)	Grant (Cambridge, UK)
Waterbath (SW 20)	Julabo Labortechnik (Seelbach, Germany)

2.2 Commercially Available Kits and Assays

Item	Supplier
Amaxa Cell Line Nucleofector Kit V	Lonza Biosystems (Basel, Switzerland)
BigDyeTerminator v1.1 Sequencing Kit	Life Technologies (Darmstadt, Germany)
Biorad Protein Assay	BioRad (München, Germany)
Cell Line Nucleofactor Kit V	Lonza Biosystems (Basel, Switzerland)
One Solution Cell Proliferation Assay (MTS)	Promega (Madison, USA)
Click-iT Protein Reaction Buffer Kit	Life Technologies (Darmstadt, Germany)
DNeasy Blood & Tissue Handbook	Qiagen (Hilden, Germany)
Endo-free Plasmid Maxi Kit	Qiagen (Hilden, Germany)
FuGENE HD Transfection Reagent	Roche (Mannheim, Germany)

High Pure PCR Product Purification Kit	Roche (Mannheim, Germany)
JetQuick Gel Extraction Kit	Genomed (Löhne, Germany)
Luciferase Assay System	Promega (Mannheim, Germany)
Mycoplasma Detection Kit for conventional PCR	Minerva Biolabs (Berlin, Germany)
NucleoSpin Plasmid Kit	Machery Nagel (Düren, Germany)
Plasmid Maxi Kit	Qiagen (Hilden, Germany)
Power SYBR Green PCR Master Mix	Life Technologies (Darmstadt, Germany)
QIAquick Gel Extraction Kit	Qiagen (Hilden, Germany)
Rapid DNA Dephos & Ligation Kit	Roche (Mannheim, Germany)
RNeasy Mini Kit	Qiagen (Hilden, Germany)
Vectastain Elite ABC Kit (Universal)	Vector Laboratories (Burlingame, USA)

2.3 Reagents

Reagent	Supplier
^{3}H-Acetyl-D-mannosamine (N-[6-^{3}H]; 1 mCi/ml) ^{3}H-L-Fucose (L-[6-^{3}H]; 1 mCi/ml)	American Radiolabeled Chemicals (St. Louis, USA)
Acetic acid	Serva (Heidelberg, Germany)
Activin A	Sigma-Aldrich (Taufkirchen, Germany)
Ammonium bicarbonate	Sigma-Aldrich (Taufkirchen, Germany)
Ampicillin sodium salt	Sigma-Aldrich (Taufkirchen, Germany)
Anti-FLAG M2 Affinity Gel	Sigma (Taufkirchen, Germany)
Bacto-Agar	Fluka Chemie (Buchs, Switzerland)
Biotin Alkyne (PEG4 carboxamide-Propargyl)	Life Technologies (Darmstadt, Germany)
Biotin Azide (PEG4 carboxamide-6-Azidohexanyl)	Life Technologies (Darmstadt, Germany)
Boric acid	Merck (Darmstadt, Germany)
Bovine Serum Albumin (BSA)	Sigma-Aldrich (Taufkirchen, Germany)
Calcium chloride ($CaCl_2$)	Merck (Darmstadt, Germany)
Chloroform	Carl Roth (Karlsruhe, Germany)
Citric acid	Merck (Darmstadt, Germany)
Click-iT AHA (L-azidohomoalanine)	Life Technologies (Darmstadt, Germany)
Click-iT Fucose Alkyne	Life Technologies (Darmstadt, Germany)
Click-iT Protein Buffer Kit	Life Technologies (Darmstadt, Germany)
Complete Protease Inhibitor Cocktail Tablets	Roche (Mannheim, Germany)

Coomassie Brilliant Blue G250	Biorad (München, Germany)
DAB-chromogene	Dako (Hamburg, Germany)
Denucleoside triphosphate (dNTP) -Mix	Life Technologies (Darmstadt, Germany)
Diaminobenzidine (Liquid DAB + substrate)	DAKO (Hamburg, Germany)
Dimethyl sulfoxide (DMSO)	Merck (Darmstadt, Germany)
Disodium hydrogen phosphate	VWR International (Bruchsal, Germany)
DMEM/Ham's F-12 (1:1) with L-Glutamine	PAA (Cölbe, Germany)
Doxycycline	Sigma-Aldrich (Taufkirchen, Germany)
Dithiothreitol (DTT)	Life Technologies (Darmstadt, Germany)
Dulbecco's PBS (1x) without Ca & Mg	PAA (Cölbe, Germany)
Dynabeads MyOne Streptavidin T1	Life Technologies (Darmstadt, Germany)
Ethanol absolute	Merck (Darmstadt, Germany)
Ethidiumbromide	Sigma-Aldrich (Taufkirchen, Germany)
Ethylenediaminetetraacetic acid (EDTA)	Merck (Darmstadt, Germany)
Fetal Bovine Serum Gold (FCS)	PAA (Cölbe, Germany)
Folin-Ciocalteus phenol reagent	Merck (Darmstadt, Germany)
Formaldehyde (37%)	Carl Roth (Karsruhe, Germany)
G-418 Sulphate	PAA (Cölbe, Germany)
Ganciclovir	Roche (Mannheim, Germany)
GelRed	Biotium (Heyward,USA)
Glycerol	Carl Roth (Karlsruhe, Germany)
HI-DI-formamide	Life Technologies (Darmstadt, Germany)
Horse serum	Vector Laboratories (Burlingame, USA)
Human TGF-ß1	Cell Signaling Tech. (Danvers, USA)
Hydrogen chloride (37%; HCl)	Merck (Darmstadt, Germany)
Hygromycin B	PAA (Cölbe, Germany)
Kodak BioMax films	Sigma-Aldrich (Taufkirchen, Germany)
Lipofectamine 2000 Reagent	Life Technologies (Darmstadt, Germany)
Methanol	AppliChem (Darmstadt, Germany)
Magnesium chloride (MgCl2)	Merck (Darmstadt, Germany)
Milk powder	Carl Roth (Karlsruhe, Germany)
NuPAGE 4 – 12% Bis-Tris Mini Gel	Life Technologies (Darmstadt, Germany)
NuPAGE Antioxidant	Life Technologies (Darmstadt, Germany)
NuPAGE MES Running buffer	Life Technologies (Darmstadt, Germany)
NuPAGE Sample Reducing Agent (10x)	Life Technologies (Darmstadt, Germany)

NuPAGE Transferbuffer (20x)	Life Technologies (Darmstadt, Germany)
Oligo (dT) Primer	Life Technologies (Darmstadt, Germany)
Ponceau S	Sigma-Aldrich (Taufkirchen, Germany)
Penicillin/Streptomycin (100x)	PAA (Cölbe, Germany)
Phosphoric acid	Carl Roth (Karlsruhe, Germany)
PhosSTOP Phosphatase Inhibitor Cocktail Tablets	Roche (Mannheim, Germany)
Potassium chloride	J.T. Baker (Deventer, Netherland)
Potassium dihydrogenphosphate	Gerbu Biochemicals (Gaisberg, Germany)
Potassium sodium tartrate	Merck (Darmstadt, Germany)
Puromycin	Sigma-Aldrich (Taufkirchen, Germany)
RNAstable	Biomatrica (San Diego, USA)
RPMI 1640 with L-Glutamine	PAA (Cölbe, Germany)
RPMI-1640 without methionine and L-glutamine	Sigma-Aldrich (Taufkirchen, Germany)
Silver nitrate	Carl Roth (Karlsruhe, Germany)
S.O.C. Medium	Life Technologies (Darmstadt, Germany)
Sodium acetate	J.T. Baker (Deventer, Netherland)
Sodium bisulfite	Merck (Darmstadt, Germany)
Sodium carbonate	J.T. Baker (Deventer, Netherland)
Sodium chloride	AppliChem (Darmstadt, Germany)
Sodium hydroxide (NaOH)	AppliChem (Darmstadt, Germany)
Streptavidin Agarose Beads	Life Technologies (Darmstadt, Germany)
Streptavidin-HRP	SouthernBiotech (Alabama, USA)
Streptavidin／R-phycoerythrin	Sigma-Aldrich (Taufkirchen, GER)
SYPRO Ruby Protein Gel Stain	Life Technologies (Darmstadt, Germany)
Trifluoroacetic acid	Thermo Scientific (Bremen, Germany)
Trichloracetic acid	Carl Roth (Karlsruhe, Germany)
Tris Base	Carl Roth (Karlsruhe, Germany)
Tris-HCl	Carl Roth (Karlsruhe, Germany)
Triton-X-100	Sigma-Aldrich (Taufkirchen, Germany)
Trypsin-EDTA (1x)	PAA (Cölbe, Germany)
Trypsin	Promega (Madison, USA)
Tween-20	Sigma-Aldrich (Taufkirchen, Germany)
Ultima Gold LSC Cocktail	Perkin Elmer (Boston, USA)
UltraPureTM Agarose	Life Technologies (Darmstadt, Germany)
Water, DNase, RNase-FREE	MP Biomedicals (Solon, USA)

Western Lightning Plus ECL	Perkin Elmer (Boston, USA)
Yeast extract	Carl Roth (Karlsruhe, Germany)

2.4 Antibodies, Enzymes and Lectins

Antibody	**Supplier**
Anti-Actin (clone C4, mouse mAb)	MP Biomedicals (Solon, USA)
Anti-ACVR2 ([149/1], mouse mAb)	Abcam (Cambridge, UK)
Anti-ACVR2 (rabbit pAb)	B.H. Jung, MD (Northwestern Memorial Hospital, Chicago, USA)
Anti-Flag (clone M2, mouse mAB)	Sigma-Aldrich (Taufkirchen, Germany)
Anti-mouse IgG (HRP conjugate, sheep)	GE Healthcare (Munich, Germany)
Anti-rabbit IgG (HRP conjugate, goat)	Promega (Madison, USA)
Anti-Phospho-Smad2 ([Ser465/467], rabbit mAb)	Cell Signaling Tech. (Danvers, USA)
Anti-Smad2 ([86F7], rabbit mAb)	Cell Signaling Tech. (Danvers, USA)
Enzyme	**Supplier**
DNase I, Amplification Grade	Life Technologies (Darmstadt, Germany)
HOT FIRE Pol DNA Polymerase	Solis Biodyne (Tartu, Estonia)
Phusion High-Fidelity DNA Polymerase	New England Biolabs (Frankfurt, Germany)
Restriction enzymes	New England Biolabs (Frankfurt, Germany)
Restriction enzymes	Promega (Mannheim, Germany)
Restriction enzymes	Roche (Mannheim, Germany)
RNase OUT	Life Technologies (Darmstadt, Germany)
Superscript II Reverse Transcriptase	Life Technologies (Darmstadt, Germany)
T4-Ligase	Roche (Mannheim, Germany)
Taq DNA Polymerase	Life Technologies (Darmstadt, Germany)
Lectin	**Supplier**
AAL (Aleuria aurantia, biotinylated)	Vector Laboratories (Burlingame, USA)
DBA (Dolichos biflorus agglutinin, biotinylated)	Vector Laboratories (Burlingame, USA)
JAC (Jacalin, biotinylated)	Vector Laboratories (Burlingame, USA)
LEA (Lycopersicon esculentum agglutinin, biotinylated)	Prof. Dr. H.J. Gabius (Ludwig-Maximilians-University, Munich, Germany)
PSA (Pisum sativum agglutinin, biotinylated)	Vector Laboratories (Burlingame, USA)
SNA (Sambucus nigra, biotinylated)	Vector Laboratories (Burlingame, USA)

VAA (Viscum album agglutinin, biotinylated)	Prof. Dr. H.J. Gabius (Ludwig-Maximilians-University, Munich, Germany)

2.5 Other Materials and Markers

Material	Supplier
0,5, 1,5, 2 ml Eppendorf Tubes	Greiner Bio-One (Frickenhausen, Germany)
15, 50 ml Falcon Tubes	Greiner Bio-One (Frickenhausen, Germany)
Cell Scraper	Greiner Bio-One (Frickenhausen, Germany)
Disposable Serological Pippette	Corning (New York, USA)
Multiwell Plates	Greiner Bio-One (Frickenhausen, Germany)
nanoAcquity C18 column	Waters GmbH (Eschborn, Germany)
Nitrocellulose Membrane Filter Paper Sandwich	Life Technologies (Darmstadt, Germany)
PCR tubes	Carl Roth (Karsruhe, Germany)
Sterile Pipet Tips	Corning (New York, USA)
Tissue Culture Dish	Orange Scientific (Braine-l'Alleud, Belgium)
Tissue Culture Flask	Greiner Bio-One (Frickenhausen, Germany)
Marker	**Supplier**
1 Kb DNA Ladder	Life Technologies (Darmstadt, Germany)
100 Bp DNA Ladder	Life Technologies (Darmstadt, Germany)
Mark12 Unstained Standard	Life Technologies (Darmstadt, Germany)
SeeBlue Plus2 Pre-Stained Standard	Life Technologies (Darmstadt, Germany)

2.6 Oligonucleotides

All oligonucleotides were synthesized by Thermo Fisher Scientific (Ulm, Germany). Primers for quantitative Real-Time reverse transcription-polymerase chain reaction (qRT-PCR) were generated by using the Universal ProbeLibrary Assay Design Center from Roche. (http://www.roche-applied-science.com/sis/rtpcr/upl/index.jsp?id=UP030000). The melting temperature of all primers was designed for 60 °C. The primers for frameshift mutation analysis were labeled with fluorescein.

Cloning	**Sequence (5’ – 3’)**
ACVR2_Stu_S	AAGAGGCCTGTTTTAAGAGATTA
ACVR2_nSmaNot_A	AAAGCGGCCGCTAGACTCCCGGGTAGACTAGATTCTTTGGGAG
ACVR2_OneStep_A	AAAGCGGCCGCTCACTTGTCATCGTCGTCCTTGTAGTCTAGACTA GATTCTTTGGGAG
ACVR2_FlagStop_S:	TTTCCCGGGGACTACAAGGACGACGATGACAAGTGAGCGGCCG CTTT
ACVR2_FlagStop_A:	AAAGGGCCCCTGATGTTCCTGCTGCTACTGTTCACTCGCCGGCG AAA
Sequencing	**Sequence (5’ – 3’)**
ACVR2_seq_F_1156	GGGATGCATTTTTGAGGATAG
M 13_rev	CAGGAAACAGCTATGAC
S2F_seq_R1	AACAAATTGGACTAATCCGGA
S2F_seq_F1	CTGGAGACGCCATCCACGC
Frameshift Analysis	**Sequence (5’ – 3’)**
ACTR2_a8b-s	GTTGCCATTTGAGGAGGAAA
ACTR 2_a8b_a	CAGCATGTTTCTGCCAATAATC
RT-PCR	**Sequence (5’ – 3’)**
ACVR2_seq_F_1156	GGGATGCATTTTTGAGGATAG
pTRE_Tight_BI_rev	GACTAGAGGATCCCCAATTCGGC
ACVR2_for_1547	CAAATGTTGACTTTCCTCCCAAA
ACVR2-3’UTR-S	GGGACTCTGAACTGGAGCTG
pTRE_Tight_BI_A2	CTGGAGATATCGTCGACAAGC
pTRE_Tight_BI_A1	CCGCGCTAGCACGCGTCAGCT
ACVR2_5’UTR_S	AGCGAGAACTTCCTCCGGATT
ACVR2_143_rev	TCACCATAACACGGTTCAACA
ACVR2_3’UTR_A	TCCCAGAGCAACATTTTTCA
qRT-PCR	**Sequence (5’ – 3’)**
ACVR2-3’UTR-S	GGGACTCTGAACTGGAGCTG
pTRE_Tight_BI_A2	CTGGAGATATCGTCGACAAGC
ACVR2_for_1547	CAAATGTTGACTTTCCTCCCAAA
S2F_Seq_R1	AACAAATTGGACTAATCCGGA
ACVR2_3’UTR_A	TCCCAGAGCAACATTTTTCA
ACVR2_seq_F_1156	GGGATGCATTTTTGAGGATAG
ACVR2 rev 1247	TACAGGTCCATCTGGTCCATCTGCAGCAG
ALG11_for	CTGGTGGAGGAGGAGAAAGA
ALG11_rev	TGACATTAACATCGCCGGTA
B4GALNT4_for	GGCCTGCAATTTGTGTACCT

B4GALNT4_rev	GAGACTCGCGGTAGAAGCAC
DPM3_for	AGTTTGGTCCGCTTTTCCTT
DPM3_rev	AGGATCGCTAGTCCCCAAAG
GALNAC4S_for	GGTTGGGCTCTATGCTGTGT
GALNAC4S_rev	ACTTGACGTTGGATGCATGA
IDS_for	GCCTCTTGAGTGCTTTGGAC
IDS_rev	CCCATTCTCCATGTTCACCT
LGALS3_for	CTGGGGAAGGGAAGAAAGAC
LGALS3_rev	CTGCAACCTTGAAGTGGTCA
MGEA5_for	CAAGAGTTTGGTGTGCCTCA
MGEA5_rev	AGGGTGCTGCAACTAAAGGA
OGT_for	TCAAGGGCACAGTTGTGGTA
OGT_rev	GGGCAGTCAAGGGTAAAACA
SLC17A5_for	CAACAACACTGGGAGGCTTT
SLC17A5_rev	TGTATTTGTGATGCCCAGGA
VCAN_for	GACAAGAAAGCAGCACCACA
VCAN_rev	GTTGGGAATCCATCAGCAGT
Smad7_for	CTCGGAAGTCAAGAGGCTGT
Smad7_rev	GCAGAGTCGGCTAAGGTGAT
c-Myc_for	CAACAACACTGGGAGGCTTT
c-Myc_rev	TGTATTTGTGATGCCCAGGA
Serpine_for	CAACTTGCTTGGGAAAGGAGC
Serpine_rev	AGTCGGGGAAGGGAGTCTTC
GAPDH_for	AGCCACATCGCTCAGACAC
GAPDH_rev	GCCCAATACGACCAAATCC
18sRNA_for	AAACGGCTACCACATCCAAG
18sRNA _rev	CCTCCAATGGATCCTCGTTA
JUNB_for	GACCAAGAGCGCATCAAAGT
JUNB_rev	AGCGTCTTCACCTTGTCCTC
PDGFB_for	AATTCAAGCACACGCATGAC
PDGFB_rev	ATAACCCTGCCCACACACTC
HES1_for	GTGAAGCACCTCCGGAAC
HES1_rev	GTCACCTCGTTCATGCACTC
LFNG_for	GTTTGAAAACAAGCGGAACG
LFNG_rev	GTGTGTCCGGGTACAGGTG
FGF9_for	GAACCAGGAAAGACCACAGC
FGF9_rev	TCCCGAGGTAGAGTCCACTG
BMP4_for	GGAGGAGGAAGAGCAGATCC
BMP4_rev	ATGTTCTTCGTGGTGGAAGC
SLC35F2_for	ATCCTGACAGCGGACCTCTA

SLC35F2_rev	AACCCCACCATGATGACAGT
KLRF1_for	GACTTGGGTGGATGGTTCTC
KLRF1_rev	GCTTTCCTTAATGGCAGCAC
FGF18_for	GGCAAGGAGACGGAATTCTA
FGF18_rev	CCAGAACCTTCTCGATGAACA
FUT1_for	AAAAGCGGACTGTGGATCTG
FUT1_rev	AGGACACAGACTAGCAGGAAGG
FUT2_for	GACCCAGAGGGAACACTGAG
FUT2_rev	GGGAAAGGAGAAAGGCATCT
FUT3_for	AGCTGTCCTCATCCACTGCT
FUT3_rev	GGCGATTTTCAGGCCTCT
FUT4_for	CCCAAGAGCATACGGAACTT
FUT4_rev	CTGTGCATCTCCTTGACTGC
FUT5_for	CGACGACCCCCTGAACTAC
FUT5_rev	AGACCATCCTGGCTAACACG
FUT6_for	AAAGGCCTGTCTCCAGATCC
FUT6_rev	GGGATCCATGGGTCAGAGTA
FUT7_for	GGAGACTGTGGATGAATAATGCT
FUT7_rev	CAGCCACAGGAGCCAGAG
FUT8_for	TGATTAACTGGACAAATTCAGCAT
FUT8_rev	TGGTAGTCCTGCAGTGAATCTTT
FUT9_for	ACGTGCTTCCCATGATATGTT
FUT9_rev	GTTTGATGTAAATGAGAAGACATGC
FUT10_for	GTGGGCTAATATCAGGCTTCA
FUT10_rev	GGGCAACTCAGGTGGGTAT
FUT11_for	CATCCCGGTAGACTCCTACG
FUT11_rev	GGGCCAAGTGGAACTTATAGC
POFUT1_for	AAGCCTCCTTTCACCAACCT
POFUT1_rev	CTGATGACCCGATGGTAAGC
POFUT2_for	AGGACAAGCACGAGTACTACAGAG
POFUT2_rev	CGGACAGACAGGAGACGTTTA
FUCA1_for	GTGCACCAGCATTGACAAGT
FUCA1_rev	CCCAAACTTACTGTCTGAACCA
FUCA2_for	TCTTGGCCTGGTTATATAATGAAAG
FUCA2_rev	AGAAGCCACCATGCTTACAGA
FUK_for	GAGATTCAGCGGTGTGTCAG
FUK_rev	TCTCCACAGAGAAGAAGACAACC
TNNI3K_for	CCCTTGGAGTTCAATATCACG
TNNI3K_rev	AGTGTTGATCCTCCATTTCCA
TSTA3_for	GCCTGTTCCGGAATATCAAA

TSTA3_rev	AGGACGTTGTCGTTCATGTG
SLC35C1_for	TGCTCACCTGCGGTATCAT
SLC35C1_rev	ACGACAGGGTGCCTTCTG
GMDS_for	GATGGTTCCTACCTGGCTGA
GMDS_rev	GACCGCCGTACAATTCCAT
GMDS_HCT116_for	GTGCAGGAAATACCCCAGAA
GMDS_HCT116_rev	ACGGAAGTTCACCACAATCC

s: sense; a: antisense; for: forward; rev: reverse

2.7 Plasmids

pcDNA3.1/His-ACVR2 and pTRE-Tight-BI-DsRed-ACVR2 plasmids were kindly provided by Dr. G. Patsos [122] and used as template DNA for the cloning of the wildtype *ACVR2* gene. The following plasmid vectors were kindly provided by Dr. I. Weidenfeld and Dr. K. Schönig (Central Health Institute, Mannheim, Germany) and used for the recombination-mediated cassette exchange (RMCE): the retroviral vector S2F-cLM2CG-FRT3 [133] contained a tetracycline (tet)-controlled bidirectional transcription unit for concurrent regulation of the two reporter genes firefly *luciferase* and red fluorescent protein *mCherry* for the screening of the stable integration site of the HCT116-mCherry cell line. The same plasmid backbone was used for replacing the *mCherry* gene by the *ACVR2* gene. The plasmid pE11.F3.HygTK.F [133] encoding a hygromycin B phosphotransferase-thymidine kinase (HygTK) translational fusion protein was used for antibiotic selection and generation of the HCT116-HygTK master cell line [134]. For retroviral assembly, the vectors pVPack-GP and pVPack-VSV-G from Stratagene (Böblingen, Germany) were used. Recombination was mediated by the enzyme flpo-recombinase that is encoded by the plasmid pCAGGS-Flpo-IRES-Puro which has been kindly provided by Dr. M. Hahn (DKFZ, Heidelberg, Germany).

2.8 Solution Recipes

Coomassie staining solution	5% Aluminium sulfate
	10% Ethanol
	0.02% Coomassie brilliant blue G250
	2% Phosphoric acid

Developing solution for silver staining	6% Sodium carbonate 4 mg/l Sodium thiosulfate 0.0074% Formaldehyde
1x DNA sample buffer (pH 7.6)	25 mM Tris 0.042% Bromophenol blue 10% Glycerol
Fixation buffer for Coomassie staining	30% Ethanol 2% Phosphoric acid
Fixation buffer 1 for Silver staining	40% Methanol 10% Acetic acid
Fixation buffer 2 for Silver staining	10% Ethanol 5% Acetic acid
Fixation buffer 3 for Silver staining	10% Ethanol
Fixation buffer for Sypro staining	50% Methanol 7% Acetic acid
LB-medium	1% Trypton 0.5% Yeast extract 1% Sodium chloride
LB-Agar with ampicillin	LB-medium with 2% Bacto-agar 50 µg/ml Ampicillin
Oxidating solution for silver staining	0.02% Sodium thiosulfate 0.0074% Formaldehyde
PBS	137 mM NaCl 27 mM KCl 100 mM Na_2HPO_4 (anhydrous) 20 mM KH_2HPO_4

PBST	PBS containing 0.1% Tween-20
1x RIPA buffer (pH 7.4)	50 mM Tris-HCl 150 mM NaCl 1% Triton X-100 1% Sodium deoxycholate 0.1% SDS 0.1 mM $CaCl_2$ 0.01 mM $MgCl_2$
Silver staining solution	2 g/l Silver nitrate 0.0185% Formaldehyde
1x TBE buffer (pH 8)	0.1 M Tris-base 83 mM Boric acid 1 mM EDTA
TBS (pH 7.5)	20 mM Tris-HCl 0.5 M NaCl
TBST (pH 7.5)	TBS containing 0.1% Tween-20
Washing buffer for Sypro staining	10% Methanol 7% Acetic acid

2.9 Biological Material

Bacterial Strain	**Characteristics**	**Supplier**
E. coli DH5α	F^- Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17(r_k^-, m_k^+) *pho*A *sup*E44 *thi*-1 *gyr*A96 *rel*A1 λ^-	Life Technologies (Darmstadt, Germany)
Cell Line	**Characteristics**	**Supplier**
HCT116	Human colorectal carcinoma cell line	American Type Culture Collection (ATCC) (Wesel, Germany)

HCT116 AWE17	HCT116 derivative expressing reverse transcriptional transactivator gene (rtTA) and EGFP fluorescent protein	Dr. A. Welman (Uni. of Edinburgh, UK) & Dr. C. Dive (Uni. of Manchester, UK)
HepG2	human liver hepatocellular carcinoma cell line	Dr. K. Breuhahn (Uni. Hospital Heidelberg, Germany)
293T	Human embryonic kidney cell line expressing the SV40 large T-antigen	ATCC (Wesel, Germany)

Cell lines were cultured in DMEM / Ham's F-12 (1:1) or RPMI 1640 with L-Glutamine supplemented with 10% heat-inactivated fetal bovine serum (FBS) Gold and 100 U/ml penicillin and 100 µg/ml streptomycin. In order to freeze the cells, freezing medium containing 10% DMSO in heat-inactivated FBS Gold was used. Transfection experiments were carried out using Fugene HD Transfection Reagent according to the manufacturer‘s instruction. Hygromycin B (Hyg, 100 µg/ml), puromycin (1.5 µg/ml) and ganciclovir (Gan, 40 µM) were used for antibiotic selection. To examine signaling effects, cells were starved for 18 h in the presence and absence of 0.5-1 µg/ml Doxycycline and subsequently incubated with 10 ng/ml recombinant Activin A for 2 h. For metabolic labeling experiments using L-Azidohomoalanine cells were cultured in methionine-free RPMI medium supplemented with 10% heat-inactivated FBS Gold and 100 U/ml penicillin and 100 mg/ml streptomycin using standard conditions.

3. METHODS

3.1 Molecular Biology Methods

3.1.1 Isolation of Genomic DNA and RNA

1-7x10^6 cells were harvested and washed by using PBS. gDNA and total RNA was extracted from the cells by using the DNeasy Blood & Tissue Kit and the RNeasy Mini Kit from Qiagen (Hilden, Germany), according to the manufacturer‘s protocol, respectively. The concentration of eluted gDNA or total RNA was subsequently determined by photometric measurement at 260 nm.

3.1.2 PCR

Standard PCR was performed by using the HOT FIRE Pol DNA Polymerase with the following cycling conditions: initial denaturation at 95 °C 15 min; 30-40 cycles at 95 °C denaturation for 30 s, 60 °C annealing for 30 s, 72 °C elongation for 30-60 s and a final elongation step at 72 °C for 2-10 min.

Inserts for cloning were amplified by using the proofreading Phusion High-Fidelity DNA Polymerase. The specificity of the PCR reaction was tested by using the same reaction batch without template DNA. The reactions were prepared on ice according to the standard protocol depicted in Table 3.1. After PCR, DNA fragments for cloning were purified by using the High Pure PCR Product Purification Kit from Roche (Mannheim, Germany), following the manufacturer’s protocol.

Table 3.1. Standard PCR reaction mixture (1x).

Component	Final Concentration	Volume [µl]
Template DNA	cDNA ~20 ng; plasmid DNA ~0.1-20 ng; gDNA ~200 ng	4
Buffer	1x	2.5
$MgCl_2$	2 mM	2
dNTP Mix	0.2 mM	2.5
Primer forward	0.5 mM	2.5
Primer reverse	0.5 mM	2.5

DNA Polymerase	0.625 U	0.125
Water, DNase, RNase-FREE	-	up to 25

3.1.3 Restriction Digest

In general, reactions were performed on ice with 30 µg of vector DNA and 60-100 µl of purified PCR product. For the restriction digestion, 1x BSA and 1x NEBuffer of the respective enzyme was added according to the manufacturer's recommendations. Reactions were filled up to an appropriate volume with double-distilled water and incubated for at least 3 h at 37 °C. In particular, for sticky end cloning, the PCR products of the *ACVR2* wildtype cDNA (insert) containing an *EcoRI* restriction site at the 5´-end and a *NotI* at the 3´-end were digested in one reaction. The vector S2F-cLM2CG-FRT3, harboring the *EcoRI*/*NotI mCherry* fragment was also digested in a double digestion reaction. Both enzymes were used each at a concentration of 40 U/µl.

In contrast to this preparative restriction digestion, analytical digestions were performed in order to verify proper ligation after cloning of the DNA fragments. Therefore, smaller amounts of plasmid DNA (0.2-1 µg) were used and incubated for a minimum of 1 h. After enzymatic digestion, the enzymes were heat-inactivated for 20 min at 65 °C.

For the PCR products, an additional purification step was performed by using the High Pure PCR Product Purification Kit, following the manufacturer's protocol. The vector DNA fragments were separated by agarose gel electrophoresis and subsequently excised from the gel with a scalpel using low energy UV light (366 nm) for visualization. Isolation of the DNA fragments was performed by using the High Pure PCR Product Purification Kit, following the manufacturer's protocol.

3.1.4 Agarose Gel Electrophoresis

The size, quality and quantity of DNA were controlled by agarose gel electrophoresis. Depending on the size of the DNA fragments, 0.8-2% agarose was heated and dissolved in 1x TBE buffer. After the gel solution had cooled down, 0.1-0.5 µg/ml ethidium bromide or 1x GelRed was added. The gel polymerized at room temperature and was mounted in an electrophoresis chamber filled with 1x TBE

buffer. The samples were diluted in 1x DNA sample buffer before loading into the wells. To determine the correct size of the DNA fragments, 5 µl of the DNA Ladder (1 Kb or 100 Bp) from Life Technologies was loaded. The gel was run at a constant voltage of 120 V and the DNA was visualized by UV light. The amount of the DNA was estimated by comparing the band intensities with the known amount of the marker bands.

3.1.5 Dephosphorylation and Ligation

After estimating the DNA concentration by using agarose gel electrophoresis and calculating the required amounts of DNA fragments in moles, assuming a molecular weight of 325 g/mol for a single nucleotide, DNA was dephosphorylated and ligated by using the Rapid DNA Dephos & Ligation Kit from Roche following the manufacturer's protocol. For sticky end cloning, a molar ratio of 1:5 (vector:insert) was applied. As a control for vector self-ligation, an additional reaction was performed in the absence of the insert DNA.

3.1.6 Transformation

For transformation, 50 µl of chemo-competent *E. coli* DH5α bacteria were thawed on ice and 1-2 µl of ligation mixture was added to the cells. The cells were incubated for 30 min on ice and subsequently heat-shocked at 42 °C for 45 s. After an incubation time of 2 min on ice, the cells were resuspended in 300 µl S.O.C. medium and incubated for 1 h while shaking at 300 rpm at 37 °C. Finally, the cells were transferred to pre-warmed LB-agar plates, including 50 µg/ml ampicillin. The agar plates were incubated overnight at 37 °C and then stored at 4 °C.

3.1.7 Colony-PCR

In order to screen for bacterial clones that carry the correct DNA insert, colony-PCR was carried out by using the grown bacterial colonies on the LB-agar plates as the template. At least 5 single colonies were subjected to colony-PCR by using primers that specifically amplify the *ACVR2* cDNA insert. The colonies were either plated onto a fresh 50 µg/ml ampicillin-containing LB-agar plate or directly inoculated into 3-

5 ml of LB-medium, supplemented with 50 µg/ml ampicillin, and grown overnight at 37 °C at 200 rpm for plasmid isolation.

3.1.8 Plasmid Isolation

Depending on the desired amount of plasmid, two alternative scales of DNA preparation were performed: for small-scale plasmid DNA preparation, single colonies were inoculated into 3-5 ml and for large-scale plasmid DNA preparation into 200 ml LB-medium containing 50 µg/ml ampicillin. The bacteria were incubated overnight at 37 °C at 200 rpm and plasmid DNA was prepared either by using the NucleoSpin Plasmid Kit for small-scale and the Plasmid Maxi Kit for large-scale plasmid isolation, according to the manufacturer's protocol. To clear the bacterial lysate of precipitated cell material prior to subjecting it to the column, folded filters were used instead of the supplied QIAfilter Cartridges. The plasmid DNA concentration was determined by a photometer at 260 nm.

3.1.9 Glycerol Stock of Bacteria

For long-term storage, 800 µl of growing overnight cultures of bacteria, containing the correct plasmid, were mixed with 200 µl of glycerol to obtain a final concentration of 20% of glycerol. After vortexing, the bacteria were stored at -80 °C.

3.1.10 Ethanol DNA Precipitation

Since the DNA fragments were eluted in a volume of ~100 µl, it was necessary to perform an ethanol precipitation step to enrich the DNA in a smaller volume suitable for ligation. To this end, 0.3 M sodium acetate, pH 5, 2 volumes of 100% ethanol (or 0.7 volume of isopropanol) and 0.3 µg/µl glycogen (carrier) was added to the sample and vortexed thoroughly. After 5 min incubation at room temperature (or 20 min at -20 °C for lower amounts of DNA), the sample was centrifuged for 20 min at 12,000 g. The pellet was then washed with 1 ml of 70% ethanol (10 min centrifugation at 12,000 g), air-dried for approximately 10 min at room temperature and resuspended in 11 µl of double-distilled water.

3.1.11 Frameshift Mutation Analysis

Frameshift mutational analysis was performed as described by Woerner et al. [56]. In brief, ~100 ng of gDNA was used for the PCR reaction by using a fluorescein labeled primer. The resulting PCR product was visualized by agarose gel electrophoresis in order to estimate the DNA amount. Approximately 2 µl of this PCR reaction was mixed with 13 µl of HI-DI-Rox and analyzed on the ABI Prism 3100 genetic analyzer by using the Sequencing Analysis software Applied Biosystems (version 3.7).

3.1.12 Sequence Analysis

For sequencing of the PCR-amplified *ACVR2* insert, ~200 ng of the cloned S2F-cLM2CG-FRT3-ACVR2 plasmid were used. PCR analysis was carried out by using the plasmid as a template and primers which are specific for the wildtype *ACVR2* transgene. The PCR products were purified by using the High Pure PCR Product Purification Kit and 6-10 µl of the eluate were subjected to the sequencing reaction. Sequencing was performed by using the BigDye Terminator v1.1 Cycle Sequencing Kit and the subsequent analysis was carried out on an ABI Prism 3100 genetic analyzer by using the Sequencing Analysis software Applied Biosystems, version 3.7.

3.1.13 cDNA Synthesis

One µg of total RNA was reverse transcribed by using Oligo (dT) primer and SuperScript II Reverse Transcriptase according to the manufacturer‘s protocol. In order to exclude amplification from remaining genomic DNA, one sample was processed in parallel in the absence of the SuperScript II Reverse Transcriptase enzyme. To validate successful cDNA synthesis, PCR was performed using 1:5 diluted cDNA as template and primer recognizing the housekeeping gene *GAPDH* (*glyceraldehyde 3-phosphate dehydrogenase*) or *18S rRNA*.

3.1.14 Real-time RT-PCR Analysis

For real-time RT-PCR experiments, primers for the respective genes and Power SYBR Green PCR Master Mix were used. Triplicates of different 1:5 diluted cDNA samples (-dox versus +dox) were analyzed in the StepOnePlus Real-Time PCR

System with the following program: 95 °C for 10 min, followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. Data were analyzed by the StepOne Software, v2.1. Analysis of the gene expression was performed by relative quantification by using the $2^{-\Delta\Delta C}_T$ method as described by Livak and Schmittgen [135]. Gene expression was normalized to the expression of the reference genes *GAPDH* or *18S rRNA*.

3.2 Biochemical Methods

3.2.1 Protein Extraction

A minimum of $5x10^6$ cells were washed once with PBS and then scraped off the plates. Cell pellets were washed with PBS and then resuspended in 1x RIPA buffer. After sonication and incubation for 1 h at 4 °C, the suspension was centrifuged (12,000 g, 20 min, 4 °C) and the protein concentration of the lysate was measured.

3.2.2 Protein Precipitation

For protein precipitation, two methods were performed by using either methanol chloroform or TCA (trichloroacetic acid).

Methanol chloroform precipitation was performed as follows: three volumes of methanol were added to the sample and mixed. Then 0.75 volume of chloroform was added and vortexed, followed by the addition of 2 volumes of double-distilled water and vortexing. Phase separation was achieved by centrifugation at 12,000 g for 5 min. Since the interface contained the proteins of interest, the upper aqueous phase was carefully removed and discarded. The sample was washed twice by using 2.25 volumes of methanol, vortexed and centrifuged at 12,000 g for 5 min. The pellet was air-dried and resuspended in required buffer.

The second method for protein precipitation was performed by using trichloroacetic acid. The sample was mixed with 1 volume of 10% TCA, vortexed and centrifuged for 10 min at 12,000 g. The pellet was dissolved in 50 µl of 0.2 N NaOH for 10 min at 56 °C or in required buffer.

3.2.3 Determination of Protein Concentration

Protein concentration was measured by using two different methods, the Bradford or the Lowry assay.

The Bradford assay was performed by using the Bio-Rad Protein Assay according to the manufacturer‘s instruction. Protein concentration was measured in a photometer at 595 nm.

For the Lowry protein assay [136], the sample was precipitated by using TCA and then dissolved in 50 µl of 0.2 N NaOH at 56 °C for 10 min. The sample was incubated for 20 min at room temperature with 500 µl of a solution, containing 0.02% copper(II) sulfate, 39.98 g/l sodium carbonate, 8 g/l NaOH and 0.4 g/l potassium sodium tartrate, followed by the addition of 500 µl of 1x Folin-Ciocalteu‘s phenol reagent and an incubation step of 30 min at room temperature. The absorbance was measured by using the photometer at 630 nm.

To obtain a standard curve, 3-5 different concentrations of BSA were also treated under the same conditions and measured to determine the correct concentration of the samples.

3.2.4 Sodium Dodecylsulfate-Polyacrylamide Gel Electrophoresis

13 µl of sample (20-50 µg protein) were mixed with 5 µl of 4x NuPAGE LDS Sample Buffer and 2 µl of 10x NuPAGE Sample Reducing Agent and incubated for 5 min at 99 °C for denaturation of the proteins. The sample was separated on a gradient NuPAGE 4-12% Bis-Tris Gel by using 1x NuPAGE MES SDS Running Buffer containing 0.25% NuPAGE Antioxidant at 200 V, 250 mA, 50 W for 35-45 min. The SeeBlue Plus2 Pre-Stained Standard protein marker was loaded when the gel was subjected to Western blot analysis. The Mark12 Unstained Standard protein marker was applied when the gel was directly used for staining the proteins in the gel.

3.2.5 Western Blot Analysis

After SDS-PAGE, the separated proteins were electroblotted onto a nitrocellulose membrane at 30 V, 400 mA and 50 W for 1 h in a buffer containing 10% methanol, 0.1% NuPAGE Antioxidant and 1x NuPAGE Transferbuffer. The membrane was briefly washed in TBST. After blocking the membrane for 30 min at room temperature by using 5% skim milk in 1x TBST (blocking solution), the following primary

antibodies were added to the blocking solution: mouse anti-ß-actin (1:20,000, room temperature, 1 h); mouse anti-ACVR2 (1:1000, 4 °C, overnight); rabbit anti-ACVR2 (1:500, 4 °C, overnight); rabbit anti-phospho-SMAD2 (1:1000, 4 °C, overnight); rabbit anti-SMAD2 (1:1000, 4 °C, overnight); mouse anti-FLAG (1:500, 4 °C, overnight). After several washing steps (10 min each at room temperature) in TBST, the blots were incubated with the secondary antibodies anti-mouse IgG-HRP (1:5000) and anti-rabbit IgG-HRP (1:2500) for 1 h at room temperature, respectively. After three washing steps (10 min each at room temperature) in TBST, the signals were detected by using Western Lightning Plus ECL. As transfer and loading control some of the blotting membranes were stained with 0.1% Ponceau S (w/v in 5% acetic acid).

3.2.6 Lectin-Western Blot Analysis

Lectin-Western blotting was performed according to the protocol for Western blot analysis as described above with the exception that biotinylated AAL (2 ng/ml) was used as primary antibodies and streptavidin-HRP (1:20,000) as secondary antibody.

3.2.7 Gel Staining Methods

Three different methods to stain proteins in the gel were performed by using coomassie, SYPRO-Ruby and silver nitrate.

The colloidal coomassie staining was performed as described by Kang et al. [137]. Briefly, after washing the gel three times with de-ionized water, the proteins were fixed by using 30% ethanol and 2% phosphoric acid for 30 min. The gel was stained by using the coomassie staining solution. Destaining of the gel by de-ionized water was performed until the bands were clearly visible.

In order to detect lower amounts of proteins (>1 ng/band), the fluorescent SYPRO-Ruby solution was used according to the manufacturer‘s instruction. The rapid protocol was performed: fixation of the proteins twice for 15 min by using 50% methanol and 7% acetic acid, staining by using the SYPRO-Ruby Gel Stain. After adding this solution, several rounds of heating in the microwave and agitating at room temperature were performed (30 s heating, 30 s agitating, 30 s heating, agitating 5 min, heating 30 s and agitating for 23 min). The gel was subsequently washed by using a mixture of 10% methanol and 7% acetic acid for 30 min. The gel

was subjected to de-ionized water and the fluorescence signal was visualized by UV light.

Silver staining was performed according to the following scheme: fixation for 1 h with 40% methanol and 10% acetic acid, followed by 30 min incubation 10% ethanol and 5% acetic acid, and 15 min with 10% ethanol. The gel was washed once with de-ionized water followed by a 1 min incubation step with an oxidating solution. After three times washing in de-ionized water for 1 min, the gel was incubated in the staining solution for 15 min, washed three times for 20 s with de-ionized water and incubated in the developing solution until the protein bands were visible. The gel was washed with de-ionized water 3x 1 min and staining was terminated by the addition of 10% acetic acid for 15 min. After 3x 5 min washing in de-ionized water, the gel was scanned for documentation.

3.3 Cell Culture Experiments

3.3.1 Maintenance of Cell Lines

All cell lines were grown in culture media (as described in 2.9) at 37 °C in an atmosphere of 5% CO_2. For subculture, confluent cells were washed once with PBS, trypsinized and split into new flasks 2-3 times per week. To ensure that the cells were free of mycoplasm contaminations, the cultures were continuously checked by conventional PCR using a Mycoplasma Detection Kit, according to the manufacturer‘s protocol.

3.3.2 Freezing and Thawing of Cells

For cell line preservation, cells were washed with PBS, trypsinized and approximately $5x10^6$ cells / cryo vial were resuspended in cold freezing medium. The cells were first frozen at -80 °C and transferred to liquid nitrogen for long-term storage. In order to thaw cells, a vial of frozen cells was placed in a 37 °C water bath, thawed and cells were directly washed with 10 ml of regular cell culture medium. Finally, the cells were centrifuged at 1000 g for 10 min, resuspended in culture medium and seeded in T75 flasks.

3.3.3 Cell Transfection

Transfection experiments were carried out by using different methods: FuGENE HD-mediated lipofection or electroporation.

The FuGENE HD Transfection Reagent was used according to the manufacturer‘s instruction. Briefly, prior to the transfection, $1x10^6$ cells were seeded on 6-well plates to reach a confluency of ~80% at the day of transfection. For HCT116 cells, a ratio of 4:2 (4 µl transfection reagent and 2 µg of total DNA) has been used.

For electroporation, $1x10^7$ cells were harvested at confluency of approximately 80% and electroporated by using 6-10 µg of DNA. Electroporation was conducted by the Amaxa Cell Line Nucleofector Kit V according to the manufacturer‘s instruction by using the program D32 of the Nucleofector I device for HCT116 cells. After electroporation, the cells were transferred to 10 cm culture dishes.

3.3.4 Recombinase-Mediated Cassette Exchange

The recombinase-mediated cassette exchange (RMCE) strategy was performed as described previously [133]. The HCT116-HygTK master cell line, which was Hyg-resistant and sensitive to Gan (Hyg^r, Gan^s), was generated by a first RMCE step as described by Lee et al. [134]. This master cell line was subjected to a second RMCE step resulting in HCT116-ACVR2 #2 clones that conferred dox-inducible expression of ACVR2 with a FLAG-tag and luciferase concurrently. Quantification of dox-inducible expression levels were determined by luciferase assays.

3.3.5 Luciferase Assay

$2x10^5$ cells / 24-well or $4x10^5$ cells / 12-well of HCT116-ACVR2 cells were seeded in duplicates one day prior to the assay and grown in the presence and absence of 1 µg/ml dox. Luciferase activity in cell lysates was measured by the Luciferase Assay System by using a luminometer according to the manufacturer‘s instruction. The luciferase activity was normalized to the protein concentration, which was determined by Bradford assay.

3.3.6 Proliferation Assay

Cells were plated in 96-well plates at densities of $1x10^3$ and grown for one to five

days. Proliferation assays were performed in triplicate by using the CellTiter 96 AQ_{ueous} One Solution Cell Proliferation Assay (MTS) Kit according to the manufacturer‘s instruction. The formation of soluble formazan product is directly proportional to the number of living cells and was determined by measuring the absorbance at 485 nm in an ELISA-Reader.

3.3.7 Glyco-Gene Chip Analysis

For the Glyco-Gene Chip analysis, HCT116-ACVR2 cells were starved overnight in the absence or presence of 1 µg/ml dox in triplicate. After incubation with 10 ng/ml activin A for 2 h RNA was isolated, RNA integrity was verified by agarose gel electrophoresis and the concentration was measured spectrophotometrically at 260 nm. 3 µg of RNA was stored in RNAstable according to the manufacturer‘s protocol and subsequently sent to the Consortium for Functional Glycomics (CFG) for Glyco-Gene Chip analysis (for detailed information see the internet webpage: www.functionalglycomics.org/static/consortium/resources/resourcecoree.shtml).
Additionally, 1 µg of the same RNA was reverse transcribed for further analyses by real-time RT-PCR.

3.3.8 Lectin-FACS Analysis

Lectin-FACS analysis was performed as described previously [122]. Cells were grown in the presence or absence of 1 µg/ml dox and 10 ng/ml activin A for 24 h or 72 h. For analysis $4x10^5$ cells were first washed with PBS to remove the residual medium (1000 g for 5-10 min at 4 °C). Then cells were resuspended in PBS containing 0.1% BSA and incubated with one of the biotinylated plant lectins for 30 min on ice. After washing the cells with PBS / 0.1% BSA, the cells were incubated with streptavidin-R-phycoerythrin (streptavidin-R-PE) in 1:40 dilution for 30 min on ice. The cells were washed once again and then analyzed in the FACSCalibur instrument. The parental HCT116 cell line (negative) and the HCT116-Tet-On cell line (EGFP) were used for adjusting the parameter settings. 10,000 counts were acquired for each data file and subsequent analysis was performed using the CellQuest Pro Software.

3.3.9 Immunoprecipitation Analysis

For immunoprecipitation of FLAG-labeled ACVR2 the anti-FLAG M2 Affinity Gel was applied according to the manufacturer`s instructions. 5 mg of HCT116-ACVR2 #2 protein dissolved in a total volume of 600 µl RIPA buffer were mixed with 130 µl of affinity gel slurry and incubated overnight on a rotor mixer. After 3 washes with 1 ml TBS, the immunoprecipitated ACVR2-FLAG protein was eluted with 70 µl of 2x SDS-PAGE sample buffer at 99 °C for 5 min. After centrifugation at 1000 g 20 µl of the supernatant was subjected to SDS-PAGE and subsequently analyzed by Western blotting.

3.3.10 Immunocytology

For immunocytology (IC) experiments, $1x10^4$ cells were seeded on diagnostic slides in 50 µl volume of medium. After 24 h incubation, the cells were fixed: the medium was discarded, the cells washed once with PBS and sparkled with a Cytofixx pump spray. The cells were additionally fixed by incubating the slides with ice-cold methanol for 8 min at -20 °C, followed by ice-cold aceton for 8 min at -20 °C. The cells were stored at -20 °C until further processing. The fixed cells were washed in PBS and then treated with methanol containing 0.6% hydrogen peroxide for 20 min at room temperature in order to block the endogenous peroxidase enzyme. The cells were washed several times in de-ionized water and the Dako Pen was used to encircle the wells. Subsequently, the cells were washed with PBS / 0.1% Tween-20 for 1 min and blocked with 10% horse serum in PBS for 30 min at room temperature in a humid chamber. The cells were washed 3x 10 min by using PBS and incubated with AAL (1:500) in PBS / 1% BSA (biotin-free) overnight at 4 °C in a humid chamber. The following day, the diagnostic slides were washed 3x 10 min with PBS / 0.1% Tween-20 and incubated for 30 min at room temperature in PBS / 1% BSA. The slides were again washed 3x 10 min with PBS / 0.1% Tween-20 and incubated with the pre-incubated (30 min 1:50 in PBS at 4 °C) AB-Complex (Vectastain Elite ABC Kit (Universal)) for 30 min at room temperature. The cells were washed twice for 3 min with PBS / 0.1% Tween-20 and incubated with the DAB-chromogene (1 drop/ml) until staining was sufficient (~15 min). De-ionized water was used to stop the reaction and fresh haemalum solution to counterstain the nuclei for 15 s. The slides were

washed again in tap water and finally in de-ionized water. The wells were covered with aquatex and cover slips and subsequently analyzed by microscopy.

3.4 Labeling Experiments

3.4.1 Radioactive Labeling

$7x10^4$ cells / well were seeded in triplicate on a 6-well plate. After 24 h, cells were re-fed with 2 ml new media containing 0.185 MBq of the respective ^{3}H-labeled saccharide (Acetyl-D-mannosamine, *N*-[mannosamine-6-^{3}H] (^{3}H-ManNAc) [185-370 GBq/mmol] or Fucose, L-[6-^{3}H] (^{3}H-L-fucose) [1.48-2.22 TBq/mmol]) and 10 ng/ml activin A. Cells were grown in the presence or absence of 0.5 µg/ml dox until they reached a confluency of 60-80%. After 72 h, the cells were washed 3 times with PBS and scraped off. The cells were centrifuged for 5 min at 1000 g at room temperature and washed with PBS. The cell pellet was solubilized in 400 µl of 0.2 N NaOH for 1 h at 56 °C. The protein concentration of two aliquots (25 µl each) was determined by the Lowry assay. 10 µl of 10 µg/µl BSA (precipitation aid) and 400 µl of 10% TCA were added to precipitate the proteins by centrifugation (10 min, 12,000 g, room temperature) and to remove unincorporated labeled saccharides. The pellet was resuspended in 400 µl of 1 N NaOH and neutralized by the addition of 200 µl of 2.5 N acetic acid and subsequently mixed with 10 ml of Ultima Gold LSC Cocktail. The samples were counted by using a liquid scintillation analyzer and dpm measurements were conducted with automatic quench correction applying the transformed Spectral Index of the External Standard / Automatic Efficiency Control (tSIE / AEC) method. Results were expressed as decays per minute (dpm) and normalized to the protein amount (mg).

3.4.2 Click-it Technology with Immunoprecipitation Analysis

3-6 x 10^5 cells / dish were plated in triplicate onto 100 mm dishes suspended in 10 ml RPMI cell media containing -/+ 0.5 µg/ml dox. After 24 h cells they were metabolically labelled with either 40 µM fucose-alkyne or azido-homoalanine (AHA). AHA-labelling was performed for 4 h in methionine-free RPMI cell medium in the absence or presence of 0.5 µg/ml doxycycline and 10 ng/ml activin A. For fucose-labelling normal RPMI cell medium with the same conditions was used for 72 h. Cells were

harvested with PBS, cell pellets were lysed in 1% SDS in 50 mM Tris-HCl, pH 8, supplemented with protease inhibitor cocktail and phosphatase inhibitor cocktail. After sonication and incubation for 30 min at 4 °C on a rotator, cell lysates were centrifuged at 12,000 g for 20 min at 4 °C. Protein concentration was determined using Bradford assay and 200 µg protein was used for detection reaction. For detecting fucose-alkyne biotin-azide and for azido-homoalanine biotin-alkyne was used together with the Click-it Protein Reaction Buffer Kit following the manufacturer's instructions. Unincorporated labeling and detection reagents were removed by methanol-chloroform-precipitation. To extract labeled proteins affinity chromatography using streptavidin-coated magnetic beads was performed. Precipitated samples were resuspended in 200 µl RIPA buffer and incubated with 80 µl streptavidin bead suspension (40 µl pure beads) for 2 h at 4 °C on a rotator. Before usage, bead slurry was washed three times with 1 ml PBST with the help of a magnetic device (6-Tube Magnetic Separation Rack). After incubation of the beads with the samples, the resin was washed three times with 1 ml PBST containing 2% SDS. Finally, protein-bound resin was stored in 100 µl PBS for subsequent mass spectrometry.

3.5 Mass Spectrometry and Data Analysis

3.5.1 Sample Preparation for Mass Spectrometry

Supernatant of samples, which were resuspended in 100 µl PBS, was discard by using a magnetic device and resin was washed with 100 µl 40 mM ammonium bicarbonate. 100 µl DTT (10mM DTT in 40 mM ammonium bicarbonate) was added to the resin and bound proteins were reduced at 45 °C for 1h in a thermomixer at 800 rpm. Free cysteine residues were alkylated with 100 µl iodoacetamide (55 mM in 40 mM ammonium bicarbonate) for 30 min at room temperature in the dark. Between each step resin was washed with 100 µl 40 mM ammonium bicarbonate. Afterwards, samples were digested at 37 °C over night using 50 µl trypsin solution (5 µl of 0,5 µg/µl trypsin in 1mM HCl diluted in 495 µl 40 mM ammonium bicarbonate). After overnight digestion the supernatant was collected into PCR tubes. After adding 5µl 1% trifluoroacetic acid, samples were sonicated for 5 min. After centrifugation at 13,000 rpm, 5 µl of each sample was transferred into new tubes and subsequently analysed by nanoLC ESIMS / MS.

3.5.2 ESI-MS/MS Analysis and Database Search

Tryptic peptide mixtures were separated using a nanoAcquity UPLC system. Peptides were trapped on a nanoAcquity C18 column, 180 µm x 20 mm, particle size 5 µm. The liquid chromatography separation was performed on a C18 column (BEH 130 C18 100 µm x 100 mm, particle size 1.7 µm) with a flow rate of 350 nl/min. For all samples, the chromatography was carried out using a 3 h gradient of solvent A (98.9% water, 1% acetonitrile, 0.1 % formic acid) and solvent B (99.9% acetonitrile and 0.1% formic acid) in the following sequence: from 0 to 4% B in 1 min, from 4 to 30% B in 140 min, from 30 to 45% B in 15 min, from 45 to 90% B in 5 min, 10 min at 90% B, from 90 to 0% B in 0.1 min, and 9.9 min at 0% B. The nanoUPLC system was coupled online to an LTQ Orbitrap XL mass spectrometer. The mass spectrometer was operated in the sensitive mode with the following parameters: capillary voltage 2400 V; capillary temperature 200 °C, normalized collision energy 35 V, activation time 30000 ms. Data were acquired by scan cycles of one FTMS scan with a resolution of 60000 and a range from 370 to 2000 m/z in parallel with six MS/MS scans in the ion trap of the most abundant precursor ions.

The mgf-files generated by Xcalibur software (Thermo Scientific, Bremen, Germany) were used for database searches with the MASCOT search engine (Matrix Science, London, UK; version 2.2) against SwissProt database (http://www.expasy.ch/sprot/sprot-top.htmlRel.2011_05). Taxonomy was set to human. The peptide mass tolerance for database searches was set to 5 ppm and fragment mass tolerance to 0,5 Da. Carbamidomethylation of C was set as fixed modification. Variable modifications included oxidation of M and deamidation of N and Q. One missed cleavage site in case of incomplete trypsin hydrolysis was allowed. Furthermore, proteins were considered as identified if more than one unique peptide had an individual ion score exceeding the MASCOT identity threshold (ion score cut-off of 30). Identification under the applied search parameters refers to False Discovery Rate (FDR) < 3,5% and a match probability of $p<0.05$, where p is the probability that the observed match is a random event.

3.5.3 Data analysis

Microsoft Office Excel was used to identify proteins, which were found at least in one sample of the triplicates. ACVR2-proficient proteins are defined as proteins that were

found in at least in one of the +dox triplicates but not in any of the -dox samples. For ACVR2-deficient proteins the reverse situation was considered. For classification and functional analysis Web-based tools were used. Blast2GO was applied for functional annotation of the identified sequences by their GO number.

4. RESULTS

4.1 Establishment of the Model Cell Lines

4.1.1 Generation of Doxycycline-inducible HCT116-ACVR2 Clones

In order to investigate whether disturbances of ACVR2 expression and signaling, which frequently occur in mismatch repair (MMR)-deficient and high-level microsatellite unstable (MSI-H) colorectal tumors might cause alterations in the protein glycosylation or even in the *"de novo"* or steady state cellular proteome, we established a cell line based model system. Since constitutive ACVR2 expression could show growth inhibiting effects, we decide to get use of a tet-inducible system in order to control ACVR2 expression. We pursued two approaches to generate ACVR2 stable transfectants, but for both we used a model system requiring HCT116 cells that contain both a constitutively-expressed transactivator gene as well as a dox-regulated *ACVR2* cDNA stably integrated into the cellular genome. This is a two-component expression system (Figure 4.1). The CRC cell line HCT116 is MMR-deficient and therefore exhibits the MSI phenotype. Furthermore, this cell line is refractory to activin A-mediated signaling due to biallelic frameshift mutations in the A8 coding microsatellite of the endogenous *ACVR2* gene (www.seltarbase.org). These HCT116 cells (HCT116 AWE17) [138] have been engineered to stably express the reverse transcriptional transactivator gene (rtTA) under the control of the chicken β-actin promoter. This first component (upper part Figure 4.1) allows monitoring of rtTA expression via IRES-based EGFP fluorescence detection and has been successfully used for expression of several different genes [139, 140]. These herein termed HCT116-tet-on cells enable dox-inducible gene expression and are used as the parental cell line. The second component (lower part Figure 4.1) containing a tet-controlled bidirectional transcription unit for simultaneous regulation of the reporter gene firefly *luciferase* or red fluorescent protein *DsRed* and the gene of interest is stably transfected into the parental HCT-tet-on cell line.

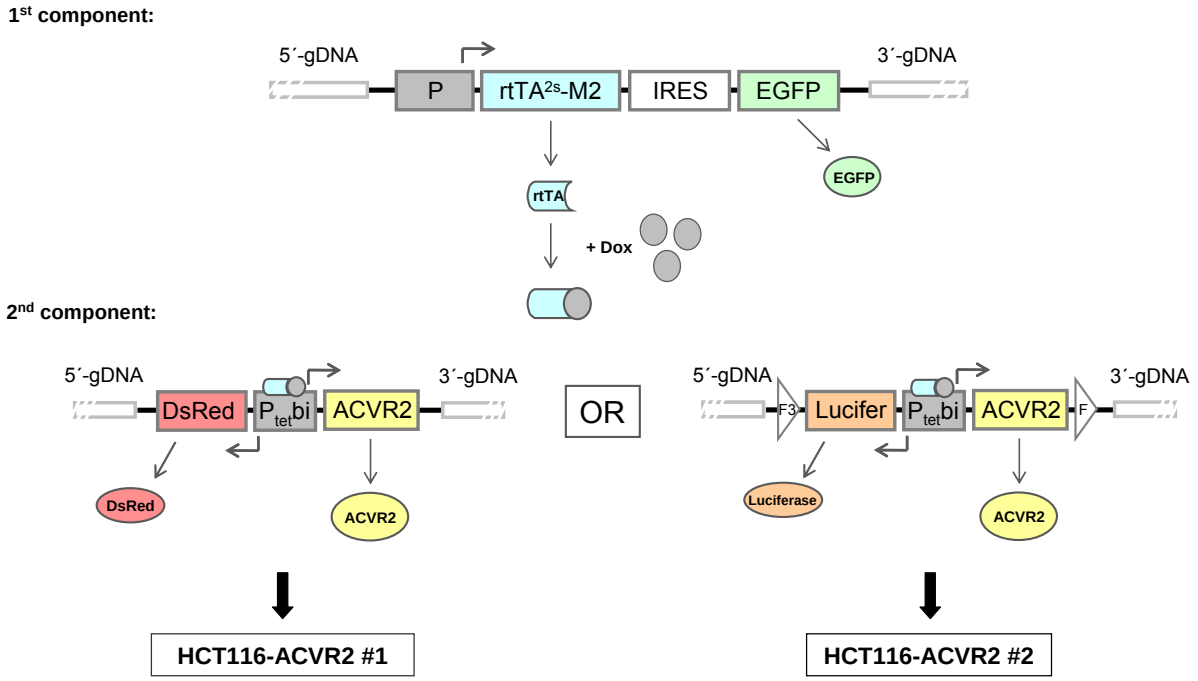


Figure 4.1: HCT116 model system. The model cell system is comprised of two components: the first includes a ß-actin promoter for constitutive expression of the tet-controlled reverse transcriptional transactivator (rtTA) followed by an internal ribosomal entry site (IRES) that allows concurrent transcription of the enhanced green fluorescence protein (EGFP) to monitor gene expression. The parental cell line encoding this component is referred to as HCT116-tet-on cell line. The second component is a construct that contains a bidirectional dox-inducible promoter (P_{tet}bi) allowing concurrent expression of two marker genes (*luciferase* or *DsRed* and *ACVR2*). The retroviral expression cassette is flanked by mutant (F3) and wildtype Flp-recombinase target sites (F). The resulting cell line is designated as HCT116-ACVR2 #1 and #2.

In the first approach ACVR2 was cloned by PCR into the dox-regulatable plasmid pTRE-Tight-BI-DsRed and verified by DNA sequencing. This expression constructs was then co-transfected into HCT116-tet-on cells with a selectable marker plasmid conferring Hygromycin B resistance. Once hyg-resistant clones have grown to reasonable size they were split and grown in the presence and absence of 1 µg/ml dox. Clones that display strong DsRed fluorescence under the fluorescent microscope in the presence of dox were also expected to express the candidate gene because both are transcribed from the same dox-regulatable bidirectional promoter. One clone, HCT116-ACVR2 #1, obtained by this approach was used for subsequent studies (left lower panel of Figure 4.1).

In the second approach we aimed to overcome the problem of stable integration of multiple copies at different genomic sites. Therefore, a retroviral vector that -

depending on the multiplicity of infection (MOI) - allows isolation of clones with single copy integration was used. This approach is based on the publication of Weidenfeld et al. [133] and was applied to establish a second ACVR2-reconstituted cell line. As recipient cells we used the recently described colorectal cancer master cell line HCT116-HygTK [134] that carry at a specific genomic site a *HygTK* marker gene flanked by two recombination sites enabling integration of any gene of interest at this particular site (Figure 4.2). For generation of the retroviral vector S2FcLM2CG-FRT3-ACVR2, the wildtype *ACVR2* cDNA was amplified by PCR from the expression plasmid pcDNA3.1/His-ACVR2 and a FLAG-tag was added at the C-terminus directly before the stop-codon of ACVR2 by PCR. Using the two recombination sites of the master cell line, in a subsequent genomic targeting approach (Figure 4.2), the dox-regulated *Luciferase-ACVR2–FLAG* expression cassette was inserted into this specific genomic site of HCT116-HygTK cells by a single recombination step thereby replacing the HygTK expression cassette and conferring ganciclovir-resistance to the recombinant clones. The resulting HCT116-ACVR2 # 2 cells showed bi-directional expression of an *ACVR2* and *luciferase* reporter gene expression in a dox-dependent manner (right lower panel of Figure 4.1).

Since this reversible expression occurs in an otherwise isogenic background, *ACVR2*-dependent effects can be specifically examined. It is important to note that results obtained from both ACVR2-reconstituted systems should reflect the inverse situation of the ACVR2-deficient status in MSI primary colorectal tumors. Both genetically modified HCT116-ACVR2 clones (#1 and #2) were further characterized and used for subsequent analyses.

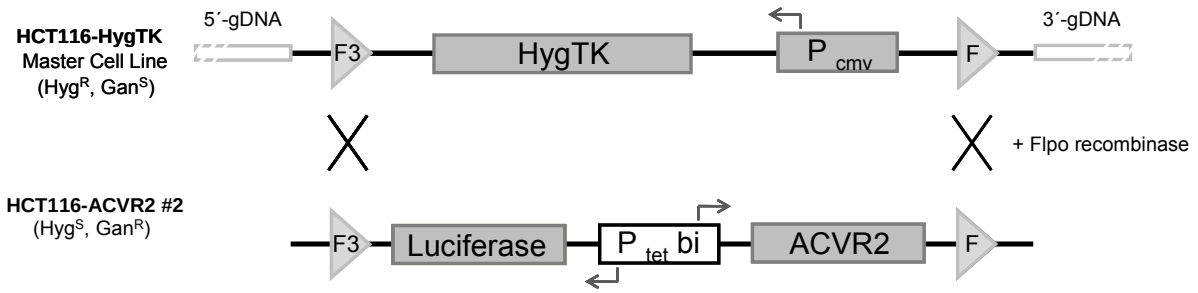


Figure 4.2: Recombination-mediated cassette exchange (RMCE) strategy. The previously described HCT116-HygTK master cell line (Hyg^R, Ganciclovir^S) carries a *HygTK* expression cassette that is regulated by a constitutive promoter (P_{CMV}) and is flanked by mutated (*F3*) and wildtype (*F*)

recombination sites. Co-transfection of these cells with the targeting vector S2F-cLM2CG-FRT3-luc-ACVR2 and a *Flpo* recombinase encoding plasmid leads to recombination-mediated replacement of the *HygTK* expression cassette by the dox-regulated bi-directional *luciferase-ACVR2* expression cassette.

4.1.2 Characterization of the HCT116-ACVR2 Clones

Next these two HCT116-ACVR2 clones were characterized in more detail. Using fluorescence microscopy HCT116-ACVR2 #1 cells showed inducible DsRed expression in a dox-dependent manner. Moreover, inducible expression of this marker was maintained for several months. Luciferase analysis of HCT116-ACVR2 #2 cells revealed an approximately 250-fold increase of luciferase activity compared to the same cells grown in the absence of dox.

Since these model cell lines lack normal DNA mismatch repair function and thus have an increased mutation rate, the A8 coding microsatellite of the reconstituted *ACVR2* gene might be affected by frameshift mutations during selection and growth of HCT116-ACVR2 cells. Therefore, we performed DNA fragment length analysis on genomic DNA of HCT116-tet-on control cells and both HCT116-ACVR2 cell clones. This type of analysis allows amplicon length resolution at the single nucleotide level. Amplification of the A8 repeat of the integrated ACVR2 cDNA from both ACVR2-reconstituted cell clones (#1 and #2) revealed two peaks (Figure 4.3). The first peak corresponds to an amplicon harboring the *ACVR2* A7 mutant coding microsatellite whereas the second peak originates from an amplicon comprised of the *ACVR2* A8 wildtype repeat (-/wt). In contrast, HCT116-tet-on control cells showed only a single peak that reflects A7 coding repeat frameshift mutations in both alleles of the endogenous *ACVR2* gene (-/-). These results demonstrate that both HCT116-ACVR2 cell clones have successfully integrated a wildtype copy of the *ACVR2* transgene into their genome.

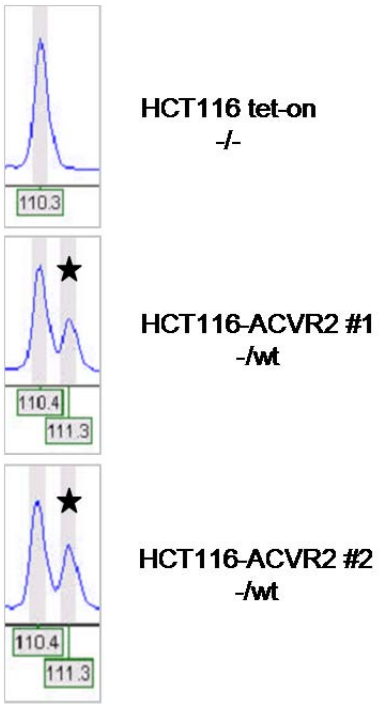


Figure 4.3: Fragment analysis of the ACVR2 A7/A8 coding repeat. Electropherograms show a single peak in HCT116-tet-on cells (homozygous mutant A7 alleles of the endogenous ACVR2 gene) in contrast to the both HCT116-ACVR2 cell clones that carry a wildtype A8 allele (asterisk) in addition to the endogenous mutant allele.

Moreover, we asked whether the integrated *ACVR2* transgene is successfully transcribed and therefore transcript-specific primers in real-time RT-PCR analysis were used to compare the expression of the endogenous A7 mutant ACVR2 transcript with the transgenic A8 ACVR2 wildtype transcript. No change in the endogenous mutant ACVR2 transcript level was observed for both reconstituted cell clones. Instead, dox treatment (1 µg/ml; 20 h) led to a strong induction of the transgenic ACVR2 wildtype transcript. HCT116-ACVR2 #1 showed a 23-fold induction and HCT116-ACVR2 #2 a 63-fold induction (Figure 4.4).

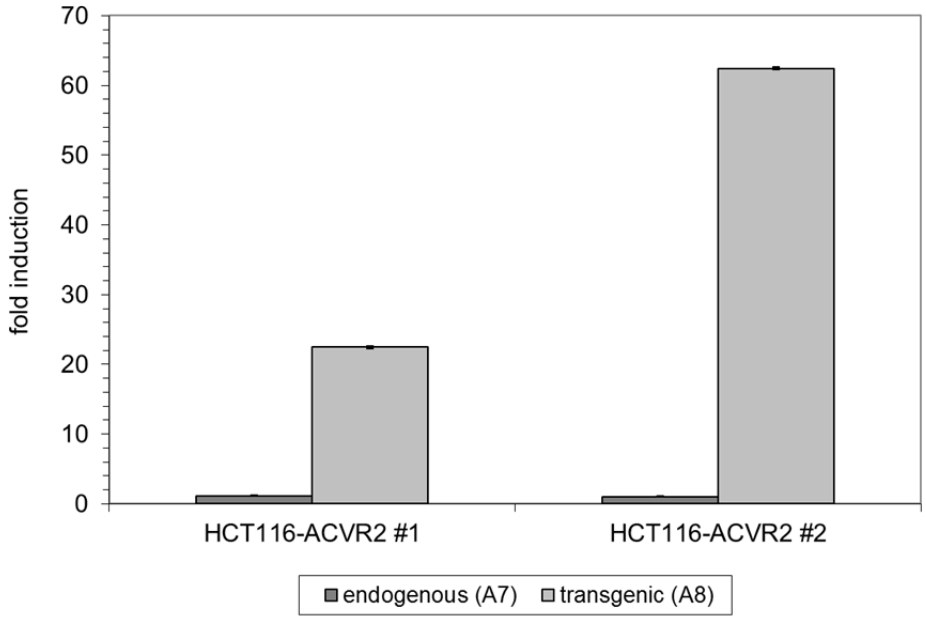


Figure 4.4: Reconstitution of ACVR2 on transcript level. Real-time RT-PCR analysis of the endogenous mutant (A7) and the transgenic wildtype (A8) ACVRR2 transcripts in the absence and presence of dox. Results represent the mean of three independent observations ± S.D.

In addition to these genetic reconstitutions of the *ACVR2* gene and transcript analyses, the inducibility and functionality of the ACVR2 protein was examined by Western blot analysis. First expression of the ACVR2 protein was analyzed for both reconstituted cell clones by using a commercially obtained anti-ACVR2 antibody as well as an antibody kindly provided by the group of Barbara H. Jung (Northwestern Memorial Hospital, Chicago, USA). Different conditions were carried out (for culturing the cells as well as for performing the Western blot), but none was successful to detect an induction of ACVR2 protein level in the presence of dox of the expected size range of 60 kDa (data not shown). Since the HCT116-ACVR2 #2 cell clone contained a C-terminal FLAG-tag, the induced ACVR2-FLAG was immune-precipitated by targeting its FLAG-tag and subsequently immunoblotting was performed. The ACVR2 protein was detected as a pattern of bands with an apparent molecular weight in the range of 62 kDa in immunoprecipitates from induced (+dox) HCT116-ACVR2 cells but not from uninduced (-dox) HCT116-ACVR2 or parental HCT116 cells (Figure 4.5). HEK293 cells transiently transfected with an ACVR2 expression plasmid showed the same band pattern and served as positive control.

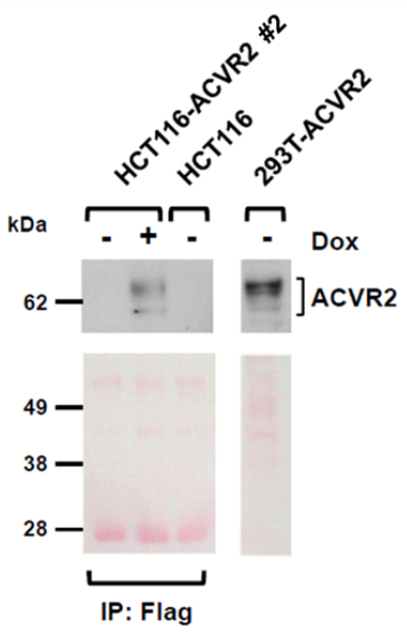


Figure 4.5: Reconstitution of ACVR2 on protein level. Inducible ACVR2 expression using immuno-precipitation (anti-Flag) followed by Western blot analysis (anti-ACVR2A) (upper section) and Ponceau S staining as a loading control (lower section). In the presence of dox for 24 h the expression of the receptor protein was detected after IP in the established HCT116-ACVR2 #2 cell line, whereas in the absence of dox no protein was visible just as in the parental HCT116 cell line. 293T cells were transiently transfected using the pcDNA3.1/His-ACVR2 construct and used as positive control.

In addition to proving induction of ACVR2 protein expression, we also examined the functional reconstitution of the ACVR2 signaling pathway as an indirect proof. Canonical ACVR2-mediated signaling is initiated by the ligand activin A (Act A), which upon ACVR1-ACVR2 receptor heterodimerization leads to recruitment and phosphorylation of receptor-regulated Smad proteins (Smad2, Smad3). Accordingly, phosphorylation of Smad2, the first downstream effector, was examined by Western blot analysis. Both HCT116-ACVR2 cell clones exposed to dox (1 µg/ml; 20 h) and the ligand Act A (2 h; 10 ng/ml) displayed higher levels of pSmad2 in comparison to cells grown in the absence of dox (Figure 4.6). In the absence of Act A and independent of ACVR2 reconstitution, almost no pSmad2 protein was observed in HCT116-ACVR2 cells. This is similar to the marginal pSmad2 protein expression in the HepG2 control cells grown in the absence of TGFß1. This low pSmad2 level was increased upon Act A exposure even in the absence of functional ACVR2 protein (+Act A, -Dox). However, the significant additional increase of pSmad2 level observed upon dox-induction was much higher than the basal level and this difference is clearly ACVR2-dependent. This ACVR2-dependent signaling activity can thus be interrogated by comparative analysis of both HCT116-ACVR2 cell clones grown in the presence of ligand and in the presence / absence of dox. The observed alterations are pSmad2-specific because overall Smad2 protein levels remained unaffected (Figure 4.6).

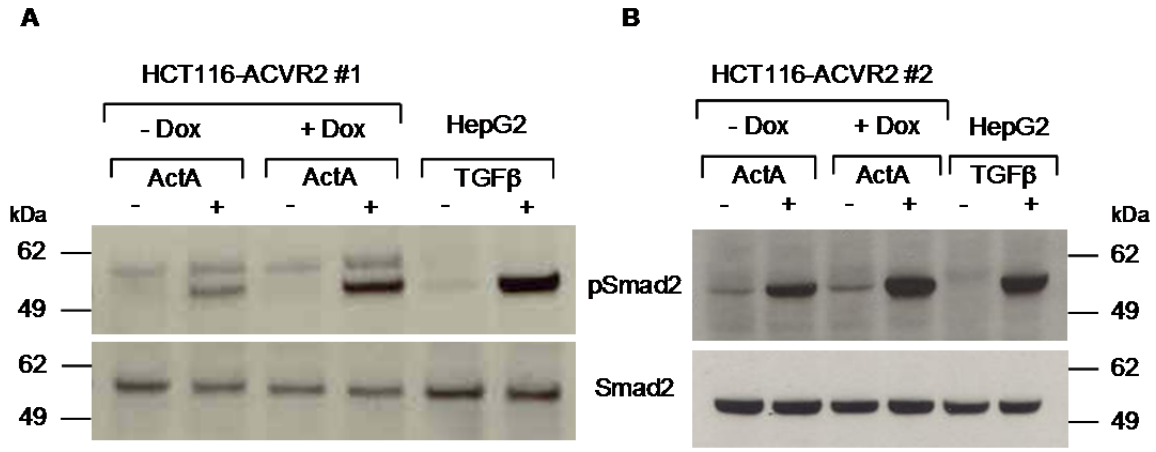


Figure 4.6: Reconstitution of ACVR2 signaling. Detection of Smad2 phosphorylation (pSmad2) by Western blot analysis. Treatment with dox (1 µg/ml) and Act A (10 ng/ml) displayed higher levels of pSmad2 in comparison to cells grown in the absence of dox. TGFß1-responsive HepG2 cells were used as a positive control. Total Smad2 levels were considered as a loading control.

To further validate proper signaling conferred by the reconstituted ACVR2 protein, we determined the transcriptional regulation of well-known ACVR2 target genes like *SMAD7*, *SERPINE and MYC-C*. Real-time RT-PCR analysis revealed dox-dependent up-regulation of Smad7 and Serpine transcripts in both reconstituted cell lines. cMyc transcript showed a modest down-regulation for HCT116-ACVR2 #2 whereas no effect was detected for HCT116-ACVR2 #1 (Figure 4.7). These results are in agreement with the literature data and confirmed the functional reconstitution of ACVR2-dependent signaling in HCT116-ACVR2 cells.

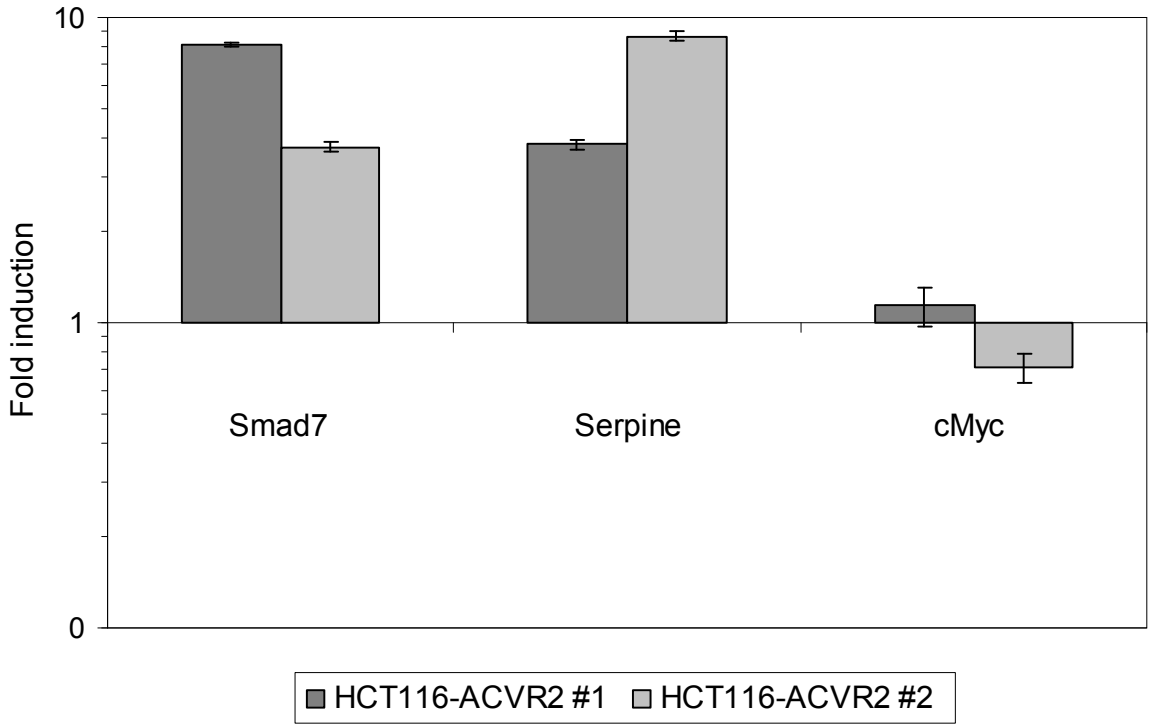


Figure 4.7: ACVR2-dependent regulation of target gene transcription. Real-time RT-PCR experiments displayed dox-dependent Smad7 and Serpine up-regulation for both ACVR2-reconstituted cell lines, whereas cMyc showed a gently dox-dependent down-regulation for the HCT116-ACVR2 #2 cell clone, however HCT116-ACVR2 #1 shows no dox-effect. Data are expressed as the mean and S.D. of three independent experiments.

In order to investigate whether ACVR2 reconstitution alters cell growth, proliferation assays were carried out. When both cell clones (#1 and #2) and the parental cell line were grown in the presence or absence of 0.5 µg/ml dox and in the presence of 10 ng/ml Act A over a period of 5 days the parental HCT116-tet-on cell line grew slightly faster than both ACVR2-reconstituted cell clones. Comparison of ACVR2 cell clones which each other showed that clone #1 grew slightly faster than clone #2. Nevertheless, a significant decrease in proliferation of both ACVR2-expressing cells after five days was observed, whereas the parental cell line displayed no growth inhibitory effect (Figure 4.8).

Overall, these results demonstrate that both ACVR2-reconstituted clones express a functionally intact ACVR2 protein that exhibits proper ACVR2-mediated signaling and confers a growth suppressive effect.

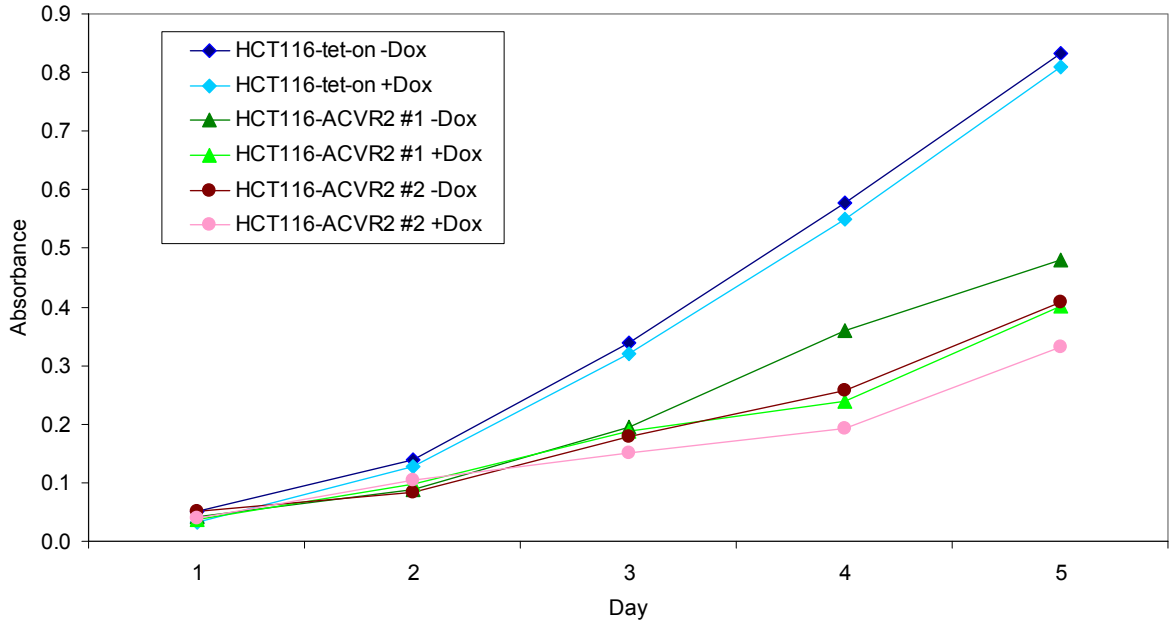


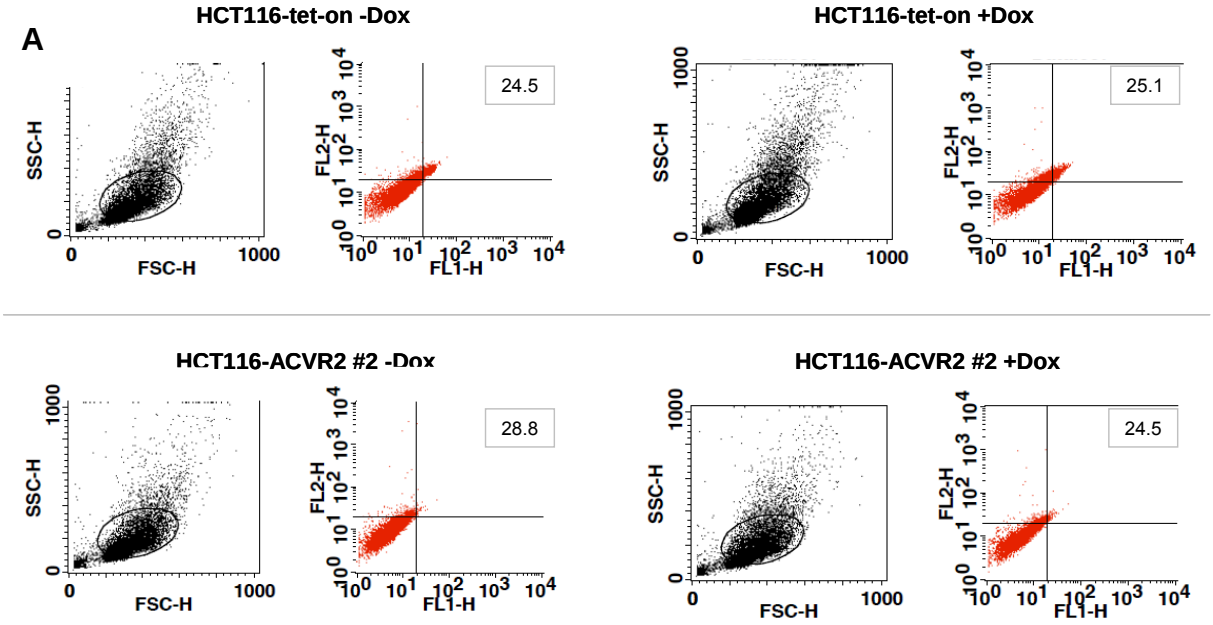
Figure 4.8: Proliferation assay of ACVR2-reconstituted cell clones. The growths of the parental HCT116-tet-on cells were compared with both HCT116-ACVR2 clones (#1 and #2) in presence and absence of dox over a period of five days. Results represent the mean of three independent observations ±S.D.

4.2 Analyses of ACVR2-dependent Glycosylation Changes

4.2.1 Glycosylation Alterations by Lectin Analysis

In recent work with ACVR2 transiently transfected cells we have obtained preliminary evidence that ACVR2 might affect the glycosylation pattern on the cell surface [122]. Therefore, and as a first approach to identify ACVR2-dependent glycosylation changes, a panel of different plant lectins was chosen. Like in the previous transient approach lectin-FACS analyses were performed. However the stable reconstituted ACVR2 cell clone #2 showed no significant alterations (Figure 4.9). Similar results were obtained for the HCT116-ACVR2 #1 cell clone (data not shown). Even upon different cell culture conditions (with / without starvation; 24, 48 or 72 h dox induction) no apparent changes between the parental cell line and the ACVR2-reconstituted cell clones could be detected. Even when lectin-based immunocytological stainings (Figure 4.10) or Western blot analyses (data not shown) were performed, we did not uncover any ACVR2-dependent alterations in protein glycosylation. Most likely, these methods lack sufficient sensitivity that would be required to uncover dynamic changes instead of steady state levels among the total cellular surface proteome.

A



B

	HCT116-tet-on		HCT116-ACVR2 #2	
Lectins	-Dox	+Dox	-Dox	+Dox
AAL	24.5	25.1	25.8	24.5
DBA	88.1	99.3	46.5	54.8
JAC	125.5	157.5	89.9	112.3
LEA	144.5	159.0	219.9	218.3
PSA	96.0	127.2	95.8	123.1
SNA	74.0	92.8	116.3	124.5
VAA	59.6	67.0	83.8	85.9

Figure 4.9: Lectin-FACS analysis. (**A**) Representative FACS analysis of the HCT116-tet-on and HCT116-ACVR2 #2 cells by using biotinylated AAL. Cells were grown in the presence or absence of dox for 72 h. (**B**) Panel of biotinylated plant lectins used for FACS analysis. Values represent the Y geometric mean fluorescence intensities of three independent experiments. FL1-H: EGFP; FL2-H: streptavidin-PE.

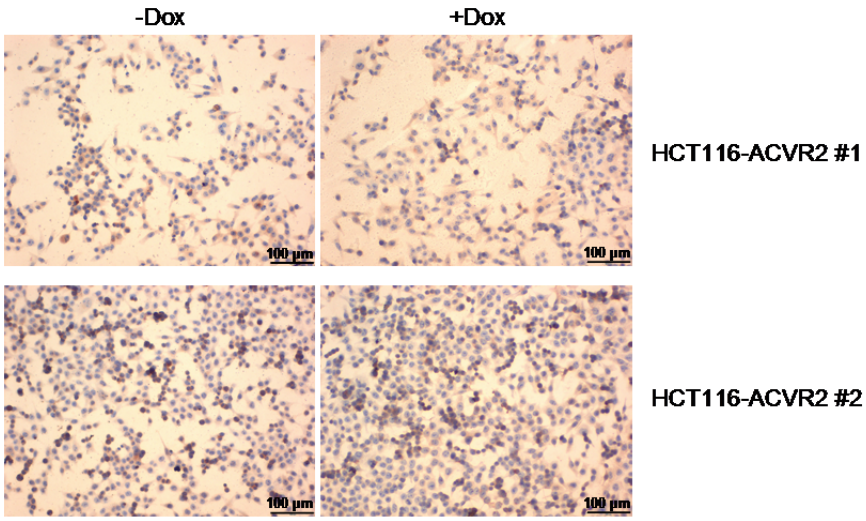


Figure 4.10: Immunocytology (IC) analysis using Aleuria Aurantia Lectin (AAL). IC Staining of both ACVR2-reconstituted cell lines (#1 and #2) in the presence or absence of dox. AAL performed diffuse staining without showing differences in treated versus untreated cells.

4.2.2 Glycosylation Alterations by Radioactive Labeling

Based on these results, we assumed that glycan variations at steady state levels might not be detectable proteins. Accordingly, radioactive labeling experiments were performed using two independent ^{3}H-labeled saccharides, L-fucose and ManNAc. With this experimental approach, measurements were focused on newly synthesized glycoproteins. As a first approach, different time periods (24 h, 48 h and 72 h) were examined. It turned out, that the incorporation of ^{3}H-ManNAc, a precursor of sialic acid, increased over time with a peak at about 72 h. Therefore measurements were performed after 72 h incorporation. Upon dox (0.5 µg/ml) induction and addition of Act A (10 ng/ml) for 72 h, a significant reduction of incorporated ^{3}H-L-fucose occurred in both ACVR2 clones but not in the parental HCT116-tet-on cell line (Figure 4.11 A). Reconstituted expression of wildtype ACVR2 led to a reduction of 11% and 25% for HCT116-ACVR2 #1 and HCT116-ACVR2 #2 cells, respectively. These results suggest that ACVR2 regulates the fucosylation of newly synthesized proteins. After induction of ACVR2 expression, incorporation of ^{3}H-ManNAc showed a significant increase in protein sialylation only in one of both clones (HCT116-ACVR2 #2) (Figure 4.11 B).

In order to identify if differential expression of one of the enzymes known to be involved in fucosylation might have caused the observed reduction of incorporated

fucose, real-time RT-PCR analysis was performed. However, neither fucosyltransferases nor fucosidases nor enzymes involved in the *de novo* or salvage pathway showed any significant alterations (< 0.5- or > 2-fold induction) at the transcript level in dox-treated versus untreated cells (Table 4.1). This was not due to short term treatment because similar results were obtained after 48 h and 72 h.

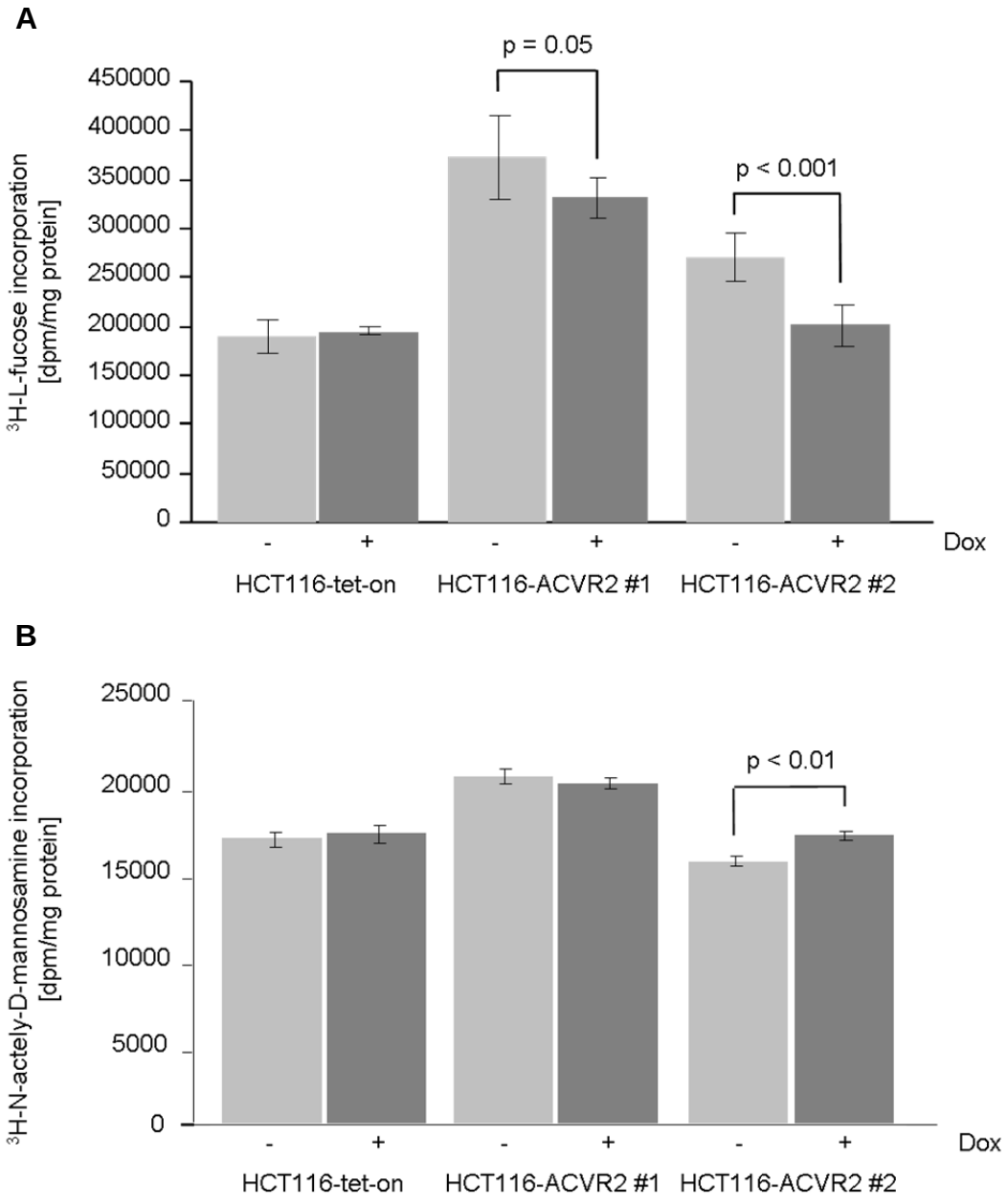


Figure 4.11: Incorporation of ^{3}H-labeled saccharides. Radioactive labeling experiments were performed with both ACVR2-reconstituted cell clones and the parental cell line. Cells were grown in the presence and absence of dox and by exposure to Act A for 72 h. (**A**) Incorporation of ^{3}H-L-fucose was reduced in presence of dox in both ACVR2 clones, in contrast to HCT116-tet-On cells. (**B**) Incubation with ^{3}H-ManNAc resulted in an increase of incorporated ManNAc in the HCT116-ACVR2

#2 cells, but neither in the HCT116-ACVR2 #1 cell clone nor in the parental HCT116-tet-on cell line. Values represent the mean of three independent experiments ±S.D.

Table 4.1: Expression analysis of fucosyltransferases, fucosidases and enzymes involved in the fucose pathway by real-time RT-PCR analysis.

Enzyme	Description	Fold induction
POFUT2	Protein O-fucosyltransferase 2	1.55
FUCA1	Fucosidase, alpha-L- 1	1.50
POFUT1	Protein O-fucosyltransferase 1	1.42
FUCA2	Fucosidase, alpha-L- 2	1.36
FUT2	Fucosyltransferase 2	1.32
FUT11	Fucosyltransferase 11	1.31
FUT 1	Fucosyltransferase 1	1.27
FUT10	Fucosyltransferase 10	1.24
FUK	Fucokinase	1.15
FUT8	Fucosyltransferase 8	1.09
TSTA3	GDP fucose synthase	0.97
TNNI3K	GDP-fucose pyrophosphorylase	0.95
GMDS	GDP-mannose 4,6-dehydratase	0.90
SLC35C1	Solute carrier family 35, member C1	0.85

4.2.3 Glycosylation Alterations by Glyco-Gene Chip Analysis

In order to get a more comprehensive view on transcript alterations induced by ACVR2 expression and possible specific mechanisms of glycan alterations in ACVR2-deficient tumor cells, a glyco-gene oligonucleotide microarray analysis was performed. At this time point only one of both cell clones, namely HCT116-ACVR2 #2, was used for further analysis, since both clones showed similar characteristics in all previous experiments. For this purpose, RNA was extracted from the HCT116-ACVR2 #2 cells and sent to the Consortium for Functional Glycomics (CFG) to compare the transcription of genes involved in glycosylation in the presence or absence of ACVR2. The Glyco-Gene Chip contained transcripts of approximately 1000 human genes, including glycosyltransferases, galectins, growth factors and receptors and glycan-degrading proteins. The results of the Glyco-Gene Chip analysis are summarized in Table 4.2 and revealed changes in several transcripts comparing dox-treated versus untreated cells after 20 h induction and exposure to Act A for 2 h. Notably, the transcript level of ACVR2 itself remained unchanged by the Glyco-Gene Chip analysis, whereas real-time PCR analysis demonstrated dox-inducible increase in ACVR2 transcript level. The microarray data also showed that transcripts of enzymes involved in fucosylation were not regulated by ACVR2 expression, confirming the previous results of the real-time RT-PCR (Table 4.1).

Strikingly, two transcripts involved in the Notch pathway (LFNG and HES1), seemed to be associated with the ACVR2 signaling pathway. The glycosyltransferase LFNG is involved in Notch receptor signaling by changing the binding affinities of the Notch ligands. LFNG transfers GlcNAc onto fucose residues of the EGF repeats of the Notch receptor. GlcNAc is the precursor substrate for the assembly of additional saccharides, such as galactose and α2,3-sialic acid. Thus, studying the interaction between Notch and ACVR2 signaling would be an interesting point for further analysis.

Table 4.2: Glyco-Gene Chip analysis of the HCT116-ACVR2 #2- cell clone.

HCT116-ACVR2 #2				
	Accession No	**Gene Title**	**Category**	**Fold Change**
upregulated	NM_002229	JUNB	Miscellaneous	2.01
	NM_005524	HES1	Notch pathway	1.95
	NM_001040167	LFNG	Glycan-transferase	1.80
	NM_002608	PDGFB	Growth Factors & Receptors	1.71
	uc001pjs	SLC35F2	Nuc. Sugar	1.38
	NM_003862	FGF18	Growth Factors & Receptors	1.32
down-regulated	NM_016523	KLRF1	CBP:C-Type Lectin	0.75
	NM_130851	BMP4	Growth Factors & Receptors	0.71
	NM_002010	FGF9	Growth Factors & Receptors	0.61

4.3 ACVR2-Dependent Alterations of Newly Synthesized Proteins

4.3.1 Analysis of the Whole *De Novo* Proteome

Since studying newly synthesized proteins seemed to be more promising than addressing the steady state levels, newly (*de novo*) synthesized proteins were analyzed as a next step. Therefore, the genetically modified cell line clone #2 was used for a click chemistry approach to determine which proteins might contribute to MSI colorectal tumorigenesis. As a basic principle, the click chemistry approach relies on metabolic labeling with azidohomoalanine, an azide-bearing analog of methionine, that upon incorporation into the nascent polypeptide chain can form a stable triazole conjugate with alkyne-activated biotin thereby enabling selective capture and covalent binding of *de novo* synthesized proteins to streptavidin-conjugated magnetic beads and subsequent identification of tryptic peptides by mass

spectrometry. Applying this approach to Act A stimulated HCT116-ACVR2 #2 cells grown either in the presence or absence of dox and using several filter and selection criteria we identified a set of 513 proteins termed herein the "*de novo* proteome" (Figure 4.12).

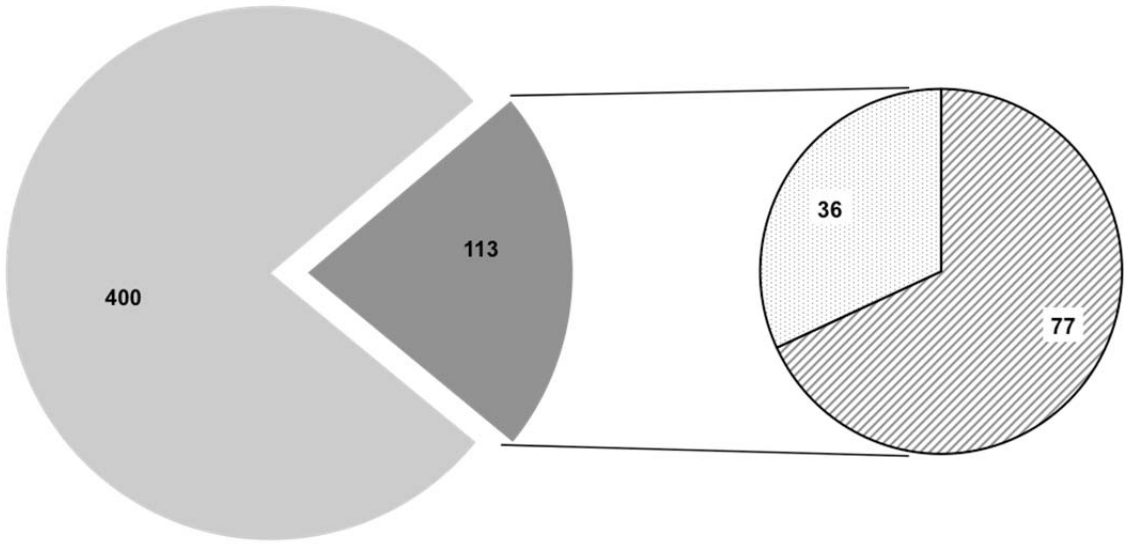


Figure 4.12: *De novo* proteome of HCT116-ACVR2 #2 cells. Individual metabolically labeled protein species detected by mass spectrometry in both ACVR2-proficient as well as in ACVR2-deficient cells (light grey area). The subset of differentially expressed proteins is indicated (dark grey) and splits up into protein species expressed exclusively in ACVR2-proficient (dashed area) or ACVR2-deficient cells (dotted area).

Only proteins recognized by at least two unique peptides and with an ion score of at least 30 were considered valid candidates. The majority of these *de novo* synthesized proteins (78%, 400/513) were detected in both dox-induced and -uninduced cells and hence remained unaffected by the ACVR2 expression status (Figure 4.12). However, a subset of proteins (22%, 113/513) showed a clear ACVR2 dependency and thus defined the ACVR2-dependent *de novo* proteome. Most of these proteins (68%, 77/113) were detected exclusively in dox-treated ACVR2-proficient HCT116-ACVR2 cells (Figure 4.12, Table 4.3 A). For some of these proteins like CALU, DEK, LETM1, IBP2, IF4B, KI67, MCM2, NP1L1, RBM8A and S10AB a link to tumorigenesis has been reported [141-150]. In contrast, a much smaller number of *de novo* synthesized proteins (32%, 36/113) was associated with

lack of ACVR2 expression in HCT116-ACVR2 cells (Figure 4.12, Table 4.3 B). Candidate proteins of this subset include the Flap endonuclease 1 (FEN1), a histone H2A variant (H2A1C), a Rho-related GTP-binding protein (RhoC), a 60S ribosomal protein (RL10a), and a subunit of the facilitator of chromatin transcription-complex (SPT16H). Human keratins detected within this set are well known contaminants resulting from sample processing. Analyzing the magnetic beads used to capture the modified proteins identified 5 proteins resulting from preblocking of the beads by the manufacturer (Table 4.4).

Table 4.3. *De Novo* Proteome of ACVR2-Proficient (A) and ACVR2-Deficient Cells (**B**). Occurrence under these conditions is mutually exclusive. Only protein hits with at least two unique peptides and with a minimal ion score of 30 were considered as identified. Experiments were performed in biological triplicates. Protein molecular weights (Mass), Mascot scores (Score), number of unique peptides (Peptides) and protein sequence coverage (Cov.[%]) are displayed for each protein with highest score annotated in UniProt.

A

ACVR2 proficient					
Accession No	**Protein Description**	**Mass [Da]**	**Score**	**Peptides**	**Cov. [%]**
1B15_HUMAN	HLA class I histocompatibility antigen, B-15 alpha chain	40648	153	3	15.5
1C17_HUMAN	HLA class I histocompatibility antigen, Cw-17 alpha chain	41612	243	5	19.9
ABHDA_HUMAN	Mycophenolic acid acyl-glucuronide esterase, mitochondrial	34253	99	2	8.8
ADT3_HUMAN	ADP/ATP translocase 3	33073	255	5	18.5
APLP2_HUMAN	Amyloid-like protein 2	87927	107	2	5.2
AT2A2_HUMAN	Sarcoplasmic/endoplasmic reticulum calcium ATPase 2	116336	80	2	2.7
CAB45_HUMAN	45 kDa calcium-binding protein	41895	75	2	11
CALU_HUMAN	Calumenin	37198	308	7	49.8
CC124_HUMAN	Coiled-coil domain-containing protein 124	25820	122	2	26
CERU_HUMAN	Ceruloplasmin	122983	126	2	3.8
CHRD1_HUMAN	Cysteine and histidine-rich domain-containing protein 1	38264	72	2	9.3
DCTN2_HUMAN	Dynactin subunit 2	44318	92	2	8.5
DDX1_HUMAN	ATP-dependent RNA helicase DDX1	83349	80	2	5.3
DEK_HUMAN	Protein DEK	42933	81	2	5.9
DNJA2_HUMAN	DnaJ homolog subfamily A member 2	46344	105	2	9.2
DX39B_HUMAN	Spliceosome RNA helicase DDX39B	49416	301	4	16.6
EIF3H_HUMAN	Eukaryotic translation initiation factor 3 subunit H	40076	77	2	8
EIF3M_HUMAN	Eukaryotic translation initiation factor 3 subunit M	42932	75	2	10.2
EMD_HUMAN	Emerin	29033	84	2	8.7
GDIR1_HUMAN	Rho GDP-dissociation inhibitor 1	23250	86	2	15.2
GRB2_HUMAN	Growth factor receptor-bound protein 2	25304	191	5	35.5
H2A2B_HUMAN	Histone H2A type 2-B	13987	93	2	12.3
H31T_HUMAN	Histone H3.1t	15613	214	3	33.1
HP1B3_HUMAN	Heterochromatin protein 1-binding protein 3	61454	103	2	5.6
IBP2_HUMAN	Insulin-like growth factor-binding protein 2	35875	117	2	7.7

IF2A_HUMAN	Eukaryotic translation initiation factor 2 subunit 1	36374	128	2	6.3
IF4B_HUMAN	Eukaryotic translation initiation factor 4B	69167	107	3	7.9
ITB4_HUMAN	Integrin beta-4	205745	125	2	1.8
JUNB_HUMAN	Transcription factor jun-B	36028	145	2	13.8
KI67_HUMAN	Antigen KI-67	360698	112	2	1
LAP2B_HUMAN	Lamina-associated polypeptide 2, isoforms beta/gamma	50696	306	5	24
LC7L3_HUMAN	Luc7-like protein 3	51834	84	2	5.8
LETM1_HUMAN	LETM1 and EF-hand domain-containing protein 1, mitochondrial	83986	122	3	6.1
MCM2_HUMAN	DNA replication licensing factor MCM2	102516	70	2	3.9
METK2_HUMAN	S-adenosylmethionine synthase isoform type-2	43975	96	2	10.9
NOP56_HUMAN	Nucleolar protein 56	66408	70	2	4
NP1L1_HUMAN	Nucleosome assembly protein 1-like 1	45631	155	2	11.8
OCAD1_HUMAN	OCIA domain-containing protein 1	27780	101	2	9.4
PDCD5_HUMAN	Programmed cell death protein 5	14276	97	2	20.8
PDIA6_HUMAN	Protein disulfide-isomerase A6	48490	163	2	12
POMP_HUMAN	Proteasome maturation protein	15836	75	2	17.7
PRS8_HUMAN	26S protease regulatory subunit 8	45768	168	3	11.8
PSA_HUMAN	Puromycin-sensitive aminopeptidase	103895	60	2	2.7
PSA5_HUMAN	Proteasome subunit alpha type-5	26565	113	2	12
PSME1_HUMAN	Proteasome activator complex subunit 1	28876	63	2	14.5
PSME2_HUMAN	Proteasome activator complex subunit 2	27555	57	2	14.6
PUR2_HUMAN	Trifunctional purine biosynthetic protein adenosine-3	108953	141	2	6
RBM8A_HUMAN	RNA-binding protein 8A	19934	92	2	10.9
RBP56_HUMAN	TATA-binding protein-associated factor 2N	62021	177	3	9.6
RCN1_HUMAN	Reticulocalbin-1	38866	120	2	11.5
RL17_HUMAN	60S ribosomal protein L17	21611	145	3	21.7
RL36_HUMAN	60S ribosomal protein L36	12303	131	2	34.3
RL6_HUMAN	60S ribosomal protein L6	32765	119	3	13.5
RS17L_HUMAN	40S ribosomal protein S17-like	15597	83	2	24.4
RS27L_HUMAN	40S ribosomal protein S27-like	9813	85	2	25
RS7_HUMAN	40S ribosomal protein S7	22113	69	2	17
RUXGL_HUMAN	Small nuclear ribonucleoprotein G-like protein	8595	64	2	25
S10A6_HUMAN	Protein S100-A6	10230	84	2	16.7
S10AB_HUMAN	Protein S100-A11	11847	116	2	30.5
SF3B3_HUMAN	Splicing factor 3B subunit 3	136575	114	2	2.1
SMD2_HUMAN	Small nuclear ribonucleoprotein Sm D2	13632	109	2	32.2

SRSF9_HUMAN	Serine/arginine-rich splicing factor 9	25640	111	2	14.9
SYNC_HUMAN	Asparagine--tRNA ligase, cytoplasmic	63758	78	2	3.5
SYTC_HUMAN	Threonine--tRNA ligase, cytoplasmic	84294	88	2	3
TBA4A_HUMAN	Tubulin alpha-4A chain	50634	566	9	39.7
TBB4A_HUMAN	Tubulin beta-4A chain	50010	794	13	44.4
TCTP_HUMAN	Translationally-controlled tumor protein	19697	135	2	23.3
TMX1_HUMAN	Thioredoxin-related transmembrane protein 1	32170	73	2	7.9
TNPO1_HUMAN	Transportin-1 O	103771	128	2	3.6
TRAP1_HUMAN	Heat shock protein 75 kDa, mitochondrial	80345	198	3	6.4
TRI25_HUMAN	E3 ubiquitin/ISG15 ligase TRIM25	72581	64	2	4.4
TXNL1_HUMAN	Thioredoxin-like protein 1	32630	110	2	19.7
TXNL1_HUMAN	Thioredoxin-like protein 1	32630	110	2	19.7
UBP2L_HUMAN	Ubiquitin-associated protein 2-like	114579	90	2	1.8
UBQL1_HUMAN	Ubiquilin-1	62479	112	2	7.8
VAPA_HUMAN	Vesicle-associated membrane protein-associated protein A	28103	150	2	17.7
VDAC3_HUMAN	Voltage-dependent anion-selective channel protein 3	30981	75	2	8.1

B

ACVR2 deficient					
Accession No	**Protein Description**	**Mass [Da]**	**Score**	**Peptides**	**Cov. [%]**
1C06_HUMAN	HLA class I histocompatibility antigen, Cw-6 alpha chain	41399	185	3	15.8
ACTBL_HUMAN	Beta-actin-like protein 2	42318	280	5	14.6
BUB3_HUMAN	Mitotic checkpoint protein BUB3	37587	83	2	7
C1TC_HUMAN	C-1-tetrahydrofolate synthase, cytoplasmic	102180	66	2	2
CBX1_HUMAN	Chromobox protein homolog 1	21519	75	2	21.6
CNN3_HUMAN	Calponin-3	36562	72	2	6.7
DHX15_HUMAN	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	91673	89	2	3.8
DJB11_HUMAN	DnaJ homolog subfamily B member 11	40774	93	2	7.5
DLDH_HUMAN	Dihydrolipoyl dehydrogenase, mitochondrial	54713	118	2	6.9
DNJB1_HUMAN	DnaJ homolog subfamily B member 1	38191	68	2	7.4
DSG2_HUMAN	Desmoglein-2	123016	187	3	4.6
EFHD2_HUMAN	EF-hand domain-containing protein D2	26794	73	2	7.5
FEN1_HUMAN	Flap endonuclease 1	42908	92	2	9.7

G6PI_HUMAN	Glucose-6-phosphate isomerase	63335	140	2	7.2
GDIB_HUMAN	Rab GDP dissociation inhibitor beta	51087	64	2	6.7
GLOD4_HUMAN	Glyoxalase domain-containing protein 4	35170	69	2	7.3
H2A1C_HUMAN	Histone H2A type 1-C	14097	189	3	32.3
HNRH3_HUMAN	Heterogeneous nuclear ribonucleoprotein H3	36960	67	2	5.2
IF4G2_HUMAN	Eukaryotic translation initiation factor 4 gamma 2	102810	107	2	4
KIF5C_HUMAN	Kinesin heavy chain isoform 5C	109997	113	2	2.6
MAGB2_HUMAN	Melanoma-associated antigen B2	35426	149	3	15
MTAP_HUMAN	S-methyl-5'-thioadenosine phosphorylase	31729	79	2	9.5
NAA50_HUMAN	N-alpha-acetyltransferase 50	19614	70	2	11.8
PSIP1_HUMAN	PC4 and SFRS1-interacting protein	60181	111	2	8.5
PSMD4_HUMAN	26S proteasome non-ATPase regulatory subunit 4	40939	152	2	16.2
RAB5C_HUMAN	Ras-related protein Rab-5C	23696	71	2	10.6
RHOC_HUMAN	Rho-related GTP-binding protein RhoC	22334	229	2	46.6
RL10A_HUMAN	60S ribosomal protein L10a	24987	80	2	11.1
RS21_HUMAN	40S ribosomal protein S21	9220	93	2	22.9
RU2B_HUMAN	U2 small nuclear ribonucleoprotein B"	25470	86	2	11.6
RUVB1_HUMAN	RuvB-like 1	50538	86	2	7.5
RUVB2_HUMAN	RuvB-like 2	51296	124	3	7.1
SP16H_HUMAN	FACT complex subunit SPT16	120409	133	2	5.2
SQSTM_HUMAN	Sequestosome-1	48455	164	4	14.8
SRP14_HUMAN	Signal recognition particle 14 kDa protein	14675	72	2	16.2
SYG_HUMAN	Glycine--tRNA ligase	83854	131	3	3.9

Table 4.4. Proteins Comprising Blocked Beads. These proteins were detected by tryptic digestion of streptavidin-coated beads washed and incubated with pure sample buffer instead of protein samples.

Accession No	Protein Description	Mass [Da]	Score	Peptides	Cov. [%]
ALBU_BOVIN	Serum albumin	71244	3326	42	82.4
APOA1_BOVIN	Apolipoprotein A-I	30258	162	2	14
CERU_SHEEP	Ceruloplasmin	120020	305	3	8.6
SAV_STRAV	Streptavidin	18822	445	4	41
TTHY_BOVIN	Transthyretin	15831	403	5	50.3

4.3.2 Analysis of Fucosylated *De Novo* Proteins

In a next step the focus was set on the fucosylated proteins, since previous radioactive labeling experiment indicate an ACVR2-dependent change in fucose incorporation. Therefore, the ACVR2-reconstituted cell line clone #2 was used again for a click chemistry approach to determine ACVR2-dependent alterations of the *de novo* synthesized fucosylated proteins. This time, the click chemistry approach relies on metabolic labeling with fucose-alkyne, which can form a stable triazole conjugate with azide-activated biotin. Identification of tryptic peptides was also performed by mass spectrometry. Applying this approach to Act A stimulated HCT116-ACVR2 #2 cells grown either in the presence (+ACVR2) or absence (-ACVR2) of dox and using several filter and selection criteria identified a set of 70 candidate proteins (Figure 4.13).

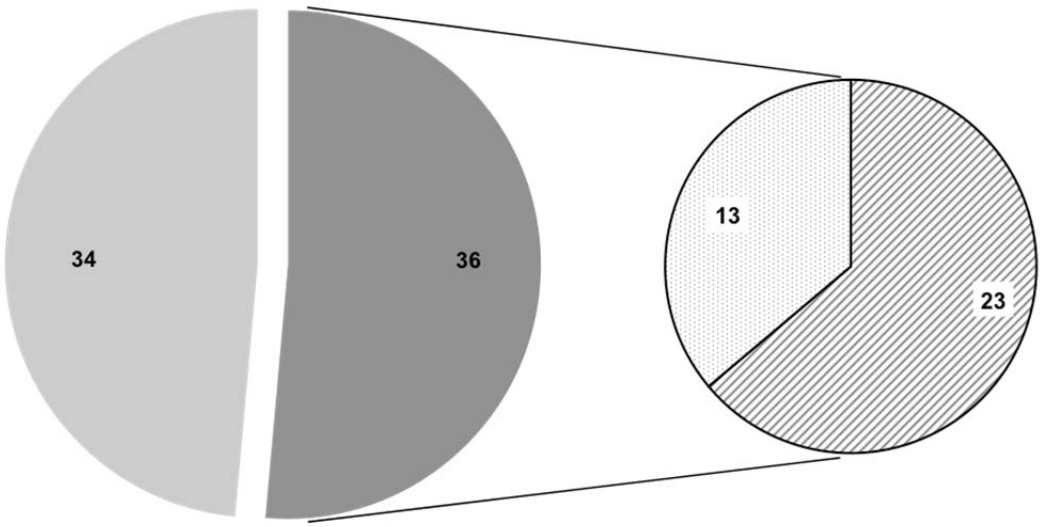


Figure 4.13: Fucosylated *de novo* proteins of HCT116-ACVR2 #2 cells. Individual metabolically labeled protein species detected by mass spectrometry in both ACVR2-proficient as well as in ACVR2-deficient cells (light grey area). The subset of differentially expressed proteins is indicated (dark grey) and splits up into protein species expressed exclusively in ACVR2-proficient (dashed area) or ACVR2-deficient cells (dotted area).

Almost half of these *de novo* synthesized proteins also including the non-specifically bound proteins (49%, 34/70) were detected in both dox-treated and -untreated cells and hence remained unaffected by the ACVR2 expression status (Figure 4.13). However, a similar number of proteins (51%, 36/70) showed a clear ACVR2 dependency and thus defined ACVR2-dependent *de novo* synthesized fucosylated proteins. Most of these proteins (64%, 23/36) were detected exclusively in dox-treated ACVR2-proficient HCT116-ACVR2 cells (Figure 4.13, Table 4.5 A). For some of these proteins like ANXA2, HMGA1, NUCL, PRIO, S10A8 and SRSF3 a link to tumorigenesis in the colon has been reported [2-6, 151]. Notably, we also identified a smaller set of newly synthesized proteins (36%, 13/36), that were detected exclusively in ACVR2-deficient HCT116-ACVR2 cells (Figure 4.13, Table 4.5 B). Proteins synthesized under this condition might be of particular clinical relevance because loss of normal ACVR2 signaling is one of the most frequent events that occurs during the development of MSI colorectal tumors. Candidate proteins of this subset including CD166, EZRI, ITA3 and PODXL are already known to be associated with colon cancer [7-9, 11].

More detailed information about all proteins identified by these mass spectrometry approaches is provided on the CD as annex to this work. Overall, these results illustrate the potential of our approach to correlate a tumor-relevant driver mutation (ACVR2) with specific changes in the tumor cell proteome. Moreover, the combination of our cellular model system with specific proteomic analysis represents a powerful novel approach to identify gene-specific cellular proteomic states and also can easily be adapted to other genes or modulators of cellular growth in a time- and dose-dependent manner.

Table 4.5. Fucosylated *De Novo* Synthesized Proteins of ACVR2-Proficient (A) and ACVR2-Deficient Cells (**B**). Occurrence under these conditions is mutually exclusive. Only proteins hits with at least two unique peptides and with a minimal ion score of 30 were considered as identified. Experiments were performed in biological triplicates. Protein molecular weights (Mass), Mascot scores (Score), number of unique peptides (Peptides) and protein sequence coverage (Cov.[%]) are displayed for each protein with highest score annotated in UniProt.

A

ACVR2-Proficient					
Accession No	**Protein Description**	**Mass [Da]**	**Score**	**Peptides**	**Cov [%]**
1B13_HUMAN	HLA class I histocompatibility antigen, B-13 alpha chain	40791	186	4	15.5
1B15_HUMAN	HLA class I histocompatibility antigen, B-15 alpha chain	40648	201	5	19.3
ANXA2_HUMAN	Annexin A2	38808	158	2	13
H2A1A_HUMAN	Histone H2A type 1-A	14225	99	2	12.2
HMGA1_HUMAN	High mobility group protein HMG-I/HMG-Y	11669	83	2	23.4
LDHA_HUMAN	L-lactate dehydrogenase A chain	36950	95	2	8.1
NOLC1_HUMAN	Nucleolar and coiled-body phosphoprotein 1	73560	301	7	13
NUCL_HUMAN	Nucleolin	76625	526	11	23.2
PCLO_HUMAN	Protein piccolo	554704	57	2	0.3
PRIO_HUMAN	Major prion protein	27871	106	2	12.3
RU17_HUMAN	U1 small nuclear ribonucleoprotein 70 kDa	51583	114	2	5.7
S10A8_HUMAN	Protein S100-A8	10885	88	2	19.4
SRS10_HUMAN	Serine/arginine-rich splicing factor 10	31339	108	2	10.3
SRSF1_HUMAN	Serine/arginine-rich splicing factor 1	27842	744	13	47.6
SRSF2_HUMAN	Serine/arginine-rich splicing factor 2	25461	249	5	20.8
SRSF3_HUMAN	Serine/arginine-rich splicing factor 3	19546	419	6	43.9
SRSF5_HUMAN	Serine/arginine-rich splicing factor 5	31359	122	2	10.7
SRSF6_HUMAN	Serine/arginine-rich splicing factor 6	39677	195	4	12.2
SRSF7_HUMAN	Serine/arginine-rich splicing factor 7	27578	340	6	28.6
SRSF9_HUMAN	Serine/arginine-rich splicing factor 9	25640	87	2	8.1
TCP4_HUMAN	Activated RNA polymerase II transcriptional coactivator p15	14386	191	3	30.7
TR150_HUMAN	Thyroid hormone receptor-associated protein 3	108658	146	2	4.1
TRA2B_HUMAN	Transformer-2 protein homolog beta	33760	143	2	12.8

B

ACVR2-Deficient					
Accession No	**Protein Description**	**Mass [Da]**	**Score**	**Peptides**	**Cov [%]**
1A02_HUMAN	HLA class I histocompatibility antigen, A-2 alpha chain	41181	82	2	6
1A33_HUMAN	HLA class I histocompatibility antigen, A-33 alpha chain	41151	235	5	21.4
1B07_HUMAN	HLA class I histocompatibility antigen, B-7 alpha chain	40777	95	2	6.1
1B44_HUMAN	HLA class I histocompatibility antigen, B-44 alpha chain	40798	183	4	13.3
1C12_HUMAN	HLA class I histocompatibility antigen, Cw-12 alpha chain	41316	197	4	19.1
CAB45_HUMAN	45 kDa calcium-binding protein	41895	86	2	9.7
CD166_HUMAN	CD166 antigen	65745	80	2	3.6
EZRI_HUMAN	Ezrin	69484	96	2	4.1
GOLM1_HUMAN	Golgi membrane protein 1	45477	269	4	24.7
ITA3_HUMAN	Integrin alpha-3	117735	100	2	2.9
PODXL_HUMAN	Podocalyxin	59055	63	2	3.8
TGON2_HUMAN	Trans-Golgi network integral membrane protein 2	51082	120	2	8.1
TMED9_HUMAN	Transmembrane emp24 domain-containing protein 9	27374	75	2	7.2

5. DISCUSSION

5.1 Inducible ACVR2 Expression in a MSI CRC Cell Line Model System

The development of colorectal cancer is a multistep process occurring in a timeframe of many years, whereby the maintenance of healthy colon tissue homeostasis which is regulated by the balance between several key signal transduction pathways, including TGF-ß, Activin, BMP, Wnt and Notch pathways [152, 153] gets disrupted. Comprehensive molecular characterization of CRC tumors at the DNA and RNA level has shaped our current knowledge about the molecular mechanisms that drive the formation of CRC. The next step is to assign molecular, cellular and biochemical functions to these predicted gene products and to explain how these products regulate complex physiological processes in cancer. In MSI CRC one main signaling pathway, mediated by the TGF-ß superfamily, controlling many major cellular processes is inactivated. In this study the signaling pathway mediated by activin A is focused. Mutations in several genes involved in the activin signaling pathway have been characterized in colon cancers. One of these genes reported to show a high frequency of coding microsatellite frameshift mutations in MSI colorectal tumors is the gene encoding the activin receptor type 2 A. ACVR2 contains polyadenine tracts at exons 3 and 10 but only its exon 10 A8 tract is mutated in ~85% of colorectal cancers with MSI [72, 73]. The biallelic frameshift mutation causes ACVR2 protein loss, signaling disruption and is associated with histologically poor grade tumors and significantly larger volume tumors [73, 74]. In the early phase of cancer development activin signaling inhibits cancer cell growth. At later stages it acts as a tumor promoter as it is coupled to MSI tumorigenesis. It has been reported that restoration of ACVR2 in colon cancer cells causes growth suppression [57]. Apart from such genetic alterations resulting primarily in proteomic changes, glycosylation changes are also a universal feature of malignant transformation and tumor progression.

Therefore, it is reasonable to assume that the mutator phenotype in MSI tumor cells also affects the cellular glycosylation machinery either directly by mutational inactivation of the enzymes or indirectly by interfering with key signaling pathways like the ACVR2 pathway that might regulate normal cellular glycosylation signatures.

As a major goal of this work, ACVR2-specific downstream signaling effects were analyzed in an ACVR2 deficient and MSI colorectal cancer cell line (HCT116) genetically modified to re-express a wildtype ACVR2 cDNA in a doxycycline-regulated manner (HCT116-ACVR2). Reconstitution of wildtype ACVR2 expression already has been reported in one previous study demonstrating growth inhibition and enhanced migration of recombinant clones [57]. However, these authors stably introduced an entire copy of chromosome 2 into HCT116 cells and thus could not exclude potential confounding effects by other genes residing on this additional chromosome.

Therefore, the first aim of this study was to generate a HCT116-derived MSI CRC cell line that allows inducible expression of an *ACVR2* transgene in a time-dependent manner. To this end, the tetracycline-controlled gene expression system (tet System) [154] was used, which drives expression of the *ACVR2* gene only in presence of doxycycline (dox). It is important to note that results obtained from this ACVR2-reconstituted model system should reflect the inverse situation of the ACVR2-deficient status in MSI primary colorectal tumors, since induction with dox activates the wildtype ACVR2 signaling pathway. While the first ACVR2-reconstituted cell clone was established by applying conventional transfection, for the second one the problem of stable integration of multiple copies at different genomic sites should be overcome by using a retroviral vector that - depending on the multiplicity of infection (MOI) – allows isolation of clones with single copy integration. This approach is based on the publication of Weidenfeld et al. published in 2009 [133]. Lee et al. [134] generated the HCT116-HygTK master cell line, being Hyg-resistant and sensitive to Gan (Hygr, Gans), which underwent a second RMCE resulting in a second ACVR2-reconstituted cell clone that conferred dox-inducible expression of ACVR2 and luciferase concurrently and had at the same time a C-terminal FLAG-tag. These two HCT116-ACVR2 cell clones were studied in detail regarding the DNA, RNA and protein level and its effect on proliferation. The integration of a wildtype copy of the *ACVR2* transgene into the HCT116 cell genome and the dox-inducible increase of the transgenic ACVR2 have been nicely displayed. However, ACVR2 protein expression in dox-induced cells could not be confirmed by analyzing the protein lysates directly to Western blot analysis. Although this strategy aims to detect physiological expression levels, this detection limit is apparently not reached for low abundant proteins like the ACVR2 receptor which is present in less than 500 copies

per cell [155]. However, the ACVR2 induction was verified by combining immunoprecipitation with Western blot analysis. Thereby, for HCT116-ACVR2 #2, containing a FLAG-tag, the reconstitution of the ACVR2 protein could be shown. Additionally, our proof of reconstituted ACVR2 signaling in the dox-induced cells demonstrates that functional amounts of the receptor are expressed. This is shown by two approaches – pSmad2 Western blot analysis and real-time RT-PCR of three known target genes of the ACVR2 pathway. In the Smad signaling experiment, increased pSmad2 levels were clearly detected upon exposure to exogenous activin A when wildtype ACVR2 protein was expressed. However, even in the absence of wildtype ACVR2 protein, there was a basal amount of pSmad2. This is consistent with the observation that activin A is also able to bind to other receptors, like ACVR2B [156]. While ACVR2B is abundant in normal sera, a significant increase in the expression level of ACVR2B on CRC cell lines and colonic mucosa was shown by immunoblotting and immunohistochemistry on tissue microarray [157]. Furthermore, increased concentrations of activin A have been observed in malignant conditions [158, 159]. Measuring the proliferation of HCT116-ACVR2 cells compared to the parental cell line showed like expected a dox-dependent slower proliferation rate. Thus, restoration of ACVR2 in colon cancer cells caused growth suppression. Since both cell clones showed similar characteristics, we decided to perform further approaches using only one clone, representing ACVR2 reconstitution.

Taken together, the experimental system established in this present study includes a variety of different advantages. First of all it enables switching the gene of interest on and off, thus overcoming the troubles of continuous gene expressing. At the same time, the consequences of short-term and long-term ACVR2 expression and signaling can be easily determined. In fact, these cells have been cultured for almost three years without losing their ability to inducibly express ACVR2. This model cell line can easily be used for large-scale production and isolation of altered proteins, including glycoproteins, which may provide novel targets or biomarkers suitable for MSI tumor diagnostics and therapy. It is important to emphasize that the HCT116-HygTK master cell, which has been used in this study to generate the HCT116-ACVR2 clones, can be used for retargeting any gene of interest at the identified genomic loci. Thus, this model system allows addressing any tumor target gene and thereby provides a broad applicability to study the functional consequences and biological relevance of a given MSI tumor-associated mutation.

5.2 ACVR2-Dependent Glycosylation Changes

The transition of healthy tissue to tumor tissue is associated with multiple changes, one of which is the glycosylation status of proteins, especially at the surface of the tumor cell. Normal regeneration of the mucosa of the gut, which is faster than in any other tissue, may cause unrestricted proliferation of colon cells and create benign colon neoplasm through mutation of oncogenes and / or tumor suppressor genes [160]. Since, the intestinal epithelial coat consists largely of various glycoconjugates, understanding the signaling events changing these glycan changes are a major task of this study. Therefore, the challenge was to examine if and how ACVR2, a key signaling receptor in normal and malignant colon epithelium, might affect the cellular glycosylation profile.

Lectins are carbohydrate-binding proteins that are useful markers for localization and characterization of glycoconjugates. They derive from plants, prokaryote and invertebrate species and have been applied to characterize some peripheral sugars. Thus, lectins have frequently been used as a tool to study glycosylation of tumor tissues. In previous work lectin-FACS analysis proved useful for detecting alterations of cell surface protein glycosylation in MSI colorectal cancer cells transiently transfected with different MSI target genes [122]. Accordingly, the same technique was applied to determine cell surface glycan changes in the stably transfected HCT116-ACVR2 cells. However, in our inducible ACVR2 expression system no alterations of overall cell surface protein glycosylation could be detected by lectin-FACS analysis. It is reasonable to assume that such glycan alterations might only be detected at high levels of ACVR2 expression that are usually obtained by transient transfections. In addition, applying lectins only detect global changes of steady state levels without considering any dynamic variations. Although lectins are powerful tools for simple oligosaccharide analyses, their cross-reactivity and low affinity restrict their use.

Therefore, in the next step a method should be applied that also accounts for such kinetic changes and at the same time addresses specific modifications of glycoproteins like sialylation and fucosylation. Such alterations of protein modification have been reported for cancers of the colon [161]. One of the most frequent alterations in the normal glycosylation pattern observed during carcinogenesis is the alteration of fucose residues of glycoproteins [162, 163]. Also sialic acids are a key component of glycoproteins and have been correlated with cancer metastasis [164].

Correlations between sialic acid with tumorgenicity in the CRC cell line HCT116 have been demonstrated [165]. However, analysis of modifications of sialic acids and fucosylated proteins in MSI colorectal tumors remains a mainly uncovered field. As a major finding of these radioactive metabolic labeling experiments, ACVR2 signaling turned out to be a regulator of fucosylation of newly synthesized glycoproteins. Hence, fucosylation of proteins appears to be a dynamically regulated process modulated by a major signaling pathway in MSI colon cancer cells. Besides the prominent effects of activin A signaling on protein fucosylation, a modest increase of sialylation upon ACVR2 reconstitution has been observed, but to a lesser extent and only in one of both cell clones. It has been recently shown that TGF-ß signaling is associated with the sialylation pattern of MSI colon cancer cells [134]. Thus, it is conceivable that ACVR2 might contribute to the regulation of protein sialylation as well, but further investigations need to be performed to resolve this issue.

Fucosylation is catalyzed by fucosyltransferases, guanosine 5'-diphosphate (GDP)-fucose synthetic enzymes and GDP-fucose transporter(s). So far thirteen fucosyltransferase genes have been identified in the human genome. GDP-fucose, which is a common donor substrate to all fucosyltransferases, is synthesized in the cytosol via two pathways, namely the salvage pathway and the *de novo* pathway (Figure 5.1). The salvage pathway synthesizes GDP-fucose from free L-fucose, derived from extracellular or lysosomal sources, whereas the *de novo* pathway transforms GDP-mannose into GDP-fucose. The salvage pathway is responsible for only about 10% of the cellular pool of GDP-fucose. Thus, cellular GDP-fucose is mainly produced by the *de novo* pathway. After GDP-fucose has been synthesized in the cytosol, it is transported to the Golgi apparatus through GDP-fucose transporter to serve as a substrate for fucosyltransferases [166]. In the *de novo* pathway GDP-mannose is transformed into GDP-fucose through 3 steps that are catalyzed by GDP-mannose-4,6-dehydratase (GMDS) [167] and GDP-4-keto-6-deoxymannose-3,5-epimerase-4-reductase (FX) [168]. FX knockout mice showed an embryonic lethality phenotype because of a virtually complete deficiency of cellular global fucosylation [169]. Thus, fucosylated oligosaccharides are involved in early growth and development as well as many pathologic conditions. It has already been shown that HCT116 cells lack fucosylation because of a GMDS mutation, consequently leading to the escape from NK cell–mediated tumor surveillance through the

acquisition of resistance to tumor necrosis factor-related apoptosis-inducing ligand (TRAIL), followed by tumor progression and metastasis [170].

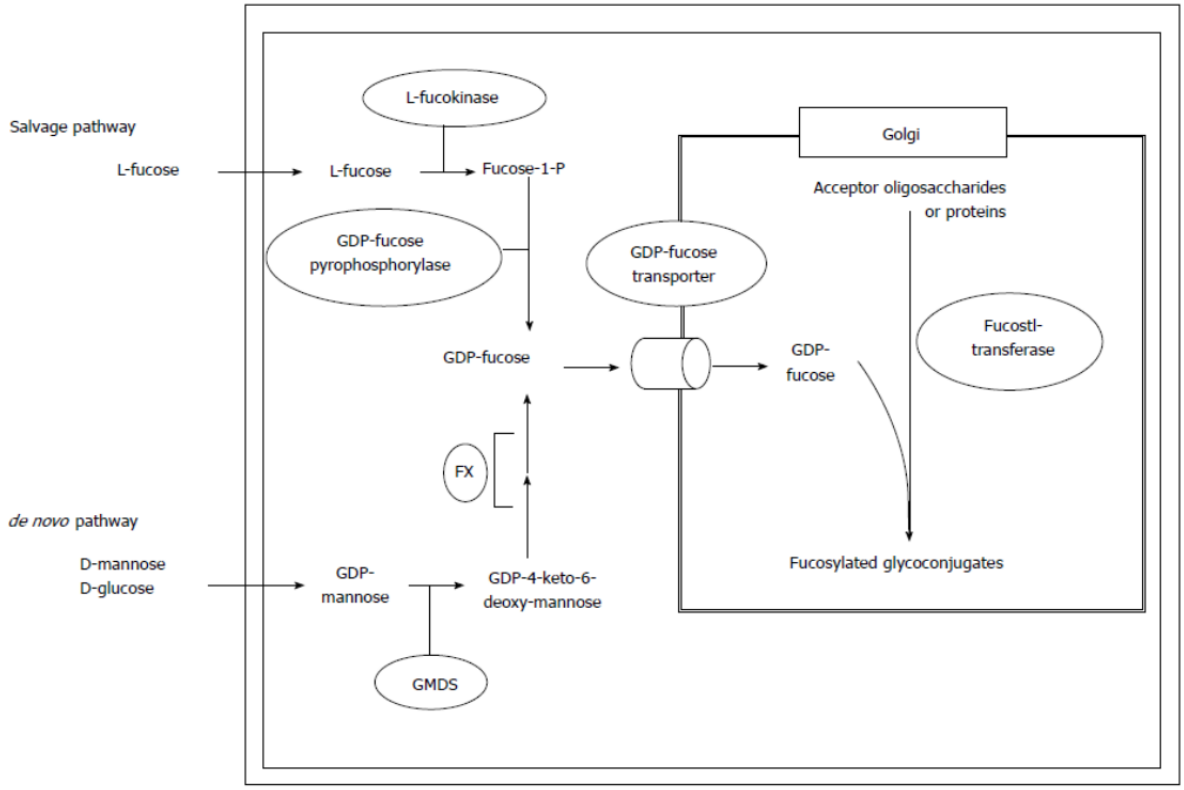


Figure 5.1: **Fucose metabolism** [166]. GDP-fucose is synthesized through two pathways, the salvage pathway and the *de novo* pathway. Labels schow GDP-4,6-dehydratase (GMDS) and GDP-4-keto-6-deoxymannose-3,5-epimerase-4-reductase (FX).

Since previous studies reported that the activity of the fucosyltransferases might be affected and that the expression of Fut8 is markedly enhanced in several types of cancer cell lines [171], analysis of the fucose pathway was initiated by examining ACVR2-dependent transcriptional regulation of candidate enzymes. Analyzing all known human fucosyltransferases, fucosidases and enzymes involved in the fucose pathway by real-time RT-PCR analysis, mRNA levels remained unaffected by reconstituted ACVR2 signaling. This result is well in agreement with our Glyco-Gene Chip data that did not reveal transcriptional alterations of any of these enzymes upon reconstituted ACVR2 signaling.

Since activin signaling is known to regulate the expression of different target genes at the transcriptional level, a Glyco-Gene Chip approach was applied to identify potential glycosylation-associated targets. By this time point we decided to use only

HCT116-ACVR2 #2, having a single copy integrated ACVR2 transgene, as a representative for the ACVR2 reconstitution. Overall, six genes (JUNB, HES1, LFNG, PDGFB, SLC35F2 and FGF18) showed ACVR2-dependent up-regulation at the transcriptional level, whereas three genes (KLRF1, BMP4 and FGF9) appeared to be ACVR2-dependent down-regulated in the model system. However, it is important to mention that this Glyco-Gene Chip represents a limited choice of about 1000 human glycosylation-associated genes, including enzymes, transporters, receptors, transcription factors, etc. Accordingly, there might be some numerical limitation leading to the small number of identified glycosylation-associated targets in our cell line system. This chip analysis enables a systematic screening approach, which might be less sensitive, however more genes in one experimental set up can be identified than in real-time RT-PCR analysis. This can be used to explain why ACVR2 itself was not found to be up-regulated in the chip analysis. Two of the identified to be up-regulated transcripts are known to be involved in the Notch signaling pathway: HES1 and LFNG. Both were also validated by real-time RT-PCR and thereby represent actual ACVR2-dependent target genes. The Notch receptors are large transmembrane proteins which are highly glycosylated at their extracellular epidermal growth factor (EGF) repeats. The Notch pathway plays a fundamental role in developmental processes. Aberrant Notch signaling was further implicated in many human disorders [172, 173]. The signaling of the four mammalian Notch receptors is modulated by binding to Delta-like and Jagged family ligands. The specificity of their interactions is determined by glycan changes mediated by fringe glycosyl-transferases, including LFNG [174]. LFNG (O-fucosylpeptide 3-beta-N-acetyl-glucosaminyltransferase or Lunatic Fringe) functions as a fucose-specific, Golgi-localized glycosyltransferase. It leads to elongation of O-linked fucose-residues on Notch, which alters Notch signaling. Thereby it is modulating Notch activity by differential glycosylation. HES1 is a well-known target gene of Notch [175]. Some reports indicate a crosstalk between TGF-ß / activin and Notch through the signaling molecule Smad3 thus explaining the increased expression of HES1 [176]. However, understanding this interaction between ACVR2 and Notch signaling needs further investigations, which could be addressed in further studies.

Taken together, so far the hypothesis that ACVR2 expression alters protein glycosylation is supported by the observation of decreased overall *de novo* fucosylation in presence of ACVR2. Furthermore, a potential ACVR2-dependent

glycosylation target gene, LFNG, was found by using the dox-inducible HCT116-ACVR2 cell line. Thus, the created model system is usable to determine ACVR2-regulated genes involved in the glycosylation machinery. Remarkably, 25% of newly synthesized fucosylated proteins were altered by ACVR2, further approaches need to be carried out to identify which glycoproteins may be affected.

5.3 ACVR2-Dependent Proteome and Glycome Analysis

As described above, metabolic labeling experiments with radioactively labeled precursors of glycan structures showed results that in contrast to the lectin-FACS experiments are not based on the analysis of the bulk of all cellular glycoproteins. Rather metabolic labeling focuses on changes in *de novo* glycoprotein biosynthesis. Thus, these changes are more specifically linked to the induced expression of reconstituted ACVR2. These experiments clearly indicate changes in newly synthesized fucosylated glycoproteins after ACVR2 induction. In order to adapt the successful metabolic labeling experiments to subsequent glycome and proteome analyses, the Click-it system, which enables to label molecules of interest in complex biological samples similar to a radiolabeled compound, was applied. Click-it technology allows robust and reliable detection and extraction of labeled molecules with high sensitivity and extremely low background. Before performing a glycome analysis, to proof the recently established technique, we decided to start by whole nascent protein labeling. Screening the proteome of tumors is a complex issue because the proteome is a dynamic and constantly changing process due to a combination of several factors [177]. These include differential splicing of the respective mRNAs, posttranslational modifications, and temporal and functional regulation of gene expression [178]. Many studies have identified colon cancer-associated proteins, however only few studies focus on the functional protein interaction networks derived from the proteomics data. Therefore, the strength of this approach lies in studying the consequences of a major signaling pathway on the proteomic assembly with emphasis on the *de novo* proteome. The incorporation of azido-homoalanine can be compared to metabolic labeling using [^{35}S]-methionine, but avoids the use of radioactive compounds and enables the specific isolation of the labeled proteins. Thus, the labeled proteins can be tagged with biotin alkyne resulting in the specific biotinylation of the metabolically labeled proteins, which enables the

extraction streptavidin beads. Our model system is not only of broad and easy applicability, but also enables highly sensitive analyses, since more than 500 metabolically labeled proteins were identified. 113 of these proteins showed a clear ACVR2 dependency. 36 of these proteins were specifically identified in ACVR2-deficient cells and are involved in key cellular processes like DNA repair and/or chromatin dynamics (e.g. FEN1, HIST1H2AC, SPTH16), protein translation (e.g. RPL10A) and signal transduction (e.g. RHOC). Some of these candidates have already been linked to colorectal cancer or play a key role in cell growth or apoptosis control. For example RhoC, a member of the Ras homologous (Rho) sub-family of low molecular-weight GTP-binding proteins, is known to be overexpressed in colon cancer and various other malignancies. It can promote post-EMT (epithelial to mesenchymal transition) cell migration and due to its correlation with tumor progression has been proposed as prognostic marker for colon carcinoma [179, 180]. Also, the DNA repair protein FEN1 has been reported to show high expression levels in proliferative tissues and its expression in tumor tissues has been correlated with increased tumor grade and invasiveness [181, 182]. Moreover, several studies suggest that expression of the histone chaperone SUPT16H – as part of the heterodimeric Facilitates Chromatin Transcription (FACT) complex – can act in a pro-tumorigenic manner because it modulates nucleosome assembly and chromatin architecture including exchange of histone variants. Since the SUPT16H subunit of FACT preferentially binds to the core histone complex H2A-H2B [183] there might be a causal link to the core histone H2A family member H2A1C that we also found to be expressed in ACVR2-deficient cells. According to these correlations it is tempting to speculate, that the *de novo* synthesized candidate proteins identified in ACVR2-deficent HCT116 cells might help to establish their tumorigenic phenotype, a hypothesis that requires further investigation.

In contrast to ACVR2-deficient HCT116 cells, a much larger number of de novo synthesized candidate proteins were observed upon reconstitution of ACVR2 signaling. The resulting proteomic changes are expected to reflect the growth suppressing and apoptosis promoting effects of this signaling pathway. In fact, several of the candidate proteins like CALU, IBP2, LETM1, PRS8, SF3B3 and TNPO1 have been reported to confer such growth inhibitory effects [141, 143, 150, 184, 185]. For example, the calcium-binding protein Calumenin (CALU) can interfere with cytokinesis when added exogenously and also has been found to be associated

with mitotic arrest and late apoptosis upon translocation from the membrane to the cytoplasm [186, 187]. Similarly, the delicate balance between insulin like growth factor receptors (IGF-IR, IGF-IIR), ligands (IGF1/2) and binding proteins (IGFBP1-6) usually ensures proper survival factor signaling in epithelial cells [188], but would be severely disturbed by for instance high levels of IBP2 and resulting inadequate low levels of IGF1/2 thereby restricting tumor cell growth and proliferation [189]. LETM1, another ACVR2-dependent candidate protein, is known to exhibit a very complex pathophysiology due to its involvement in Wolf-Hirschhorn Syndrome, ion homeostasis and mitochondrial biogenesis [190], but also can induce programmed tumor cell death upon over-expression [150]. Likewise, the 26S proteasomal component PRS8, also known as PSMC5, has been demonstrated to suppress colon cancer stem cell growth in mouse xenografts and its yeast ortholog is capable to initiate caspase-8-dependent cell death [184, 191]. Overall, these examples illustrate the potential of this approach to detect tumor relevant proteomic changes. Nevertheless it is also important to emphasize that candidate proteins not yet linked to tumor-relevant functions should not be completely ignored, since some yet unknown mechanism or interaction with other proteins might be effective.

To focus again on fucosylation, also the glycome analysis of the CRC model cell line was established. To this end [^{3}H]-L-fucose applied in the initial metabolic labeling experiments was substituted by fucose-alkyne. Thus, the model system allowed the ACVR2-dependent incorporation of the alkyne-containing fucose and thereby also the simple and efficient isolation of newly synthesized, labeled proteins from a complex proteomic background by click chemistry. Overall, 70 proteins were identified, 36 of these indicating a clear ACVR2-dependency. Since, the salvage pathway is responsible for only about 10% of the fucose metabolism, the amount of detected newly synthesized protein was low. Five of the thirteen proteins specifically identified in ACVR2-deficient cells belong to the HLA class I histocompatibility antigen, which are involved in the presentation of foreign antigens to the immune system. A reduction or loss of HLA-A, B, C antigens in colorectal carcinoma has been reported with an inverse correlation with the degree of differentiation [192]. Some others of the identified proteins are involved in key processes like inter- or intra-cellular activities and / or dynamics (CAB45, CD166, EZRI, ITA3, TGON2 and TMED9) and cellular response to viral infection (GOLM1). A last ACVR2-deficient identified protein was podocalyxin (PODXL), a sialoglycoprotein thought to be the

major constituent of the glycocalyx of podocytes. Podocalyxin-like 1 has been found to be expressed in several CRC cell lines [193] and its over-expression is an independent factor of poor prognosis in colorectal cancer [194]. Furthermore, PODXL has been shown to interact with other proteins such as ezrin (EZRI), an established mediator of metastasis [195]. In contrast a much larger number of fucosylated candidate proteins were observed upon reconstitution of ACVR2 signaling. Also in this case some interesting proteins involved in key cellular processes like DNA repair and / or chromatin dynamics (H2A1A, NOLC1 and NUCL), transcription or translation regulation (HMGA1, RU17, TCP4, TR150, TRA2B, Serine / arginine-rich splicing factors) and signal transduction (ANXA2) were found. Some of these candidates have already been linked to colorectal cancer or play a key role in cell growth or apoptosis control. For example Annexin A2 (ANXA2), which is a calcium-dependent, phospholipid-binding protein found on various cell types. It is associated with various tumor types and plays multiple roles in regulating cellular functions, including angiogenesis, proliferation, apoptosis, cell migration, invasion and adhesion [196]. It has been indicated to be abundantly expressed in many cancer tissues and is believed to play an important role in tumorigenesis and breast cancer progression [197]. ANXA2 consists of phosphorylation sites for different kinases and it has been reported that Tyr23 phosphorylation of Annexin A2 mediates cell scattering and branching morphogenesis [198], and regulates Rho-mediated actin rearrangement and cell adhesion. The dynamic remodeling of the actin cytoskeleton is required for cell spreading, motility, and migration [199]. Overall, these examples illustrate the potential correlation to ACVR2-signaling, but nevertheless, further investigations are needed to understand the underlying mechanisms of these linked processes.

Based on current knowledge, several of the whole *de novo* protein candidates as well as of the fucosylated glycoproteins identified in ACVR2-deficient as well as in ACVR2-proficient HCT116 cells can be functionally assigned to hallmark processes of cancer development (Figure 5.2). Altogether this study provides a versatile platform to analyze the glycomic and proteomic *de novo* assembly of any MSI-relevant target gene of interest. This isogenic MSI tumor cell line model system provided analyzing the ACVR2 signaling mechanisms that contribute to the *de novo* protein synthesis in MSI tumor cells and identifying the proteomic assembly. Thus, this approach might facilitate the identification of tumor markers that could be used for diagnostic and therapeutic applications.

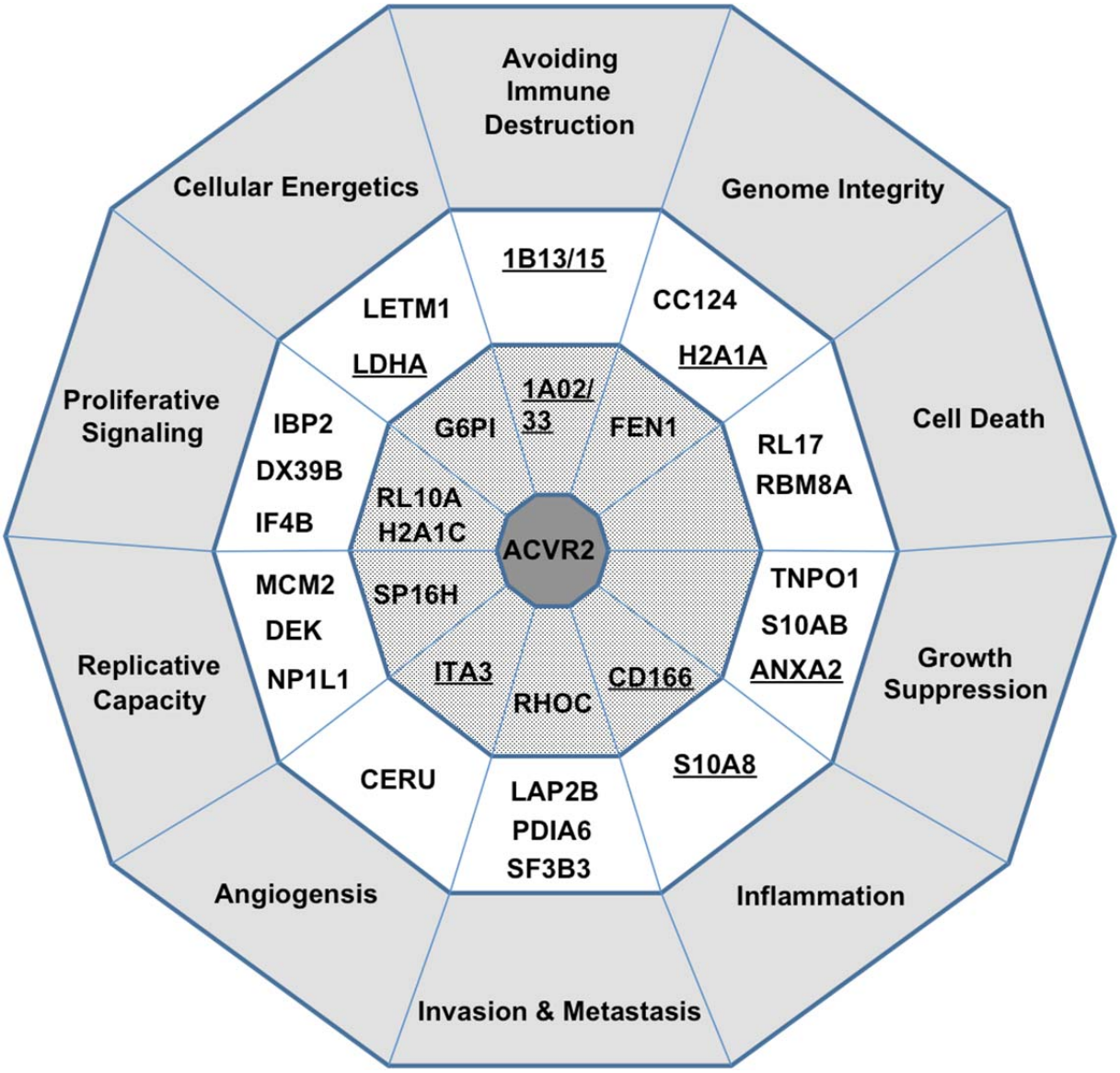


Figure 5.2: Functional association of *de novo* candidate proteins (modified according to [200]). Several proteins labeled with AHA or fucose-alkyne (underlined) expressed in ACVR2-deficient (stippled ring) or in ACVR2-proficient (white ring) HCT116-ACVR2 #2 cells were assigned to hallmark processes (grey ring) of cancer development.

5.4 Perspectives

This study was largely focused on ACVR2-dependent activin signaling and its role in altering glycan signatures in MSI CRC cells. Some key findings have been described, including the ACVR2-dependent fucosylation changes of newly synthesized proteins in the HCT116 MSI colon cell line. A further finding of this study was the interaction with LFNG in ACVR2 expressing cells. The importance of this finding is underlined by the involvement of this glycosyltransferases in the regulation of Notch signaling. This leads to the assumption that Notch itself may be affected by ACVR2-dependent glycoprotein alteration. It could be speculated that the LFNG enzyme, which transfers GlcNAc onto fucose residues of the Notch receptor, might not be able to elicit its function, when the incorporation of the fucose residue is altered. Finally, the pathologic role of fucosylation in this model cell line remains to be elucidated in detail.

The second main focus of this approach was directed on the screening of the proteome und glycome. A universal feature of cancer cells are direct proteome alterations and the change in their glycosylation phenotype, with several effects on this tumor cells behavior [201, 202]. In this sense, glycosylation analysis has become an important target for proteomic research and has reached great interest to understand the molecular events associated with tumor development and progression. Although demonstrating the principle utility of our approach in detecting tumor-relevant proteomic and glycomic changes, there is also potential for further development and refinement of the methodology. The present study focused on soluble proteins. Since the extraction of the labeled proteins with high concentrations of detergents or chaotropic substances would not interfere with subsequent binding of biotin-labeled proteins to streptavidin beads, it would be possible to also apply it to membrane-bound or other difficult-to-solubilize proteins. An attractive refinement would be the combination of our strategy with quantitative proteomic analytical methods. Thus, this approach could be combined with the SILAC- or ITRAQ-quantification. Finally a combination with various prefractionation methods, including electrophoretic and chromatographic methods is feasible.

Most analyses within this study were performed with two independent HCT116-ACVR2 clones. In order to strengthen the ACVR2-dependent glycosylation changes, the results need to be extended to additional MSI colorectal cancer cell lines. Most importantly, the results gained from this *in vitro* cell culture model system need to be

confirmed in primary tissues. This requires MSI and MSS colorectal cancer tissues enabling immunohistochemical analyses of glycan and / or glycoprotein alterations. At the same time glycan-specific tools like antibodies need to be developed. Moreover, the altered glycan structures need to be identified and characterized. Thus, new tumor markers could be established, which are not only monitors for diagnosis or therapy, but also represent the biological characters of cancer cells.

In summary, the established MSI CRC cell line model system provides a unique experimental platform to determine ACVR2-dependent alterations of the MSI tumor cell proteome and glycome.

6. REFERENCES

1. van Steenbergen, L.N., et al., *Increasing incidence and decreasing mortality of colorectal cancer due to marked cohort effects in southern Netherlands.* Eur J Cancer Prev, 2009. **18**(2): p. 145-52.

2. Weitz, J., et al., *Colorectal cancer.* Lancet, 2005. **365**(9454): p. 153-65.

3. Levin, B., et al., *Screening and surveillance for the early detection of colorectal cancer and adenomatous polyps, 2008: a joint guideline from the American Cancer Society, the US Multi-Society Task Force on Colorectal Cancer, and the American College of Radiology.* CA Cancer J Clin, 2008. **58**(3): p. 130-60.

4. Lieberman, D., *Colon cancer screening and surveillance controversies.* Curr Opin Gastroenterol, 2009. **25**(5): p. 422-7.

5. Hoff, G. and J.A. Dominitz, *Contrasting US and European approaches to colorectal cancer screening: which is best?* Gut, 2010. **59**(3): p. 407-14.

6. Hol, L., et al., *Screening for colorectal cancer: randomised trial comparing guaiac-based and immunochemical faecal occult blood testing and flexible sigmoidoscopy.* Gut, 2010. **59**(1): p. 62-8.

7. Oort, F.A., et al., *Colonoscopy-controlled intra-individual comparisons to screen relevant neoplasia: faecal immunochemical test vs. guaiac-based faecal occult blood test.* Aliment Pharmacol Ther, 2010. **31**(3): p. 432-9.

8. Lieberman, D., *Progress and challenges in colorectal cancer screening and surveillance.* Gastroenterology, 2010. **138**(6): p. 2115-26.

9. Kinzler, K.W. and B. Vogelstein, *Lessons from hereditary colorectal cancer.* Cell, 1996. **87**(2): p. 159-70.

10. Jasperson, K.W., et al., *Hereditary and familial colon cancer.* Gastroenterology, 2010. **138**(6): p. 2044-58.

11. Cunningham, D., et al., *Colorectal cancer.* Lancet, 2010. **375**(9719): p. 1030-47.

12. Brosens, L.A., et al., *Gastrointestinal polyposis syndromes.* Curr Mol Med, 2007. **7**(1): p. 29-46.

13. Bertario, L., et al., *Predictors of metachronous colorectal neoplasms in sporadic adenoma patients.* Int J Cancer, 2003. **105**(1): p. 82-7.

14. Newcomb, P.A., et al., *Long-term efficacy of sigmoidoscopy in the reduction of colorectal cancer incidence.* J Natl Cancer Inst, 2003. **95**(8): p. 622-5.

15. Thiis-Evensen, E., et al., *Population-based surveillance by colonoscopy: effect on the incidence of colorectal cancer. Telemark Polyp Study I.* Scand J Gastroenterol, 1999. **34**(4): p. 414-20.

16. Vogelstein, B. and K.W. Kinzler, *The Genetic Basis of Human Cancer*. 2002, New York: McGraw Hill Professional.

17. Vogelstein, B., et al., *Genetic alterations during colorectal-tumor development.* N Engl J Med, 1988. **319**(9): p. 525-32.

18. Muto, T., H.J. Bussey, and B.C. Morson, *The evolution of cancer of the colon and rectum.* Cancer, 1975. **36**(6): p. 2251-70.

19. Risio, M., *The natural history of adenomas.* Best Pract Res Clin Gastroenterol, 2010. **24**(3): p. 271-80.

20. Wong, W.M., et al., *Histogenesis of human colorectal adenomas and hyperplastic polyps: the role of cell proliferation and crypt fission.* Gut, 2002. **50**(2): p. 212-7.

21. Worthley, D.L., et al., *Colorectal carcinogenesis: road maps to cancer.* World J Gastroenterol, 2007. **13**(28): p. 3784-91.

22. Yarze, J. and G. Falls, *Clinical Vignettes - Colon.* American Journal of Gastroenterology, 2009(104): p. 318-351.

23. Grady, W.M., *Genomic instability and colon cancer.* Cancer Metastasis Rev, 2004. **23**(1-2): p. 11-27.

24. Knudson, A.G., Jr., *Mutation and cancer: statistical study of retinoblastoma.* Proceedings of the National Academy of Sciences of the United States of America, 1971. **68**(4): p. 820-3.

25. Knudson, A.G., *Antioncogenes and human cancer.* Proceedings of the National Academy of Sciences of the United States of America, 1993. **90**(23): p. 10914-21.

26. Smalley, M.J. and T.C. Dale, *Wnt signaling and mammary tumorigenesis.* J Mammary Gland Biol Neoplasia, 2001. **6**(1): p. 37-52.

27. Fodde, R., *The APC gene in colorectal cancer.* Eur J Cancer, 2002. **38**(7): p. 867-71.

28. Henderson, B.R., *Nuclear-cytoplasmic shuttling of APC regulates beta-catenin subcellular localization and turnover.* Nat Cell Biol, 2000. **2**(9): p. 653-60.

29. Bienz, M. and H. Clevers, *Linking colorectal cancer to Wnt signaling.* Cell, 2000. **103**(2): p. 311-20.

30. Polakis, P., *The adenomatous polyposis coli (APC) tumor suppressor.* Biochim Biophys Acta, 1997. **1332**(3): p. F127-47.

31. Huber, O., et al., *Nuclear localization of beta-catenin by interaction with transcription factor LEF-1.* Mech Dev, 1996. **59**(1): p. 3-10.

32. He, T.C., et al., *Identification of c-MYC as a target of the APC pathway.* Science, 1998. **281**(5382): p. 1509-12.

33. Shtutman, M., et al., *The cyclin D1 gene is a target of the beta-catenin/LEF-1 pathway.* Proceedings of the National Academy of Sciences of the United States of America, 1999. **96**(10): p. 5522-7.

34. Leslie, A., et al., *The colorectal adenoma-carcinoma sequence.* Br J Surg, 2002. **89**(7): p. 845-60.

35. Liu, F., *SMAD4/DPC4 and pancreatic cancer survival. Commentary re: M. Tascilar et al., The SMAD4 protein and prognosis of pancreatic ductal adenocarcinoma. Clin. Cancer Res., 7: 4115-4121, 2001.* Clin Cancer Res, 2001. **7**(12): p. 3853-6.

36. Lane, D.P., *Cancer. p53, guardian of the genome.* Nature, 1992. **358**(6381): p. 15-6.

37. Boland, C.R., et al., *A National Cancer Institute Workshop on Microsatellite Instability for cancer detection and familial predisposition: development of international criteria for the determination of microsatellite instability in colorectal cancer.* Cancer research, 1998. **58**(22): p. 5248-57.

38. Hampel, H., et al., *Cancer risk in hereditary nonpolyposis colorectal cancer syndrome: later age of onset.* Gastroenterology, 2005. **129**(2): p. 415-21.

39. Lynch, H.T. and A. de la Chapelle, *Genetic susceptibility to non-polyposis colorectal cancer.* J Med Genet, 1999. **36**(11): p. 801-18.

40. Thibodeau, S.N., G. Bren, and D. Schaid, *Microsatellite instability in cancer of the proximal colon.* Science, 1993. **260**(5109): p. 816-9.

41. Hoeijmakers, J.H., *Genome maintenance mechanisms for preventing cancer.* Nature, 2001. **411**(6835): p. 366-74.

42. Soreide, K., et al., *Microsatellite instability in colorectal cancer.* Br J Surg, 2006. **93**(4): p. 395-406.

43. Kim, I.J., et al., *Development and applications of a beta-catenin oligonucleotide microarray: beta-catenin mutations are dominantly found in the proximal colon cancers with microsatellite instability.* Clin Cancer Res, 2003. **9**(8): p. 2920-5.

44. Markowitz, S., et al., *Inactivation of the type II TGF-beta receptor in colon cancer cells with microsatellite instability.* Science, 1995. **268**(5215): p. 1336-8.

45. Ellegren, H., *Microsatellites: simple sequences with complex evolution.* Nat Rev Genet, 2004. **5**(6): p. 435-45.

46. Shah, S.N., S.E. Hile, and K.A. Eckert, *Defective mismatch repair, microsatellite mutation bias, and variability in clinical cancer phenotypes.* Cancer Res, 2010. **70**(2): p. 431-5.

47. Kloor, M., M. von Knebel Doeberitz, and J.F. Gebert, *Molecular testing for microsatellite instability and its value in tumor characterization.* Expert Rev Mol Diagn, 2005. **5**(4): p. 599-611.

48. Boland, C.R. and A. Goel, *Microsatellite instability in colorectal cancer.* Gastroenterology, 2010. **138**(6): p. 2073-2087 e3.

49. Rajagopalan, H., et al., *Tumorigenesis: RAF/RAS oncogenes and mismatch-repair status.* Nature, 2002. **418**(6901): p. 934.

50. McGivern, A., et al., *Promoter hypermethylation frequency and BRAF mutations distinguish hereditary non-polyposis colon cancer from sporadic MSI-H colon cancer.* Fam Cancer, 2004. **3**(2): p. 101-7.

51. Peltomaki, P. and H. Vasen, *Mutations associated with HNPCC predisposition -- Update of ICG-HNPCC/INSiGHT mutation database.* Dis Markers, 2004. **20**(4-5): p. 269-76.

52. Linnebacher, M., et al., *Frameshift peptide-derived T-cell epitopes: a source of novel tumor-specific antigens.* Int J Cancer, 2001. **93**(1): p. 6-11.

53. Schwitalle, Y., et al., *Immune response against frameshift-induced neopeptides in HNPCC patients and healthy HNPCC mutation carriers.* Gastroenterology, 2008. **134**(4): p. 988-97.

54. Woerner, S.M., et al., *Pathogenesis of DNA repair-deficient cancers: a statistical meta-analysis of putative Real Common Target genes.* Oncogene, 2003. **22**(15): p. 2226-35.

55. Duval, A. and R. Hamelin, *Mutations at coding repeat sequences in mismatch repair-deficient human cancers: toward a new concept of target genes for instability.* Cancer research, 2002. **62**(9): p. 2447-54.

56. Woerner, S.M., et al., *Systematic identification of genes with coding microsatellites mutated in DNA mismatch repair-deficient cancer cells.* International journal of cancer. Journal international du cancer, 2001. **93**(1): p. 12-9.

57. Jung, B.H., et al., *Activin type 2 receptor restoration in MSI-H colon cancer suppresses growth and enhances migration with activin.* Gastroenterology, 2007. **132**(2): p. 633-44.

58. Attisano, L., et al., *Activation of signalling by the activin receptor complex.* Mol Cell Biol, 1996. **16**(3): p. 1066-73.

59. Chen, Y.G., et al., *Regulation of cell proliferation, apoptosis, and carcinogenesis by activin.* Exp Biol Med (Maywood), 2002. **227**(2): p. 75-87.

60. Attisano, L., et al., *Identification of human activin and TGF beta type I receptors that form heteromeric kinase complexes with type II receptors.* Cell, 1993. **75**(4): p. 671-80.

61. Ebner, R., et al., *Determination of type I receptor specificity by the type II receptors for TGF-beta or activin.* Science, 1993. **262**(5135): p. 900-2.

62. Laiho, M., M.B. Weis, and J. Massague, *Concomitant loss of transforming growth factor (TGF)-beta receptor types I and II in TGF-beta-resistant cell mutants implicates both receptor types in signal transduction.* J Biol Chem, 1990. **265**(30): p. 18518-24.

63. Wrana, J.L., et al., *TGF beta signals through a heteromeric protein kinase receptor complex.* Cell, 1992. **71**(6): p. 1003-14.

64. Mathews, L.S. and W.W. Vale, *Expression cloning of an activin receptor, a predicted transmembrane serine kinase.* Cell, 1991. **65**(6): p. 973-82.

65. Ebisawa, T., et al., *Characterization of bone morphogenetic protein-6 signaling pathways in osteoblast differentiation.* J Cell Sci, 1999. **112 (Pt 20)**: p. 3519-27.

66. Yamashita, H., et al., *Osteogenic protein-1 binds to activin type II receptors and induces certain activin-like effects.* J Cell Biol, 1995. **130**(1): p. 217-26.

67. Hilden, K., et al., *Expression of type II activin receptor genes during differentiation of human K562 cells and cDNA cloning of the human type IIB activin receptor.* Blood, 1994. **83**(8): p. 2163-70.

68. Tsuchida, K., et al., *Signal transduction pathway through activin receptors as a therapeutic target of musculoskeletal diseases and cancer.* Endocr J, 2008. **55**(1): p. 11-21.

69. Lebrun, J.J., et al., *Roles of pathway-specific and inhibitory Smads in activin receptor signaling.* Mol Endocrinol, 1999. **13**(1): p. 15-23.

70. Deacu, E., et al., *Activin type II receptor restoration in ACVR2-deficient colon cancer cells induces transforming growth factor-beta response pathway genes.* Cancer Res, 2004. **64**(21): p. 7690-6.

71. Yagi, K., et al., *c-myc is a downstream target of the Smad pathway.* J Biol Chem, 2002. **277**(1): p. 854-61.

72. Hempen, P.M., et al., *Evidence of selection for clones having genetic inactivation of the activin A type II receptor (ACVR2) gene in gastrointestinal cancers.* Cancer Res, 2003. **63**(5): p. 994-9.

73. Jung, B., et al., *Loss of activin receptor type 2 protein expression in microsatellite unstable colon cancers.* Gastroenterology, 2004. **126**(3): p. 654-9.

74. Jung, B., et al., *Influence of target gene mutations on survival, stage and histology in sporadic microsatellite unstable colon cancers.* Int J Cancer, 2006. **118**(10): p. 2509-13.

75. Grady, W.M. and J.M. Carethers, *Genomic and epigenetic instability in colorectal cancer pathogenesis.* Gastroenterology, 2008. **135**(4): p. 1079-99.

76. Bauer, J., et al., *Effects of activin and TGFbeta on p21 in colon cancer.* PLoS One, 2012. **7**(6): p. e39381.

77. Jung, B., et al., *Activin signaling in microsatellite stable colon cancers is disrupted by a combination of genetic and epigenetic mechanisms.* PLoS One, 2009. **4**(12): p. e8308.

78. Davis, B.G., *Biochemistry. Mimicking posttranslational modifications of proteins.* Science, 2004. **303**(5657): p. 480-2.

79. Hunter, T., *Signaling--2000 and beyond.* Cell, 2000. **100**(1): p. 113-27.

80. Nadolski, M.J. and M.E. Linder, *Protein lipidation.* FEBS J, 2007. **274**(20): p. 5202-10.

81. Hochstrasser, M., *Origin and function of ubiquitin-like proteins.* Nature, 2009. **458**(7237): p. 422-9.

82. Apweiler, R., H. Hermjakob, and N. Sharon, *On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database.* Biochim Biophys Acta, 1999. **1473**(1): p. 4-8.

83. Olofsson, S. and T. Bergstrom, *Glycoconjugate glycans as viral receptors.* Ann Med, 2005. **37**(3): p. 154-72.

84. Chen, H., et al., *Handbook of carbohydrate engineering; Mammalian glycosylation: An overview of carbohydrate biosynthesis.* CRC Press, Taylor & Francis group, 2005.

85. Silbert, J.E. and G. Sugumaran, *Biosynthesis of chondroitin/dermatan sulfate.* IUBMB Life, 2002. **54**(4): p. 177-86.

86. !!! INVALID CITATION !!!

87. Spiro, R.G., *Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds.* Glycobiology, 2002. **12**(4): p. 43R-56R.

88. Varki, A., et al., *Essentials of glycobiology.* Cold Spring harbor laboratory press, 2008. **2nd ed.**

89. Davies, G.J., T.M. Gloster, and B. Henrissat, *Recent structural insights into the expanding world of carbohydrate-active enzymes.* Curr Opin Struct Biol, 2005. **15**(6): p. 637-45.

90. Varki, A., et al., *Essentials of Glycobiology*. 2009, Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY).

91. Hakomori, S., *Aberrant glycosylation in cancer cell membranes as focused on glycolipids: overview and perspectives.* Cancer Res, 1985. **45**(6): p. 2405-14.

92. Hakomori, S., *Tumor malignancy defined by aberrant glycosylation and sphingo(glyco)lipid metabolism.* Cancer Res, 1996. **56**(23): p. 5309-18.

93. Orntoft, T.F. and E.M. Vestergaard, *Clinical aspects of altered glycosylation of glycoproteins in cancer.* Electrophoresis, 1999. **20**(2): p. 362-71.

94. Hakomori, S., *Glycosylation defining cancer malignancy: new wine in an old bottle.* Proc Natl Acad Sci U S A, 2002. **99**(16): p. 10231-3.

95. Kim, Y.J. and A. Varki, *Perspectives on the significance of altered glycosylation of glycoproteins in cancer.* Glycoconj J, 1997. **14**(5): p. 569-76.

96. Lau, K.S., et al., *Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation.* Cell, 2007. **129**(1): p. 123-34.

97. Dennis, J.W., et al., *Beta 1-6 branching of Asn-linked oligosaccharides is directly associated with metastasis.* Science, 1987. **236**(4801): p. 582-5.

98. Granovsky, M., et al., *Suppression of tumor growth and metastasis in Mgat5-deficient mice.* Nat Med, 2000. **6**(3): p. 306-12.

99. Guo, H.B., et al., *N-acetylglucosaminyltransferase V expression levels regulate cadherin-associated homotypic cell-cell adhesion and intracellular signaling pathways.* J Biol Chem, 2003. **278**(52): p. 52412-24.

100. Takano, R., E. Muchmore, and J.W. Dennis, *Sialylation and malignant potential in tumour cell glycosylation mutants.* Glycobiology, 1994. **4**(5): p. 665-74.

101. Takada, A., et al., *Contribution of carbohydrate antigens sialyl Lewis A and sialyl Lewis X to adhesion of human cancer cells to vascular endothelium.* Cancer Res, 1993. **53**(2): p. 354-61.

102. Renkonen, J., T. Paavonen, and R. Renkonen, *Endothelial and epithelial expression of sialyl Lewis(x) and sialyl Lewis(a) in lesions of breast carcinoma.* Int J Cancer, 1997. **74**(3): p. 296-300.

103. Fuster, M.M. and J.D. Esko, *The sweet and sour of cancer: glycans as novel therapeutic targets.* Nat Rev Cancer, 2005. **5**(7): p. 526-42.

104. Brockhausen, I., *Mucin-type O-glycans in human colon and breast cancer: glycodynamics and functions.* EMBO Rep, 2006. **7**(6): p. 599-604.

105. Lugli, A., et al., *Prognostic significance of mucins in colorectal cancer with different DNA mismatch-repair status.* J Clin Pathol, 2007. **60**(5): p. 534-9.

106. Nakamori, S., et al., *MUC1 mucin expression as a marker of progression and metastasis of human colorectal carcinoma.* Gastroenterology, 1994. **106**(2): p. 353-61.

107. Boland, C.R. and G.D. Deshmukh, *The carbohydrate composition of mucin in colonic cancer.* Gastroenterology, 1990. **98**(5 Pt 1): p. 1170-7.

108. Capon, C., et al., *Oligosaccharide structures of mucins secreted by the human colonic cancer cell line CL.16E.* J Biol Chem, 1992. **267**(27): p. 19248-57.

109. Yuan, M., et al., *Comparison of T-antigen expression in normal, premalignant, and malignant human colonic tissue using lectin and antibody immunohistochemistry.* Cancer Res, 1986. **46**(9): p. 4841-7.

110. Itzkowitz, S.H., et al., *Expression of Tn, sialosyl-Tn, and T antigens in human colon cancer.* Cancer Res, 1989. **49**(1): p. 197-204.

111. Said, I.T., et al., *Comparison of different techniques for detection of Gal-GalNAc, an early marker of colonic neoplasia.* Histol Histopathol, 1999. **14**(2): p. 351-7.

112. Saitoh, O., et al., *Differential glycosylation and cell surface expression of lysosomal membrane glycoproteins in sublines of a human colon cancer exhibiting distinct metastatic potentials.* J Biol Chem, 1992. **267**(8): p. 5700-11.

113. Seelentag, W.K., et al., *Prognostic value of beta1,6-branched oligosaccharides in human colorectal carcinoma.* Cancer Res, 1998. **58**(23): p. 5559-64.

114. Qiu, Y., et al., *Plasma glycoprotein profiling for colorectal cancer biomarker identification by lectin glycoarray and lectin blot.* J Proteome Res, 2008. **7**(4): p. 1693-703.

115. Hittelet, A., et al., *Upregulation of galectins-1 and -3 in human colon cancer and their role in regulating cell migration.* Int J Cancer, 2003. **103**(3): p. 370-9.

116. Nakamura, M., et al., *Involvement of galectin-3 expression in colorectal cancer progression and metastasis.* Int J Oncol, 1999. **15**(1): p. 143-8.

117. Legendre, H., et al., *Prognostic values of galectin-3 and the macrophage migration inhibitory factor (MIF) in human colorectal cancers.* Mod Pathol, 2003. **16**(5): p. 491-504.

118. Kim, H., et al., *Different gene expression profiles between microsatellite instability-high and microsatellite stable colorectal carcinomas.* Oncogene, 2004. **23**(37): p. 6218-25.

119. Ramasamy, V., et al., *Oligosaccharide preferences of beta1,4-galactosyltransferase-I: crystal structures of Met340His mutant of human beta1,4-galactosyltransferase-I with a pentasaccharide and trisaccharides of the N-glycan moiety.* J Mol Biol, 2005. **353**(1): p. 53-67.

120. Giordanengo, V., et al., *Cloning and expression of cDNA for a human Gal(beta1-3)GalNAc alpha2,3-sialyltransferase from the CEM T-cell line.* Eur J Biochem, 1997. **247**(2): p. 558-66.

121. Roeckel, N., et al., *High frequency of LMAN1 abnormalities in colorectal tumors with microsatellite instability.* Cancer research, 2009. **69**(1): p. 292-9.

122. Patsos, G., et al., *Compensation of loss of protein function in microsatellite-unstable colon cancer cells (HCT116): a gene-dependent effect on the cell surface glycan profile.* Glycobiology, 2009. **19**(7): p. 726-34.

123. Andre, S., et al., *Tumor suppressor p16INK4a--modulator of glycomic profile and galectin-1 expression to increase susceptibility to carbohydrate-dependent induction of anoikis in pancreatic carcinoma cells.* FEBS J, 2007. **274**(13): p. 3233-56.

124. Seales, E.C., et al., *A protein kinase C/Ras/ERK signaling pathway activates myeloid fibronectin receptors by altering beta1 integrin sialylation.* J Biol Chem, 2005. **280**(45): p. 37610-5.

125. Vogelstein, B., et al., *Cancer genome landscapes.* Science, 2013. **339**(6127): p. 1546-58.

126. Rahn, J.J., G.M. Adair, and R.S. Nairn, *Use of gene targeting to study recombination in mammalian cell DNA repair mutants.* Methods Mol Biol, 2012. **920**: p. 445-70.

127. Kohli, M., et al., *Facile methods for generating human somatic cell gene knockouts using recombinant adeno-associated viruses.* Nucleic Acids Res, 2004. **32**(1): p. e3.

128. Santiago, Y., et al., *Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases.* Proc Natl Acad Sci U S A, 2008. **105**(15): p. 5809-14.

129. Baum, C., et al., *Mutagenesis and oncogenesis by chromosomal insertion of gene transfer vectors.* Hum Gene Ther, 2006. **17**(3): p. 253-63.

130. Ong, S.E., *The expanding field of SILAC.* Anal Bioanal Chem, 2012. **404**(4): p. 967-76.

131. Dieterich, D.C., et al., *Labeling, detection and identification of newly synthesized proteomes with bioorthogonal non-canonical amino-acid tagging.* Nat Protoc, 2007. **2**(3): p. 532-40.

132. Mori, Y., et al., *Instabilotyping: comprehensive identification of frameshift mutations caused by coding region microsatellite instability.* Cancer Res, 2001. **61**(16): p. 6046-9.

133. Weidenfeld, I., et al., *Inducible expression of coding and inhibitory RNAs from retargetable genomic loci.* Nucleic acids research, 2009. **37**(7): p. e50.

134. Lee, J., et al., *Transforming growth factor beta receptor 2 (TGFBR2) changes sialylation in the microsatellite unstable (MSI) Colorectal cancer cell line HCT116.* PLoS One, 2013. **8**(2): p. e57074.

135. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.

136. Lowry, O.H., et al., *Protein measurement with the Folin phenol reagent.* J Biol Chem, 1951. **193**(1): p. 265-75.

137. Kang, D., et al., *Highly Sensitive and Fast Protein Detection with Coomassie Brilliant Blue in Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis.* Bull. Korean Chem. Soc., 2002. **23**(11): p. 1511-1512.

138. Welman, A., J. Barraclough, and C. Dive, *Generation of cells expressing improved doxycycline-regulated reverse transcriptional transactivator rtTA2S-M2.* Nat Protoc, 2006. **1**(2): p. 803-11.

139. Welman, A., et al., *Construction and characterization of multiple human colon cancer cell lines for inducibly regulated gene expression.* J Cell Biochem, 2005. **94**(6): p. 1148-62.

140. Barraclough, J., et al., *Increases in c-Yes expression level and activity promote motility but not proliferation of human colorectal carcinoma cells.* Neoplasia, 2007. **9**(9): p. 745-54.

141. Galamb, O., et al., *[Identification of colorectal cancer, adenoma, and inflammatory bowel disease specific gene expression patterns using whole genomic oligonucleotide microarray system].* Orv Hetil, 2007. **148**(44): p. 2067-79.

142. Lin, L., et al., *DEK over expression as an independent biomarker for poor prognosis in colorectal cancer.* BMC Cancer, 2013. **13**: p. 366.

143. Diehl, D., et al., *IGFBP-2 overexpression reduces the appearance of dysplastic aberrant crypt foci and inhibits growth of adenomas in chemically induced colorectal carcinogenesis.* Int J Cancer, 2009. **124**(9): p. 2220-5.

144. Chen, H.J., et al., *Dual inhibition of EGFR and c-Met kinase activation by MJ-56 reduces metastasis of HT29 human colorectal cancer cells.* Int J Oncol, 2013. **43**(1): p. 141-50.

145. Oshima, C.T., K. Iriya, and N.M. Forones, *Ki-67 as a prognostic marker in colorectal cancer but not in gastric cancer.* Neoplasma, 2005. **52**(5): p. 420-4.

146. Giaginis, C., et al., *Clinical significance of MCM-2 and MCM-5 expression in colon cancer: association with clinicopathological parameters and tumor proliferative capacity.* Dig Dis Sci, 2009. **54**(2): p. 282-91.

147. Line, A., et al., *Characterisation of tumour-associated antigens in colon cancer.* Cancer Immunol Immunother, 2002. **51**(10): p. 574-82.

148. Solier, S., et al., *Genome-wide analysis of novel splice variants induced by topoisomerase I poisoning shows preferential occurrence in genes encoding splicing factors.* Cancer Res, 2010. **70**(20): p. 8055-65.

149. Wang, G., et al., *Colorectal cancer progression correlates with upregulation of S100A11 expression in tumor tissues.* Int J Colorectal Dis, 2008. **23**(7): p. 675-82.

150. Piao, L., et al., *Association of LETM1 and MRPL36 contributes to the regulation of mitochondrial ATP production and necrotic cell death.* Cancer Res, 2009. **69**(8): p. 3397-404.

151. Zhang, Z., et al., *Regulation of growth and prostatic marker expression by activin A in an androgen-sensitive prostate cancer cell line LNCAP.* Biochem Biophys Res Commun, 1997. **234**(2): p. 362-5.

152. Crosnier, C., D. Stamataki, and J. Lewis, *Organizing cell renewal in the intestine: stem cells, signals and combinatorial control.* Nat Rev Genet, 2006. **7**(5): p. 349-59.

153. Levy, L. and C.S. Hill, *Alterations in components of the TGF-beta superfamily signaling pathways in human cancer.* Cytokine Growth Factor Rev, 2006. **17**(1-2): p. 41-58.

154. Gossen, M., et al., *Transcriptional activation by tetracyclines in mammalian cells.* Science, 1995. **268**(5218): p. 1766-9.

155. Beck, M., et al., *The quantitative proteome of a human cell line.* Mol Syst Biol, 2011. **7**: p. 549.

156. Mathews, L.S. and W.W. Vale, *Characterization of type II activin receptors. Binding, processing, and phosphorylation.* J Biol Chem, 1993. **268**(25): p. 19013-8.

157. Babel, I., et al., *Identification of tumor-associated autoantigens for the diagnosis of colorectal cancer in serum using high density protein microarrays.* Mol Cell Proteomics, 2009. **8**(10): p. 2382-95.

158. Harada, K., et al., *Serum immunoreactive activin A levels in normal subjects and patients with various diseases.* J Clin Endocrinol Metab, 1996. **81**(6): p. 2125-30.

159. Leto, G., et al., *Activin A circulating levels in patients with bone metastasis from breast or prostate cancer.* Clin Exp Metastasis, 2006. **23**(2): p. 117-22.

160. Alberts, B., et al., *Molecular Biology of the Cell, 4th ed.* Garland Science, N.Y, 2002: p. 1313-1362.

161. Adamczyk, B., T. Tharmalingam, and P.M. Rudd, *Glycans as cancer biomarkers.* Biochim Biophys Acta, 2012. **1820**(9): p. 1347-53.

162. Campion, B., et al., *Presence of fucosylated triantennary, tetraantennary and pentaantennary glycans in transferrin synthesized by the human hepatocarcinoma cell line Hep G2.* Eur J Biochem, 1989. **184**(2): p. 405-13.

163. Taketa, K., M. Izumi, and E. Ichikawa, *Distinct molecular species of human alpha-fetoprotein due to differential affinities to lectins.* Ann N Y Acad Sci, 1983. **417**: p. 61-8.

164. Yogeeswaran, G. and P.L. Salk, *Metastatic potential is positively correlated with cell surface sialylation of cultured murine tumor cell lines.* Science, 1981. **212**(4502): p. 1514-6.

165. Morgenthaler, J., W. Kemmner, and R. Brossmer, *Sialic acid dependent cell adhesion to collagen IV correlates with in vivo tumorigenicity of the human colon carcinoma sublines HCT116, HCT116a and HCT116b.* Biochemical and biophysical research communications, 1990. **171**(2): p. 860-6.

166. Moriwaki, K. and E. Miyoshi, *Fucosylation and gastrointestinal cancer.* World J Hepatol, 2010. **2**(4): p. 151-61.

167. Sullivan, F.X., et al., *Molecular cloning of human GDP-mannose 4,6-dehydratase and reconstitution of GDP-fucose biosynthesis in vitro.* J Biol Chem, 1998. **273**(14): p. 8193-202.

168. Tonetti, M., et al., *Synthesis of GDP-L-fucose by the human FX protein.* J Biol Chem, 1996. **271**(44): p. 27274-9.

169. Smith, P.L., et al., *Conditional control of selectin ligand expression and global fucosylation events in mice with a targeted mutation at the FX locus.* J Cell Biol, 2002. **158**(4): p. 801-15.

170. Moriwaki, K., et al., *Deficiency of GMDS leads to escape from NK cell-mediated tumor surveillance through modulation of TRAIL signaling.* Gastroenterology, 2009. **137**(1): p. 188-98, 198 e1-2.

171. Miyoshi, E., et al., *Expression of alpha1-6 fucosyltransferase in rat tissues and human cancer cell lines.* Int J Cancer, 1997. **72**(6): p. 1117-21.

172. Fortini, M.E., *Notch signaling: the core pathway and its posttranslational regulation.* Dev Cell, 2009. **16**(5): p. 633-47.

173. Kopan, R. and M.X. Ilagan, *The canonical Notch signaling pathway: unfolding the activation mechanism.* Cell, 2009. **137**(2): p. 216-33.

174. Haines, N. and K.D. Irvine, *Glycosylation regulates Notch signalling.* Nat Rev Mol Cell Biol, 2003. **4**(10): p. 786-97.

175. Jarriault, S., et al., *Signalling downstream of activated mammalian Notch.* Nature, 1995. **377**(6547): p. 355-8.

176. Bordonaro, M., et al., *The Notch ligand Delta-like 1 integrates inputs from TGFbeta/Activin and Wnt pathways.* Exp Cell Res, 2011. **317**(10): p. 1368-81.

177. Tyers, M. and M. Mann, *From genomics to proteomics.* Nature, 2003. **422**(6928): p. 193-7.

178. Van Eyk, J.E., *Proteomics: unraveling the complexity of heart disease and striving to change cardiology.* Curr Opin Mol Ther, 2001. **3**(6): p. 546-53.

179. Bellovin, D.I., et al., *Reciprocal regulation of RhoA and RhoC characterizes the EMT and identifies RhoC as a prognostic marker of colon carcinoma.* Oncogene, 2006. **25**(52): p. 6959-67.

180. Wang, H.B., et al., *Expression of RhoA and RhoC in colorectal carcinoma and its relations with clinicopathological parameters.* Clin Chem Lab Med, 2009. **47**(7): p. 811-7.

181. Kim, I.S., et al., *Gene expression of flap endonuclease-1 during cell proliferation and differentiation.* Biochim Biophys Acta, 2000. **1496**(2-3): p. 333-40.

182. Lam, J.S., et al., *Flap endonuclease 1 is overexpressed in prostate cancer and is associated with a high Gleason score.* BJU Int, 2006. **98**(2): p. 445-51.

183. Winkler, D.D., et al., *Histone chaperone FACT coordinates nucleosome interaction through multiple synergistic binding events.* J Biol Chem, 2011. **286**(48): p. 41883-92.

184. Kumar, Y., V. Radha, and G. Swarup, *Interaction with Sug1 enables Ipaf ubiquitination leading to caspase 8 activation and cell death.* Biochem J, 2010. **427**(1): p. 91-104.

185. Putker, M., et al., *Redox-dependent control of FOXO/DAF-16 by transportin-1.* Mol Cell, 2013. **49**(4): p. 730-42.

186. Ostergaard, M., et al., *Proteomic profiling of fibroblasts reveals a modulating effect of extracellular calumenin on the organization of the actin cytoskeleton.* Proteomics, 2006. **6**(12): p. 3509-19.

187. Bull, V.H., et al., *Temporal proteome profiling of taxol-induced mitotic arrest and apoptosis.* Electrophoresis, 2010. **31**(11): p. 1873-85.

188. Samani, A.A., et al., *The role of the IGF system in cancer growth and metastasis: overview and recent insights.* Endocr Rev, 2007. **28**(1): p. 20-47.

189. Michell, N.P., M.J. Langman, and M.C. Eggo, *Insulin-like growth factors and their binding proteins in human colonocytes: preferential degradation of insulin-like growth factor binding protein 2 in colonic cancers.* Br J Cancer, 1997. **76**(1): p. 60-6.

190. Nowikovsky, K., et al., *Perspectives on: SGP symposium on mitochondrial physiology and medicine: the pathophysiology of LETM1.* J Gen Physiol, 2012. **139**(6): p. 445-54.

191. Zhao, Y., et al., *Inhibition of the transcription factor Sp1 suppresses colon cancer stem cell growth and induces apoptosis in vitro and in nude mouse xenografts.* Oncol Rep, 2013. **30**(4): p. 1782-92.

192. Stein, B., et al., *Reduction or loss of HLA-A,B,C antigens in colorectal carcinoma appears not to influence survival.* Br J Cancer, 1988. **57**(4): p. 364-8.

193. Ito, T., et al., *Extracellular and transmembrane region of a podocalyxin-like protein 1 fragment identified from colon cancer cell lines.* Cell Biol Int, 2007. **31**(12): p. 1518-24.

194. Larsson, A., et al., *Overexpression of podocalyxin-like protein is an independent factor of poor prognosis in colorectal cancer.* Br J Cancer, 2011. **105**(5): p. 666-72.

195. Casey, G., et al., *Podocalyxin variants and risk of prostate cancer and tumor aggressiveness.* Hum Mol Genet, 2006. **15**(5): p. 735-41.

196. Zheng, L., et al., *Tyrosine 23 phosphorylation-dependent cell-surface localization of annexin A2 is required for invasion and metastases of pancreatic cancer.* PLoS One, 2011. **6**(4): p. e19390.

197. Lokman, N.A., et al., *The role of annexin A2 in tumorigenesis and cancer progression.* Cancer Microenviron, 2011. **4**(2): p. 199-208.

198. de Graauw, M., et al., *Annexin A2 phosphorylation mediates cell scattering and branching morphogenesis via cofilin Activation.* Mol Cell Biol, 2008. **28**(3): p. 1029-40.

199. Rescher, U., et al., *Tyrosine phosphorylation of annexin A2 regulates Rho-mediated actin rearrangement and cell adhesion.* J Cell Sci, 2008. **121**(Pt 13): p. 2177-85.

200. Hanahan, D. and R.A. Weinberg, *Hallmarks of cancer: the next generation.* Cell, 2011. **144**(5): p. 646-74.

201. Hakomori, S., *Aberrant glycosylation in tumors and tumor-associated carbohydrate antigens.* Adv Cancer Res, 1989. **52**: p. 257-331.

202. Srinivas, P.R., et al., *Proteomics for cancer biomarker discovery.* Clin Chem, 2002. **48**(8): p. 1160-9.