

Review

# Proteogenomic Approaches to Understand Gene Mutations and Protein Structural Alterations in Colon Cancer

Soumyadev Sarkar 

Center for Fundamental and Applied Microbiomics, Biodesign Institute, Arizona State University, Tempe, AZ 85281, USA; ssarka54@asu.edu

**Abstract:** With colon cancer being one of the deadliest and most common cancers, understanding the mechanisms behind colon cancer is crucial in improving therapies. One of the newest approaches in cancer research is the concept of proteogenomics. While genomic data is not sufficient to understand cancer, the integration of multi-omics data including proteomics in conjunction with protein modeling has a better potential to elucidate protein structural alterations and characterize tumors. This enables a more efficient diagnosis of cancer and improves remedial strategies. In this review, we aim to discuss the linkage between gene mutations and protein structural alterations that lead to colon cancer. Topics include alterations in the glycoproteome and structures of proteases that impact colon cancer development. Additionally, we highlight the importance of precision oncology with an emphasis on proteogenomic approaches, protein modeling, and the potential impact on colon cancer therapy.

**Keywords:** colon; colorectal; cancer; proteogenomics; protein; structure; modeling; function; precision; oncology

## 1. Introduction

The projected annual incidence of cancer is around 18 million, with colorectal cancers (CRC) ranking third among male cancers and second among female cancers [1]. The second most deadly cancer in the world, CRC is also the third most common cancer overall and has a high mortality rate of 9.2% of all cancer deaths [2]. More recently, the age group with higher incidence rates has fallen below the usual recommendation for starting colon cancer screenings [3]. This pattern is concerning since more colon cancers are now likely to occur in individuals that have not yet been screened and remain undiscovered for longer. In addition to being deadly and frequently occurring, CRC is a very complicated illness with many interrelated variables. The many environmental and genetic factors that affect the etiology of CRC contribute to its complexity.

Nearly 42% of all cancer cases have risk factors that may be preventable, such as smoking, eating poorly, getting too little fiber, being overweight, and not getting enough exercise [4]. However, only around 10% of CRC cases are due to family history or genetic ties, and the majority of CRC incidences are indirectly caused by a number of reasons which are primarily sporadic in combination with some hereditary factors [5]. The effect of the gut microbiota, including its activities, byproducts, and interactions with the host, on the risk of developing CRC is also intertwined with these factors. When there is dysbiosis or a lack of certain functions required for homeostasis, the microbiota can negatively or positively affect the immune system and colon [6]. Therefore, the complete picture of CRC development is probably not best captured by genomic-only techniques because CRC risk factors are complicated. Multi-omics methods are a fantastic tool for capturing many biological vantage points and can reveal new facets of complex diseases like CRC, which involve a wide variety of risk factors [7]. To progress diagnoses and the finding of biomarkers, other methods must be taken into consideration because genetics and familial history only have a minor impact on the total caseload.



**Citation:** Sarkar, S. Proteogenomic Approaches to Understand Gene Mutations and Protein Structural Alterations in Colon Cancer. *Physiologia* **2023**, *3*, 11–29. <https://doi.org/10.3390/physiologia3010002>

Academic Editor: Philip J. Atherton

Received: 29 October 2022

Revised: 13 December 2022

Accepted: 19 December 2022

Published: 4 January 2023



**Copyright:** © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Proteomics deal with the identification, localization, regulation, and quantification of proteins in a biological system. This method can be used to investigate significant protein modifications that genomic-only approaches overlook, like posttranslational alterations that significantly change a protein’s functional capabilities [8]. Proteogenomics, which combines genomic techniques with a proteomics approach, has the potential to identify more protein markers for CRC, allowing for earlier identification of risk factors and the implementation of therapy or preventative measures. In order to understand the functions associated with these proteins, proteogenomics has been employed in the field of colon cancer to identify the protein abundance profiles of CRC samples and link the abundance to functional genomic data [9,10]. Even with recent developments in treatment, improvements in the early diagnosis of CRC are essential since they increase survival rates [11].

CRC is leading contributor to morbidity, difficult to treat, and has significant long-term health impacts. Therefore, it is essential to further the study of CRC by considering several viewpoints, such as protein modifications and their role in the emergence of cancer. The significance of protein mutations in CRC that result in altered functionalities will be briefly covered in this review, along with when and how proteogenomics can be used to provide high resolution information on protein status and function and how it can be optimally used to diagnose or treat CRC.

## 2. Colorectal Cancer and Proteogenomics

### 2.1. Mutations That Lead to CRC

CRC tumorigenesis is associated with (1) microsatellite instability (MIN) (2) chromosomal instability (CIN), and (3) mutations involving a wide range of tumor-suppressor genes, proto-oncogenes, and epigenetic changes [12].

Chromosomal instability refers to the loss of one allele which might be associated with tumor-suppressor genes [12]. The genomic changes associated with the chromosomal instability pathway include activation of proto-oncogenes like K-Ras and loss of p53 and Adenomatous polyposis coli (APC) [13]. Microsatellite instability is a consequence of mutations in DNA mismatch repair genes, and that fail to perform repairs during the process of DNA replication. That often results in frameshift mutations, ultimately showing hallmarks of CRC such as angiogenesis and limitless replicative potential which is also the characteristic of stem cells [14,15]. In addition, there are point mutations of different oncogenes associated with CRC. The genes that are susceptible to mutations are *KRAS* [12,16,17], *TP53* [12,18,19], *APC* [12], *BRAF* [12,20], *SMAD4* [12,21–24], *β-Catenin* [12], and *AXIN* [12,25] (Table 1).

**Table 1.** Molecular mutations that lead to colorectal cancer. The genes listed are susceptible to mutations in colorectal cancer along with the locations of genes in the chromosome, the pattern of mutations, and the outcomes.

Genes	Locations	Function	Mutation Outcomes
<i>KRAS</i>	12p12	Proto-oncogenes have intrinsic GTPase activity	Trigger the transduction of differentiative signals, even without any extracellular stimuli [12,16,17]
<i>TP53</i>	short (p) arm of chromosome 17	Ensures cell cycle arrest and apoptosis to maintain genomic integrity	Results in the formation of a stable protein that no more can bind the DNA and activates target genes [12,18,19]
<i>APC</i>	5q21	Controls transcription of several cell proliferation genes	Increases transcription of $\beta$ -Catenin targets including cyclin D, ephrins, caspases, and C-myc [12]
<i>BRAF</i>	3p22-p21.3	Proto-oncogene	Results in being constitutively active in a RAS independent manner [12,20]

Table 1. Cont.

Genes	Locations	Function	Mutation Outcomes
<i>SMAD4</i>	long arm (q) of chromosome 18 at band 21.1	Regulate transcription of target genes, and act as a tumor-suppressor gene	Unable to regulate gene transcription, disrupt TGF- $\beta$ signaling [12,21–24]
$\beta$ -Catenin	3p22-p21.3	Transactivate target genes that inhibit apoptosis or encourage cell proliferation	Wnt-signaling activation [12]
<i>AXIN1 and AXIN2</i>	16p13.3 and 17q24.1	Down-regulate WNT pathway	Unable to regulate targeted pathways [12,25]

This review will focus on the structural abnormalities in the proteins translated from these mutated genes.

## 2.2. Functional Alterations from Molecular Mutations

Molecular mutations in the genes alter the normal functioning of the resultant proteins. RAS proteins can regulate several pathways such as apoptosis, differentiation, and cell growth [26]. Molecular mutations in the *KRAS* genes deregulate the protein to constitutive nature and are active even when there are no external stimuli [17,27]. *TP53* is a transcription factor having pro-apoptotic activity, enabling cell-cycle arrest [28–31]. Mutations in the *TP53* gene lead to a stable mutant protein that fails to bind to the DNA and triggers a set of target genes [19,32]. *BRAF* gene encodes for a protein belonging to the RAF family, and mutations lead to the constitutive activation of the RAS pathway [33]. Targets of *APC* include proteins such as C-myc, cyclin D, caspase, and ephrins. *APC* controls the transcription of these cell proliferation genes [34]. *APC* can also control microtubules [12]. Mutation in the C-terminal sequence of the *APC* can lead to the deregulation of *APC* and initiate colon tumorigenesis [35].  $\beta$ -Catenin also transactivates a set of target genes which may induce proliferation of the cells or is inhibitory towards apoptosis [36]. Mutations in the  $\beta$ -Catenin lead to its stabilization and ultimately lead to activation of the WNT-signaling [37]. *SMAD* and *AXIN* are also tumor suppressor genes, and mutations of these genes would lead to the activation of CRC [12]. It is also worth mentioning another layer of complexity that complicates CRC is the post-translational modifications (PTM) in the form of glycosylation [38], phosphorylation, acetylation, and ubiquitination [39]. PTMs resulting in amino acid modifications lead to the structural and functional diversity of proteins [40]. In CRC, PTMs can regulate a wide range of cellular processes such as transduction of cell signaling, energy generation and consumption, and DNA repair [41].

To sum up, we deduce the impact of molecular mutations on the functions of the different players regulating CRC. A current knowledge gap exists on how these mutations can cause structural aberrations in the proteins, which in turn give rise to functional abnormalities. The tumor-suppressor proteins have a target set of genes and understanding the anomalies of the structure-function relationship can be challenging.

## 2.3. Protein Modifications in CRC

### 2.3.1. Post-Translational Modifications in CRC

Post-translational modification (PTM) has a significant role in the development of cancer [42]. Such alterations can lead to the structural and functional diversity of proteins [40]. In CRC, PTMs like phosphorylation, ubiquitination, and acetylation are modifications of high biological significance [39]. For CRC development, the modifications were identified on the surface of proteins like Plasma protease C1 inhibitor (IC1), vitamin D-binding protein (VDBP), albumin (ALBU), X-ray repair cross-complementing protein 6 (XRCC6), and complement C4-A (CO4A) [39].

### 2.3.2. Alterations in the Patterns of Glycoproteins and Proteases That Impact CRC Development

Protein glycosylation has received considerable interest in cancer research owing to its relation to cancer development [43]. Plasma glycoproteins have been used to screen different types of cancer such as cancer antigen 125 (CA-125) in ovarian cancer, cancer antigen 15-3 (CA15-3) in breast cancer, and prostate-specific antigen (PSA) in prostate cancer [44–46]. Although, there was a knowledge gap regarding the impact of altered patterns of glycoproteins in CRC until recently [47]. Similarly, cysteine proteases like cathepsin L (CATL) and Cathepsin B (CATB), and serine proteases like tissue-type plasminogen activator (TPA) and urokinase-(UPA) and their inhibitor type-1 (PAI-1) have prominent functions in the CRC development [48]. Such proteins have been upregulated in CRC [48].

The impact of protein glycosylation and proteases is well-known in CRC development, however, there are outstanding questions about the structural alterations in the proteins that might lead to changes in protein functions.

### 2.4. Proteogenomics Approaches in Cancer, Specifically in CRC

CRC is intricately complex and relying only on the information from genomics can be insufficient for cancer diagnosis and treatment. This gives a half-cooked story of the events happening. Information about proteins is welcoming since it can help us to comprehend completely the underlying molecular pathology of cancer [49]. Transcriptomic profiling could be an upgrade that could add to the genome information. Moreover, transcriptomic information could improve the characterization of tumors that could facilitate specific cancer treatments [50]. However, as the target of most anticancer drugs are proteins, the limitation of the transcriptome lies in the fact that it fails to identify the changes in the functional status of the proteins involved in cancer [49]. That said, genomic information could be very useful in deciphering the somatic genomic and epigenomic modifications in the tumor cells [51]. However, there should be cohesiveness in preparing a catalog of these modifications along with systematic functional investigations to uncover the role of these modifications in inducing malignant transformation [51]. One of the large collaborative projects that exist currently is The Clinical Proteomic Tumor Analysis Consortium (CPTAC) [49]. The network initiated by the National Cancer Institute enhanced the comprehension of the molecular basis of cancer [52]. Although CRC is one of the focus areas of this network, it is now populated by ovarian, breast, and other types of cancer as well [52].

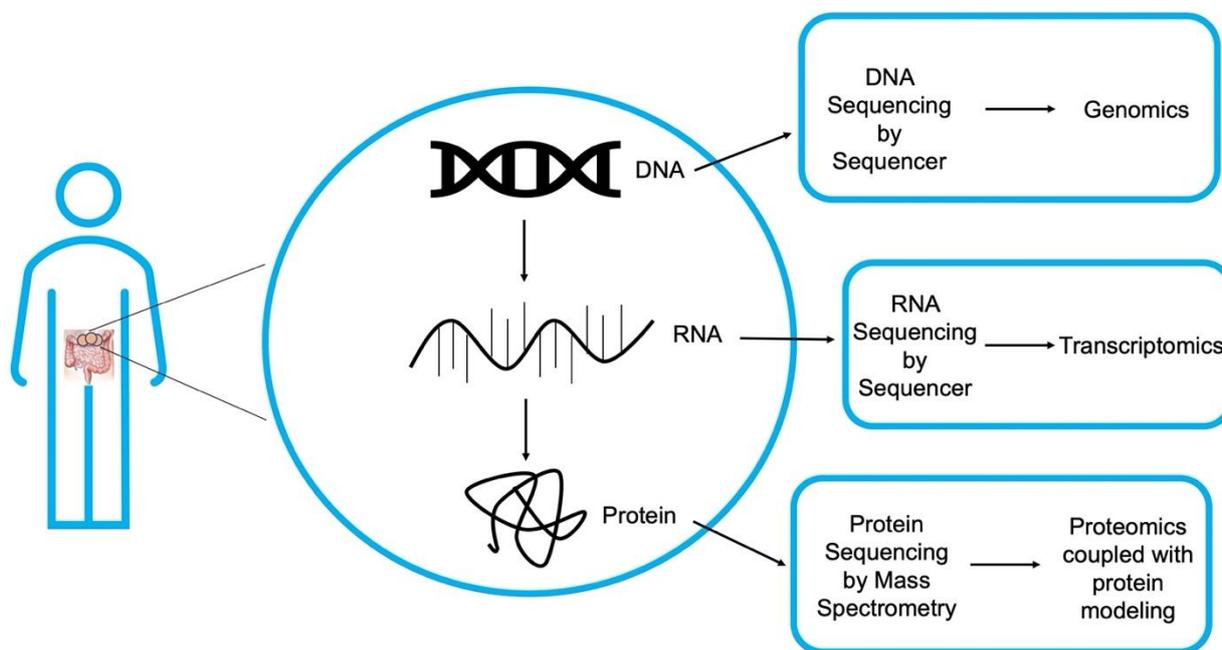
PTMs such as ubiquitinylation, phosphorylation, and glycosylation can impact protein stability in CRC [53]. Additionally, PTMs have the potential to alter antibody recognition and affinity. Proteogenomics could yield holistic means to address issues related to CRC by correcting both gene and protein sequences and circumvents the limitations of only genomic and transcriptomic investigations. Proteogenomics approaches in CRC could be used as a useful technique in biomarker discovery. In CRC, searching for predictive biomarkers has been a difficult task [54]. Proper biomarker discovery for a cohort affected by CRC would be helpful for the physicians to recommend specific targeted treatment for that group, thereby reducing the overall health expenses [55]. It has been shown previously that in cancer, the mRNA transcript abundance does not correlate with protein abundance [9]. This study along with several other studies [55–58] highlights the need for proteogenomics in resolving issues associated with cancer. Moreover, these studies also indicate that proteins in CRC are involved in the apoptotic process regulation and cellular protein metabolic process.

The molecular function of proteins is governed by the interaction selectivity with the partner molecules [39]. Such type of interactions often needs a stable and rigid structure of a protein. PTMs may induce small structural changes that would lead to complete loss or switching in the biological activity of the protein. Therefore, the comprehension of the tertiary environment of protein modification is required to understand the function and contributions of such altered proteins in pathophysiological processes.

Identification of global proteomic differences between normal and CRC tumor tissues is necessary to unravel cancer biomarkers and neoantigens. Yet, the current understanding is lacking in CRC cohorts. Global phosphoproteomics, ubiquitomics, and protein glycosylation analyses on human CRC are inadequate at this moment. Phosphoproteomics data suggested that Rb is amplified in CRC and Rb phosphorylation could be a target in CRC [10]. Likewise, phosphoproteomics data from CRC can help in targeting signaling proteins and pathways, providing information to generate therapeutic targets. In CRC, altered ubiquitination of CDK1 could be a pro-metastatic factor in colon adenocarcinoma [59]. Likewise, aberrant protein glycosylation, which resulted in pathological alterations that are widespread in CRC, and the underlying mechanisms for the contributions to CRC tumor progression are largely unknown. Enrichment of genes such as *B3GNT2*, *B4GALT2*, *ST6GALNAC2* has been associated with biosynthesis of N- and cores 1-3 O-linked glycans in the colon, and accounts for 16% of CRCs evaluated [60]. However, such studies are few and need more concerted efforts to understand CRC. Moreover, it is worth noting that the colon cancer-associated proteins and phosphites had little similarities and overlaps with already known cancer genes available in the Cancer Gene Census, further warranting future proteogenomics investigations of proteins, PTMs and associated pathways involved in CRC, so that novel biomarkers and neoantigens are deciphered [10].

2.5. Information from Proteogenomic Approaches and Precision Oncology

Proteogenomics is the area of research at the interface of genomics and proteomics [61] (Figure 1).



**Figure 1.** The concept of proteogenomics workflow. DNA, RNA and protein extracted from the colorectal cancer patients to be used to understand the genomics, transcriptomics, and proteomics. Protein modeling will provide solution to resolve the structure-function relationship of the proteins involved in CRC.

The field integrates information from genomics, transcriptomics, and proteomics [62]. Proteomics aims at producing a quantitative and complete map of the proteomes for a species [61]. Proteomics can yield a range of information on the cellular localization of proteins, entail details about the protein’s interaction partners and networks, post-translational modifications in proteins, whereas phosphoproteomics is able to generate a substantial amount of information about the signaling pathways [63]. This makes proteomics a perfect tool to understand cancer biology as most of the abnormalities will be at the protein levels.

As previously discussed, genomics would not be able to comprehend the events associated with alterations of the protein structures, functions and post-translational modifications. On the contrary, the proteome has the potential to bridge the analytical distance between the cancer genotype and the phenotype [49]. A typical approach to applying the concept of the proteome to cancer is to use the mass-spectrometry-dependent proteomics data and then pairing up with a proper database search algorithm such as MASCOT (Matrix Science Inc., Boston, MA, USA) [49,64]. It is worth mentioning that MASCOT algorithm has been successfully applied to identify novel diagnostic protein biomarkers in CRC [65,66]. There is also the need to include genomics and transcriptomics with proteomics while studying CRC. The introduction of the genomic and transcriptomic information would suffice the limitation of the proteome information alone on the novel proteins that would have no reference in the reference protein database [49]. Genomics, transcriptomics, and proteomics together could capture activity patterns in proteins driven by events like chromosomal deletion and amplification events, DNA copy number, and alterations in the expression of microRNA [49].

This brings us to the discussion of precision oncology—a newer approach to tackling cancer [67]. Precision oncology is the concept of implementing targeted treatments that could be customized based on individual tumor signatures and takes genomic, transcriptomic, and proteomic information into account [67]. There has been significant progress on targeted customized therapies in breast, lung, and melanoma tumors [68]. However, there have been only a few attempts to characterize the CRC using the proteogenomics approach, which could encourage precision oncology [9,10,69]. Till now, proteogenomics of CRC revealed distinct mutational profiles, therapeutic approaches, and candidate driver genes [7,9,10]. In CRC, proteogenomics approaches have the potential to promote customized drug development, however, need improving functional annotation resulting from genomic aberrations [49]. The multidimensional approach is promising for application to precision oncology [9,10,70–74].

Proteogenomics can drive therapeutic hypothesis generation and encourage precision oncology [75]. We propose that predictive biomarkers discovery through proteomics could generate the most efficient drugs to upgrade the current treatment paradigm of CRC. Additionally, newer and novel targets of CRC could provide a rationale for future drug development [75]. However, these approaches in CRC experience a “large p, small n” issue because the number of genes, mutations, proteins, and modification sites that could be evaluated in a study is several times more than the number of samples that can be evaluated in a possible timeline using techniques like LC-MS/MS. These putative candidates for therapeutic applications must be further examined before implementing clinical applications. This review hypothesizes the need for proper, efficient downstream analysis through incorporating proteogenomics approaches coupled with protein modeling.

## 2.6. Proteomics to Understand Structures of Proteins

Combination of Two-dimensional difference gel electrophoresis (2D DIGE) and Matrix-Assisted Laser Desorption/Ionization—Time of Flight (MALDI-TOF) mass spectrometry was successfully applied to identify biomarkers of CRC [76]. Proteins in extracellular vesicles (EVs) of CRC were evaluated using combination of absolute quantification labeling (iTRAQ) [77] and multidimensional liquid-chromatography-tandem mass spectrometry (MS/MS) [78]. Moreover, CRC biomarkers have been successfully identified using liquid chromatography-tandem mass spectrometry (LC-MS/MS) [79]. So, it is evident that mass spectrometry has a successful and considerable application in the identification of proteins associated with CRC.

Native Mass Spectrometry (NMS) has recently evolved and has the potential to be applied in cancer research [80]. NMS can successfully dissect the native contacts between the proteins, protein complexes [81], and protein kinetics [80]. NMS can yield information about protein conformation, cofactor content, topography, and protein complex stoichiometry [82,83]. As reported before, CRC is due to mutations of oncogenic mutants including

*KRAS*. *KRAS* mutations are prevalent in CRC amounting to 45% in the United States and as high as 49% in China [84]. NMS could be ideal to study energetics and kinetics of GTP hydrolysis of such oncogenic mutants and makes NMS suitable particularly for CRC [80]. *KRAS* protein is a membrane-bound GTPase (GTP hydrolase) which can control a range of cellular signaling pathways like the MAPK and PI3K pathways [84]. When bound to the GDP form, *KRAS* remains in an inactive state, and upon GTP-bound, the protein becomes active. In CRC oncogenic status, it is often observed that the *KRAS* exists in active status as its intrinsic GTPase function and GTPase activating proteins are hindered [85,86]. So, GTP hydrolysis status is key in CRC patients, and the effectiveness of NMS could help in the detailed investigation of such GTP hydrolysis. Mutant alleles with strongly hindered GTP hydrolysis including G12R and Q61R/K/L mutants, are less effectively degraded by direct pan-*KRAS* proteolysis targeting chimeras (PROTACs) targeting the GDP-bound form [84]. More investigations into such and other mutation forms would potentially help to establish newer *KRAS*-targeting drugs as a concept of precision therapy and can substitute for existing chemotherapeutic approaches.

Hydrogen/deuterium exchange (HDX) is another technique that can reveal substantial information about the dynamics of the protein [81]. HDX-MS thus has the potential to be applied in elucidating the PTMs associated with CRC by giving a clear account of the conformational dynamics of proteins and protein complexes [87]. Although this technique is nascent in being applied to the field of cancer investigation but could be beneficial if it could be applied. HDX-MS can easily bridge the existing knowledge gap of the protein structure-function relationship [81].

We emphasize on proteogenomics approach that uses the information from DNA sequencing, expressed sequence tags (ESTs), RNA sequencing, and ribosome profiling to generate customized protein sequence databases. This information would then help to interpret the proteomics data from MALDI-TOF-MS, LC-MS/MS, and iTRAQ-LC. This customized data would then encourage validation at the protein level of the gene expression data, and in turn refine gene models [61].

A typical proteogenomics approach workflow involves the extraction of DNA, RNA, and proteins from the tumor samples [88]. The DNA extracted will be used to obtain the genomic information, mRNA for transcriptome profiling, and proteins to understand the proteome [89]. For DNA and RNA sequence data generated, the first approach will be to trim the low-quality sequences. DNA data could be used to detect mutations, somatic copy number alterations, and microsatellite instability prediction, while RNA data can be used to obtain transcriptome profiles [89]. Workflow for proteome involves extraction of proteins, tryptic digestion, labeling of proteins by TMT-10, peptide fractionation by liquid chromatography, metal affinity chromatography to perform the phosphopeptide enrichment, and implementation of a mass spectrometry technique such as LC-MS/MS [90]. High-Performance Liquid Chromatography and a combination of Mass Spectrometry (HPLC-MS/MS) have been successfully applied to identify and characterize the PTMs associated with CRC when compared with normal patients [39]. Specifically, a bioinformatics approach was put forward for CRC patients and compared with healthy samples to identify acetylation, phosphorylation, and ubiquitination which eventually aimed at determining the altered biological activity of the proteins [39]. Acetylation, phosphorylation, and ubiquitination have significant contributions to CRC. Phosphorylated retinoblastoma proteins (Rb) have been linked with reduced apoptosis in CRC cells [91]. Phosphorylated Rbs interact with different protein partners and show different Rb functions in CRC [91]. Protein lysine acetylation could impact CRC metastasis through several pathways. PTMs involving proteins like isocitrate dehydrogenase (IDH1) are seen to influence hypoxia-inducible factor 1-alpha (HIF1 $\alpha$ ) dependent transcription of steroid receptor co-activator (SRC) transcription which further control CRC progression [92]. A very recent approach for cancer treatment is to eliminate oncogenic proteins by regulating the ubiquitin-proteasome system (UPS) [93]. Deubiquitinating enzymes (DUBs) tend to play important role in CRC formation and development by increasing the oncogenes stability [94]. Specifically, in CRC,

the Wnt-signaling pathway has been observed to be impacted by USP14 by enhancing Wnt signaling pathway [95]. USP4 and USP7 enhance the  $\beta$ -catenin and results in CRC tumor progression by modulating Wnt signaling [96,97]. This reiterates the study of such PTMs in CRC and forms the theoretical background for the application of such protein PTMs.

### 2.7. Approach to Protein Structure Modeling, and In Silico Mutations

One of the resources the scientific community has is the rich depository of crystal structures of the various tumor-suppressor proteins [98–101]. This review proposes to optimally utilize the crystal structure depositories so that the investigators can generate the necessary tertiary structure of various proteins linked to CRC. One popular way to generate the wild-type and mutated structure of the proteins is to use the power of protein structure modeling and in silico mutations using the information from these protein crystal structure depositories such as RCSB Protein Data Bank.

Protein mutations can lead to six possible outcomes—(1) protein activity alterations [102–107] (2) impacting protein–protein interactions [108–111] (3) affecting protein folding [103–105] (4) modifications in protein localization [112,113], (5) changing the half-life [114], and (6) combination of all these effects [102,103,108,109,112–114]. The knowledge from protein structure modeling and in silico mutations can address the knowledge gaps involving altered protein activity, structural abnormalities, protein–protein interactions, and protein folding.

Protein modeling, molecular dynamics simulation, and in silico mutations have been successfully employed to gain insights into the tumor-suppressor proteins in various cancer studies [115–117]. Additionally, this approach has even been applied to proteins involved in CRC [118–120]. However, there should be more concerted efforts to apply these in silico approaches to solve the mysteries involving cancer, and with a pool of proteomics data coming up, that should provide an ideal platform to implement this concept in near future [121].

### 2.8. Protein Modeling to Assess the Tertiary Structure (3D) of Proteins in CRC

Mass Spectrometry over the years has yielded a wealth of information on the identification of biomarkers of CRC [122,123]. Biomarkers can differentiate between a healthy patient with a cancer patient as biomarkers are only detected in the patient's blood or body fluids [124]. However, we want to highlight the importance of the structural aberrations in the proteins associated with CRC. The proteogenomics approach will generate insights into the gene levels, mRNA levels, and protein levels as well. A well-planned approach to applying mass spectrometry and in silico protein modeling could open new frontiers in cancer research, including CRC. Protein modeling would provide insights into the mechanisms of the functions of the altered proteins [125].

The usual approach from the mass spectrometry technique is to generate a spectrum from which peptide sequences can be detected [126]. Then, the sequences are searched against the NCBI database to identify the candidate proteins from which the sequences have been obtained [126]. As discussed previously, this approach could be problematic when novel proteins need to be identified. The absence of the necessary information in the database might pose challenges to identify those novel candidates. This is where proteogenomics could be useful, and conjugating proteogenomics with protein modeling would not only identifying the novel proteins but could also entail insights into the detailed mechanisms of such proteins [49,127].

Future cancer research investigations should focus on the tertiary structures of the proteins rather than on the sequences. The peptide sequences coming out of the mass spectrometry studies could be used as an input for the protein modeling techniques. One of the typical workflows could include the involvement of standard crosslinking techniques in which a crosslinker reacts with a definite residue and the second one form the crosslink [128]. A crosslinker is a reagent that has two functional groups separated by a spacer region. Two types of cross-linkers could be used for protein structural modeling—(1)

Standard crosslinkers, (2) Photo-crosslinkers. The crosslinker form covalent bonds when the crosslinker reacts with protein. For using photo-crosslinkers, the photoreactive groups are required to be activated with ultraviolet light [128]. The user then needs to comprehend the upper distance bound of the involved crosslinked residues. The two residues of the proteins can only react if the distance of the reactive groups is within the range of the crosslinker. The reactive crosslinkers would then store the spatial information. To elucidate the spatial information, the experimenter then digests the protein using enzymes like trypsin or other proteases [128]. Then, mass spectrometry of the generated peptides is carried out. After that, specialized and customized database search software evaluates the crosslinks resulting from the mass spectrometry data. The crosslinks finally dictate the input data to data-driven protein structure modeling [128].

Primarily, three protein modeling methods exist currently (Table 2):

- (1) Homology modeling: This method also known as comparative modeling could be used when a protein with a crystal structure is available in the database [129]. The query protein must possess >30% sequence identity with the protein available in the database [130]. The homology model could be built using efficient tools like MODELLER [131]. Previously, a wide range of proteins associated with cancer has been studied by this method [132–134].
- (2) Modeling by threading/fold recognition: Information on the protein folds based on similar proteins is used in predicting the structure of the proteins the users want to model. I-TASSER online server [135–137] can be used for modeling where different databases are used and the workflow is user-friendly.
- (3) Ab initio strategy: This is a powerful approach to predict protein structures when an appropriate homolog structure is unavailable in the database. The model is initiated and built using the information on the most favorable energy conformations of the participating amino acids, and also calculates the potential chemical interactions among the amino acid sequences [138]. However, this technique can be time-consuming and computationally intensive [139]. I-TASSER can apply the ab initio modeling when an appropriate template is absent [135]. QUARK [140,141] and CONFOLD2 [142] are other useful web servers for generating ab initio protein structures from amino acid sequences. While QUARK uses Monte Carlo simulation under the influence of an atomic-level-knowledge-based force field, CONFOLD2 uses a subset of input contacts to understand the protein fold space guided by a soft square energy function.

**Table 2.** Three protein modeling strategies with advantages and disadvantages of each system.

Protein Modeling System	Advantages	Disadvantages	Reference
Homology modeling	High-resolution structures can be generated	Physicochemical principle of protein modeling cannot be deciphered	[143]
Modeling by threading/ fold recognition	Works better for proteins when templates available are of distant homologies	The structures are less reliable than homology modeling	[144]
Ab initio strategy	Answers on how the protein takes a specific structure out of many structural possibilities	Less reliable for larger protein structures composed of more than 150 residues	[145]

To model the proteins with PTMs, PyTMs, a plugin devised for PyMOL is very useful [146]. The workflow enables easy standardization techniques, and ensures fast and easy user controls [146].

Although the techniques are independent, these can be intertwined to solve the protein structures. For example, the homology modeling approach and ab initio strategies could be used simultaneously. The ab initio can predict the areas of the proteins without homology, while the homology modeling approach can be applied to the other parts of the protein where there is a template to work with [81]. The generated structure can then be validated

using chemical cross-linking coupled to mass spectrometry (XL-MS) and surface labeling coupled to MS (SL-MS) [81]. The downsides of the protein modeling approach are its incapability of predicting multimeric protein complexes [81]. The downside however can be solved by using protein docking tools that can predict the models of protein complexes [81].

Often studies related to cancer might demand insights into various protein complexes, and so a definite strategy should be important. In CRC development, it is often observed that multiple candidates or protein complexes are involved either independently or in cascades, and so a definite strategy should be in place to deal with such instances [147]. Even the involvement of protein complexes in chemotherapy resistance for CRC has been observed [148]. However, such findings on the involvement of multimeric protein complexes and cascades are still in infancy and the involvement of proteogenomics approaches could be useful to understand CRC more extensively. A tool like ClusPro2.0 utilizes a thermodynamics-based approach to predict the lowest energy well for interacting proteins [149]. During the protein–protein docking, one protein is kept static and another forms several conformations. Similar conformations are kept together, and the conformations with the highest frequency are selected. The selected candidates are then refined and energy-minimized, and could be further validated using solution techniques like SL-MS and XL-MS [149]. If multiple subunits are required to generate a protein complex, this strategy can be repeated several times to consider additional subunits one at a time.

In cancer, cofactors often play roles in protein functioning and these need to be considered while generating protein 3D structures [150,151]. A usual approach could be to use SwissDock to locate the cofactor position within the protein and then perform the energy minimization to obtain the best possible confirmation of the protein with the cofactor [152].

For visualizing the monomeric and multimeric proteins, Chimera is a productive software with the plugin Xlink Analyzer that analyzes the protein structure validation using cross-linking data [153,154].

The target sites of colon cancer-related proteins are often nucleic acids, which then guide the downstream processes [155]. Modeling such protein-nucleic acids can be challenging, although there have been some recent breakthroughs [156,157]. Structure-based methods are preferred over sequence-based methods, and PRIME 2.0.1 is useful in predicting such complex protein-DNA structures [158].

### *2.9. Protein Docking of Mutated Proteins to the Substrates to Understand the Impact on Functions, the Need for Energy Minimization and Molecular Dynamics Simulation*

Proteins that have structural alterations should have a profound effect leading to cancer [159]. So, understanding the mutational implications of the structural stabilities of the proteins is critical. There are reports of understanding the mutants of the key proteins [160,161], however, there should be more initiatives from the scientific community that could decipher the mutated protein's functional mechanisms. One approach to induce amino acid alterations is to use Swiss PDB viewer [162], and then use appropriate protein docking software to dock the involved partners of the mutated proteins. There are several protein docking tools available currently with Autodock [163] and Autodock Vina [164,165] being the most reliable ones. The users will be able to easily compare the wild-type protein with the mutated counterparts and how the mutation impacts the overall binding affinities or specificities with the interacting partners. The change in the functions could also be tracked by such an approach [161].

Energy minimization, structural refinements, and molecular dynamics simulations are critical steps that need to be followed during protein molecular docking. Energy minimization of protein structures ensures the proper molecular arrangement of amino acids in space as the initial structures might not be energetically favorable [166]. Structural refinements of the protein structure are then carried out by adding missing atoms and neutralizing charges [167]. Energy minimization and structural refinements are steps under molecular dynamics simulation studies. Trajectory data from simulation should be

analyzed carefully to select the best energy-minimized structure of proteins, interacting partners and the protein-ligand structures [167]. Energy-minimized refined structures with appropriate simulation techniques ensure the stability and near-naiveness of the protein structures and interactions [168]. Currently, several useful tools are available to perform the simulation of proteins, with AMBER [169], NAMD [170], GROMACS [171], and CHARMM [172] being the most popular ones.

#### 2.10. Combining Hydrogen Exchange Mass Spectrometry and Protein Modeling to Understand 3D Structures

Hydrogen deuterium exchange mass spectrometry (HDX-MS) is a sophisticated and powerful technique that can elucidate the behavior of proteins by resolving the structure, dynamics, and function [173]. One of the interesting ways to optimize the HDX-MS experimental data is by conjugating the data to guide computational modeling tools like molecular docking and molecular dynamics [174,175]. In the past, molecular dynamics simulations have been used to decipher the structural properties obtained from HDX-MS data [176–182], even complementing the HDX-MS data by applying short time scales [183–186]. So going forward, there is a huge scope for integrating the strength of HDX-MS and modeling techniques to study the protein structure alterations in CRC [187].

### 3. Conclusions

Studies carried out in microorganisms like bacteria and yeast indicated a 50% correlation at the mRNA and protein levels [188]. However, the correlation significantly decreases when the genome complexity increases as in the case of humans, where only 30% of variations in protein levels could be explained by corresponding changes in the levels of mRNA [189]. This is due to various post-transcriptional (variable mRNA translational efficiency, siRNA regulation) and post-translational regulations (phosphorylation, glycosylation) which impact the protein stability [190]. So, it is difficult to comprehend the proteome dynamics from functional genomics data alone, and there is a need to complement the proteomics approach to genomics and transcriptomics resulting in the complete proteogenomics approach [190].

CRC is one of the deadliest cancers in the world [191]. Human genome sequencing has enabled the scientific community to fathom the genetic alterations in CRC in a comprehensive manner [192]. However, interpreting CRC only at the genomic level will not give the full picture, instead, the review suggests incorporating the information from the RNA and protein levels to gauge the functional alterations.

Proteogenomics approaches in CRC can facilitate the concept of precision oncology and encourage the applications of customized targeted treatments. In CRC, the proteogenomics approaches have been successfully applied to reveal new therapeutic avenues [10]. Specifically, personalized neoantigens for CRC patients were identified that could be used for the generation of therapeutic hypothesis [10].

The review highlights the importance of proteomics, and the way it could be integrated with protein modeling to grasp the underlying knowledge gaps in understanding the structural alterations in CRC, and also cancer in general. The major takeaways are: (1) From the clinical perspective, CRC is caused due to genetic changes involving *KRAS*, *TP53*, *APC*, *BRAF*, *SMAD4* and several other alterations (2) These alterations are not only a phenomenon occurring at the genomic or transcriptomic level, but does show up at the protein level owing to mutated structures and changed functions of such proteins (3) So, it is the best bet to correlate the genotype-phenotype attributes to capture the full extent of the CRC diagnosis, biomarker discovery and treatment (4) Proteomics has the potential to resolve the existing issues of understanding CRC by providing new insights regarding protein profiling, biomarker discovery and PTMs (5) The MS-based proteomics data would be coupled with protein modeling, protein docking, in silico mutations to understand the differences in protein's behavior between healthy and CRC patients. This knowledge would

then complement the information from clinical studies that relies on targeted proteomics to yield a better understanding of CRC. We not only discussed the existing information and questions associated with understanding CRC but also provided the avenues that can be explored to meet the current challenges.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Acknowledgments:** We are grateful to Sonny TM Lee and Tanner Richie of Division of Biology, Kansas State University for critically reviewing and editing the article. We would like to acknowledge the Lee lab, Division of Biology, Kansas State University for helpful insights, inputs, and discussions. Soumyadev Sarkar would also like to acknowledge the support of Ferran Garcia-Pichel of Center for Fundamental and Applied Microbiomics, Arizona State University and all the members of Garcia-Pichel Lab. The authors are thankful to Srijoni Basu and Bidisha Ghosh for the help with the inputs for the preparation of the figure.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [[CrossRef](#)] [[PubMed](#)]
2. Ferlay, J.; Colombet, M.; Soerjomataram, I.; Mathers, C.; Parkin, D.M.; Piñeros, M.; Znaor, A.; Bray, F. Estimating the Global Cancer Incidence and Mortality in 2018: GLOBOCAN Sources and Methods. *Int. J. Cancer* **2019**, *144*, 1941–1953. [[CrossRef](#)] [[PubMed](#)]
3. Siegel, R.L.; Miller, K.D.; Goding Sauer, A.; Fedewa, S.A.; Butterly, L.F.; Anderson, J.C.; Cercek, A.; Smith, R.A.; Jemal, A. Colorectal Cancer Statistics, 2020. *CA Cancer J. Clin.* **2020**, *70*, 145–164. [[CrossRef](#)] [[PubMed](#)]
4. Islami, F.; Goding Sauer, A.; Miller, K.D.; Siegel, R.L.; Fedewa, S.A.; Jacobs, E.J.; McCullough, M.L.; Patel, A.V.; Ma, J.; Soerjomataram, I.; et al. Proportion and Number of Cancer Cases and Deaths Attributable to Potentially Modifiable Risk Factors in the United States. *CA Cancer J. Clin.* **2018**, *68*, 31–54. [[CrossRef](#)] [[PubMed](#)]
5. Gausman, V.; Dornblaser, D.; Anand, S.; Hayes, R.B.; O’Connell, K.; Du, M.; Liang, P.S. Risk Factors Associated With Early-Onset Colorectal Cancer. *Clin. Gastroenterol. Hepatol.* **2020**, *18*, 2752–2759.e2. [[CrossRef](#)]
6. Birt, D.F.; Phillips, G.J. Diet, Genes, and Microbes: Complexities of Colon Cancer Prevention. *Toxicol. Pathol.* **2014**, *42*, 182–188. [[CrossRef](#)]
7. Imperial, R.; Ahmed, Z.; Toor, O.M.; Erdoğan, C.; Khaliq, A.; Case, P.; Case, J.; Kennedy, K.; Cummings, L.S.; Melton, N.; et al. Comparative Proteogenomic Analysis of Right-Sided Colon Cancer, Left-Sided Colon Cancer and Rectal Cancer Reveals Distinct Mutational Profiles. *Mol. Cancer* **2018**, *17*, 177. [[CrossRef](#)]
8. Aslam, B.; Basit, M.; Nisar, M.A.; Khurshid, M.; Rasool, M.H. Proteomics: Technologies and Their Applications. *J. Chromatogr. Sci.* **2017**, *55*, 182–196. [[CrossRef](#)]
9. Zhang, B.; Wang, J.; Wang, X.; Zhu, J.; Liu, Q.; Shi, Z.; Chambers, M.C.; Zimmerman, L.J.; Shaddox, K.F.; Kim, S.; et al. Proteogenomic Characterization of Human Colon and Rectal Cancer. *Nature* **2014**, *513*, 382–387. [[CrossRef](#)]
10. Vasaikar, S.; Huang, C.; Wang, X.; Petyuk, V.A.; Savage, S.R.; Wen, B.; Dou, Y.; Zhang, Y.; Shi, Z.; Arshad, O.A.; et al. Proteogenomic Analysis of Human Colon Cancer Reveals New Therapeutic Opportunities. *Cell* **2019**, *177*, 1035–1049.e19. [[CrossRef](#)]
11. Aguiar Junior, S.; de Oliveira, M.M.; Silva, D.R.M.E.; de Mello, C.A.L.; Calsavara, V.F.; Curado, M.P. Survival of patients with colorectal cancer in a cancer center. *Arq. Gastroenterol.* **2020**, *57*, 172–177. [[CrossRef](#)] [[PubMed](#)]
12. Sameer, A.S. Colorectal Cancer: Molecular Mutations and Polymorphisms. *Front. Oncol.* **2013**, *3*, 114. [[CrossRef](#)] [[PubMed](#)]
13. Armaghany, T.; Wilson, J.D.; Chu, Q.; Mills, G. Genetic Alterations in Colorectal Cancer. *Gastrointest. Cancer Res.* **2012**, *5*, 19–27. [[PubMed](#)]
14. Boland, C.R.; Goel, A. Microsatellite Instability in Colorectal Cancer. *Gastroenterology* **2010**, *138*, 2073–2087.e3. [[CrossRef](#)] [[PubMed](#)]
15. Mathonnet, M.; Perraud, A.; Christou, N.; Akil, H.; Melin, C.; Battu, S.; Jauberteau, M.-O.; Denizot, Y. Hallmarks in Colorectal Cancer: Angiogenesis and Cancer Stem-like Cells. *World J. Gastroenterol.* **2014**, *20*, 4189–4196. [[CrossRef](#)] [[PubMed](#)]
16. Fearon, E.R. Molecular Genetic Studies of the Adenoma-Carcinoma Sequence. *Adv. Intern. Med.* **1994**, *39*, 123–147. [[PubMed](#)]
17. Schubert, S.; Shannon, K.; Bollag, G. Hyperactive Ras in Developmental Disorders and Cancer. *Nat. Rev. Cancer* **2007**, *7*, 295–308. [[CrossRef](#)] [[PubMed](#)]
18. Levine, A.J.; Momand, J.; Finlay, C.A. The P53 Tumour Suppressor Gene. *Nature* **1991**, *351*, 453–456. [[CrossRef](#)]
19. Soussi, T.; Bérout, C. Assessing TP53 Status in Human Tumours to Evaluate Clinical Outcome. *Nat. Rev. Cancer* **2001**, *1*, 233–240. [[CrossRef](#)]

20. Wan, P.T.C.; Garnett, M.J.; Mark Roe, S.; Lee, S.; Niculescu-Duvaz, D.; Good, V.M.; Cancer Genome Project; Michael Jones, C.; Marshall, C.J.; Springer, C.J.; et al. Mechanism of Activation of the RAF-ERK Signaling Pathway by Oncogenic Mutations of B-RAF. *Cell* **2004**, *116*, 855–867. [[CrossRef](#)]
21. Shi, Y. Structural Insights on Smad Function in TGF $\beta$  Signaling. *BioEssays* **2001**, *23*, 223–232. [[CrossRef](#)] [[PubMed](#)]
22. Attisano, L.; Wrana, J.L. Smads as Transcriptional Co-Modulators. *Curr. Opin. Cell Biol.* **2000**, *12*, 235–243. [[CrossRef](#)] [[PubMed](#)]
23. Massagué, J.; Blain, S.W.; Lo, R.S. TGF $\beta$  Signaling in Growth Control, Cancer, and Heritable Disorders. *Cell* **2000**, *103*, 295–309. [[CrossRef](#)] [[PubMed](#)]
24. Attisano, L.; Lee-Hoeflich, S.T. The Smads. *Genome Biol.* **2001**, *2*, REVIEWS3010. [[CrossRef](#)] [[PubMed](#)]
25. Lammi, L.; Arte, S.; Somer, M.; Jarvinen, H.; Lahermo, P.; Thesleff, I.; Pirinen, S.; Nieminen, P. Mutations in AXIN2 Cause Familial Tooth Agenesis and Predispose to Colorectal Cancer. *Am. J. Hum. Genet.* **2004**, *74*, 1043–1050. [[CrossRef](#)]
26. Khosravi-Far, R.; Der, C.J. The Ras Signal Transduction Pathway. *Cancer Metastasis Rev.* **1994**, *13*, 67–89. [[CrossRef](#)]
27. Fearon, E.R.; Vogelstein, B. A Genetic Model for Colorectal Tumorigenesis. *Cell* **1990**, *61*, 759–767. [[CrossRef](#)]
28. Sionov, R.V.; Haupt, Y. The Cellular Response to P53: The Decision between Life and Death. *Oncogene* **1999**, *18*, 6145–6157. [[CrossRef](#)]
29. Prives, C.; Hall, P.A. The P53 Pathway. *J. Pathol.* **1999**, *187*, 112–126. [[CrossRef](#)]
30. Vousden, K.H.; Lu, X. Live or Let Die: The Cell's Response to P53. *Nat. Rev. Cancer* **2002**, *2*, 594–604. [[CrossRef](#)]
31. Lacroix, M.; Toillon, R.-A.; Leclercq, G. P53 and Breast Cancer, an Update. *Endocr. Relat. Cancer* **2006**, *13*, 293–325. [[CrossRef](#)] [[PubMed](#)]
32. Rivlin, N.; Brosh, R.; Oren, M.; Rotter, V. Mutations in the P53 Tumor Suppressor Gene: Important Milestones at the Various Steps of Tumorigenesis. *Genes Cancer* **2011**, *2*, 466–474. [[CrossRef](#)] [[PubMed](#)]
33. Houben, R.; Becker, J.C.; Kappel, A.; Terheyden, P.; Bröcker, E.-B.; Goetz, R.; Rapp, U.R. Constitutive Activation of the Ras-Raf Signaling Pathway in Metastatic Melanoma Is Associated with Poor Prognosis. *J. Carcinog.* **2004**, *3*, 6. [[CrossRef](#)] [[PubMed](#)]
34. Polakis, P. Wnt Signaling and Cancer. *Genes Dev.* **2000**, *14*, 1837–1851. [[CrossRef](#)]
35. Zhang, L.; Shay, J.W. Multiple Roles of APC and Its Therapeutic Implications in Colorectal Cancer. *J. Natl. Cancer Inst.* **2017**, *109*, 1–10. [[CrossRef](#)]
36. Saffroy, R.; Lemoine, A.; Debuire, B.; Brousse, P. Atlas of Genetics and Cytogenetics in Oncology and Haematology. 2004. Available online: <https://www.semanticscholar.org/paper/Atlas-of-Genetics-and-Cytogenetics-in-Oncology-and-Saffroy-Lemoine/9e01730f96d96c58e43cd06c0317df77af65e306> (accessed on 15 August 2022).
37. Behrens, J. The Role of the Wnt Signaling Pathway in Colorectal Tumorigenesis. *Biochem. Soc. Trans.* **2005**, *33*, 672–675. [[CrossRef](#)]
38. Zhang, L.; Ten Hagen, K.G. Pleiotropic Effects of O-Glycosylation in Colon Cancer. *J. Biol. Chem.* **2018**, *293*, 1315–1316. [[CrossRef](#)]
39. Tikhonov, D.; Kulikova, L.; Kopylov, A.; Malsagova, K.; Stepanov, A.; Rudnev, V.; Kaysheva, A. Super Secondary Structures of Proteins with Post-Translational Modifications in Colon Cancer. *Molecules* **2020**, *25*, 3144. [[CrossRef](#)]
40. Karve, T.M.; Cheema, A.K. Small Changes Huge Impact: The Role of Protein Posttranslational Modifications in Cellular Homeostasis and Disease. *J. Amino Acids* **2011**, *2011*, 207691. [[CrossRef](#)]
41. Vidal, C.J. *Post-Translational Modifications in Health and Disease*; Springer: New York, NY, USA, 2010.
42. Sharma, B.S.; Prabhakaran, V.; Desai, A.P.; Bajpai, J.; Verma, R.J.; Swain, P.K. Post-Translational Modifications (PTMs), from a Cancer Perspective: An Overview. *Oncogene* **2019**, *2*, 1–11. [[CrossRef](#)]
43. Pinho, S.S.; Reis, C.A. Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nat. Rev. Cancer* **2015**, *15*, 540–555. [[CrossRef](#)] [[PubMed](#)]
44. Kirwan, A.; Utratna, M.; O'Dwyer, M.E.; Joshi, L.; Kilcoyne, M. Glycosylation-Based Serum Biomarkers for Cancer Diagnostics and Prognostics. *Biomed Res. Int.* **2015**, *2015*, 490531. [[CrossRef](#)] [[PubMed](#)]
45. Kailemia, M.J.; Park, D.; Lebrilla, C.B. Glycans and Glycoproteins as Specific Biomarkers for Cancer. *Anal. Bioanal. Chem.* **2017**, *409*, 395–410. [[CrossRef](#)]
46. Ho, W.-L.; Hsu, W.-M.; Huang, M.-C.; Kadomatsu, K.; Nakagawara, A. Protein Glycosylation in Cancers and Its Potential Therapeutic Applications in Neuroblastoma. *J. Hematol. Oncol.* **2016**, *9*, 100. [[CrossRef](#)] [[PubMed](#)]
47. Chantaraamporn, J.; Champattanachai, V.; Khongmanee, A.; Verathamjamras, C.; Prasongsook, N.; Mingkwan, K.; Luevisadpibul, V.; Chutipongtanate, S.; Svasti, J. Glycoproteomic Analysis Reveals Aberrant Expression of Complement C9 and Fibronectin in the Plasma of Patients with Colorectal Cancer. *Proteomes* **2020**, *8*, 26. [[CrossRef](#)] [[PubMed](#)]
48. Herszényi, L.; Plebani, M.; Carraro, P.; De Paoli, M.; Roveroni, G.; Cardin, R.; Tulassay, Z.; Naccarato, R.; Farinati, F. The Role of Cysteine and Serine Proteases in Colorectal Carcinoma. *Cancer* **1999**, *86*, 1135–1142. [[CrossRef](#)]
49. Rodriguez, H.; Zenklusen, J.C.; Staudt, L.M.; Doroshow, J.H.; Lowy, D.R. The next Horizon in Precision Oncology: Proteogenomics to Inform Cancer Diagnosis and Treatment. *Cell* **2021**, *184*, 1661–1670. [[CrossRef](#)]
50. Rodon, J.; Soria, J.-C.; Berger, R.; Miller, W.H.; Rubin, E.; Kugel, A.; Tsimberidou, A.; Saintigny, P.; Ackerstein, A.; Braña, I.; et al. Genomic and Transcriptomic Profiling Expands Precision Cancer Medicine: The WINTHER Trial. *Nat. Med.* **2019**, *25*, 751–758. [[CrossRef](#)]
51. Chin, L.; Hahn, W.C.; Getz, G.; Meyerson, M. Making Sense of Cancer Genomic Data. *Genes Dev.* **2011**, *25*, 534–555. [[CrossRef](#)]
52. Macklin, A.; Khan, S.; Kislinger, T. Recent Advances in Mass Spectrometry Based Clinical Proteomics: Applications to Cancer Research. *Clin. Proteom.* **2020**, *17*, 17. [[CrossRef](#)]

53. Claydon, A.J.; Beynon, R. Proteome Dynamics: Revisiting Turnover with a Global Perspective. *Mol. Cell. Proteom.* **2012**, *11*, 1551–1565. [[CrossRef](#)] [[PubMed](#)]
54. Chauvin, A.; Boisvert, F.-M. Clinical Proteomics in Colorectal Cancer, a Promising Tool for Improving Personalised Medicine. *Proteomes* **2018**, *6*, 49. [[CrossRef](#)] [[PubMed](#)]
55. Chauvin, A.; Wang, C.-S.; Geha, S.; Garde-Granger, P.; Mathieu, A.-A.; Lacasse, V.; Boisvert, F.-M. The Response to Neoadjuvant Chemoradiotherapy with 5-Fluorouracil in Locally Advanced Rectal Cancer Patients: A Predictive Proteomic Signature. *Clin. Proteom.* **2018**, *15*, 1–16. [[CrossRef](#)]
56. Martin, P.; Noonan, S.; Mullen, M.P.; Scaife, C.; Tosetto, M.; Nolan, B.; Wynne, K.; Hyland, J.; Sheahan, K.; Elia, G.; et al. Predicting Response to Vascular Endothelial Growth Factor Inhibitor and Chemotherapy in Metastatic Colorectal Cancer. *BMC Cancer* **2014**, *14*, 887. [[CrossRef](#)] [[PubMed](#)]
57. Croner, R.S.; Sevim, M.; Metodiev, M.V.; Jo, P.; Ghadimi, M.; Schellerer, V.; Brunner, M.; Geppert, C.; Rau, T.; Stürzl, M.; et al. Identification of Predictive Markers for Response to Neoadjuvant Chemoradiation in Rectal Carcinomas by Proteomic Isotope Coded Protein Label (ICPL) Analysis. *Int. J. Mol. Sci.* **2016**, *17*, 209. [[CrossRef](#)]
58. Gong, F.-M.; Peng, X.-C.; Tan, B.-X.; Ge, J.; Chen, X.; Chen, Y.; Xu, F.; Bi, F.; Hou, J.-M.; Liu, J.-Y. Comparative Proteomic Analysis of Irinotecan-Sensitive Colorectal Carcinoma Cell Line and Its Chemoresistant Counterpart. *Anticancer Drugs* **2011**, *22*, 500–506. [[CrossRef](#)] [[PubMed](#)]
59. Zhang, Y.; Chen, C.; Yu, T.; Chen, T. Proteomic Analysis of Protein Ubiquitination Events in Human Primary and Metastatic Colon Adenocarcinoma Tissues. *Front. Oncol.* **2020**, *10*, 1684. [[CrossRef](#)]
60. Venkitachalam, S.; Revoredo, L.; Varadan, V.; Fecteau, R.E.; Ravi, L.; Lutterbaugh, J.; Markowitz, S.D.; Willis, J.E.; Gerken, T.A.; Guda, K. Biochemical and Functional Characterization of Glycosylation-Associated Mutational Landscapes in Colon Cancer. *Sci. Rep.* **2016**, *6*, 23642. [[CrossRef](#)]
61. Nesvizhskii, A.I. Proteogenomics: Concepts, Applications and Computational Strategies. *Nat. Methods* **2014**, *11*, 1114–1125. [[CrossRef](#)]
62. Mudge, J.M.; Frankish, A.; Harrow, J. Functional Transcriptomics in the Post-ENCODE Era. *Genome Res.* **2013**, *23*, 1961–1973. [[CrossRef](#)]
63. Mann, M.; Kulak, N.A.; Nagaraj, N.; Cox, J. The Coming Age of Complete, Accurate, and Ubiquitous Proteomes. *Mol. Cell* **2013**, *49*, 583–590. [[CrossRef](#)] [[PubMed](#)]
64. Perkins, D.N.; Pappin, D.J.; Creasy, D.M.; Cottrell, J.S. Probability-Based Protein Identification by Searching Sequence Databases Using Mass Spectrometry Data. *Electrophoresis* **1999**, *20*, 3551–3567. [[CrossRef](#)]
65. Ludvigsen, M.; Thorlacius-Ussing, L.; Vorum, H.; Moyer, M.P.; Stender, M.T.; Thorlacius-Ussing, O.; Honoré, B. Proteomic Characterization of Colorectal Cancer Cells versus Normal-Derived Colon Mucosa Cells: Approaching Identification of Novel Diagnostic Protein Biomarkers in Colorectal Cancer. *Int. J. Mol. Sci.* **2020**, *21*, 3466. [[CrossRef](#)] [[PubMed](#)]
66. Huang, C.-Y.; Lee, K.-C.; Tung, S.-Y.; Huang, W.-S.; Teng, C.-C.; Lee, K.-F.; Hsieh, M.-C.; Kuo, H.-C. 2D-DIGE-MS Proteomics Approaches for Identification of Gelsolin and Peroxiredoxin 4 with Lymph Node Metastasis in Colorectal Cancer. *Cancers* **2022**, *14*, 3189. [[CrossRef](#)]
67. Hodson, R. Precision Oncology. *Nature* **2020**, *585*, S1. [[CrossRef](#)]
68. Di Nicolantonio, F.; Vitiello, P.P.; Marsoni, S.; Siena, S.; Tabernero, J.; Trusolino, L.; Bernards, R.; Bardelli, A. Precision Oncology in Metastatic Colorectal Cancer—From Biology to Medicine. *Nat. Rev. Clin. Oncol.* **2021**, *18*, 506–525. [[CrossRef](#)]
69. Ma, Y.-S.; Huang, T.; Zhong, X.-M.; Zhang, H.-W.; Cong, X.-L.; Xu, H.; Lu, G.-X.; Yu, F.; Xue, S.-B.; Lv, Z.-W.; et al. Proteogenomic Characterization and Comprehensive Integrative Genomic Analysis of Human Colorectal Cancer Liver Metastasis. *Mol. Cancer* **2018**, *17*, 139. [[CrossRef](#)]
70. Archer, T.C.; Ehrenberger, T.; Mundt, F.; Gold, M.P.; Krug, K.; Mah, C.K.; Mahoney, E.L.; Daniel, C.J.; LeNail, A.; Ramamoorthy, D.; et al. Proteomics, Post-Translational Modifications, and Integrative Analyses Reveal Molecular Heterogeneity within Medulloblastoma Subgroups. *Cancer Cell* **2018**, *34*, 396–410.e8. [[CrossRef](#)]
71. Mertins, P.; Mani, D.R.; Ruggles, K.V.; Gillette, M.A.; Clauser, K.R.; Wang, P.; Wang, X.; Qiao, J.W.; Cao, S.; Petralia, F.; et al. Proteogenomics Connects Somatic Mutations to Signalling in Breast Cancer. *Nature* **2016**, *534*, 55–62. [[CrossRef](#)]
72. Mundt, F.; Rajput, S.; Li, S.; Ruggles, K.V.; Mooradian, A.D.; Mertins, P.; Gillette, M.A.; Krug, K.; Guo, Z.; Hoog, J.; et al. Mooradian Mass Spectrometry–Based Proteomics Reveals Potential Roles of NEK9 and MAP2K4 in Resistance to PI3K Inhibition in Triple-Negative Breast Cancers. *Cancer Res.* **2018**, *78*, 2732–2746. [[CrossRef](#)]
73. Huang, K.-L.; Li, S.; Mertins, P.; Cao, S.; Gunawardena, H.P.; Ruggles, K.V.; Mani, D.R.; Clauser, K.R.; Tanioka, M.; Usary, J.; et al. Proteogenomic Integration Reveals Therapeutic Targets in Breast Cancer Xenografts. *Nat. Commun.* **2017**, *8*, 14864. [[CrossRef](#)] [[PubMed](#)]
74. Matsunuma, R.; Chan, D.W.; Kim, B.-J.; Singh, P.; Han, A.; Saltzman, A.B.; Cheng, C.; Lei, J.T.; Wang, J.; Roberto da Silva, L.; et al. DPYSL3 Modulates Mitosis, Migration, and Epithelial-to-Mesenchymal Transition in Claudin-Low Breast Cancer. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, E11978–E11987. [[CrossRef](#)] [[PubMed](#)]
75. Lei, J.T.; Zhang, B. Proteogenomics Drives Therapeutic Hypothesis Generation for Precision Oncology. *Br. J. Cancer* **2021**, *125*, 1–3. [[CrossRef](#)] [[PubMed](#)]

76. Albulescu, R.; Jose Petrescu, A.; Sarbu, M.; Grigore, A.; Ica, R.; Munteanu, C.V.A.; Albulescu, A.; Militaru, I.V.; Zamfir, A.-D.; Petrescu, S.; et al. Mass Spectrometry for Cancer Biomarkers. In *Proteomics Technologies and Applications*; IntechOpen: London, UK, 2019; ISBN 9781789846102.
77. Melchior, K.; Tholey, A.; Heisel, S.; Keller, A.; Lenhof, H.-P.; Meese, E.; Huber, C.G. Proteomic Study of Human Glioblastoma Multiforme Tissue Employing Complementary Two-Dimensional Liquid Chromatography- and Mass Spectrometry-Based Approaches. *J. Proteome Res.* **2009**, *8*, 4604–4614. [[CrossRef](#)]
78. Jimenez, L.; Yu, H.; McKenzie, A.J.; Franklin, J.L.; Patton, J.G.; Liu, Q.; Weaver, A.M. Quantitative Proteomic Analysis of Small and Large Extracellular Vesicles (EVs) Reveals Enrichment of Adhesion Proteins in Small EVs. *J. Proteome Res.* **2019**, *18*, 947–959. [[CrossRef](#)]
79. Lee, C.-H.; Im, E.-J.; Moon, P.-G.; Baek, M.-C. Discovery of a Diagnostic Biomarker for Colon Cancer through Proteomic Profiling of Small Extracellular Vesicles. *BMC Cancer* **2018**, *18*, 1058. [[CrossRef](#)]
80. Moghadamchargari, Z.; Huddleston, J.; Shirzadeh, M.; Zheng, X.; Clemmer, D.E.; M Raushel, F.; Russell, D.H.; Laganowsky, A. Intrinsic GTPase Activity of K-RAS Monitored by Native Mass Spectrometry. *Biochemistry* **2019**, *58*, 3396–3405. [[CrossRef](#)]
81. Tokmina-Lukaszewska, M.; Patterson, A.; Berry, L.; Scott, L.; Balasubramanian, N.; Bothner, B. The Role of Mass Spectrometry in Structural Studies of Flavin-Based Electron Bifurcating Enzymes. *Front. Microbiol.* **2018**, *9*, 1397. [[CrossRef](#)]
82. Kirshenbaum, N.; Michaelevski, I.; Sharon, M. Analyzing Large Protein Complexes by Structural Mass Spectrometry. *J. Vis. Exp.* **2010**, *40*, PMC3149987. [[CrossRef](#)]
83. Laganowsky, A.; Reading, E.; Hopper, J.T.S.; Robinson, C.V. Mass Spectrometry of Intact Membrane Protein Complexes. *Nat. Protoc.* **2013**, *8*, 639–651. [[CrossRef](#)]
84. Hofmann, M.H.; Gerlach, D.; Misale, S.; Petronczki, M.; Kraut, N. Expanding the Reach of Precision Oncology by Drugging All KRAS Mutants. *Cancer Discov.* **2022**, *12*, 924–937. [[CrossRef](#)] [[PubMed](#)]
85. Ratner, N.; Miller, S.J. A RASopathy Gene Commonly Mutated in Cancer: The Neurofibromatosis Type 1 Tumour Suppressor. *Nat. Rev. Cancer* **2015**, *15*, 290–301. [[CrossRef](#)] [[PubMed](#)]
86. Targeting, K. G12C: From Inhibitory Mechanism to Modulation of Antitumor Effects in Patients. *Cell* **2020**, *183*, 850–859.
87. Liu, H.; Wang, D.; Zhang, Q.; Zhao, Y.; Mamonova, T.; Wang, L.; Zhang, C.; Li, S.; Friedman, P.A.; Xiao, K. Parallel Post-Translational Modification Scanning Enhancing Hydrogen-Deuterium Exchange-Mass Spectrometry Coverage of Key Structural Regions. *Anal. Chem.* **2019**, *91*, 6976–6980. [[CrossRef](#)] [[PubMed](#)]
88. De Marchi, T.; Pyl, P.T.; Sjöström, M.; Klasson, S.; Sartor, H.; Tran, L.; Pekar, G.; Malmström, J.; Malmström, L.; Niméus, E. Proteogenomic Workflow Reveals Molecular Phenotypes Related to Breast Cancer Mammographic Appearance. *J. Proteome Res.* **2021**, *20*, 2983–3001. [[CrossRef](#)] [[PubMed](#)]
89. Krug, K.; Jaehnig, E.J.; Satpathy, S.; Blumenberg, L.; Karpova, A.; Anurag, M.; Miles, G.; Mertins, P.; Geffen, Y.; Tang, L.C.; et al. Proteogenomic Landscape of Breast Cancer Tumorigenesis and Targeted Therapy. *Cell* **2020**, *183*, 1436–1456.e31. [[CrossRef](#)]
90. Mertins, P.; Tang, L.C.; Krug, K.; Clark, D.J.; Gritsenko, M.A.; Chen, L.; Clauser, K.R.; Clauss, T.R.; Shah, P.; Gillette, M.A.; et al. Reproducible Workflow for Multiplexed Deep-Scale Proteome and Phosphoproteome Analysis of Tumor Tissues by Liquid Chromatography–Mass Spectrometry. *Nat. Protoc.* **2018**, *13*, 1632–1661. [[CrossRef](#)]
91. Chen, L.; Liu, S.; Tao, Y. Regulating Tumor Suppressor Genes: Post-Translational Modifications. *Signal Transduct. Target. Ther.* **2020**, *5*, 90. [[CrossRef](#)]
92. Wang, B.; Ye, Y.; Yang, X.; Liu, B.; Wang, Z.; Chen, S.; Jiang, K.; Zhang, W.; Jiang, H.; Mustonen, H.; et al. SIRT2-Dependent IDH1 Deacetylation Inhibits Colorectal Cancer and Liver Metastases. *EMBO Rep.* **2020**, *21*, e48183. [[CrossRef](#)]
93. D’Arcy, P.; Wang, X.; Linder, S. Deubiquitinase Inhibition as a Cancer Therapeutic Strategy. *Pharmacol. Ther.* **2015**, *147*, 32–54. [[CrossRef](#)]
94. Yun, S.-I.; Hong, H.K.; Yeo, S.-Y.; Kim, S.-H.; Cho, Y.B.; Kim, K.K. Ubiquitin-Specific Protease 21 Promotes Colorectal Cancer Metastasis by Acting as a Fra-1 Deubiquitinase. *Cancers* **2020**, *12*, 207. [[CrossRef](#)] [[PubMed](#)]
95. Jung, H.; Kim, B.-G.; Han, W.H.; Lee, J.H.; Cho, J.-Y.; Park, W.S.; Maurice, M.M.; Han, J.-K.; Lee, M.J.; Finley, D.; et al. Deubiquitination of Dishevelled by Usp14 Is Required for Wnt Signaling. *Oncogenesis* **2013**, *2*, e64. [[CrossRef](#)] [[PubMed](#)]
96. Yun, S.-I.; Kim, H.H.; Yoon, J.H.; Park, W.S.; Hahn, M.-J.; Kim, H.C.; Chung, C.H.; Kim, K.K. Ubiquitin Specific Protease 4 Positively Regulates the WNT/ $\beta$ -Catenin Signaling in Colorectal Cancer. *Mol. Oncol.* **2015**, *9*, 1834–1851. [[CrossRef](#)] [[PubMed](#)]
97. Ma, P.; Yang, X.; Kong, Q.; Li, C.; Yang, S.; Li, Y.; Mao, B. The Ubiquitin Ligase RNF220 Enhances Canonical Wnt Signaling through USP7-Mediated Deubiquitination of  $\beta$ -Catenin. *Mol. Cell. Biol.* **2014**, *34*, 4355–4366. [[CrossRef](#)]
98. Cho, Y.; Gorina, S.; Jeffrey, P.D.; Pavletich, N.P. Crystal Structure of a P53 Tumor Suppressor-DNA Complex: Understanding Tumorigenic Mutations. *Science* **1994**, *265*, 346–355. [[CrossRef](#)]
99. Xing, Y.; Clements, W.K.; Le Trong, I.; Hinds, T.R.; Stenkamp, R.; Kimelman, D.; Xu, W. Crystal Structure of a Beta-Catenin/APC Complex Reveals a Critical Role for APC Phosphorylation in APC Function. *Mol. Cell* **2004**, *15*, 523–533. [[CrossRef](#)]
100. Xing, Y.; Clements, W.K.; Kimelman, D.; Xu, W. Crystal Structure of a Beta-Catenin/Axin Complex Suggests a Mechanism for the Beta-Catenin Destruction Complex. *Genes Dev.* **2003**, *17*, 2753–2764. [[CrossRef](#)]
101. Pantsar, T. The Current Understanding of KRAS Protein Structure and Dynamics. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 189–198. [[CrossRef](#)]

102. Ode, H.; Matsuyama, S.; Hata, M.; Neya, S.; Kakizawa, J.; Sugiura, W.; Hoshino, T. Computational Characterization of Structural Role of the Non-Active Site Mutation M36I of Human Immunodeficiency Virus Type 1 Protease. *J. Mol. Biol.* **2007**, *370*, 598–607. [[CrossRef](#)]
103. Lorch, M.; Mason, J.M.; Sessions, R.B.; Clarke, A.R. Effects of Mutations on the Thermodynamics of a Protein Folding Reaction: Implications for the Mechanism of Formation of the Intermediate and Transition States. *Biochemistry* **2000**, *39*, 3480–3485. [[CrossRef](#)]
104. Lorch, M.; Mason, J.M.; Clarke, A.R.; Parker, M.J. Effects of Core Mutations on the Folding of a Beta-Sheet Protein: Implications for Backbone Organization in the I-State. *Biochemistry* **1999**, *38*, 1377–1385. [[CrossRef](#)] [[PubMed](#)]
105. Alfalah, M.; Keiser, M.; Leeb, T.; Zimmer, K.-P.; Naim, H.Y. Compound Heterozygous Mutations Affect Protein Folding and Function in Patients with Congenital Sucrase-Isomaltase Deficiency. *Gastroenterology* **2009**, *136*, 883–892. [[CrossRef](#)] [[PubMed](#)]
106. Koukouritaki, S.B.; Poch, M.T.; Henderson, M.C.; Siddens, L.K.; Krueger, S.K.; VanDyke, J.E.; Williams, D.E.; Pajewski, N.M.; Wang, T.; Hines, R.N. Identification and Functional Analysis of Common Human Flavin-Containing Monooxygenase 3 Genetic Variants. *J. Pharmacol. Exp. Ther.* **2007**, *320*, 266–273. [[CrossRef](#)] [[PubMed](#)]
107. Cristofaro, R.; Carotti, A.; Akhavan, S.; Palla, R.; Peyvand, F.; Altomare, C.; Mannucci, P.M. The Natural Mutation by Deletion of Lys9 in the Thrombin A-Chain Affects the PKa Value of Catalytic Residues, the Overall Enzyme's Stability and Conformational Transitions Linked to Na Binding. *FEBS J.* **2006**, *273*, 159–169. [[CrossRef](#)] [[PubMed](#)]
108. Jones, R.; Ruas, M.; Gregory, F.; Moulin, S.; Delia, D.; Manoukian, S.; Rowe, J.; Brookes, S.; Peters, G. A CDKN2A Mutation in Familial Melanoma That Abrogates Binding of P16INK4a to CDK4 but Not CDK6. *Cancer Res.* **2007**, *67*, 9134–9141. [[CrossRef](#)] [[PubMed](#)]
109. Ung, M.-U.; Lu, B.; McCammon, J.A. E230Q Mutation of the Catalytic Subunit of CAMP-Dependent Protein Kinase Affects Local Structure and the Binding of Peptide Inhibitor. *Biopolymers* **2006**, *81*, 428–439. [[CrossRef](#)] [[PubMed](#)]
110. Rignall, T.R.; Baker, J.O.; McCarter, S.L.; Adney, W.S.; Vinzant, T.B.; Decker, S.R.; Himmel, M.E. Effect of Single Active-Site Cleft Mutation on Product Specificity in a Thermostable Bacterial Cellulase. *Biotechnol. Fuels Chem.* **2002**, *98*, 383–394.
111. Hardt, M.; Laine, R.A. Mutation of Active Site Residues in the Chitin-Binding Domain ChBDChIA1 from Chitinase A1 of *Bacillus Circulans* Alters Substrate Specificity: Use of a Green Fluorescent Protein Binding Assay. *Arch. Biochem. Biophys.* **2004**, *426*, 286–297. [[CrossRef](#)]
112. Tiede, S.; Cantz, M.; Spranger, J.; Bräulke, T. Missense Mutation in The N-Acetylglucosamine-1-Phosphotransferase Gene (GNPTA) in a Patient with Mucopolysaccharidosis II Induces Changes in the Size and Cellular Distribution of GNPTG. *Hum. Mutat.* **2006**, *27*, 830–831. [[CrossRef](#)]
113. Krumbholz, M.; Koehler, K.; Huebner, A. Cellular Localization of 17 Natural Mutant Variants of ALADIN Protein in Triple A Syndrome—Shedding Light on an Unexpected Splice Mutation. *Biochem. Cell Biol.* **2006**, *84*, 243–249. [[CrossRef](#)]
114. Kiel, C.; Serrano, L. Complexities in Quantitative Systems Analysis of Signaling Networks. *Comput. Syst. Biol.* **2014**, *65*–88.
115. Cabrera, H.S.; Medina, I.C.; Tayo, L.L. In Silico Screening of Inhibitors of P53-MDM2 Protein Complex through Homology Modelling and Molecular Docking. *AIP Conf. Proc.* **2018**, *2045*, 020075.
116. e Zahra, S.N.; Khattak, N.A.; Mir, A. Comparative Modeling and Docking Studies of P16ink4/Cyclin D1/Rb Pathway Genes in Lung Cancer Revealed Functionally Interactive Residue of RB1 and Its Functional Partner E2F1. *Theor. Biol. Med. Model.* **2013**, *10*, 1–9. [[CrossRef](#)] [[PubMed](#)]
117. Hossain, M.S.; Roy, A.S.; Islam, M.S. In Silico Analysis Predicting Effects of Deleterious SNPs of Human RASSF5 Gene on Its Structure and Functions. *Sci. Rep.* **2020**, *10*, 14542. [[CrossRef](#)] [[PubMed](#)]
118. Govindarasu, M.; Ganeshan, S.; Ansari, M.A.; Alomary, M.N.; AlYahya, S.; Alghamdi, S.; Almeahadi, M.; Rajakumar, G.; Thiruvengadam, M.; Vaiyapuri, M. In Silico Modeling and Molecular Docking Insights of Kaempferitrin for Colon Cancer-Related Molecular Targets. *J. Saudi Chem. Soc.* **2021**, *25*, 101319. [[CrossRef](#)]
119. Rosita, A.S.; Begum, T.N. Molecular Docking Analysis of the TNIK Receptor Protein with a Potential Inhibitor from the NPACT Databas. *Bioinformatics* **2020**, *16*, 387–392. [[CrossRef](#)] [[PubMed](#)]
120. Selvaraj, J.; Ponnulakshmi, R.; Abilasha, S.; Nalini, D.; Vijayalakshmi, P.; Vishnupriya, V.; Mohan, S.K. Docking Analysis of Importin-11 Homology Model with the Phyto Compounds towards Colorectal Cancer Treatment. *Bioinformatics* **2020**, *16*, 153–159. [[CrossRef](#)]
121. Edelman, L.B.; Eddy, J.A.; Price, N.D. In Silico Models of Cancer. *Wiley Interdiscip. Rev. Syst. Biol. Med.* **2010**, *2*, 438–459. [[CrossRef](#)]
122. Ang, C.S.; Baker, M.S.; Nice, E.C. Chapter Thirteen—Mass Spectrometry-Based Analysis for the Discovery and Validation of Potential Colorectal Cancer Stool Biomarkers. In *Methods in Enzymology*; Shukla, A.K., Ed.; Academic Press: Cambridge, MA, USA, 2017; Volume 586, pp. 247–274.
123. Atak, A.; Khurana, S.; Gollapalli, K.; Reddy, P.J.; Levy, R.; Ben-Salmon, S.; Hollander, D.; Donyo, M.; Heit, A.; Hotz-Wagenblatt, A.; et al. Quantitative Mass Spectrometry Analysis Reveals a Panel of Nine Proteins as Diagnostic Markers for Colon Adenocarcinomas. *Oncotarget* **2018**, *9*, 13530–13544. [[CrossRef](#)]
124. Liang, S.-L.; Chan, D.W. Enzymes and Related Proteins as Cancer Biomarkers: A Proteomic Approach. *Clin. Chim. Acta* **2007**, *381*, 93–97. [[CrossRef](#)]
125. Wang, J.; Luttrell, J., 4th; Zhang, N.; Khan, S.; Shi, N.; Wang, M.X.; Kang, J.-Q.; Wang, Z.; Xu, D. Exploring Human Diseases and Biological Mechanisms by Protein Structure Prediction and Modeling. *Adv. Exp. Med. Biol.* **2016**, *939*, 39–61. [[PubMed](#)]

126. Ocak, S.; Chaurand, P.; Massion, P.P. Mass Spectrometry-Based Proteomic Profiling of Lung Cancer. *Proc. Am. Thorac. Soc.* **2009**, *6*, 159–170. [[CrossRef](#)] [[PubMed](#)]
127. Kuhlman, B.; Bradley, P. Advances in Protein Structure Prediction and Design. *Nat. Rev. Mol. Cell Biol.* **2019**, *20*, 681–697. [[CrossRef](#)] [[PubMed](#)]
128. Schneider, M.; Belsom, A.; Rappsilber, J. Protein Tertiary Structure by Crosslinking/Mass Spectrometry. *Trends Biochem. Sci.* **2018**, *43*, 157–169. [[CrossRef](#)] [[PubMed](#)]
129. Kelley, L.A.; Mezulis, S.; Yates, C.M.; Wass, M.N.; Sternberg, M.J.E. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [[CrossRef](#)] [[PubMed](#)]
130. Pearson, W.R. An Introduction to Sequence Similarity (“homology”) Searching. *Curr. Protoc. Bioinform.* **2013**, *42*, 3.1.1–3.1.8. [[CrossRef](#)]
131. Webb, B.; Sali, A. Comparative Protein Structure Modeling Using MODELLER. *Curr. Protoc. Bioinform.* **2016**, *54*, 5–6. [[CrossRef](#)]
132. Hazai, E.; Bikádi, Z. Homology Modeling of Breast Cancer Resistance Protein (ABCG2). *J. Struct. Biol.* **2008**, *162*, 63–74. [[CrossRef](#)] [[PubMed](#)]
133. Shehadi, I.A.; Rashdan, H.R.M.; Abdelmonsef, A.H. Homology Modeling and Virtual Screening Studies of Antigen MLLA-42 Protein: Identification of Novel Drug Candidates against Leukemia—An In Silico Approach. *Comput. Math. Methods Med.* **2020**, *2020*, 8196147. [[CrossRef](#)]
134. Chandrasekaran, G.; Hwang, E.C.; Kang, T.W.; Kwon, D.D.; Park, K.; Lee, J.-J.; Lakshmanan, V.-K. Computational Modeling of Complete HOXB13 Protein for Predicting the Functional Effect of SNPs and the Associated Role in Hereditary Prostate Cancer. *Sci. Rep.* **2017**, *7*, 43830. [[CrossRef](#)]
135. Yang, J.; Yan, R.; Roy, A.; Xu, D.; Poisson, J.; Zhang, Y. The I-TASSER Suite: Protein Structure and Function Prediction. *Nat. Methods* **2014**, *12*, 7–8. [[CrossRef](#)] [[PubMed](#)]
136. Roy, A.; Kucukural, A.; Zhang, Y. I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction. *Nat. Protoc.* **2010**, *5*, 725–738. [[CrossRef](#)] [[PubMed](#)]
137. Zhang, Y. I-TASSER Server for Protein 3D Structure Prediction. *BMC Bioinform.* **2008**, *9*, 40. [[CrossRef](#)] [[PubMed](#)]
138. Hardin, C.; Pogorelov, T.V.; Luthey-Schulten, Z. Ab Initio Protein Structure Prediction. *Curr. Opin. Struct. Biol.* **2002**, *12*, 176–181. [[CrossRef](#)] [[PubMed](#)]
139. Ovchinnikov, S.; Park, H.; Kim, D.E.; DiMaio, F.; Baker, D. Protein Structure Prediction Using Rosetta in CASP12. *Proteins* **2018**, *86* (Suppl. 1), 113–121. [[CrossRef](#)] [[PubMed](#)]
140. Xu, D.; Zhang, Y. Ab Initio Protein Structure Assembly Using Continuous Structure Fragments and Optimized Knowledge-Based Force Field. *Proteins* **2012**, *80*, 1715–1735. [[CrossRef](#)]
141. Mortuza, S.M.; Zheng, W.; Zhang, C.; Li, Y.; Pearce, R.; Zhang, Y. Improving Fragment-Based Ab Initio Protein Structure Assembly Using Low-Accuracy Contact-Map Predictions. *Nat. Commun.* **2021**, *12*, 5011. [[CrossRef](#)]
142. Adhikari, B.; Cheng, J. CONFOLD2: Improved Contact-Driven Ab Initio Protein Structure Modeling. *BMC Bioinform.* **2018**, *19*, 22. [[CrossRef](#)]
143. Lee, J.; Wu, S.; Zhang, Y. Ab Initio Protein Structure Prediction. In *From Protein Structure to Function with Bioinformatics*; Springer: Dordrecht, The Netherlands, 2008; ISBN 9781402090578.
144. Rost, B.; Schneider, R.; Sander, C. Protein Fold Recognition by Prediction-Based Threading. *J. Mol. Biol.* **1997**, *270*, 471–480. [[CrossRef](#)]
145. Zhang, Y. Progress and Challenges in Protein Structure Prediction. *Curr. Opin. Struct. Biol.* **2008**, *18*, 342–348. [[CrossRef](#)]
146. Warnecke, A.; Sandalova, T.; Achour, A.; Harris, R.A. PyTMs: A Useful PyMOL Plugin for Modeling Common Post-Translational Modifications. *BMC Bioinform.* **2014**, *15*, 370. [[CrossRef](#)] [[PubMed](#)]
147. Berney, C.R.; Fisher, R.J.; Yang, J.; Russell, P.J.; Crowe, P.J. Protein Markers in Colorectal Cancer: Predictors of Liver Metastasis. *Ann. Surg.* **1999**, *230*, 179–184. [[CrossRef](#)] [[PubMed](#)]
148. Pan, R.; Zhang, Z.; Jia, H.; Ma, J.; Wu, C.; Xue, P.; Cai, W.; Zhang, X.; Sun, J. CAMTA1-PPP3CA-NFATc4 Multi-Protein Complex Mediates the Resistance of Colorectal Cancer to Oxaliplatin. *Cell Death Discov.* **2022**, *8*, 129. [[CrossRef](#)] [[PubMed](#)]
149. Kozakov, D.; Hall, D.R.; Xia, B.; Porter, K.A.; Padhorney, D.; Yueh, C.; Beglov, D.; Vajda, S. The ClusPro Web Server for Protein-Protein Docking. *Nat. Protoc.* **2017**, *12*, 255–278. [[CrossRef](#)]
150. Zhang, X.-P.; Liu, F.; Wang, W. Regulation of the DNA Damage Response by P53 Cofactors. *Biophys. J.* **2012**, *102*, 2251–2260. [[CrossRef](#)]
151. Schmidt, S.; Denk, S.; Wiegering, A. Targeting Protein Synthesis in Colorectal Cancer. *Cancers* **2020**, *12*, 1298. [[CrossRef](#)]
152. Grosdidier, A.; Zoete, V.; Michielin, O. SwissDock, a Protein-Small Molecule Docking Web Service Based on EADock DSS. *Nucleic Acids Res.* **2011**, *39*, W270–W277. [[CrossRef](#)]
153. Pettersen, E.F.; Goddard, T.D.; Huang, C.C.; Couch, G.S.; Greenblatt, D.M.; Meng, E.C.; Ferrin, T.E. UCSF Chimera?A Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [[CrossRef](#)]
154. Kosinski, J.; von Appen, A.; Ori, A.; Karius, K.; Müller, C.W.; Beck, M. Xlink Analyzer: Software for Analysis and Visualization of Cross-Linking Data in the Context of Three-Dimensional Structures. *J. Struct. Biol.* **2015**, *189*, 177–183. [[CrossRef](#)]
155. Li, X.-L.; Zhou, J.; Chen, Z.-R.; Chng, W.-J. P53 Mutations in Colorectal Cancer—Molecular Pathogenesis and Pharmacological Reactivation. *World J. Gastroenterol.* **2015**, *21*, 84–93. [[CrossRef](#)]

156. Xie, J.; Zheng, J.; Hong, X.; Tong, X.; Liu, X.; Song, Q.; Liu, S.; Liu, S. Protein-DNA Complex Structure Modeling Based on Structural Template. *Biochem. Biophys. Res. Commun.* **2021**, *577*, 152–157. [CrossRef]
157. Zhao, H.; Wang, J.; Zhou, Y.; Yang, Y. Predicting DNA-Binding Proteins and Binding Residues by Complex Structure Prediction and Application to Human Proteome. *PLoS ONE* **2014**, *9*, e96694. [CrossRef] [PubMed]
158. Zheng, J.; Xie, J.; Hong, X.; Liu, S. RMAalign: An RNA Structural Alignment Tool Based on a Novel Scoring Function RMscore. *BMC Genom.* **2019**, *20*, 276. [CrossRef]
159. Lazar, I.M.; Karcini, A.; Ahuja, S.; Estrada-Palma, C. Proteogenomic Analysis of Protein Sequence Alterations in Breast Cancer Cells. *Sci. Rep.* **2019**, *9*, 10381. [CrossRef] [PubMed]
160. Abbasi, M.; Sadeghi-Aliabadi, H.; Hassanzadeh, F.; Amanlou, M. Prediction of Dual Agents as an Activator of Mutant P53 and Inhibitor of Hsp90 by Docking, Molecular Dynamic Simulation and Virtual Screening. *J. Mol. Graph. Model.* **2015**, *61*, 186–195. [CrossRef] [PubMed]
161. Tan, Y.S.; Mhoumadi, Y.; Verma, C.S. Roles of Computational Modelling in Understanding P53 Structure, Biology, and Its Therapeutic Targeting. *J. Mol. Cell Biol.* **2019**, *11*, 306–316. [CrossRef]
162. Guex, N.; Peitsch, M.C. SWISS-MODEL and the Swiss-PdbViewer: An Environment for Comparative Protein Modeling. *Electrophoresis* **1997**, *18*, 2714–2723. [CrossRef] [PubMed]
163. Morris, G.M.; Huey, R.; Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexibility. *J. Comput. Chem.* **2009**, *30*, 2785–2791. [CrossRef] [PubMed]
164. Eberhardt, J.; Santos-Martins, D.; Tillack, A.F.; Forli, S. AutoDock Vina 1.2.0: New Docking Methods, Expanded Force Field, and Python Bindings. *J. Chem. Inf. Model.* **2021**, *61*, 3891–3898. [CrossRef]
165. Trott, O.; Olson, A.J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [CrossRef] [PubMed]
166. Roy, K.; Kar, S.; Das, R.N. Chapter 5—Computational Chemistry. In *Understanding the Basics of QSAR for Applications in Pharmaceutical Sciences and Risk Assessment*; Roy, K., Kar, S., Das, R.N., Eds.; Academic Press: Boston, UK, 2015; ISBN 9780128015056.
167. Sarkar, S.; Gupta, S.; Chakraborty, W.; Senapati, S.; Gachhui, R. Homology Modeling, Molecular Docking and Molecular Dynamics Studies of the Catalytic Domain of Chitin Deacetylase from *Cryptococcus Laurentii* Strain RY1. *Int. J. Biol. Macromol.* **2017**, *104*, 1682–1691. [CrossRef] [PubMed]
168. Hospital, A.; Goñi, J.R.; Orozco, M.; Gelpí, J.L. Molecular Dynamics Simulations: Advances and Applications. *Adv. Appl. Bioinform. Chem.* **2015**, *8*, 37–47. [PubMed]
169. Case, D.A.; Darden, T.A.; Cheatham, T.E.; Simmerling, C.; Wang, J. Others AMBER 12 San Francisco: University of California. 2012. Available online: <https://ambermd.org/doc12/Amber12.pdf> (accessed on 21 August 2022).
170. Nelson, M.T.; Humphrey, W.; Gursoy, A.; Dalke, A.; Kalé, L.V.; Skeel, R.D.; Schulten, K. NAMD: A Parallel, Object-Oriented Molecular Dynamics Program. *Int. J. Supercomput. Appl. High Perform. Comput.* **1996**, *10*, 251–268. [CrossRef]
171. Hess, B.; Kutzner, C.; van der Spoel, D.; Lindahl, E. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *J. Chem. Theory Comput.* **2008**, *4*, 435–447. [CrossRef]
172. Brooks, B.R.; Brooks, C.L., 3rd; Mackerell, A.D., Jr.; Nilsson, L.; Petrella, R.J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; et al. CHARMM: The Biomolecular Simulation Program. *J. Comput. Chem.* **2009**, *30*, 1545–1614. [CrossRef]
173. Weis, D.D. *Hydrogen Exchange Mass Spectrometry of Proteins: Fundamentals, Methods, and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2016; ISBN 9781118616499.
174. Vallat, B.; Webb, B.; Westbrook, J.; Sali, A.; Berman, H.M. Archiving and Disseminating Integrative Structure Models. *J. Biomol. NMR* **2019**, *73*, 385–398. [CrossRef]
175. Dokholyan, N.V. Experimentally-Driven Protein Structure Modeling. *J. Proteom.* **2020**, *220*, 103777. [CrossRef]
176. Huang, L.; So, P.-K.; Yao, Z.-P. Protein Dynamics Revealed by Hydrogen/Deuterium Exchange Mass Spectrometry: Correlation between Experiments and Simulation. *Rapid Commun. Mass Spectrom.* **2019**, *33* (Suppl. 3), 83–89. [CrossRef]
177. Makepeace, K.A.T.; Brodie, N.I.; Popov, K.I.; Gudavicius, G.; Nelson, C.J.; Petrotchenko, E.V.; Dokholyan, N.V.; Borchers, C.H. Ligand-Induced Disorder-to-Order Transitions Characterized by Structural Proteomics and Molecular Dynamics Simulations. *J. Proteom.* **2020**, *211*, 103544. [CrossRef]
178. Frame, N.M.; Kumanan, M.; Wales, T.E.; Bandara, A.; Fändrich, M.; Straub, J.E.; Engen, J.R.; Gursky, O. Structural Basis for Lipid Binding and Function by an Evolutionarily Conserved Protein, Serum Amyloid A. *J. Mol. Biol.* **2020**, *432*, 1978–1995. [CrossRef]
179. Paço, L.; Zarate-Perez, F.; Clouser, A.F.; Atkins, W.M.; Hackett, J.C. Dynamics and Mechanism of Binding of Androstenedione to Membrane-Associated Aromatase. *Biochemistry* **2020**, *59*, 2999–3009. [CrossRef] [PubMed]
180. Pacheco, S.; Widjaja, M.A.; Gomez, J.S.; Crowhurst, K.A.; Abrol, R. The Complex Role of the N-Terminus and Acidic Residues of HdeA as PH-Dependent Switches in Its Chaperone Function. *Biophys. Chem.* **2020**, *264*, 106406. [CrossRef] [PubMed]
181. Devaurs, D.; Antunes, D.A.; Kavraki, L.E. Computational Analysis of Complement Inhibitor Compstatin Using Molecular Dynamics. *J. Mol. Model.* **2020**, *26*, 231. [CrossRef] [PubMed]
182. Jia, R.; Martens, C.; Shekhar, M.; Pant, S.; Pellowe, G.A.; Lau, A.M.; Findlay, H.E.; Harris, N.J.; Tajkhorshid, E.; Booth, P.J.; et al. Hydrogen-Deuterium Exchange Mass Spectrometry Captures Distinct Dynamics upon Substrate and Inhibitor Binding to a Transporter. *Nat. Commun.* **2020**, *11*, 6162. [CrossRef]
183. Redhair, M.; Hackett, J.C.; Pelletier, R.D.; Atkins, W.M. Dynamics and Location of the Allosteric Midazolam Site in Cytochrome P4503A4 in Lipid Nanodiscs. *Biochemistry* **2020**, *59*, 766–779. [CrossRef]

184. Huang, L.; So, P.-K.; Chen, Y.W.; Leung, Y.-C.; Yao, Z.-P. Conformational Dynamics of the Helix 10 Region as an Allosteric Site in Class A  $\beta$ -Lactamase Inhibitory Binding. *J. Am. Chem. Soc.* **2020**, *142*, 13756–13767. [[CrossRef](#)]
185. Lee, S.; Jeong, Y.; Simms, J.; Warner, M.L.; Poyner, D.R.; Chung, K.Y.; Pioszak, A.A. Calcitonin Receptor N-Glycosylation Enhances Peptide Hormone Affinity by Controlling Receptor Dynamics. *FASEB J.* **2020**, *34*, 1.
186. Medina, E.; Villalobos, P.; Hamilton, G.L.; Komives, E.A.; Sanabria, H.; Ramírez-Sarmiento, C.A.; Babul, J. Intrinsically Disordered Regions of the DNA-Binding Domain of Human FoxP1 Facilitate Domain Swapping. *J. Mol. Biol.* **2020**, *432*, 5411–5429. [[CrossRef](#)]
187. Devaurs, D.; Antunes, D.A.; Borysik, A.J. Computational Modeling of Molecular Structures Guided by Hydrogen-Exchange Data. *J. Am. Soc. Mass Spectrom.* **2022**, *33*, 215–237. [[CrossRef](#)]
188. Gygi, S.P.; Rochon, Y.; Franza, B.R.; Aebersold, R. Correlation between Protein and mRNA Abundance in Yeast. *Mol. Cell. Biol.* **1999**, *19*, 1720–1730. [[CrossRef](#)]
189. Vogel, C.; Marcotte, E.M. Insights into the Regulation of Protein Abundance from Proteomic and Transcriptomic Analyses. *Nat. Rev. Genet.* **2012**, *13*, 227–232. [[CrossRef](#)] [[PubMed](#)]
190. Faulkner, S.; Dun, M.D.; Hondermarck, H. Proteogenomics: Emergence and Promise. *Cell. Mol. Life Sci.* **2015**, *72*, 953–957. [[CrossRef](#)] [[PubMed](#)]
191. Rawla, P.; Sunkara, T.; Barsouk, A. Epidemiology of Colorectal Cancer: Incidence, Mortality, Survival, and Risk Factors. *Prz. Gastroenterol.* **2019**, *14*, 89–103. [[CrossRef](#)] [[PubMed](#)]
192. Munteanu, I.; Mastalier, B. Genetics of Colorectal Cancer. *J. Med. Life* **2014**, *7*, 507–511.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.