Structural Changes in Proteins With Post-translational Modifications in Female Oncopathologies

Dmitry Tikhonov
Institute of Applied Mathematics of Russian Academy of Sciences
Liudmila Kulikova
Institute of Applied Mathematics of Russian Academy of Sciences
Arthur T. Kopylov (a.t.kopylov@gmail.com)
V.N. Orekhovich Institute of Biomedical Chemistry
Vladimir Rudnev
Institute of Theoretical and Experimental Biophysics
Alexander Stepanov
V.N. Orekhovich Institute of Biomedical Chemistry
Kristina Malsagova
V.N. Orekhovich Institute of Biomedical Chemistry
Alexander Izotov
V.N. Orekhovich Institute of Biomedical Chemistry
Dmitry Kulikov
Moscow Regional Research and Clinical Institute, Russian Federation
Alexey Zulkarnaev
Moscow Regional Research and Clinical Institute, Russian Federation
Dmitry Enikeev
Sechenov University
Natalia Potoldykova
Sechenov University
Anna L. Kaysheva
V.N. Orekhovich Institute of Biomedical Chemistry

Research Article

Keywords: breast cancer, ovarian cancer, post-translational modifications, protein structural motifs, active environment of amino acid, molecular dynamics, ultrahigh-resolution mass spectrometry

Posted Date: May 17th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-505949/v1

Abstract

Post-translational processing leads to conformational changes in protein structure that modulate molecular functions and change the signature of metabolic transformations and immune responses. Some post-translational modifications (PTMs), such as phosphorylation and acetylation, are strongly related to oncogenic processes and malignancy. This study investigated a PTM pattern in patients with gender-specific ovarian or breast cancer. Proteomic profiling and analysis of cancer-specific PTM patterns were performed using high-resolution UPLC-MS/MS. Structural analysis, topology, and stability of PTMs associated with sex-specific cancers were analyzed using molecular dynamics modeling. We identified highly specific PTMs, of which 12 modified peptides from eight distinct proteins derived from patients with ovarian cancer and 6 peptides of three proteins favored patients from the group with breast cancer. We found that all defined PTMs were localized in the compact and stable structural motifs exposed outside the solvent environment. PTMs increase the solvent-accessible surface area of the modified moiety and its active environment. The observed conformational changes are still inadequate to activate the structural degradation and enhance protein elimination/clearance; however, it is sufficient for the significant modulation of protein activity.

Introduction

Epithelial ovarian cancer accounts for up to 90% of all malignant ovarian neoplasms. Moreover, breast cancer most commonly affects milk ducts and lobules (also known as invasive ductal carcinoma). It can also originate from glandular tissue (invasive lobular carcinoma) and germ cells [1]. Breast cancer is the most prevalent type of tumor in women and is the second most common cancer type in general [2]. Determination of molecular events (signatures) associated with the onset and progression is a major task due to the lack of efficient early diagnostics and limited success in treatment strategy. As a result, high disability and mortality levels are arising[2].

The transformation of normal, regular cells into neoplastic cells is accompanied by various endogenous molecular events, which are generally orchestrated by the comprehensive and dynamic network of post-translational modifications (PTM) [3]. PTMs are thought to play a pivotal role in the maintenance of different biological processes; thus, disruption of PTM crosstalk may initiate a complex of oncogenic events. Several PTM moieties can be represented on the protein surface and create a “PTM code.” The proper “code” is important for the organization of intracellular signaling through interaction with various effectors. Therefore, different strategies and proteomic tools, including affinity-based separation, TiO2 enrichment, and the HDX technique, have been recently developed to decipher distinct PTMs and determine their localization on the protein surface. To date, up to 450 different PTMs have been distinguished and annotated. The most prevalent PTMs are phosphorylation, acetylation, methylation, and ubiquitination [4]. Less known citrullination and SUMOylation can be more favorable due to alterations in $CD4^+$ T cells to cellular stress, immune activation [5], sensing of DNA damage response, and telomere maintenance [6].

The main source of PTM identification is mass spectrometry-based proteomics, supplied by interatomic and transcriptomic integrative approaches. However, mass spectrometry data are highly redundant, and search engines are imperfect, whereas validation of the putatively identified PTMs by immunochemistry is limited. This leads to several misidentifications and significantly obstructs the investigation of the exact role of specific PTMs in oncogenesis. Therefore, only a small fraction of PTMs is well validated and is related to several types of cancer, including glycosylation of COX-2 in colorectal cancer [7], citrullination of fibronectin during renal cancer [8], phosphorylation of PKM2 in thyroid cancer [9], and deSUMOylation mediated by SENP1 in prostate cancer [10].

Dysregulation of the SUMOylation pathway has been reported to be associated with breast cancer via Mel-18 activity, which controls ESR1 expression and governs SENP1 activity [11,12]. Ubiquitination of POU class 5 homeobox 1 protein (OCT4) leads to decreased survival of breast cancer patients[13]. Despite the wide scrutiny of ovarian cancer tumorigenesis, little is known about the interplay of PTMs in this phenotype. According to a meta-analysis, most information is related to histone acetylation and ubiquitination [14]. Other studies have reported the specific role of TRIM71 ubiquitin ligase-mediated regulation of p53, which can retard ovarian cancer development [15]. Irregular N-glycosylation triggers the IRE1α-XBP1 pathway, which causes a stress response of the endoplasmic reticulum to unfolded proteins in T cells and enhances ovarian cancer tumorigenesis and malignancy [16]. In addition, repression of HDAC6, which deacetylates p53 on Lys120, suppresses ovarian tumorigenesis [17].

In this report, we would like to highlight that breast and ovarian cancer-specific molecular events can be considered through the prism of post-translational actions and the induced changes in protein structure stability. Structural biology has provided a good understanding of motif elements determined by the specific folding of secondary structures. Such motifs are typically tightly packed and accommodated by the neighboring or closest segments of the polypeptide chain, helixes, and β-strands and maintain the originality of spatial folding regardless of the homology of proteins where the elements were being observed in [18,19]. Owing to the high stability of the protein globule, helical pairs are the most attractive for structural analysis. Therefore, in silico structural molecular analysis and modeling of PTM moieties can be accomplished on isolated sustained helical pairs instead of ordered protein molecules, which improves calculations without losing information.

In this study, we performed a structural analysis of the geometry and topology of motives with PTM moieties accommodated in targeted peptides derived from cancer-associated proteins. We found that the cancer-specific serological proteome significantly differed from those of the aligned phenotype of healthy volunteers. We identified 74 and 25 proteins specific for ovarian cancer and breast cancer groups, respectively, and 50 proteins were shared between assayed groups but omitted in the control group. The study determined that 65 and 88 proteins in ovarian cancer and breast cancer groups, respectively, significantly differed in their abundance compared to those in the control group. We identified cancer-specific PTMs and

conducted a structural analysis in support of AMBER software. Among the identified PTMs, 12 modified peptides were localized in eight proteins of ovarian cancer patients, whereas six modified peptides were localized in three proteins of breast cancer patients.

We suppose, that most oncopathologies can be induced by uncharacterized PTMs even before sensitive alterations in gene expression. Furthermore, based on the presented data, a particular PTMs pattern can be featured for a certain cancer phenotype.

Results

Proteomic analysis of plasma samples by ovarian and breast cancer

Identification of serological protein signatures that define ovarian cancer and breast cancer phenotypes was achieved by categorizing qualitative and quantitative features. The mean number of plasma protein identifications for the studied groups is shown in **Figure 1A**.

Symmetry comparative analysis showed a wide cluster of proteins (n=147) shared between all studied groups (including the control group), a smaller collection for both types of cancer (n=50; breast cancer and ovarian cancer), and two clusters distinct for patients with either ovarian (n=25) or breast (n=74) cancer (**Figure 1A**).

A comparative analysis of the protein content in the blood samples of the study participants revealed that 65 and 88 proteins ($p<0,05$) in the ovarian cancer and breast cancer groups, respectively, significantly differed in their abundance compared to those in the control group (**Supplementary Table S1**). Furthermore, the abundancy-based principal component analysis (PCA), taken for n=147 proteins, returned satisfactory segregation of all three studied groups by the projection of the first two principal compounds with PC1 explaining 11.87% variance, and PC2 of 6.46% (**Figure 1B**). Differentially expressed proteins and their fold-changes in patients with oncopathologies compared to the control group are listed in **Supplementary Table S1**.

Functional annotation

Among the mutual fractions of proteins detected and identified in all studied groups (n=147), most were attributed as extracellular and involved in the activation of the complement system, including its terminal phase (n=40, HSA-166665, HSA-174577), platelet degranulation (n=27, HSA-114608), innate immune system (n=49, HSA-168249), hemostasis (n=37, HSA-109582), and regulation of post-translational phosphorylation (n=18, HSA-8957275). Proteins shared between the cancerous groups are also characterized as extracellular and regulate the immune response (n=10, HSA-168256), metabolism of proteins (n=10, HSA-392499), innate immune system (n=8, HSA-168249), and hemostasis (n=10, HSA-109582). Cancer-specific proteins are implicated in vesicle-mediated transport (n=11, HSA-5653656), activation of the immune system (n=35, HSA-168249, HSA-168256), and membrane trafficking (n=8, HSA-199991).

Post-translational modifications of proteins

Modification moieties were analyzed for the five most prevalent PTMs, including phosphorylation of serine (pS), threonine (pT), tyrosine (pY), N-terminal acetylation of lysine, and ubiquitination of lysine (**Supplementary Materials Figure S1**). After validation, the resulting PTMs were populated and extracted into a separate list of PTMs associated with and found explicitly in the considered oncopathologies (**Table 1**). We observed PTMs among 12 peptides derived from eight proteins in patients with ovarian cancer and six peptides, carrying modified moieties, from three proteins found in breast cancer patients.

There were also a few overlapping PTMs revealed in albumin and Ig heavy chain V-III region CAM. Sequence coverage of peptides carrying PTMs consisted of 12% to 79%, and at least three distinct and proteotypic peptides were identified with a high confidence rate for each protein. A high scoring function for the identified peptides was provided by satisfying the number of b- and y-type ions with marginal measurement error (**Supplementary Materials Table S1**).

Structural analysis of proteins carrying PTMs

Several well-known structural motifs are characterized by specific spatial folding and geometry in structural biology. Such motifs are typically tightly molded and shaped by the adjustment segments of polypeptide chains, helices, and β-strands [18,19].

In this study, we investigated motifs molded by secondary structures and carrying different PTMs. Such motifs are represented as helical pairs designed by two consequent helices bridged by irregular linkers of different lengths and conformations [18]. Recently, we demonstrated that modifying moieties are frequently observed in α-α-corners, α-α-hairpins, and L- and V-structures. The geometry rules and topological classification of the helical pairs are shown in **Supplementary Materials Figure S2**.

Based on the rules for recognizing helical pairs (**Supplementary Materials Figure S1**), we selected protein structures from the Protein Data Bank (PDB) matching the targeted peptides that carry empirically identified PTMs [18]. Each selection sampled from 1 to 172 protein structures fit the recognition and specific polypeptide chain rules. (**Table 2**, column “PDB structures”).

A comparative analysis of the geometric features of the assayed protein structures was performed according to the following criteria:

1. Similarity of the defined helical pairs among selected polypeptide chains.
2. Type of spatial conformation for motifs shaped from selected secondary structures.
3. Solvent accessibility to the local folded motif
4. Influence of modified amino acid residues on protein structure stability.

Molecular dynamics experiments determined the stability of ovarian and breast cancer-specific protein structures bearing PTM moieties.

Table 2 summarizes the structural analysis data of the selected proteins with PTM. The solvent-accessible surface area of the intact amino acid residue exposed to the solvent was always less than that of the modified residue (**Table 2**, column "Active environment"). Thus, the modifiable amino acid residues are continuously exposed to the solvent, and the modification process is associated with the enlargement of the solvent-accessible surface area. The surface area of the modified residue exceeded that of the intact (unmodified) amino acid residue by 50%.

The adjacent environment (neighboring amino acid residues) was evaluated as an alteration of the total surface area accessible to the surrounding solvent (data are represented in column "Active environment," **Table 2**). The mean surface area of the modified environment frequently exceeds that of the intact (unmodified) environment. However, the difference is not as explicit as that for the separate modified amino acid residue. In some cases, the PTM moiety can significantly increase the accessible surface area. However, the total surface area of the exposed active environment is equal to or even less than that of the unmodified polypeptide chain. Hence, it can be assumed that the neighboring amino acid residues eliminate the increased surface area caused by the mounted PTM.

The results of the solvent-accessible surface area alterations observed for the modified amino acid residue and its active environment are shown in **Figure 2**. The active environment is defined as the amino acid surrounding the modified residue and is capable of changing the solvent-accessible surface area of a certain motif. We determined the active environment within each motif and identified the number of constituent amino acids and their spatial coordinates within the affected polypeptide chain.

Excluding four cases, the total solvent-accessible surface area of the modified residue and its active environment increased, in contrast to that of the intact residue. The notable instances (peptides EQL-***ac(K)***-AVMDDFAAFVEK (ALBU), ***ac(K)***-VPQVSTPTLVEVSR (ALBU), RHPYF-***p(Y)***-APELLFFAK (ALBU), and YF-***ac(K)***-PGMPFDLMVFVTNPDGSPAYR (CO3)) are highlighted by the larger solvent-accessible surface area of the modified residue compared to the intact residue. However, the solvent-accessible surface area of its active environment is smaller than that of the unchanged form of residue.

The distribution of the protein population bearing the identified PTMs (**Table 2**) depends on the solvent-accessible area for the particular amino acid residue and its active environment before and after mounting the PTM. The modified moiety increased exhibition of the surface area of the modified amino acid residue to solvent and its active environment (**Figure 3**). However, the summed solvent-accessible surface areas do not always exceed those before the modification and are characterized by generous scattering.

Molecular dynamics simulation of protein molecules containing PTM associated with the development of ovarian and breast cancer

In this study, we analyzed the similarity, stability, and influence of the PTM moiety on the degree of the solvent-accessibility local area for the modified polypeptide chain of protein molecules. The results of molecular dynamics showed that the geometry and topology features of motifs before and after mounting the PTM are kept within acceptable ranges over the time of the simulation experiment. Particular attention has been drawn to motifs with strongly interacting and axially intercepting helices, typically α-α-corners. The between-distance is negligible (contacting helices), and the area and projection perimeter are distinct from the null value. We did not scrutinize L- and V-structures because helices do not intercept ($d \neq r$, $d < r$). Moreover, the area and perimeter for such motifs are close to a null value. Simultaneously, the contribution of motifs comprising more than two helices allowed the extension of the range of observable structures. The main criteria for inclusion in the MD simulation was the actual contact between helices; however, the linker length and structure did not act as a significant input.

By employing these selection criteria, we expanded the number of motifs that can be sampled for the analysis. In the previous study, only helix pairs comprising two sequential helices were utilized. In this study, we investigated closely contacted helices (from 1 to 30 helices) with intersecting axes. Simultaneously, the connection may vary in length and conformation and may include secondary structures (**Table 2**, column "Motifs"). We detected five PTM moieties in albumin (ALBU), three in IGHA1 and CO3, four in APOA2 and A1AT, and eight in APOA1 (**Supplementary Materials Table S2** and **Figure 4**). The calculated coordinates for the detected PTMs, α- and θ-angles, minimal (r), and inter-planar (d) distances between helices, areas, perimeters, and standard deviations (sd, sr, sα, sθ, sS, and sP) for the PTM-containing motifs are presented in **Supplementary Materials Table S2**.

The molecular dynamics results showed that during the set time (0.5 ns), the geometrical features of the studied motifs (intact and modified) were within the acceptable ranges. However, we have identified conformational changes that distinguish modified from intact forms. (**Figure 4**). The molecular dynamics experiment has demonstrated that the stability of structural blocks directly depends on the strength of interactions between helices (as indicated by the values of d, r, S, P, etc.). It has also been demonstrated that the PTM moiety (mounting of modification function) does not induce complete rupture between interacting helices and, consequently, does not disrupt the motif with accommodated PTM function.

Molecular dynamics simulation revealed that mounting of the modifying moiety did not lead to the complete rupture of intramolecular bonds and motif disruption. However, some protein molecules showed an ambiguous attitude (**Figure 4** and **Supplemental Materials**). The two distinct motifs containing polypeptide chain EQL-ac(K)-AVMDDFAAFVEK with acetylated lysine (bolded) should be noted. The topology of the first modified helix was 396-410 (400-414 in the intact molecule), and the location of the second modified helix was 538-555 (542-559 in the intact molecule). The motif comprises seven helices, although there are five helices between the considered contacting helices and the intersecting axes. At the beginning of the molecular simulation, the initial geometry met the following features: the minimal and inter-planar distances were equal and made 11.7Å, while the area and polygonal perimeter were 110.2Å^2 and 43Å, respectively, and the axial angles were α=55^o and θ= -57^o. Following-up the molecular simulation, the distances (r and d) increased, indicating misalignment between helices. The axial angles, area, and perimeter also decreased. The inspected predisposition holds for both modified and unmodified motifs; however, the modified structure inclines to a more remarkable alteration in conformational geometry (inter-planar and minimal distances were 12.9Å and 15.6Å; area and perimeter were S=22.9Å^2 and P=12.1Å, correspondingly; torsion angle θ= -109^o).

A counter situation was observed with the motif with two consequent helices localized at positions 514-530 and 538-555 (positions 518-534 and 5420559 in the unmodified molecule). Initially, the motif was attributed to the following customized settings: d=r=10.6Å, S=194.2Å^2, P=59.6Å, α=41^o, θ= -35^o. The geometry changed insignificantly during the molecular simulation as for the intact molecule and made d= r=11.8Å, S=171.8Å^2, P=55.5Å, α=46^o, θ= -40^o for the intact molecule; and d=r=10.6Å, S=169.8Å^2, P=55.4Å, α=50^o, θ= -41^o for the molecule with a mounted PTM moiety, which indicated higher stability than the previously examined helical pair.

Discussion

Since PTMs change the physical and chemical properties of proteins, they regulate catalytic activity and molecular function; however, the newly acquired properties and functions of the non-specifically modified proteins are unknown in addition to whether the essential molecular property is completely lost. Further, such PTMs can accelerate DNA mutations through inappropriate regulation of gene expression and misleading signaling pathways. Such non-specifically modified proteins are beneficial for cancer phenotypes and promote cancer onset, invasion, and progression. Therefore, we suggest that finding such PTMs and modified proteins can be a central point for deeper insights into metabolic processes in cancer cells.

In this study, we attempted to specify PTM-induced conformational changes in proteins found in patients with ovarian or breast cancers. We suggest that such PTMs, whether mounted at an unspecific topology or cause conformation instability, can change and modulate protein activities and functional properties, contributing meaningfully to oncogenesis. Evidence exists regarding the cautioning role for such PTMs in biological processes, which provides an excellent opportunity for the utility of cancer-specific PTM patterns.

Due to cancer is typically accompanied by immune response and acute inflammation, alpha-1-antitrypsin (A1AT) is one of the most investigative markers of dire reactions on tumorigenesis. It is and essential tissues and cells protecting element against neutrophil elastase and proteinase-3 activity increased during inflammation, and participates in various metabolic activities through binding to different ligands (e.g., cytokines, lipoproteins, lipids, plasmin, trypsin) with moderate to high reversible affinity [20]. Because of its vigorous conformational flexibility, A1AT can take on different forms (polymeric, oxidized, cleft) adapted for specific biological processes and ligand types [21,22].

Reports indicate that PTM of A1AT induces conformational changes in A1AT, which modulates structural sustainability and manages molecular catalytic activities [20]. Such PTMs enhance protein stability and sterically protect the molecule against proteolytic activity and conformation-induced aggregation and degradation, thus increasing the protein half-life [23–25]. The known A1AT PTMs play a central role in the realization of its immunomodulatory and catalytic activities. Glycosylation of K^{342} and V^{75} is essential for protein polymerization [26,27]. Recent studies showed that different glycosylation patterns are specifically characterized for patients with non-small-cell lung cancer and lung adenocarcinoma and may serve as a promising marker for non-invasive differentiation and early diagnostics [28]. Unspecific bi-antennary di-sialylated glycosylation and increased fucosylation of A1AT has been demonstrated in patients with aggressive form of ovarian and breast cancers and assumingly may contribute in rapid malignancy due to partial inactivation of A1AT activity [29]. A high vulnerability of A1AT to modification has been detected in non-cancer patients with rheumatoid arthritis. Non-specific carbamoylation of K^{359} was screened in both seropositive and seronegative rheumatic patients and suggestively enhances autoimmune response since the same amino acid residue can be targeted for citrullination [29]. In contrast, oxidation of M^{351} and M^{358} is responsible for binding A1AT with the inhibitor when the protein manifests protease activity during inflammation [30], and S-nitrosylation of C^{232} plays a vital role in the regulation of apoptosis [31].

However, we discovered two novel PTMs (ac-K^{25} and ac-K^{125}) that are distant from known modifiable sites (**Table 3**). The propensity for PTM-induced conformational polymorphism emphasizes a comprehensive pattern of A1AT molecular functions. Thus, the two newly discovered PTMs may possess unknown functions and have a high potency in predicting therapy response. To date, there are no appropriate methods that utilize a sign of A1AT modification and types of recognized PTMs for the evaluation of disease progression or management.

Albumin is one of the major serological proteins and is essential for the transportation of ligands with various properties (e.g., hormones, lipids, xenobiotics) and maintenance of osmotic blood pressure [43]. Albumin is highly attractive as a transporting tool in pharmaceutics because of its ability to bind with and transfer a wide variety of functional molecules, including exogenous origin. There is evidence of many albumin modifications with vague or explicit responsibilities, but they change its affinity. Glycation of C^{34} is the most scrutinized and known type of albumin PTM, which induces

conformational change and favors the binding capacity with various ligands, including warfarin, tolazamide, acetohexamide, and tolbutamide [44,45]. In contrast, S-nitrosylation of C^{34} is responsible for transporting anionic organic compounds and heavy cations and negatively regulates affinity to fatty acids [35,46]. The role of another type of albumin modification, cysteination of C^{34}, has not yet been determined; however, it may diminish the affinity binding with bilirubin, tryptophane, warfarin, and diazepam [34].

Previously, we identified several types of albumin PTMs (acetylation and phosphorylation) in patients with colorectal cancer [18]. We hypothesized that albumin phosphorylation at the warfarin-binding site (C^{34}) might be regulated by impairments of the cellular signaling network and contribute meaningfully to tumor growth and progression [18]. We suggest that the presently discovered unusual PTMs are associated with the regulation and dealing with the affinity to low-molecular-weight compounds because the indicated PTMs are localized at structural domain II (topological position 196-383) and domain III (topological position 384-585) close to the fatty acids (S^{342} and R^{348}) and drugs (R^{348}-E^{450}) binding sites [47].

Apolipoproteins (APO) bind lipids and act as ligands for cell surface receptors and co-factors for some enzymes [48]. APOA2 is the major compound constituting high-density lipoprotein (HDL) particles and is involved in vesicle remodeling through direct interaction with APOA1. However, the possible role of the modified or intact form of APOA2 in oncogenesis is still unclear, since its expression level is frequently controversial in different cancer types (**Table 3**).

The complement system is believed to be part of the innate immune response. It can be initiated in three different pathways (classical, alternative, and lectin), distinguished by the manner of continual activation. While the classical pathway is triggered after binding with either circulating or surface-immobilized immune-complexes, lectin and alternative pathways are primed by pathogen-associated molecular patterns, which are well-recognized microbial antigens or their conserved motifs [49]. Regardless of the activation pathway type, the complement starts by fission of the C3-factor consisting of 13 structural domains, which following by several consequent conformational transformations essential to exhibit multiple interaction sites with various immune-stimulating effectors [50]. The proteolytic fission of C3 produces a C3b-factor that induces conformational relocations to uncover ligand-binding sites and arrange specific thioester segments needed for covalent binding with the target antigen surface. Due to tremendous and supreme influence on angiogenesis, positive regulation of VEGF, cell migration, and extracellular matrix reorganization, the C3 is frequently mentioned beyond the immune system and in the context of malignancy origination and oncogenic properties [38]. The post-translational processing of the C3 factor primes the cleavage of the tetra-arginine linker and generates α-chain (110 kDa) and β-chain (75 kDa) adjusted by a disulfide bond and containing well-recognized glycosylation sites (N^{917} and N^{63} in the α- and β-chains, respectively) [39]. Recently, the phosphorylation of pT^{1009} has been detected as a new yet uncharacterized type of PTM, apparently produced after the catalytic processing of the cleaved C3-factor by kinases [40]. The presence of pT^{1009} was confirmed in patients with breast or ovarian cancer, but was absent in the control group (**Table 3**). The C3-factor can be phosphorylated at numerous different sites, which expectedly affects its molecular properties and controls the functional activity; however, the consequences of such processing are inexact due to insufficient attention to the role of C3 phosphorylation in cancer pathophysiology. Based on the obtained data, we suggest the high oncogenic potency of this PTM moiety because of the phosphorylation of pT^{1009}. Increased fucosylation at N^{85} and sialylation at N^{939} patterns provides opportunity in distinguishing of patient with colorectal cancer from adenoma, when changes in protein level are not sufficient [41]. Specific glycoforms pattern has been proposed as utility in early diagnostic tool of hepatocellular carcinoma, despite the cancer-specific role of the detected glycosylation sited did not examined [51]. Global profiling of human serum revealed several new proposed glycans in patient with liver cancer [52], among which the glycan S1H5N was the most amenable to distinguish cancer patients from healthy donors and may improve the level of diagnostics. As has been shown by molecular dynamics simulation, the observed in our study cancer-specific modification (Ac-K^{283}) induces a drastic conformational change. It keeps the C3-factor stable, which probably reflects the acquisition of new activities and functional properties.

Changes in iron status was found in a prevalent of patients with advanced cancer. It is suggested, the cancer-related anemia can be caused by chronic inflammation or as an effect of treatment of such patients [53]. In this respect, changes in the expression of serotransferrin and its receptors might be targeted for the anti-cancer therapy. It has been discussed, that cancer patients are characterized by the distinct patterns of TRFE glycosylation, which may influence the binding capacity and restrict interaction with TRFE receptors [54]. High glycosylation levels at N^{432} and N^{630} has been shown specifically in patients with the advanced pancreatic cancer and attributed to acute inflammation [42] and, thus, inversely promoting expression of HIF1-α expression [55]. Although, the number of versatile glycosylation sites in TRFE for cancer phenotypes is rigorously investigated, their roles remain mostly suggestive and outcome from the iron transporting function of the protein. In this study we identified two new PTM sites at Ac-K^{331} and Ac-K^{453}. It has been examined that despite newly attached PTMs, the stability of TRFE is sufficient, however the Ac-K^{453} is the closest site to iron-binding Y^{445} residue and the presence of acetyl-moiety may prevent tight interaction of the second iron ion, hence, limiting binding capacity of TRFE.

Conclusions

In this study, we reported PTMs found at non-specific localizations on molecular surfaces that have never been annotated previously and not detected in normal plasma or regular cells/tissues; therefore, we can suggest that they are atypical. Assumingly, the origin of such PTMs is featured for cancer phenotype and can progress tumorigenesis. However, the majority of protein molecules maintain stability after accommodation of non-specifically localized PTM moiety, as has been shown by molecular dynamics experiment.

We suggest a new approach for analyzing conformational changes and structural re-ordering caused by PTM mounting. The analysis targets a small part of the protein molecule (motif) and the active environment organized by adjacent amino acid residues. It includes secondary structural elements that are most affected after PTM of the target residue. We determined the main structural features of the studied motifs, including spatial coordinates,

distances between helices, torsion angle, length, area, and polygonal perimeter of helices within the accounted motifs. It was demonstrated that the modifying moiety was always exhibited outside the protein globule in the solvent environment. Following the molecular dynamic simulation analysis, we calculated the protein distribution depending on the accommodated PTM type, the end-up solvent-accessible area before and after mounting the PTM function, and the active environment that also mutates exhibition to the surrounding solvent. It has been established that the solvent-accessible area of the modified amino acid residue always exceeds that of the intact residue. However, the total accessible area of the active environment, combined with the modified residue, can be equal or marginally smaller than that of the incorporated intact residue. The molecular dynamics simulation showed that the accommodated PTM moiety did not dramatically change the molecular stability and did not approximate the molten globule state. Despite the high structural stability, the discovered molecular conformations might be irreversible and can contribute significantly to the supervision of protein functional activity. The elaborated approach can provide an excellent opportunity to evaluate the influence of PTM-caused functional flexibility on the onset and growth of tumors and improve the relevance of the currently under-evaluated PTMs.

Methods

Demography and Ethical Consideration

The study population comprised of the group of patients with stage II–III breast cancer (*n*=24, aged 48±11 years old) and the group of patients with ovarian cancer at stages II-III (*n*=53, aged 52±12 years old), who had been inpatients at the M.F. Vladimirsky Research Clinics (Moscow). The control group comprised of 30 healthy volunteers (aged 36±11 years old) with no previous oncological disease history. All groups under consideration were sex-aligned (only women participated in the study). Patients and healthy volunteers provided written Informed consent to participate in the study, and the study was approved by the local ethical committee of M.F. Vladimirsky Research Clinics (protocol no. 18 of December 24, 2020) and Sechenov University (protocol no 10-19 July 17, 2019) according to the WMA Declaration of Helsinki for ethical principles for medical research involving human subjects.

Sample Preparation for MS Analysis

Following overnight fasting, peripheral blood samples (up to 5 mL) were collected in pre-chilled EDTA-2K$^+$ vacuum tubes between 9 and 11 a.m. The collected blood samples were centrifuged for 10 min at 1,500 × *g* and t=4^{o}C. The obtained plasma (supernatant) fraction was carefully transferred into clean cryotubes with a nominal volume of 2 mL.

Blood plasma (40 µL) was mixed with a denaturation solution (0.1% deoxycholic acid sodium salt, 6% acetonitrile, and 75 mM triethylammonium bicarbonate, pH 8.5) to 500 µL.

Digestion was performed with trypsin (200 ng/µL supplemented in 30 mM acetic acid) in two stages: at the first stage, trypsin was added at a ratio of 1:50 (*w/w*) and the reaction was incubated for 3 h at 37°C. In the second stage, the enzyme was added at a ratio of 1:100 (*w/w*), and the reaction was incubated at 37°C for 12 h.

Mass Spectrometry Protein Registration

The analysis was conducted on a high-resolution Q Exactive-HF mass spectrometer (Thermo Scientific, Waltham, MA, USA) with an installed introduced nano-spray ionization (NSI) ionization source (Thermo Scientific). The selection of mass spectrometry parameters for data acquisition was the requirements of the Human Proteome Organization (HUPO Guidelines, bullet point 9, version 3.0.0, released October 15, 2019) for the minimal length of the detected peptide for consideration and justification of PE1 proteins (according to the Uniprot KB Classification).

Data acquisition was performed in a positive ionization mode in the range of 420–1250 *m/z* for precursor ions (with resolution *R* = 60 K) and in a range with the first recorded mass of 110 *m/z* for fragment ions (with resolution of *R* = 15 K). Precursor ions were accumulated for a maximum integration time of 15 ms, and fragment ions were accumulated for a maximum integration time of 85 ms. Top 20 precursor ions with a charge state between *z* = 2 + and *z* = 4+ were collected in the ion trap and pushed to the collision cell for fragmentation in high-energy collision dissociation mode. The activation energy was normalized at 27% for *m/z* = 524, *z* = 2+, and ramped within ±20% of the installed value.

Analytical separation was performed using an Ultimate 3000 RSLC Nano UPLC system (Thermo Scientific). Samples were quantitatively (2 µg) loaded onto the enrichment column Acclaim Pepmap® (5 × 0.3 mm, 300 Å pore size, 5 µm particle size) and washed at a flow rate of 20 µL/min for 4 min using a loading solvent (2.5% acetonitrile, 0.1% formic acid, and 0.03% acetic acid). Following the loading stage, peptides were separated on an Acclaim Pepmap® analytical column (75 µm × 150 mm, 1.8 µm particle size, 60 Å pore size) in a linear gradient of mobile phases A (water with 0.1% formic acid and 0.03% acetic acid) and B (acetonitrile with 0.1% formic acid and 0.03% acetic acid) at a flow rate of 0.3 µL/min. The following elution scheme was applied: the gradient started at 2.5% of B for 3 min and raised to 12% of B for the next 15 min, then to 37% of B for the next 27 min, and to 50% for the next 3 min. The gradient was rapidly increased to 90% of B for 2 min and was maintained for 8 min at a flow rate of 0.45 µL/min. Enrichment and analytical columns were equilibrated in the initial gradient conditions for the next 13 min at a flow rate of 0.3 µL/min before the following sample run. Mass spectrometric measurements were performed using the equipment of “Human Proteome” Core Facility (IBMC, Moscow, Russia).

Protein Identification and Criteria Selection for Post-translational Modifications

Adapted peak lists were searched using the OMSSA (version 2.1.9, Proteomics Resource, Seattle, WA, USA) search engine against a concatenated target/decoy protein sequence database UniProtKB (88703 (target) sequences with a restricted taxonomy (Homo sapiens). The decoy sequences were populated by the reverse sequence algorithm of the SearchGUI engine (release 3.1.16, Compomics, Gent-Zwijnaarde, Belgium).

Peptides were parsed with mass tolerance of 10.0 ppm for the MS1 (precursor) level and with a tolerance of 0.01 Da for the MS2 (fragment ions) level. Trypsin was set as a specific protease, and a maximum of two missed (internal) cleavages were allowed. Modifications of acetyl (K), phospho (S), phospho (T), phospho (Y), and Gly-Gly (K) were selected as flexible. Peptides and proteins were identified using PeptideShaker version 1.16.11 (Compomics, Gent-Zwijnaarde, Belgium) and validated at a 1.0% false discovery rate estimated as a decoy hit distribution.

To eliminate probable false positive results due to the concatenated search of several PTMs, we curated only those results that fit the following requirements: (a) at least 98% confidence for peptide identification, (b) at least 80% of peptide sequence coverage by fragmentation spectra, and (c) at least 10 units of D-score for PTM probability. Furthermore, the extracted data were manually curated to avoid possible false identifications. During this step, we removed those spectra that do not contain proper y/b fragment ion pairs designation that mark and locate the exact amino acid residue carrying PTM.

To consider the detectable PTM moieties as relevant to cancer phenotype and for the structural and molecular dynamic analysis, they should meet the following criteria: (a) the total set of PTM moieties must be detectable and identified in at least 50% of each cancer phenotype; (b) each PTM moiety should be identified in at least 20% of PTM-carrying subjects.

Analysis of Post-translational Protein Modifications

In the present study, proteins containing tryptic peptides of a certain type of modification were selected from the PDB [56]. The selected proteins belong to the class of alpha-helical and globular proteins. For each peptide, a sample was created to determine the conformational template. Furthermore, all motifs containing the tryptic peptides were selected from the database. Our previously developed method was used to recognize and select structural motifs [57–59].

The secondary structure of proteins was determined using the Kabsch and Sander DSSP methods [60]. Using the same program, the available contact surfaces with the solvent were determined. Definitions of important structural motif characteristics have been described in previous studies [57,59,61]. Visual analysis of the structures was performed using the RasMol molecular graphics program [62].

In total, the DSSP program distinguishes three types of helices: α-helix, π-helix, and helix 3/10. The DSSP program also solves the problem of determining the beginning and end of the helix. A candidate for the desired structure (helical pair) is a protein region that contains two helices and a region of a protein thread between the helices, which is called connection. For each helix of the structure, the axis of the cylinder on which this spiral was wound was determined using the least square method. The axis of the cylinder will be found more precisely when the closer the helix is to the ideal. The quality of the axis assessment was characterized by the value of the standard deviation. We selected those helices (and, accordingly, structures) for which the accuracy of the axis estimation satisfies a predetermined criterion. The two axes of the helices define the spatial structure completely. It is known that two parallel planes can be drawn through two non-intersecting straight lines in space such that the first axis belongs to the first plane and the second to the second plane. An axis lying in one plane can be projected onto another plane. Thus, the spatial structure is fully described by the distance between the parallel planes and the projections of the axes of the helices onto the plane.

To establish the stability of structural motifs without and considering the modification, a numerical experiment of molecular dynamics was planned and carried out. Molecular dynamic simulation experiments were performed using the AMBER software (version 11) [63]. To calculate the molecular trajectories for the selected proteins, the experiments performed with no consideration of the water environment and at a field strength of AMBER ff03 at 300 K [64]. Therefore, the complete energy of the considering system was minimized at the fixed atomic coordinates, which deployed the condition to organize the atomic interaction order. The molecular system was further heated up to the selected temperature (300 K), and molecular trajectories were recorded during 0.5 nanoseconds every 0.005 nanoseconds. The resulting molecular trajectories were visualized using the VMD (version 1.9.1) software [65]. The free energy for the binding profile of the molecular complexes was estimated by Born's method [66], and the distance between atoms was calculated using CPPTRAJ software (a part of AMBER (version 11) software). The spatial geometry of polypeptide chains was characterized according to the categorized rules of Kabsch and Sander [60].

The following geometric features [18] of the assayed motifs with mounted PTMs were recorded every 0.005 ns: minimal (**r**) and inter-planar (**d**) distances between helices, α- and torsion (**θ**) angles between axes of helices, area (**S**), and perimeter (**P**) of polygons of the helices projections intersection. Means and standard deviations were recorded and utilized to defined the stability of protein molecule relatively to the initial intact condition. The geometric features were recorded for motifs that met the following criteria: (**1**) the presence of a PTM moiety, (**2**) the PTM must be localized on or close to the helix or at least on the polypeptide chain lining the affected helices; (**3**) helices should be in tight contact and the minimal and inter-planar distances must be equal ($r = d \leq 16$Å), but the area and the perimeter should have a non-null and not close to zero [57,59–61].

Statistical analysis

Mass spectrometric intensities below the instrumental-adjusted threshold were converted by OMSSA to null values and imputated with the minimal threshold value of 10^5 counts. Proteins were quantified with Intensity values representing normalized summed peptide intensities correlating with protein abundances.

Principal component analysis was performed for the set of proteins shared between the groups of study. To assess the similarity of identified proteomes an upset plot was generated using the UpSetR function in UpSetR [67]. Proteins attributed to certain pathology groups and featured by an arbitrary fold change cut-off of >2 and significance p-values of <0.05 (Wilcoxon test) were considered as meaningfully different in quantitative property. Statistical analyses were performed using an in-house script in R.

Significantly altered proteins were submitted for functional and pathways annotation analysis at a q-value threshold less than q <0.01 using PANTHER Overrepresentation Test of Gene Ontology toolset [68], and Bonferroni correction for multiple testing has been applied. The enriched terms were refined with similarity coefficient of >0.7.

Declarations

Acknowledgments

The work was done in the framework of the Russian Federation fundamental research program for the long-term period for 2021–2030 and Institute of Mathematical Problems of Biology RAS-the Branch of Keldysh Institute of Applied Mathematics of Russian Academy of Sciences 0017-2019-0009 and State Assignment of Institute of Theoretical and Experimental Biophysics of Russian Academy of Science 075-00381-21-00.

Author Contributions

Conceptualization, D.T. and A.L.K.; methodology, V.R., A.S. and A.T.K.; software, A.S. and D.T.; validation, A.L.K., A.T.K. and D.K.; formal analysis, A.T.K., A.I. and A.Z; investigation, A.T.K., A.L.K., K.M. and L.K.; resources, D.K., A.Z., N.P. and D.E.; data curation, A.L.K. and A.S.; writing—original draft preparation, A.T.K.; writing—review and editing, A.T.K and A.L.K. All authors have read and agreed to the published version of the manuscript.

Competing Interests

The authors declare no conflict of interest

References

(1) Breast cancer - Symptoms and causes - Mayo Clinic https://www.mayoclinic.org/diseases-conditions/breast-cancer/symptoms-causes/syc-20352470 (accessed Jan 25, 2021).

(2) Breast cancer statistics | World Cancer Research Fund https://www.wcrf.org/dietandcancer/cancer-trends/breast-cancer-statistics (accessed Jan 25, 2021).

(3) 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. *Oncogen***2019**, *2* (3). https://doi.org/10.35702/onc.10012.

(4) Venne, A. S.; Kollipara, L.; Zahedi, R. P. The next Level of Complexity: Crosstalk of Posttranslational Modifications. *Proteomics***2014**, *14* (4–5), 513–524. https://doi.org/10.1002/pmic.201300344.

(5) Brentville, V. A.; Vankemmelbeke, M.; Metheringham, R. L.; Durrant, L. G. Post-Translational Modifications Such as Citrullination Are Excellent Targets for Cancer Therapy. *Semin. Immunol.***2020**, *47*, 101393. https://doi.org/10.1016/j.smim.2020.101393.

(6) Han, Z.-J.; Feng, Y.-H.; Gu, B.-H.; Li, Y.-M.; Chen, H. The Post-Translational Modification, SUMOylation, and Cancer (Review). *Int. J. Oncol.***2018**, *52* (4), 1081–1094. https://doi.org/10.3892/ijo.2018.4280.

(7) Prieto, P.; Jaén, R. I.; Calle, D.; Gómez-Serrano, M.; Núñez, E.; Fernández-Velasco, M.; Martín-Sanz, P.; Alonso, S.; Vázquez, J.; Cerdán, S.; Peinado, M. Á.; Boscá, L. Interplay between Post-Translational Cyclooxygenase-2 Modifications and the Metabolic and Proteomic Profile in a Colorectal Cancer Cohort. *World J. Gastroenterol.***2019**, *25* (4), 433–446. https://doi.org/10.3748/wjg.v25.i4.433.

(8) Zeltz, C.; Gullberg, D. Post-Translational Modifications of Integrin Ligands as Pathogenic Mechanisms in Disease. *Matrix Biol. J. Int. Soc. Matrix Biol.***2014**, *40*, 5–9. https://doi.org/10.1016/j.matbio.2014.08.001.

(9) Wang, P.; Sun, C.; Zhu, T.; Xu, Y. Structural Insight into Mechanisms for Dynamic Regulation of PKM2. *Protein Cell***2015**, *6* (4), 275–287. https://doi.org/10.1007/s13238-015-0132-x.

(10) Cheng, J.; Bawa, T.; Lee, P.; Gong, L.; Yeh, E. T. H. Role of Desumoylation in the Development of Prostate Cancer. *Neoplasia N. Y. N***2006**, *8* (8), 667–676. https://doi.org/10.1593/neo.06445.

(11) Lee, J.-Y.; Won, H.-Y.; Park, J.-H.; Kim, H.-Y.; Choi, H.-J.; Shin, D.-H.; Kang, J.-H.; Woo, J.-K.; Oh, S.-H.; Son, T.; Choi, J.-W.; Kim, S.; Kim, H.-Y.; Yi, K.; Jang, K.-S.; Oh, Y.-H.; Kong, G. MEL-18 Loss Mediates Estrogen Receptor-α Downregulation and Hormone Independence. *J. Clin. Invest.***2015**, *125* (5), 1801–1814. https://doi.org/10.1172/JCI73743.

(12) Heo, K.-S. Regulation of Post-Translational Modification in Breast Cancer Treatment. *BMB Rep.***2019**, *52* (2), 113–118. https://doi.org/10.5483/BMBRep.2019.52.2.017.

(13) Cho, Y.; Kang, H. G.; Kim, S.-J.; Lee, S.; Jee, S.; Ahn, S. G.; Kang, M. J.; Song, J. S.; Chung, J.-Y.; Yi, E. C.; Chun, K.-H. Post-Translational Modification of OCT4 in Breast Cancer Tumorigenesis. *Cell Death Differ.***2018**, *25* (10), 1781–1795. https://doi.org/10.1038/s41418-018-0079-6.

(14) Marsh, D. J.; Shah, J. S.; Cole, A. J. Histones and Their Modifications in Ovarian Cancer - Drivers of Disease and Therapeutic Targets. *Front. Oncol.***2014**, *4*, 144. https://doi.org/10.3389/fonc.2014.00144.

(15) Chen, Y.; Hao, Q.; Wang, J.; Li, J.; Huang, C.; Zhang, Y.; Wu, X.; Lu, H.; Zhou, X. Ubiquitin Ligase TRIM71 Suppresses Ovarian Tumorigenesis by Degrading Mutant P53. *Cell Death Dis.***2019**, *10* (10), 737. https://doi.org/10.1038/s41419-019-1977-3.

(16) Song, M.; Sandoval, T. A.; Chae, C.-S.; Chopra, S.; Tan, C.; Rutkowski, M. R.; Raundhal, M.; Chaurio, R. A.; Payne, K. K.; Konrad, C.; Bettigole, S. E.; Shin, H. R.; Crowley, M. J. P.; Cerliani, J. P.; Kossenkov, A. V.; Motorykin, I.; Zhang, S.; Manfredi, G.; Zamarin, D.; Holcomb, K.; Rodriguez, P. C.; Rabinovich, G. A.; Conejo-Garcia, J. R.; Glimcher, L. H.; Cubillos-Ruiz, J. R. IRE1α-XBP1 Controls T Cell Function in Ovarian Cancer by Regulating Mitochondrial Activity. *Nature***2018**, *562* (7727), 423–428. https://doi.org/10.1038/s41586-018-0597-x.

(17) Bitler, B. G.; Wu, S.; Park, P. H.; Hai, Y.; Aird, K. M.; Wang, Y.; Zhai, Y.; Kossenkov, A. V.; Vara-Ailor, A.; Rauscher, F. J.; Zou, W.; Speicher, D. W.; Huntsman, D. G.; Conejo-Garcia, J. R.; Cho, K. R.; Christianson, D. W.; Zhang, R. ARID1A-Mutated Ovarian Cancers Depend on HDAC6 Activity. *Nat. Cell Biol.***2017**, *19* (8), 962–973. https://doi.org/10.1038/ncb3582.

(18) 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* (14). https://doi.org/10.3390/molecules25143144.

(19) Efimov, A. V. Standard Structures in Proteins. *Prog. Biophys. Mol. Biol.***1993**, *60* (3), 201–239. https://doi.org/10.1016/0079-6107(93)90015-C.

(20) Lechowicz, U.; Rudzinski, S.; Jezela-Stanek, A.; Janciauskiene, S.; Chorostowska-Wynimko, J. Post-Translational Modifications of Circulating Alpha-1-Antitrypsin Protein. *Int. J. Mol. Sci.***2020**, *21* (23). https://doi.org/10.3390/ijms21239187.

(21) Xu, H.; Wang, Y.; Lin, S.; Deng, W.; Peng, D.; Cui, Q.; Xue, Y. PTMD: A Database of Human Disease-Associated Post-Translational Modifications. *Genomics Proteomics Bioinformatics***2018**, *16* (4), 244–251. https://doi.org/10.1016/j.gpb.2018.06.004.

(22) Xu, Y.; Yang, Y.; Wang, Z.; Li, C.; Shao, Y. A Systematic Review on Posttranslational Modification in Proteins: Feature Construction, Algorithm and Webserver. *Protein Pept. Lett.***2018**, *25* (9), 807–814. https://doi.org/10.2174/0929866525666180925151720.

(23) Lilkova, E.; Petkov, P.; Ilieva, N.; Krachmarova, E.; Nacheva, G.; Litov, L. Molecular Modeling of the Effects of Glycosylation on the Structure and Dynamics of Human Interferon-Gamma. *J. Mol. Model.***2019**, *25* (5), 127. https://doi.org/10.1007/s00894-019-4013-8.

(24) Stelzl, T.; Geillinger-Kästle, K. E.; Stolz, J.; Daniel, H. Glycans in the Intestinal Peptide Transporter PEPT1 Contribute to Function and Protect from Proteolysis. *Am. J. Physiol. Gastrointest. Liver Physiol.***2017**, *312* (6), G580–G591. https://doi.org/10.1152/ajpgi.00343.2016.

(25) Shental-Bechor, D.; Levy, Y. Folding of Glycoproteins: Toward Understanding the Biophysics of the Glycosylation Code. *Curr. Opin. Struct. Biol.***2009**, *19* (5), 524–533. https://doi.org/10.1016/j.sbi.2009.07.002.

(26) Ekeowa, U. I.; Freeke, J.; Miranda, E.; Gooptu, B.; Bush, M. F.; Pérez, J.; Teckman, J.; Robinson, C. V.; Lomas, D. A. Defining the Mechanism of Polymerization in the Serpinopathies. *Proc. Natl. Acad. Sci. U. S. A.***2010**, *107* (40), 17146–17151. https://doi.org/10.1073/pnas.1004785107.

(27) Miranda, E.; Ferrarotti, I.; Berardelli, R.; Laffranchi, M.; Cerea, M.; Gangemi, F.; Haq, I.; Ottaviani, S.; Lomas, D. A.; Irving, J. A.; Fra, A. The Pathological Trento Variant of Alpha-1-Antitrypsin (E75V) Shows Nonclassical Behaviour during Polymerization. *FEBS J.***2017**, *284* (13), 2110–2126. https://doi.org/10.1111/febs.14111.

(28) Liang, Y.; Ma, T.; Thakur, A.; Yu, H.; Gao, L.; Shi, P.; Li, X.; Ren, H.; Jia, L.; Zhang, S.; Li, Z.; Chen, M. Differentially Expressed Glycosylated Patterns of α-1-Antitrypsin as Serum Biomarkers for the Diagnosis of Lung Cancer. *Glycobiology***2015**, *25* (3), 331–340. https://doi.org/10.1093/glycob/cwu115.

(29) Verheul, M. K.; Yee, A.; Seaman, A.; Janssen, G. M.; van Veelen, P. A.; Drijfhout, J. W.; Toes, R. E. M.; Mahler, M.; Trouw, L. A. Identification of Carbamylated Alpha 1 Anti-Trypsin (A1AT) as an Antigenic Target of Anti-CarP Antibodies in Patients with Rheumatoid Arthritis. *J. Autoimmun.***2017**, *80*, 77–84. https://doi.org/10.1016/j.jaut.2017.02.008.

(30) Matheson, N. R.; Wong, P. S.; Schuyler, M.; Travis, J. Interaction of Human Alpha-1-Proteinase Inhibitor with Neutrophil Myeloperoxidase. *Biochemistry***1981**, *20* (2), 331–336. https://doi.org/10.1021/bi00505a016.

(31) Kaner, Z.; Engelman, R.; Schuster, R.; Rider, P.; Greenberg, D.; Av-Gay, Y.; Benhar, M.; Lewis, E. C. S-Nitrosylation of A1-Antitrypsin Triggers Macrophages Toward Inflammatory Phenotype and Enhances Intra-Cellular Bacteria Elimination. *Front. Immunol.***2019**, *10*, 590. https://doi.org/10.3389/fimmu.2019.00590.

(32) Takaki, T. EEG Enhancement during Electrosurgery Using an Optical Fibre. *Med. Biol. Eng. Comput.***1988**, *26* (4), 397–403. https://doi.org/10.1007/BF02442298.

(33) Goodarzi, M. T.; Turner, G. A. Decreased Branching, Increased Fucosylation and Changed Sialylation of Alpha-1-Proteinase Inhibitor in Breast and Ovarian Cancer. *Clin. Chim. Acta Int. J. Clin. Chem.***1995**, *236* (2), 161–171. https://doi.org/10.1016/0009-8981(95)06049-j.

(34) Nagumo, K.; Tanaka, M.; Chuang, V. T. G.; Setoyama, H.; Watanabe, H.; Yamada, N.; Kubota, K.; Tanaka, M.; Matsushita, K.; Yoshida, A.; Jinnouchi, H.; Anraku, M.; Kadowaki, D.; Ishima, Y.; Sasaki, Y.; Otagiri, M.; Maruyama, T. Cys34-Cysteinylated Human Serum Albumin Is a Sensitive Plasma Marker in Oxidative Stress-Related Chronic Diseases. *PloS One***2014**, *9* (1), e85216. https://doi.org/10.1371/journal.pone.0085216.

(35) Kashiba-Iwatsuki, M.; Miyamoto, M.; Inoue, M. Effect of Nitric Oxide on the Ligand-Binding Activity of Albumin. *Arch. Biochem. Biophys.***1997**, *345* (2), 237–242. https://doi.org/10.1006/abbi.1997.0258.

(36) Watanabe, H.; Imafuku, T.; Otagiri, M.; Maruyama, T. Clinical Implications Associated With the Posttranslational Modification-Induced Functional Impairment of Albumin in Oxidative Stress-Related Diseases. *J. Pharm. Sci.***2017**, *106* (9), 2195–2203. https://doi.org/10.1016/j.xphs.2017.03.002.

(37) Tagliabracci, V. S. A Single Kinase Generates the Majority of the Secreted Phosphoproteome. 15.

(38) Rutkowski, M. J.; Sughrue, M. E.; Kane, A. J.; Mills, S. A.; Parsa, A. T. Cancer and the Complement Cascade. *Mol. Cancer Res.***2010**, *8* (11), 1453–1465. https://doi.org/10.1158/1541-7786.MCR-10-0225.

(39) Ricklin, D.; Reis, E. S.; Mastellos, D. C.; Gros, P.; Lambris, J. D. Complement Component C3 - The “Swiss Army Knife” of Innate Immunity and Host Defense. *Immunol. Rev.***2016**, *274* (1), 33–58. https://doi.org/10.1111/imr.12500.

(40) Phosphorylation of plasma proteins with emphasis on complement component C3 - PubMed https://pubmed.ncbi.nlm.nih.gov/10403476/ (accessed Jan 25, 2021).

(41) Qiu, Y.; Patwa, T. H.; Xu, L.; Shedden, K.; Misek, D. E.; Tuck, M.; Jin, G.; Ruffin, M. T.; Turgeon, D. K.; Synal, S.; Bresalier, R.; Marcon, N.; Brenner, D. E.; Lubman, D. M. Plasma Glycoprotein Profiling for Colorectal Cancer Biomarker Identification by Lectin Glycoarray and Lectin Blot. *J. Proteome Res.***2008**, *7* (4), 1693–1703. https://doi.org/10.1021/pr700706s.

(42) Sarrats, A.; Saldova, R.; Pla, E.; Fort, E.; Harvey, D. J.; Struwe, W. B.; de Llorens, R.; Rudd, P. M.; Peracaula, R. Glycosylation of Liver Acute-Phase Proteins in Pancreatic Cancer and Chronic Pancreatitis. *Proteomics Clin. Appl.***2010**, *4* (4), 432–448. https://doi.org/10.1002/prca.200900150.

(43) Lee, S.-J.; Kim, S.-J.; Seo, H.-H.; Shin, S.-P.; Kim, D.; Park, C.-S.; Kim, K.-T.; Kim, Y.-H.; Jeong, J.-S.; Kim, I.-H. Over-Expression of MiR-145 Enhances the Effectiveness of HSVtk Gene Therapy for Malignant Glioma. *Cancer Lett.***2012**, *320* (1), 72–80. https://doi.org/10.1016/j.canlet.2012.01.029.

(44) Ascoli, G. A.; Domenici, E.; Bertucci, C. Drug Binding to Human Serum Albumin: Abridged Review of Results Obtained with High-Performance Liquid Chromatography and Circular Dichroism. *Chirality***2006**, *18* (9), 667–679. https://doi.org/10.1002/chir.20301.

(45) Bar-Or, R.; Rael, L. T.; Bar-Or, D. Dehydroalanine Derived from Cysteine Is a Common Post-Translational Modification in Human Serum Albumin. *Rapid Commun. Mass Spectrom. RCM***2008**, *22* (5), 711–716. https://doi.org/10.1002/rcm.3421.

(46) Gaston, B. M.; Carver, J.; Doctor, A.; Palmer, L. A. S-Nitrosylation Signaling in Cell Biology. *Mol. Interv.***2003**, *3* (5), 253–263. https://doi.org/10.1124/mi.3.5.253.

(47) Al-Harthi, S.; Lachowicz, J. I.; Nowakowski, M. E.; Jaremko, M.; Jaremko, Ł. Towards the Functional High-Resolution Coordination Chemistry of Blood Plasma Human Serum Albumin. *J. Inorg. Biochem.***2019**, *198*, 110716. https://doi.org/10.1016/j.jinorgbio.2019.110716.

(48) Ren, L.; Yi, J.; Li, W.; Zheng, X.; Liu, J.; Wang, J.; Du, G. Apolipoproteins and Cancer. *Cancer Med.***2019**, *8* (16), 7032–7043. https://doi.org/10.1002/cam4.2587.

(49) Role of the complement system in the tumor microenvironment | Cancer Cell International | Full Text https://cancerci.biomedcentral.com/articles/10.1186/s12935-019-1027-3 (accessed Jan 25, 2021).

(50) Gros, P.; Milder, F. J.; Janssen, B. J. C. Complement Driven by Conformational Changes. *Nat. Rev. Immunol.***2008**, *8* (1), 48–58. https://doi.org/10.1038/nri2231.

(51) Zhu, J.; Warner, E.; Parikh, N. D.; Lubman, D. M. Glycoproteomic Markers of Hepatocellular Carcinoma-Mass Spectrometry Based Approaches. *Mass Spectrom. Rev.***2019**, *38* (3), 265–290. https://doi.org/10.1002/mas.21583.

(52) Isailovic, D.; Kurulugama, R. T.; Plasencia, M. D.; Stokes, S. T.; Kyselova, Z.; Goldman, R.; Mechref, Y.; Novotny, M. V.; Clemmer, D. E. Profiling of Human Serum Glycans Associated with Liver Cancer and Cirrhosis by IMS-MS. *J. Proteome Res.***2008**, *7* (3), 1109–1117. https://doi.org/10.1021/pr700702r.

(53) Macciò, A.; Madeddu, C.; Gramignano, G.; Mulas, C.; Tanca, L.; Cherchi, M. C.; Floris, C.; Omoto, I.; Barracca, A.; Ganz, T. The Role of Inflammation, Iron, and Nutritional Status in Cancer-Related Anemia: Results of a Large, Prospective, Observational Study. *Haematologica***2015**, *100* (1), 124–132. https://doi.org/10.3324/haematol.2014.112813.

(54) Shen, Y.; Li, X.; Dong, D.; Zhang, B.; Xue, Y.; Shang, P. Transferrin Receptor 1 in Cancer: A New Sight for Cancer Therapy. *Am. J. Cancer Res.***2018**, *8* (6), 916–931.

(55) Biswas, S.; Tapryal, N.; Mukherjee, R.; Kumar, R.; Mukhopadhyay, C. K. Insulin Promotes Iron Uptake in Human Hepatic Cell by Regulating Transferrin Receptor-1 Transcription Mediated by Hypoxia Inducible Factor-1. *Biochim. Biophys. Acta***2013**, *1832* (2), 293–301. https://doi.org/10.1016/j.bbadis.2012.11.003.

(56) Berman, H. M. The Protein Data Bank. *Nucleic Acids Res.***2000**, *28* (1), 235–242. https://doi.org/10.1093/nar/28.1.235.

(57) D. A. Tikhonov, L. I. Kulikova, A. V. Efimov, “Statistical analysis of the internal distances of helical pairs in protein molecules”, Mat. Biolog. Bioinform., 11:2 (2016), 170–190 http://www.mathnet.ru/php/archive.phtml?wshow=paper&jrnid=mbb&paperid=276&option_lang=eng (accessed Jan 25, 2021).

(58) (PDF) The Study of Interhelical Angles in the Structural Motifs Formed by Two Helices | D. Tikhonov - Academia.Edu.

(59) Tikhonov, D. A.; Kulikova, L. I.; Efimov, A. V. Analysis of the Torsion Angles between Helical Axes in Pairs of Helices in Protein Molecules. *Math. Biol. Bioinforma.***2018**, *13* (S), t17–t28. https://doi.org/10.17537/2018.13.t17.

(60) Kabsch, W.; Sander, C. Dictionary of Protein Secondary Structure: Pattern Recognition of Hydrogen-Bonded and Geometrical Features. *Biopolymers***1983**, *22* (12), 2577–2637. https://doi.org/10.1002/bip.360221211.

(61) (PDF) The Study of Interhelical Angles in the Structural Motifs Formed by Two Helices | D. Tikhonov - Academia.edu https://www.academia.edu/40785746/The_Study_of_Interhelical_Angles_in_the_Structural_Motifs_Formed_by_Two_Helices (accessed Jan 25, 2021).

(62) Sayle, R. A.; Milner-White, E. J. RASMOL: Biomolecular Graphics for All. *Trends Biochem. Sci.***1995**, *20* (9), 374. https://doi.org/10.1016/s0968-0004(00)89080-5.

(63) The Amber biomolecular simulation programs - PubMed https://pubmed.ncbi.nlm.nih.gov/16200636/ (accessed Jan 25, 2021).

(64) Duan, Y.; Wu, C.; Chowdhury, S.; Lee, M. C.; Xiong, G.; Zhang, W.; Yang, R.; Cieplak, P.; Luo, R.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. A Point-Charge Force Field for Molecular Mechanics Simulations of Proteins Based on Condensed-Phase Quantum Mechanical Calculations. *J. Comput. Chem.***2003**, *24* (16), 1999–2012. https://doi.org/10.1002/jcc.10349.

(65) VMD: visual molecular dynamics - PubMed https://pubmed.ncbi.nlm.nih.gov/8744570/ (accessed Jan 25, 2021).

(66) Modification of the Generalized Born Model Suitable for Macromolecules | The Journal of Physical Chemistry B https://pubs.acs.org/doi/10.1021/jp994072s (accessed Jan 25, 2021).

(67) Conway, J.; Lex, A.; Gehlenborg, N. UpSetR: An R Package for the Visualization of Intersecting Sets and Their Properties. *Bioinforma. Oxf. Engl.***2017**, *33*. https://doi.org/10.1093/bioinformatics/btx364.

(68) Szklarczyk, D.; Gable, A. L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N. T.; Morris, J. H.; Bork, P.; Jensen, L. J.; Mering, C. von. STRING V11: Protein-Protein Association Networks with Increased Coverage, Supporting Functional Discovery in Genome-Wide Experimental Datasets. *Nucleic Acids Res.***2019**, *47* (D1), D607–D613. https://doi.org/10.1093/nar/gky1131.

Tables

Table 1. The list of proteins carrying the defined PTMs and found specifically in patients with ovarian cancer or patients with breast cancer. Peptides are listed with the main accompanying mass-spectrometric characterizations: PSM – peptide spectra match (is the number of spectra matching the theoretical peptide sequence with high score); b- and y-type of fragment ions are C- and N-terminal sequential fragment ions populated after peptide (precursor ion) decay. Complete results are available in the *Supplemental Materials Table S2*.

PTMs OBSERVED IN PATIENTS WITH OVARIAN CANCER (OC)						
Gene	**Protein**	**Sequence of peptide carrying PTM**	**PSM**	**Number of peptide identifications**	**Seq., %**	**B/Y**
A1AT	Alpha-1-antitrypsin	***ac(K)***-QINDYVEK	131	17	55	1B/8Y
		TDTSHHDQDHPTFN***ac(K)***ITPNLAEFAFSLYR	131	17	55	0B/6Y
ALBU	Albumin	EQL***ac(K)***AVMDDFAAFVEK	94	17	32	0B/7Y
		NYAEA***ac(K)***DVFLGMFLYEYAR	454	25	31	0B/6Y
		RHPYF***p(Y)***APELLFFAK	618	46	76	3B/6Y
		VFDE***ac(K)***PLVEEPQNLIK	618	46	76	0B/5Y
APOA2	Apolipoprotein A-II	EPCVE***p(S)***LVSQYFQTVTDYGK	29	6	63	1B/6Y
CO3	Complement C3	***ac(K)***VLLDGVQNPR	161	32	31	0B/5Y
HV307	Ig heavy chain V-III region CAM	DDS***ac(K)***NTLYLQMNSLR	11	4	12	0B/9Y
IGHA 1	Ig alpha-1 chain C region	SGNTFRPEVHLLPPP***p(S)***EELALNELVTLTCLAR	90	13	63	0B/7Y
KNG1	Kininogen-1	E***p(T)***TCSKESNEELTESCETK	13	9	21	7B/9Y
TRFE	Serotransferrin	MDA***ac(K)***MYLGYEYVTAIR	145	19	30	0B/8Y
PTMs OBSERVED IN PATIENTS WITH BREAST CANCER (BC)						
ALBU	Albumin	EQL***ac(K)***AVMDDFAAFVEK	465	27	43	1B/7Y
		ac(K)VPQVSTPTLVEVSR	465	27	43	0B/5Y
		NYAEA***ac(K)***DVFLGMFLYEYAR	687	48	79	0B/8Y
APOA1	Apolipoprotein A-I	QLNL***ac(K)***LLDNWDSVTSTFSK	43	16	73	4B/10Y
TRFE	Serotransferrin	***ac(K)***SASDLTWDNLK	357	38	66	0B/9Y
		MDA***ac(K)***MYLGYEYVTAIR	159	17	34	2B/8Y

Ac(K) – acetylation of lysine; p(Y), p(T), p(S) – phosphorylation of tyrosine, threonine and serine, correspondingly; PSM – peptide spectra; Seq., % – amino acids coverage for the identified peptides; b/y– the number of revealed and attained b- and y-type fragment ions.

Table 2. Solvent accessible area alterations of amino acid residue and its active environment before and after modification (mounting of PTM) in breast and ovarian cancer-specific proteins.

Gene (Protein name)	Sequence of peptide carrying PTM	PDB structures	Average area accessible to the solvent, Å^2				Helices	Motives
			Amino acid		Active environment			
			intact	PTM	intact	PTM		
OVARIAN CANCER								
A1AT (Alpha-1-antitrypsin)	***ac(K)***-QINDYVEK	35	119,99	167,54	335,14	364,17	4	1
	TDTSHHDQDHPTFN-***ac(K)***-ITPNLAEFAFSLYR	23	109,78	153,36	346,66	377,32		
ALBU (Albumin)	EQL-***ac(K)***-AVMDDFAAFVEK	156	113,03	135,71	441,31	431,30	30	4
	ac(k)-VPQVSTPTLVEVSR	172	21,5	37,29	320,08	312,11		3
	NYAEa-***ac(K)***-DVFLGMFLYEYAR	162	82,02	108,97	294,16	305,17		3
	RHPYF-***p(Y)***-APELLFFAK	163	22,49	31,01	165,83	154,03		3
	VFDEF-***ac(K)***-PLVEEPQNLIK	162	131,78	179,83	278,90	316,22		2
APOA2 (Apolipoprotein A-II)	EPCVE-***p(S)***-LVSQYFQTVTDYGK	56	69,47	101,31	498,39	501,38	4	n/d*
CO3 (Complement C3)	YF-***ac(k)***-PGMPFDLMVFVTNPDGSPAYR	51	13,86	16,87	134,78	110,5	1	n/d
HV307 (Ig heavy chain V-III)	DDS-***ac(K)***-NTLYLQMNSLR	2	126,05	177,7	316	349,85	0	0
IGHA1 (Ig alpha-1 chain)	SGNTFRPEVHLLPPP-***p(S)***-EELALNELVTLTCLAR	34	103,55	168,88	318,55	361,30	3	1
BREAST CANCER								
ALBU (Albumin)	NYAEA-***ac(K)***-DVFLGMFLYEYA	162	82,02	108,97	294,16	305,17	30	3
	EQL-***ac(K)***-AVMDDFAAFVEK	156	113,03	135,71	441,31	431,30		4
APOA1 (Apolipoprotein A-I)	QLNL ***ac(K)***-LLDNWDSVTSTFSK	6	136,75	184,78	317,5	348,78	8	n/d
TRFE (Serotransferrin)	***ac(K)***-SASDLTWDNLK	32	111,32	155,53	322,75	349,13	22	12
	MDA-***ac(K)***-MYLGYEYVTAIR	72	64,48	104,60	260,24	279,48	22	12

n/d* – motives with interacting helices are not found; PDB structures – the number of PDB structures matching the target polypeptide chain; PTM SEQQ – amino acid sequence of polypeptide chain carrying the detected and identified PTM.

Table 3. Functional amendments of proteins under the impact of the selected PTMs reported in the literature and observed in the present study.

Protein	Our experiment	Biological role of protein	PTM, literature data	Biological role of the allocated PTM	Ref.
A1AT	ac-K^{25}, ac-K^{125}	Tissue protection from the hydrolytic catalytic activity	Glu-K^{342} и Glu-V^{75}	The dire role of polymers aggregation is uncertain but observable.	[26,27]
			Ox-M^{351} и ox-M^{358}	Signs of oxidative stress and pro-inflammatory activity.	[32]
			NO-C^{232}	Regulation of apoptosis.	[31]
			Galactosylation and Fucosylation patterns	Specific distinguishing of non-small-cell lung cancer from lung adenocarcinoma	[28]
			3N-glycosylation sites	Increase outer arm fucosylation in ovarian cancer and breast cancer patients.	[33]
			Carbamoyl-K^{359}	Enhances autoimmune response in patients with rheumatoid arthritis.	[29]
ALBU	ac-K^{545}, ac-K^{323}, ac-K^{343}, ac-K^{378}	Regulation of affinity binding with various low-molecular-weight ligands	Glu-C^{34}	Reduction of affinity binding with tryptophane, bilirubin, warfarin, and diazepam.	[34]
			NO-C^{34}	Regulation of organic anions and metal transportation; reduction of affinity binding with fatty acids.	[35]
			SH-C^{34}	Marker of oxidative stress.	[36]
APOA2	p-S^{9}	Lipid binding and homeostasis [37]	N/D	N/D	[37]
CO3	Ac-K^{283}	May act as pro-oncogenic factor due to positive regulation of VEGF expression level; involved in angiogenesis, organization and clearance of extracellular matrix, and cell migration [38]	Glc-Asn^{917} (α-chain); Glc-Asn^{63} (β-chain)	N/D	[39]
			p-T^{1009} (intact C3)	N/D	[40]
			Fuc-N^{85} and Sia-N^{939}	Increase in α-2,6-sialic acid and in fucosylation in colorectal cancers patients	[41]
TRFE	Ac-K^{331}, Ac-K^{453}	Binding and transfer of iron ions; regulation of iron level, especially during pregnancy.	High level of glycosylation of N^{432} and N^{630}	Acute inflammation in pancreatic patients in advance to tumorigenesis.	[42]
KNG1	p-T^{326}	Regulation of blood coagulation, blood glucose level and induction of nociceptors.	Non-specific glycosylation of N^{294}	Unknown role; detected in colorectal cancer patients	[41]

Figures

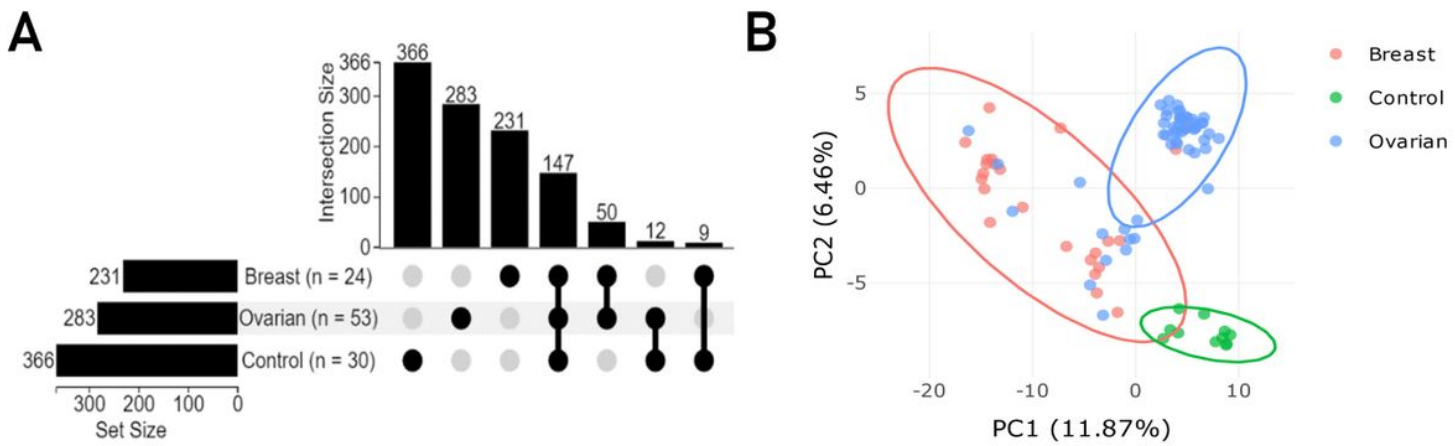

Figure 1

The upset diagram showing the distribution and the intercept size of protein identifications amongst groups of breast and ovarian cancer phenotypes, and the control group (A). PCA analysis carried out for the mutual fraction of proteins (n=147) with no imputation and reflecting the quantitative alteration of the identified proteins between groups: breast cancer (BC) - red ellipse, ovarian cancer (OC) - blue ellipse and the control (CNT) - green ellipse. The abundancy (quantitative item) was fit to the log-scale (B).

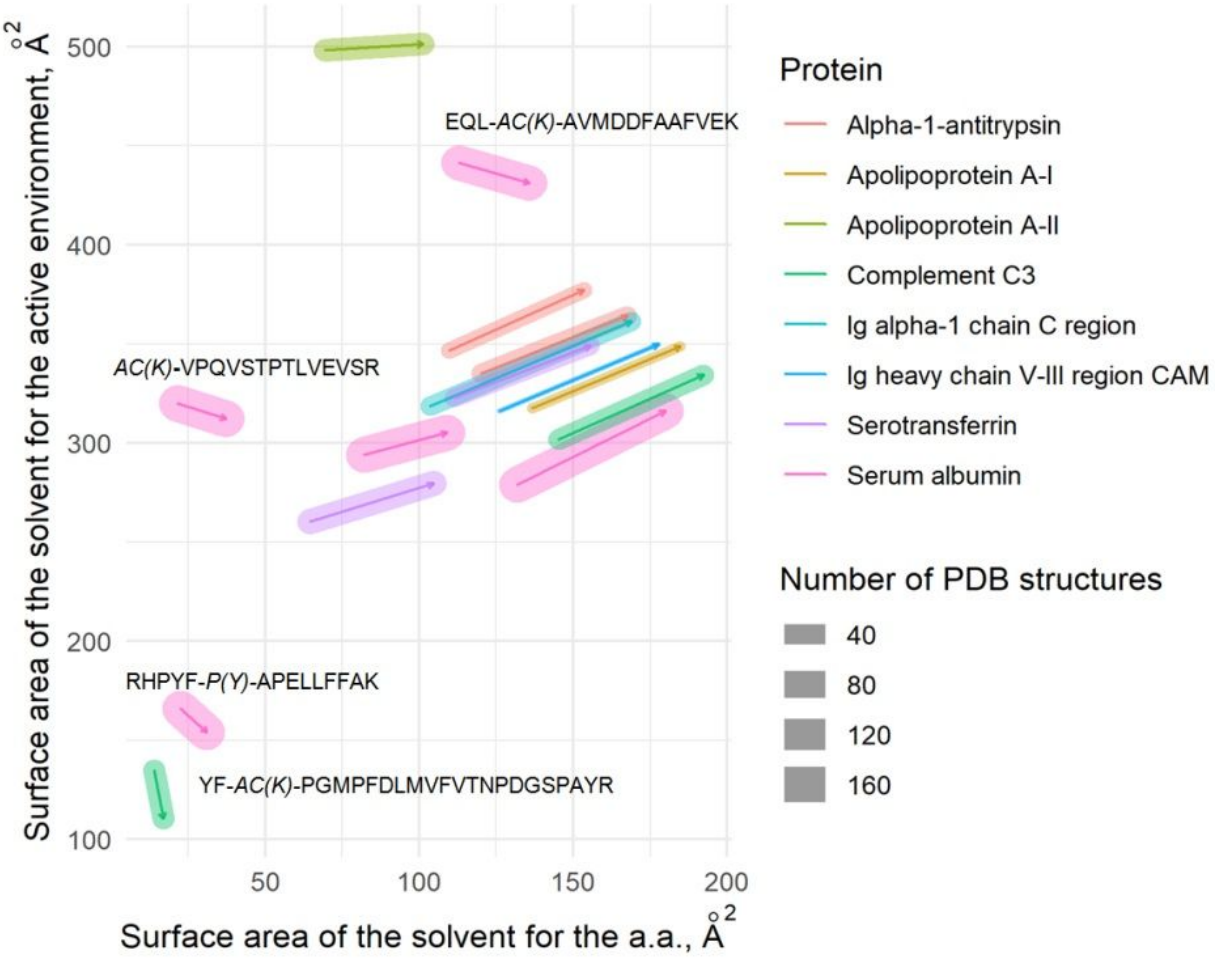

Figure 2

Variation of the total solvent-accessible area (in square angstroms) of active environment after modification of amino acid residue (mounting of PTM). The modified proteins found in plasma samples of patients with breast cancer and ovarian cancer are color-coded. The number of PDB matched

structures used for the analysis are designated by bar size.

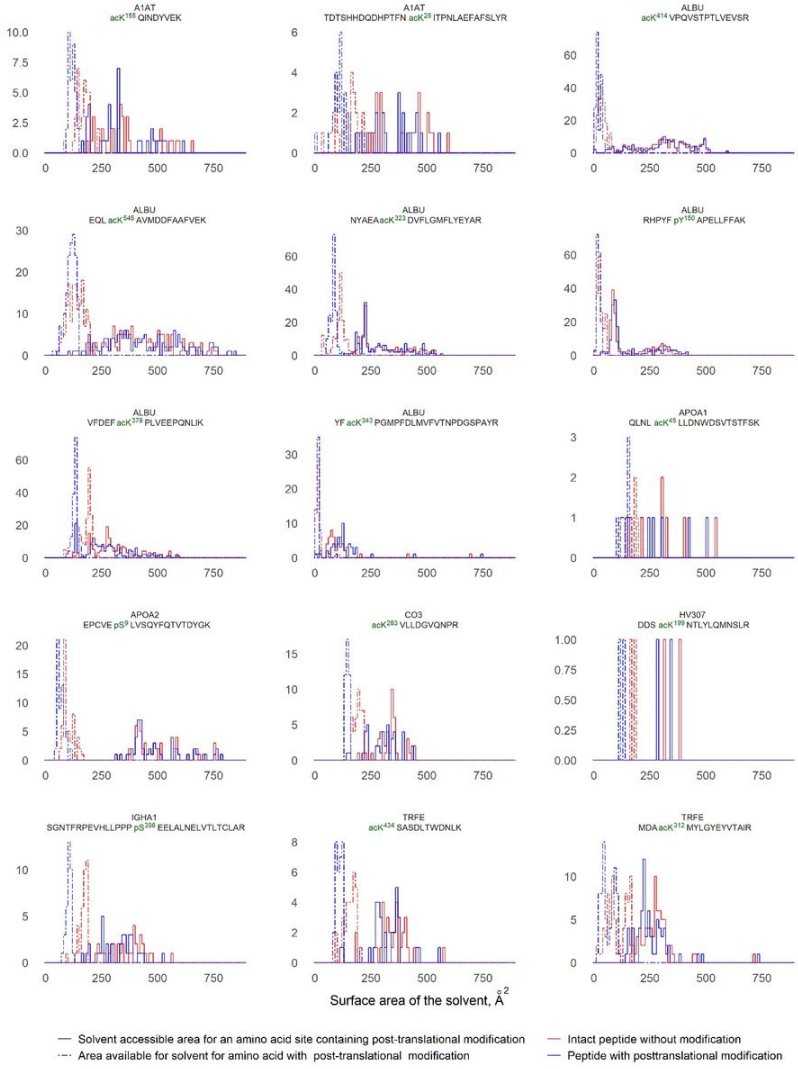

Figure 3

The distribution of proteins population with the identified PTMs depending on the solvent-accessible area for the certain amino acid residue and its active environment before and after mounting of PTM. The OX axis indicated the surface area, accessible for the surrounding solvent (in Å2); the OY axis indicated the number of the affected protein molecules. Color-code defines amino acids before modification (blue dashed line), amino acids after modification (pink dashed line), active environment before modification (blue solid line) and active environment after modification (pink solid line).

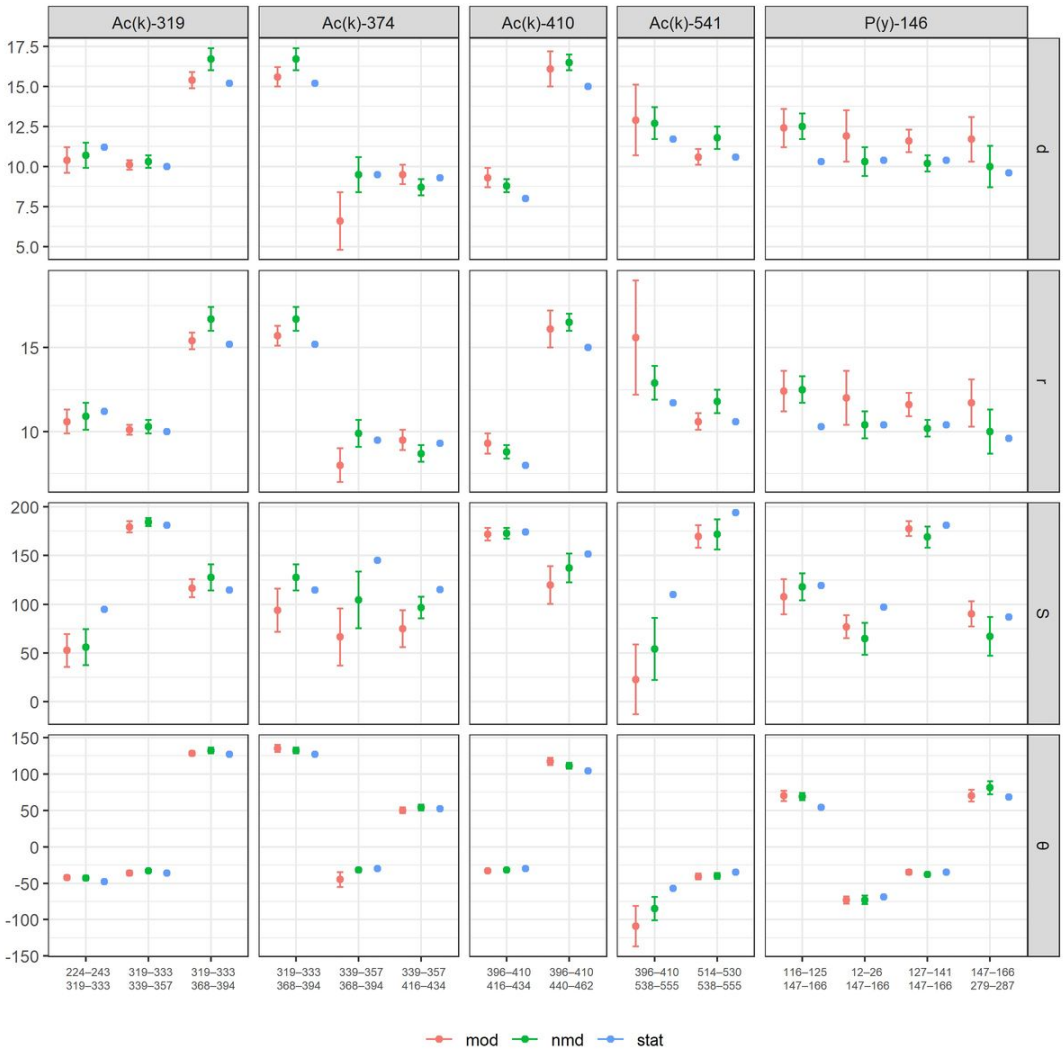

Figure 4

Calculated geometrical features of albumin (ALBU) motives with PTM moiety. Each motif is characterized by the estimated geometrical features for intact molecule (stat; blue color), calculated results of molecular dynamics for the unmodified molecule (nmd; green color), and calculated features for the modified molecule (mod; red color). The right axis indicates (d) distances between helices, (r) inter-planar distance between helices, S – area of the spatial projection, and θ – torsion angle.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryMaterials.pdf