HHS Public Access

Author manuscript

Curr Opin Struct Biol. Author manuscript; available in PMC 2021 June 01.

Published in final edited form as:

Curr Opin Struct Biol. 2020 June ; 62: 102–111. doi:10.1016/j.sbi.2019.12.002.

Emerging patterns of Tyrosine sulfation and O-glycosylation cross-talk and co-localization

Akul Y. Mehta[1], **Jamie Heimburg-Molinaro**[1], **Richard D. Cummings**[1], **Christoffer K. Goth**[1]

[1]Department of Surgery, Beth Israel Deaconess Medical Center, National Center for Functional Glycomics, Harvard Medical School, Boston, MA, 02215, U.S.A.

Abstract

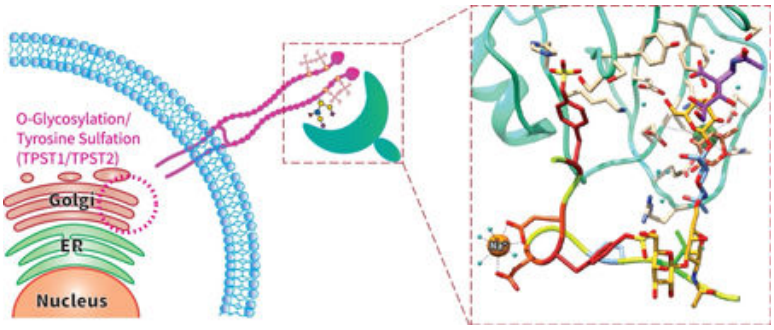
Post-translational modifications (PTMs) drive the diversity of the proteome and broadly regulate protein function. Interplay between different types of PTMs further enables tight and dynamic fine-tuning of molecular functions. O-glycosylation on serine, threonine, and tyrosine residues is a major PTM with diverse roles in development, differentiation, pathogenesis, and proteolytic processing. Other examples of cross-talk between PTMs also exists, such as PSGL-1, where the combined presence of N-terminal sulfotyrosines and O-glycans is pivotal for selectin binding. A handful of other related examples of O-glycans and sulfotyrosine co-localization has been described but it is not yet recognized as a general regulatory phenomenon. In this review, we highlight the emerging global pattern of co-localization of cell-surface and extracellular sulfotyrosines with O-glycans, which we term ‘multi-motif’ interactions, from a wide range of protein classes. We also discuss the barriers, and existing and future tools needed to dissect the biological impact and biomedical potential.

Graphical Abstract

Mehta et al, Author Contribution Statement:
Akul Y. Mehta- Conceptualization, Visualization, Roles/Writing - original draft; Writing - review & editing
Jamie Heimburg-Molinaro- Project administration, Roles/Writing - original draft; Writing - review & editing
Richard D. Cummings- Conceptualization, Supervision, Roles/Writing - original draft; Writing - review & editing, Funding acquisition
Christoffer K. Goth- Conceptualization, Visualization, Roles/Writing - original draft; Writing - review & editing

The authors declare no conflicts of interest.

Declaration of interests- Mehta et al.
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.



Keywords

O-glycosylation; tyrosine sulfation; post-translational modifications; glycoproteins; functional glycomics; glycan-binding proteins

Introduction

Glycans are involved in a broad spectrum of biological functions driven by their physiochemical properties. Glycans can be specifically recognized by a vast collection of proteins, including enzymes, antibodies, and lectins collectively termed Glycan-binding proteins (GBPs). Glycoproteins may have one or more glycans and some glycans, such as glycosaminoglycans, are exceptionally long polysaccharides. Conceptually, the simplest interaction is monovalent between a carbohydrate binding domain and a single glycan. However, biological interactions are often much more complex, for example due to glycan multivalency and avidity effects caused by interactions with other domains of the glycoprotein and the GBP (1). Glycans can also play important functional roles more indirectly, i.e. by affecting protein structure or other interactions, as observed for N-glycosylation in affecting protein folding, and as observed for GalNAc type O-glycosylation (hereafter O-glycosylation) in affecting proteolytic cleavage or dimerization (2). Glycans may also occupy acceptor sites otherwise utilized by other functional groups, as in O-GlcNAc and phosphorylation of serine and threonine (3). Here we summarize the growing amount of evidence that O-glycosylation is often co-localized with tyrosine sulfation in functionally important protein domains and regions, and such multi-motif interactions may serve as co-regulatory PTMs in promoting binding of GBPs.

O-glycosylation and tyrosine sulfation- important post-translational modifications (PTMs)

Sulfation of tyrosine occurs in the Golgi (Figure 1) and is directed by two protein-tyrosine sulfotransferases (TPST1/TPST2), each with different specificities and expression patterns (4). The TPSTs are believed to preferentially act on protein substrates possessing contiguous stretches of residues containing Glu/Asp clustered around tyrosines, although no consensus sequence has been determined to date. Sulfation is heterogeneous and there is no evidence of regulated dynamic hydrolysis of sulfotyrosines by human enzymes, although bacterial sulfatases are known (5) and tyrosine sulfation may also hydrolyze spontaneously under low pH. Tyrosine sulfation affects many biological functions including: host-pathogen interactions (6), proteolytic processing (7), and ligand binding (8). Proteins which recognize

phosphotyrosine do not recognize sulfotyrosines, due to differences in charge, binding geometry and size (9). Likewise, antibodies raised against sulfotyrosines do not recognize phosphotyrosine (10). Computational simulations have also suggested that electrostatic potentials for sulfotyrosine are very different in comparison with phosphotyrosine (11). *In toto,* these reports suggest that sulfotyrosine-containing proteins would most likely have unique natural binding partners, different from phosphotyrosine containing proteins.

Currently, 487 proteins from all organisms in the UniProt database (12) have been annotated to contain sulfotyrosines (Figure 2). Among them, 54 human proteins are identified with 32 proteins having direct publications linked to them. Interestingly, these 32 proteins are found to be localized either on the cell membrane or are secreted extracellularly. Additionally, the majority of these proteins are involved in important molecular functions of signaling (e.g. receptor, transducer, GPCR) or enzyme activity (e.g. protease, serine protease, hydrolase). These proteins also play vital roles in a number of biological processes, including cell adhesion, hormones, hemostasis, host-pathogen/virus interactions, inflammation and immunity.

O-glycosylation is initiated in the Golgi by up to 20 different polypeptide GalNAc-transferases (GalNAc-Ts), which add a GalNAc residue to the hydroxyl group of a serine, threonine or, less often, tyrosine (possibly competing with the TPSTs) in the protein backbone (13). The structures of O-glycans are built in a step-wise process and can be remodeled throughout the Golgi and trans Golgi network (TGN) during trafficking by a complex orchestration of specific glycosyltransferases and chaperones. There are eight different O-glycan core structures, with the most common one being core-1 (Galβ1–3GalNAc) also named the T-antigen. Core-1 to −4 can be extended into complex O-glycans, which can carry specific antigens such as ABO or Lewis structures (14). A major functional attribute of O-glycans are their diverse structures and dynamic terminal capping by a variety of monosaccharides, including sialic acids, through sialylation by up to 20 sialyltransferases; desialylation can also be driven by membrane bound and secreted neuraminidases (Neu1–4) in different cellular compartments (15). Sialyl Lewis x is an example of an important functional terminal structure, critical in normal lymphocyte activity as noted below.

P-Selectin Glycoprotein Ligand-1 (PSGL-1)

A well-characterized example of the multi-motif interaction involving O-glycosylation and tyrosine sulfation is that of PSGL-1. PSGL-1 is an adhesion receptor with a central regulatory role in immune cell trafficking (16) and is crucial for facilitating rolling of circulating leukocytes on blood vessels, platelets and endothelial cells. PSGL-1 is known to bind all selectins, although with highest affinity to P-selectin. The binding domain of PSGL-1 for P- and L-selectin is located in the N-terminus and selectin binding is dependent on the presence of a core 2 O-glycan carrying the fucosylated and sialylated epitope termed sialyl Lewis x (sLe^x) on residue threonine 57 (Thr57). The expression of sLe^x is dependent on a specific set of enzymes, which are constitutively expressed in neutrophils and many granulocytes, and also expressed in T cells upon appropriate stimuli, ensuring that only PSGL-1 on activated leukocytes can be bound by P-selectin. The relative positioning of sulfotyrosines and O-glycans is critical for interactions with P- and L-selectin, as

demonstrated by synthetic glycopeptides in which the sLe^x-containing O-glycan was moved to another site at Thr44, thereby abolishing interactions with P-selectin. Such results illustrate the importance of the position of the glycan relative to important sulfated tyrosine residues on Tyr46, 48 and 51 and other peptide determinants (17,18) (Figure 3). Several studies have also reported that PSGL-1 is capable of binding chemokines, including CCL19, CCL21 and CCL27, and that this binding is required for optimal T-cell trafficking (19); interesting, PSGL-1 and the CC chemokine receptor 7 (CCR7) have overlapping binding sites for CCL19 (20). In contrast to P-selectin binding, chemokines appear to only bind PSGL-1 on resting and not activated T-cells, potentially due to differences in glycan structures on PSGL-1 on different T-cell subsets. It is currently not known how individual sulfation sites on PSLG-1 can affect chemokine binding and how these are affected by varying glycan structures, but this creates the exciting possibility that the glycosylation on Thr44 of PSGL-1 regulates more than one biological function.

Endoglycan

Endoglycan, also known as Podocalyxin-like protein 2, is part of the CD34 family and is expressed in vascular epithelium, smooth muscle cells, and hematopoietic precursor cells (21). It is extensively O-glycosylated, has several N-glycosylation motifs and contains Ser-linked chondroitin sulfate. Unlike the other family members, endoglycan, in some similarity to PSGL-1, also has an acidic N-terminal domain carrying two tyrosine sulfation sites at position Tyr97 and Tyr118 and one O-glycosylation site at Thr124. Endoglycan is a known L-selectin ligand and the binding is dependent on the PTMs (22) with sTyr118 being more critical than sTyr 97 (23). Interestingly, endoglycan has also been suggested to be a P- and E- selectin ligand and may be an important ligand on memory and germinal center B-cells for vascular selectins (24). However, P-selectin binds with considerably higher affinity to PSGL-1 based glycosulfopeptides compared to endoglycan based homologs (23). The two other CD34 family members, CD34 and Podocalyxin, do not show evidence of carrying sulfotyrosine, but both are heavily O-glycosylated and all members have been suggested to be able to bind chemokines, although to our knowledge this is yet to be confirmed experimentally.

Chemokine receptors

The chemokine system is a central regulator of the immune system, controlling the migration and positioning of immune cells in development, homeostasis, and inflammation (25). The interaction between chemokines and chemokine receptors (CKRs) is traditionally described via a two-step model in which the CKR N-terminus interacts with the chemokine, followed by docking into the receptor binding pocket. Consequently, the N-terminal domains of CKRs are indispensable for ligand binding and receptor function. There is a high degree of promiscuity in the chemokine system with ligands, and receptors shared amongst each other and differential tissue-specific expression of both chemokine and CKRs, further add to this complexity. All CKRs contain multiple potential tyrosine sulfation sites in their extracellular N-terminus, and several of these have been experimentally identified (reviewed in (8)). Tyrosine sulfation of CKRs can enhance affinity for chemokine binding as shown for several different CKRs (26–28). However, other studies have demonstrated that the effect of

tyrosine sulfation can be chemokine specific, providing a basis for fine-tuning and biased signaling (29–31).

Most CKRs, and especially the subfamily of CC chemokine receptors (CCRs), also contain potential O-glycosylation sites in their extracellular N-terminus. For example, CCR5 which contains four sialylated O-glycans in the N-terminus, which were shown to modulate chemokine binding (32). Moreover, sialylation of CCR7 has inhibitory effects on CCL19 and CCL21 binding (33), while polysialylation on CCR7 is crucial for CCL21 binding (29). Recently, glycoproteomics approaches were used to identify 3 exact glycosylation sites in CCR7 from a T-cell line (CEM-T4), and a site in the N-terminus of CCR4 was also identified (34). Taken together, these results indicate a clear emerging pattern of O-glycosylation and tyrosine sulfation in the N-termini of all CCRs. Based on the few existing studies, these patterns could potentially be a general regulatory mechanism for chemokine binding/specificity and may represent an important missing clue to understanding the regulation and apparent promiscuity of this system.

Pathogen encoded proteins

Pathogen encoded proteins, such as viral chemokine binding proteins (CKBP) or tick-derived evasins provide additional clues that multi-motif interactions involving tyrosine sulfation and glycosylation may play a general role in chemokine interactions. Several virally encoded CKBP are known to be glycoproteins (35,36), but generally the glycosylation sites and the effect of glycosylation are not known. Recent studies reported the exact glycosylation sites of several different virus strains, revealing interesting patterns comparable to the human CKR (37). The cytomegalovirus (CMV) CKRs US28 and UL78 were both observed to carry several N-terminal O-glycosylation sites and both likewise carry N-terminal sulfotyrosines important for ligand binding (38). Similarly, several O-glycosylation sites were detected in the CMV CKBP CL22a and these were in close proximity to a known functionally important tyrosine sulfation site (39). Evasins are produced and secreted by ticks, together with several other factors in order to modulate the hosts' immune system, inhibit leukocyte recruitment, and prolong attachment (40). A systematic analysis recently identified more than 250 putative evasins, and sequence alignment revealed that they generally contain predicted tyrosine sulfation sites and N-glycosylation sequons (41). O-glycosylation was not investigated, but several evasins carry proline-, serine- and threonine-rich stretches predicted to be O-glycosylated. The thrombin inhibitor Mandanin-1, also produced by ticks, exhibited a similar pattern with in which two tyrosine residues at 32 and 35 are sulfated, along with the presence of O-glycans (42). Although the positions of the glycosylation sites were not determined, close-by residues 40, 41, and 43 are all predicted to be O-glycosites using the NetOGlyc 4.0 server (43). Another thrombin inhibitor Hirudin P6 produced by the leech *Hirudinaria manillensis,* contains both sulfotyrosine and O-glycans (44). A recent study deconvoluted structural aspects of thrombin inhibition and demonstrated that, whereas addition of O-glycans increases the thrombin inhibitory activity compared to unmodified peptide, tyrosine sulfation exhibits two-orders of magnitude increase in inhibitory activity; however, the addition of glycans to the peptide with tyrosine sulfation did not lead to an improvement in activity (45). It is

possible in such cases that the glycosylation plays an extrinsic role in the stability of the peptide, rather than an intrinsic role of improving binding affinity.

Hemostasis System

The hemostatic system plays a critical role in preventing bleeding during injury, together with the immune and inflammatory responses. It is not surprising therefore, that the hemostatic system would be a part of the nexus tightly regulated by several PTMs. There is strong evidence to suggest the presence of sulfotyrosines on a number of proteins in the coagulation cascade and hemostasis pathway, e.g., coagulation factor VIII (46) (47), factor IX (48), heparin cofactor 2 (49), alpha-2-antiplasmin (50) and fibrinogen gamma chain (51); and in platelet proteins such as glycoprotein Ib alpha (GPIbα) (52). Interestingly, many of these proteins have binding partners which also have been shown to bind heparin; hence, the interplay between heparin and sulfotyrosines interactions could be critical in several hemostatic pathways. In particular, the interaction between GPIbα and thrombin has been shown via crystal structure and NMR studies to be heavily dependent on sulfotyrosine residues binding to exosite 2 of thrombin (53), a site on thrombin also known to bind heparin. GPIbα has three sulfotyrosine sites clustered at Tyr292, Tyr294, and Tyr295, and there is a known glycosylation site nearby at Thr308 (54). Yet structural studies often omit examination of the role of glycans due to the intrinsic difficulty of studying large glycoprotein structures. An earlier study had found “altered glycosylation” to have no effect on thrombin and von Willebrand factor binding, however the change in glycosylation was not sufficiently characterized (55). Additionally, binding of sulfotyrosine residues on thrombin exerts an allosteric effect on the enzyme. A comprehensive glycoproteomics study on human plasma, platelets and endothelial cells, revealed the exact glycosylation sites in many central players of the hemostatic system (56) and observed co-localization with known tyrosine sulfation sites in several examples.

Multi-PTM crossroads?

As presented in Figure 4, additional PTMs may also be present in the regions where O-glycosylation and tyrosine sulfation co-localize. Extracellular phosphorylation on serines and threonines is carried out by the Fam20 family of kinases (57) whereas extracellular tyrosine sulfation can be catalysed by vertebrate lonesome kinase (VLK) (58) and several examples of direct overlap - i.e. residues that can be either O-glycosylated or phosphorylated - have been identified (2). The chemokine receptors generally have a conserved cysteine in their N-termini, which forms a disulfide bridge with ECL3. N-glycosylation is also observed in several examples as evident for Factor IX and motifs exists in several CCR N-termini especially in the CXCR family of chemokine receptors and in the tick-encoded evasins. Proteolytic cleavage also occur in close proximity to glyco-sulfo multi-motif patterns in several of these regions including in PSGL-1, which can be cleaved by neutrophil elastase and cathepsin, abolishing P-selectin binding (59). Osteopontin is a particular complex example and can be cleaved by multiple proteases at slightly different positions including thrombin, metalloproteases, cathepsin and plasmin (60), Osteopontin is O-glycosylated and tyrosine sulfated (61), but furthermore contains tyrosine/serine/threonine phosphorylation (58,61) all within the same region. The snake venom metalloproteinase mocarhagin generally recognizes sulfotyrosine and also has lectin activity (62). It is known to cleave

PSGL-1(63) and several key glycoproteins in the hemostatic system (64). Collectively, these observations point to even more complex regulation by the added layer of multiple PTMs, and the potential combinatorial effect of these modifications exponentially increase the number of possible proteoforms.

Methods for studying glyco-sulfo multi-motifs

Both sulfation and glycosylation are often difficult to define in terms of their positions and structures. Tyrosine sulfates are labile by nature, have a small size, and acceptor site heterogeneity. Differentiation between sulfotyrosine and phosphotyrosine required advanced mass spectrometric (MS) techniques. O-glycosylation exhibits both macro- and microheterogeneity and differs between individual cell types, tissues, and differentiation states. Historically, sulfation of tyrosine has been studied by used of metabolic labelling potentially combined with mutation of tyrosine to phenylalanine. More recently, an elegant method was developed by use of an expanded genetic code and modified tRNA synthase to incorporate tyrosine sulfation in specific residues in a bacterial system (65), and novel inhibitors of the TPSTs may also aid our understanding (66). Synthetic sulfopeptides have been applied to decipher the function of individual sulfation sites as well as in structural studies (reviewed in (67)). Sulfopeptides are usually generated enzymatically by use of recombinant TPSTs (68) or by *in vitro* synthesis using protected tyrosine sulfate building blocks (30). The latter has the advantage of accurate control of individual sites if the peptide sequence encompasses multiple tyrosine acceptor sites. The size of the sulfoproteome is still unknown, but has been estimated that up to 2% of the entire proteome may contain sulfotyrosine, which is likely to expand as MS methods and sensitivity advances. Enrichment of sulfopeptides has been accomplished with antibodies and titanium oxide columns, and new approaches to distinguish between tyrosine sulfation and the nominally isobaric tyrosine phosphorylation have also been developed by taking advantage of the instability of sulfation during fragmentation (69).

Recent advances in genetic engineering and MS have accelerated studies in glycobiology. An immense amount of information on localization of O-glycosylation sites is being generated, driving the emergence of specific patterns and functions, and has improved our ability to predict where sites are located (70). The ability to present proteins of interest in engineered cell systems with defined glycosylation capacities by knockin, knockout (71) or even tunable enzyme levels (72), are highly applicable tools. Yet, even these approaches do not permit complete control on modifications of individual sites, either in terms of occupancies or glycan structure. Synthetic glyco- and sulfopeptides have provided important insights into molecular interactions and biological functions of individual residues and modifications. Still, a relevant concern is whether secondary structures of such small peptides reflect that of the full length proteins from which they are designed. Several structures of sulfopeptides derived from chemokine receptors bound to different chemokines have been published and the interface of the interactions generally corresponds well to what is known from the full- length receptors. However, the orientation, at least in one case, was incompatible with docking of the bound chemokine to the second binding site within the second extracellular loop, although this could be explained by down-stream structural re-arrangements (73). Several groups have used modified PSGL-1 peptides to decipher the role

of both sulfation and glycosylation. The crystal structure of a glyco-sulfopeptide (GSP) in complex with P-selectin revealed how the sulfated tyrosine at position 10 interacts with neighboring leucines to facilitate salt bridge formation with residues in P-selectin (74) (Figure 3). Other studies revealed the importance of individual tyrosines (18), and a glycosylated and sulfonated PSGL-1 structure was shown to be similar to native PSGL-1 by molecular dynamics simulation (75). Collectively, glyco- and sulfopeptides continue to be a powerful tool and can be extended to include other modifications to accurately dissect the roles of individual PTMs on single residues. Furthermore, such peptides can aid in understanding the structural aspects of these interactions and may serve as MS standards to address another barrier - the structures and occupancies of PTMs on individual proteins from primary cells and tissues.

Conclusions

There is now overwhelming evidence that the co-localization of O-glycosylation and tyrosine sulfation in a multi-motif presentation is a much more prevalent phenomenon than previously appreciated. O-glycoproteomic methods have advanced in recent years although there are still challenges and the exploration of the ‘sulfome’ is in its infancy. However, technological advances now allow for the field to begin addressing these patterns in more detail. The co-localization seems to occur in functionally important domains and is often involved in protein-protein interactions or ligand binding. Consequently, the understanding of these patterns could lead to new fundamental insights and open new avenues of drug development, by either mimicking the glyco-sulfo patterns or by targeting PTM specific subsets of individual proteins.

Acknowledgements

Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311 and Independent Research Fund Denmark (7025-00083B) for CKG, and grant P41GM103694 to RDC.

References

1. Cummings RD (2019) “Stuck on sugars - how carbohydrates regulate cell adhesion, recognition, and signaling”. Glycoconj. J 36, 241–257 [PubMed: 31267247]
2. Goth CK, Vakhrushev SY, Joshi HJ, Clausen H, and Schjoldager KT (2018) Fine-Tuning Limited Proteolysis: A Major Role for Regulated Site-Specific O-Glycosylation. Trends Biochem. Sci 43, 269–284 [PubMed: 29506880]
3. Hart GW, Greis KD, Dong LY, Blomberg MA, Chou TY, Jiang MS, Roquemore EP, Snow DM, Kreppel LK, Cole RN, and et al. (1995) O-linked N-acetylglucosamine: the “yin-yang” of Ser/Thr phosphorylation? Nuclear and cytoplasmic glycosylation. Adv. Exp. Med. Biol 376, 115–123 [PubMed: 8597237]
4. Mishiro E, Sakakibara Y, Liu MC, and Suiko M (2006) Differential enzymatic characteristics and tissue-specific expression of human TPST-1 and TPST-2. J. Biochem 140, 731–737 [PubMed: 17028309]
5. Wilkins PP, Moore KL, McEver RP, and Cummings RD (1995) Tyrosine sulfation of P-selectin glycoprotein ligand-1 is required for high affinity binding to P-selectin. J. Biol. Chem 270, 22677–22680 [PubMed: 7559387]

6. Farzan M, Mirzabekov T, Kolchinsky P, Wyatt R, Cayabyab M, Gerard NP, Gerard C, Sodroski J, and Choe H (1999) Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. Cell 96, 667–676 [PubMed: 10089882]
7. Bundgaard JR, Vuust J, and Rehfeld JF (1995) Tyrosine O-sulfation promotes proteolytic processing of progastrin. EMBO J 14, 3073–3079 [PubMed: 7621822]
8. Ludeman JP, and Stone MJ (2014) The structural role of receptor tyrosine sulfation in chemokine recognition. Br. J. Pharmacol 171, 1167–1179 [PubMed: 24116930]
9. Ju T, Niu W, Cerny R, Bollman J, Roy A, and Guo J (2013) Molecular recognition of sulfotyrosine and phosphotyrosine by the Src homology 2 domain. Mol Biosyst 9, 1829–1832 [PubMed: 23567872]
10. Hoffhines AJ, Damoc E, Bridges KG, Leary JA, and Moore KL (2006) Detection and purification of tyrosine-sulfated proteins using a novel anti-sulfotyrosine monoclonal antibody. J Biol Chem 281, 37877–37887 [PubMed: 17046811]
11. Rapp C, Klerman H, Levine E, and McClendon CL (2013) Hydrogen bond strengths in phosphorylated and sulfated amino acid residues. PLoS One 8, e57804
12. UniProt C (2019) UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res 47, D506–D515 [PubMed: 30395287]
13. Bennett EP, Mandel U, Clausen H, Gerken TA, Fritz TA, and Tabak LA (2012) Control of mucin-type O-glycosylation: a classification of the polypeptide GalNAc-transferase gene family. Glycobiology 22, 736–756 [PubMed: 22183981]
14. Brockhausen I, and Stanley P (2015) O-GalNAc Glycans in Essentials of Glycobiology (rd, Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, and Seeberger PH eds.), Cold Spring Harbor (NY) pp 113–123
15. Varki A, Schnaar RL, and Schauer R (2015) Sialic Acids and Other Nonulosonic Acids in Essentials of Glycobiology (rd, Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, and Seeberger PH eds.), Cold Spring Harbor (NY) pp 179–195
16. Tinoco R, Carrette F, Barraza ML, Otero DC, Magana J, Bosenberg MW, Swain SL, and Bradley LM (2016) PSGL-1 Is an Immune Checkpoint Regulator that Promotes T Cell Exhaustion. Immunity 44, 1470 [PubMed: 27332735]
17. Leppanen A, Mehta P, Ouyang YB, Ju T, Helin J, Moore KL, van Die I, Canfield WM, McEver RP, and Cummings RD (1999) A novel glycosulfopeptide binds to P-selectin and inhibits leukocyte adhesion to P- selectin. J. Biol. Chem 274, 24838–24848 [PubMed: 10455156]
18. Leppanen A, White SP, Helin J, McEver RP, and Cummings RD (2000) Binding of glycosulfopeptides to P-selectin requires stereospecific contributions of individual tyrosine sulfate and sugar residues. J. Biol. Chem 275, 39569–39578 [PubMed: 10978329]
19. Veerman KM, Williams MJ, Uchimura K, Singer MS, Merzaban JS, Naus S, Carlow DA, Owen P, Rivera-Nieves J, Rosen SD, and Ziltener HJ (2007) Interaction of the selectin ligand PSGL-1 with chemokines CCL21 and CCL19 facilitates efficient homing of T cells to secondary lymphoid organs. Nat. Immunol 8, 532–539 [PubMed: 17401367]
20. Veldkamp CT, Kiermaier E, Gabel-Eissens SJ, Gillitzer ML, Lippner DR, DiSilvio FA, Mueller CJ, Wantuch PL, Chaffee GR, Famiglietti MW, Zgoba DM, Bailey AA, Bah Y, Engebretson SJ, Graupner DR, Lackner ER, LaRosa VD, Medeiros T, Olson ML, Phillips AJ, Pyles H, Richard AM, Schoeller SJ, Touzeau B, Williams LG, Sixt M, and Peterson FC (2015) Solution Structure of CCL19 and Identification of Overlapping CCR7 and PSGL-1 Binding Sites. Biochemistry 54, 4163–4166 [PubMed: 26115234]
21. Sassetti C, Van Zante A, and Rosen SD (2000) Identification of endoglycan, a member of the CD34/podocalyxin family of sialomucins. J. Biol. Chem 275, 9001–9010 [PubMed: 10722749]
22. Fieger CB, Sassetti CM, and Rosen SD (2003) Endoglycan, a member of the CD34 family, functions as an L-selectin ligand through modification with tyrosine sulfation and sialyl Lewis x. J. Biol. Chem 278, 27390–27398 [PubMed: 12889478]
23. Leppanen A, Parviainen V, Ahola-Iivarinen E, Kalkkinen N, and Cummings RD (2010) Human L-selectin preferentially binds synthetic glycosulfopeptides modeled after endoglycan and containing

tyrosine sulfate residues and sialyl Lewis x in core 2 O-glycans. Glycobiology 20, 1170–1185 [PubMed: 20507883]

24. Kerr SC, Fieger CB, Snapp KR, and Rosen SD (2008) Endoglycan, a member of the CD34 family of sialomucins, is a ligand for the vascular selectins. J. Immunol 181, 1480–1490 [PubMed: 18606703]
25. Solari R, Pease JE, and Begg M (2015) "Chemokine receptors as therapeutic targets: Why aren't there more drugs?". Eur. J. Pharmacol 746, 363–367 [PubMed: 25016087]
26. Phillips AJ, Taleski D, Koplinski CA, Getschman AE, Moussouras NA, Richard AM, Peterson FC, Dwinell MB, Volkman BF, Payne RJ, and Veldkamp CT (2017) CCR7 Sulfotyrosine Enhances CCL21 Binding. International journal of molecular sciences 18
27. Tan JH, Ludeman JP, Wedderburn J, Canals M, Hall P, Butler SJ, Taleski D, Christopoulos A, Hickey MJ, Payne RJ, and Stone MJ (2013) Tyrosine sulfation of chemokine receptor CCR2 enhances interactions with both monomeric and dimeric forms of the chemokine monocyte chemoattractant protein-1 (MCP-1). J. Biol. Chem 288, 10024–10034 [PubMed: 23408426]
28. Veldkamp CT, Seibert C, Peterson FC, Sakmar TP, and Volkman BF (2006) Recognition of a CXCR4 sulfotyrosine by the chemokine stromal cell-derived factor-1alpha (SDF-1alpha/CXCL12). J. Mol. Biol 359, 1400–1409 [PubMed: 16725153]
29. Kiermaier E, Moussion C, Veldkamp CT, Gerardy-Schahn R, de Vries I, Williams LG, Chaffee GR, Phillips AJ, Freiberger F, Imre R, Taleski D, Payne RJ, Braun A, Forster R, Mechtler K, Muhlenhoff M, Volkman BF, and Sixt M (2016) Polysialylation controls dendritic cell trafficking by regulating chemokine recognition. Science 351, 186–190 [PubMed: 26657283]
30. Simpson LS, Zhu JZ, Widlanski TS, and Stone MJ (2009) Regulation of chemokine recognition by site-specific tyrosine sulfation of receptor peptides. Chem. Biol 16, 153–161 [PubMed: 19246006]
31. Zhu JZ, Millard CJ, Ludeman JP, Simpson LS, Clayton DJ, Payne RJ, Widlanski TS, and Stone MJ (2011) Tyrosine sulfation influences the chemokine binding selectivity of peptides derived from chemokine receptor CCR3. Biochemistry 50, 1524–1534 [PubMed: 21235238]
32. Bannert N, Craig S, Farzan M, Sogah D, Santo NV, Choe H, and Sodroski J (2001) Sialylated O-glycans and sulfated tyrosines in the NH2-terminal domain of CC chemokine receptor 5 contribute to high affinity binding of chemokines. J. Exp. Med 194, 1661–1673 [PubMed: 11733580]
33. Hauser MA, Kindinger I, Laufer JM, Spate AK, Bucher D, Vanes SL, Krueger WA, Wittmann V, and Legler DF (2016) Distinct CCR7 glycosylation pattern shapes receptor signaling and endocytosis to modulate chemotactic responses. J. Leukoc. Biol 99, 993–1007 [PubMed: 26819318]
34. Yang W, Ao M, Hu Y, Li QK, and Zhang H (2018) Mapping the O-glycoproteome using site-specific extraction of O-linked glycopeptides (EXoO). Mol. Syst. Biol 14, e8486
35. Van de Walle GR, May ML, Sukhumavasi W, von Einem J, and Osterrieder N (2007) Herpesvirus chemokine-binding glycoprotein G (gG) efficiently inhibits neutrophil chemotaxis in vitro and in vivo. J. Immunol 179, 4161–4169 [PubMed: 17785855]
36. Viejo-Borbolla A, Martinez-Martin N, Nel HJ, Rueda P, Martin R, Blanco S, Arenzana-Seisdedos F, Thelen M, Fallon PG, and Alcami A (2012) Enhancement of chemokine function as an immunomodulatory strategy employed by human herpesviruses. PLoS Pathog 8, e1002497
37. Bagdonaite I, Norden R, Joshi HJ, King SL, Vakhrushev SY, Olofsson S, and Wandall HH (2016) Global Mapping of O-Glycosylation of Varicella Zoster Virus, Human Cytomegalovirus, and Epstein-Barr Virus. J. Biol. Chem 291, 12014–12028 [PubMed: 27129252]
38. Casarosa P, Waldhoer M, LiWang PJ, Vischer HF, Kledal T, Timmerman H, Schwartz TW, Smit MJ, and Leurs R (2005) CC and CX3C chemokines differentially interact with the N terminus of the human cytomegalovirus-encoded US28 receptor. J. Biol. Chem 280, 3275–3285 [PubMed: 15546882]
39. Wang X, Sanchez J, Stone MJ, and Payne RJ (2017) Sulfation of the Human Cytomegalovirus Protein UL22A Enhances Binding to the Chemokine RANTES. Angew. Chem. Int. Ed. Engl 56, 8490–8494 [PubMed: 28488292]
40. Bonvin P, Power CA, and Proudfoot AE (2016) Evasins: Therapeutic Potential of a New Family of Chemokine-Binding Proteins from Ticks. Front. Immunol 7, 208 [PubMed: 27375615]

41. Hayward J, Sanchez J, Perry A, Huang C, Rodriguez Valle M, Canals M, Payne RJ, and Stone MJ (2017) Ticks from diverse genera encode chemokine-inhibitory evasin proteins. J. Biol. Chem 292, 15670–15680 [PubMed: 28778927]
42. Thompson RE, Liu X, Ripoll-Rozada J, Alonso-Garcia N, Parker BL, Pereira PJB, and Payne RJ (2017) Tyrosine sulfation modulates activity of tick-derived thrombin inhibitors. Nature chemistry 9, 909–917
43. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-Silva L, Gupta R, Bennett EP, Mandel U, Brunak S, Wandall HH, Levery SB, and Clausen H (2013) Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. EMBOJ 32, 1478–1488
44. Steiner V, Knecht R, Bornsen KO, Gassmann E, Stone SR, Raschdorf F, Schlaeppi JM, and Maschler R (1992) Primary structure and function of novel O-glycosylated hirudins from the leech Hirudinaria manillensis. Biochemistry 31, 2294–2298 [PubMed: 1540584]
45. Hsieh YS, Wijeyewickrema LC, Wilkinson BL, Pike RN, and Payne RJ (2014) Total synthesis of homogeneous variants of hirudin P6: a post-translationally modified anti-thrombotic leech-derived protein. Angew. Chem. Int. Ed. Engl 53, 3947–3951 [PubMed: 24615823]
46. Severs JC, Carnine M, Eguizabal H, and Mock KK (1999) Characterization of tyrosine sulfate residues in antihemophilic recombinant factor VIII by liquid chromatography electrospray ionization tandem mass spectrometry and amino acid analysis. Rapid Commun Mass Spectrom 13, 1016–1023 [PubMed: 10368977]
47. Leyte A, van Schijndel HB, Niehrs C, Huttner WB, Verbeet MP, Mertens K, and van Mourik JA (1991) Sulfation of Tyr1680 of human blood coagulation factor VIII is essential for the interaction of factor VIII with von Willebrand factor. J. Biol. Chem 266, 740–746 [PubMed: 1898735]
48. Arruda VR, Hagstrom JN, Deitch J, Heiman-Patterson T, Camire RM, Chu K, Fields PA, Herzog RW, Couto LB, Larson PJ, and High KA (2001) Posttranslational modifications of recombinant myotube-synthesized human factor IX. Blood 97, 130–138 [PubMed: 11133752]
49. Baglin TP, Carrell RW, Church FC, Esmon CT, and Huntington JA (2002) Crystal structures of native and thrombin-complexed heparin cofactor II reveal a multistep allosteric mechanism. Proc Natl Acad Sci U S A 99, 11079–11084 [PubMed: 12169660]
50. Hortin G, Fok KF, Toren PC, and Strauss AW (1987) Sulfation of a tyrosine residue in the plasmin-binding domain of alpha 2-antiplasmin. J Biol Chem 262, 3082–3085 [PubMed: 2434496]
51. Meh DA, Siebenlist KR, Brennan SO, Holyst T, and Mosesson MW (2001) The amino acid sequence in fibrin responsible for high affinity thrombin binding. Thromb Haemost 85, 470–474 [PubMed: 11307817]
52. Uff S, Clemetson JM, Harrison T, Clemetson KJ, and Emsley J (2002) Crystal structure of the platelet glycoprotein Ib(alpha) N-terminal domain reveals an unmasking mechanism for receptor activation. J Biol Chem 277, 35657–35663 [PubMed: 12087105]
53. Lechtenberg BC, Freund SM, and Huntington JA (2014) GpIbalpha interacts exclusively with exosite II of thrombin. Journal of molecular biology 426, 881–893 [PubMed: 24316004]
54. Titani K, Takio K, Handa M, and Ruggeri ZM (1987) Amino acid sequence of the von Willebrand factor-binding domain of platelet membrane glycoprotein Ib. Proc Natl Acad Sci U S A 84, 5610–5614 [PubMed: 3497398]
55. Li CQ, Ye P, Cao Z, Wang H, Lu L, Nicastro P, Wood E, Robert JJ, Ouwehand WH, Hill F, Lopez JA, and Wardell MR (2001) Expression of the amino-terminal domain of platelet glycoprotein Ib alpha: exploitation of a calmodulin tag for determination of its functional activity. Protein Expr Purif 22, 200–210 [PubMed: 11437595]
56. King SL, Joshi HJ, Schjoldager KT, Halim A, Madsen TD, Dziegiel MH, Woetmann A, Vakhrushev SY, and Wandall HH (2017) Characterizing the O-glycosylation landscape of human plasma, platelets, and endothelial cells. Blood Adv 1, 429–442 [PubMed: 29296958]
57. Tagliabracci VS, Wiley SE, Guo X, Kinch LN, Durrant E, Wen J, Xiao J, Cui J, Nguyen KB, Engel JL, Coon JJ, Grishin N, Pinna LA, Pagliarini DJ, and Dixon JE (2015) A Single Kinase Generates the Majority of the Secreted Phosphoproteome. Cell 161, 1619–1632 [PubMed: 26091039]

58. Bordoli MR, Yum J, Breitkopf SB, Thon JN, Italiano JE Jr., Xiao J, Worby C,Wong SK, Lin G, Edenius M, Keller TL, Asara JM, Dixon JE, Yeo CY, and Whitman M (2014) A secreted tyrosine kinase acts in the extracellular environment. Cell 158, 1033–1044 [PubMed: 25171405]
59. Gardiner EE, De Luca M, McNally T, Michelson AD, Andrews RK, and Berndt MC (2001) Regulation of P-selectin binding to the neutrophil P-selectin counter-receptor P-selectin glycoprotein ligand-1 by neutrophil elastase and cathepsin G. Blood 98, 1440–1447 [PubMed: 11520793]
60. Christensen B, Schack L, Klaning E, and Sorensen ES (2010) Osteopontin is cleaved at multiple sites close to its integrin-binding motifs in milk and is a novel substrate for plasmin and cathepsin D. J. Biol. Chem 285, 7929–7937 [PubMed: 20071328]
61. Christensen B, Petersen TE, and Sorensen ES (2008) Post-translational modification and proteolytic processing of urinary osteopontin. Biochem. J 411, 53–61 [PubMed: 18072945]
62. Ward CM, Vinogradov DV, Andrews RK, and Berndt MC (1996) Characterization of mocarhagin, a cobra venom metalloproteinase from Naja mocambique mocambique, and related proteins from other Elapidae venoms. Toxicon 34, 1203–1206 [PubMed: 8931262]
63. De Luca M, Dunlop LC, Andrews RK, Flannery JV Jr., Ettling R, Cumming DA, Veldman GM, and Berndt MC (1995) A novel cobra venom metalloproteinase, mocarhagin, cleaves a 10-amino acid peptide from the mature N terminus of P-selectin glycoprotein ligand receptor, PSGL-1, and abolishes P-selectin binding. J. Biol. Chem 270, 26734–26737 [PubMed: 7592904]
64. Ward CM, Andrews RK, Smith AI, and Berndt MC (1996) Mocarhagin, a novel cobra venom metalloproteinase, cleaves the platelet von Willebrand factor receptor glycoprotein Ibalpha. Identification of the sulfated tyrosine/anionic sequence Tyr-276-Glu-282 of glycoprotein Ibalpha as a binding site for von Willebrand factor and alpha-thrombin. Biochemistry 35, 4929–4938 [PubMed: 8664285]
65. Li X, and Liu CC (2018) Site-Specific Incorporation of Sulfotyrosine Using an Expanded Genetic Code. Methods Mol. Biol 1728, 191–200 [PubMed: 29404999]
66. Byrne DP, Li Y, Ngamlert P, Ramakrishnan K, Eyers CE, Wells C, Drewry DH, Zuercher WJ, Berry NG, Fernig DG, and Eyers PA (2018) New tools for evaluating protein tyrosine sulfation: tyrosylprotein sulfotransferases (TPSTs) are novel targets for RAF protein kinase inhibitors. Biochem. J 475, 2435–2455 [PubMed: 29934490]
67. Szpakowska M, Fievez V, Arumugan K, van Nuland N, Schmit JC, and Chevigne A (2012) Function, diversity and therapeutic potential of the N-terminal domain of human chemokine receptors. Biochem. Pharmacol 84, 1366–1380 [PubMed: 22935450]
68. Seibert C, Sanfiz A, Sakmar TP, and Veldkamp CT (2016) Preparation and Analysis of N-Terminal Chemokine Receptor Sulfopeptides Using Tyrosylprotein Sulfotransferase Enzymes. Methods Enzymol 570, 357–388 [PubMed: 26921955]
69. Chen G, Zhang Y, Trinidad JC, and Dann C 3rd. (2018) Distinguishing Sulfotyrosine Containing Peptides from their Phosphotyrosine Counterparts Using Mass Spectrometry. J. Am. Soc. Mass Spectrom 29, 455–462 [PubMed: 29313205]
70. Steentoft C, Bennett EP, Schjoldager KT, Vakhrushev SY, Wandall HH, and Clausen H (2014) Precision genome editing: a small revolution for glycobiology. Glycobiology 24, 663–680 [PubMed: 24861053]
71. Narimatsu Y, Joshi HJ, Nason R, Van Coillie J, Karlsson R, Sun L, Ye Z, Chen YH, Schjoldager KT, Steentoft C, Furukawa S, Bensing BA, Sullam PM, Thompson AJ, Paulson JC, Bull C, Adema GJ, Mandel U, Hansen L, Bennett EP, Varki A, Vakhrushev SY, Yang Z, and Clausen H (2019) An Atlas of Human Glycosylation Pathways Enables Display of the Human Glycome by Gene Engineered Cells. Mol. Cell 75, 394–407 e395 [PubMed: 31227230]
72. Hintze J, Ye Z, Narimatsu Y, Madsen TD, Joshi HJ, Goth CK, Linstedt A, Bachert C, Mandel U, Bennett EP, Vakhrushev SY, and Schjoldager KT (2018) Probing the contribution of individual polypeptide GalNAc-transferase isoforms to the O-glycoproteome by inducible expression in isogenic cell lines. J. Biol. Chem 293, 19064–19077 [PubMed: 30327431]
73. Kufareva I, Salanga CL, and Handel TM (2015) Chemokine and chemokine receptor structure and interactions: implications for therapeutic strategies. Immunol. Cell Biol 93, 372–383 [PubMed: 25708536]

74. Somers WS, Tang J, Shaw GD, and Camphausen RT (2000) Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to SLe(X) and PSGL-1. Cell 103, 467–479 [PubMed: 11081633]

75. Krishnamurthy VR, Sardar MY, Ying Y, Song X, Haller C, Dai E, Wang X, Hanjaya-Putra D, Sun L, Morikis V, Simon SI, Woods RJ, Cummings RD, and Chaikof EL (2015) Glycopeptide analogues of PSGL-1 inhibit P-selectin in vitro and in vivo. Nature communications 6, 6387

76. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, and Ferrin TE (2004) UCSF Chimera--a visualization system for exploratory research and analysis. J. Comput. Chem 25, 1605–1612 [PubMed: 15264254]

Highlights

- A growing body of evidence shows O-glycosylation and tyrosine sulfation co-localize in important functional protein domains in a broad range of protein classes.
- Recent advances in methods and technologies now allow for more detailed dissection of these patterns.
- The co-localization of O-glycosylation and tyrosine sulfation occur in functionally important domains making it highly relevant for future drug discovery.

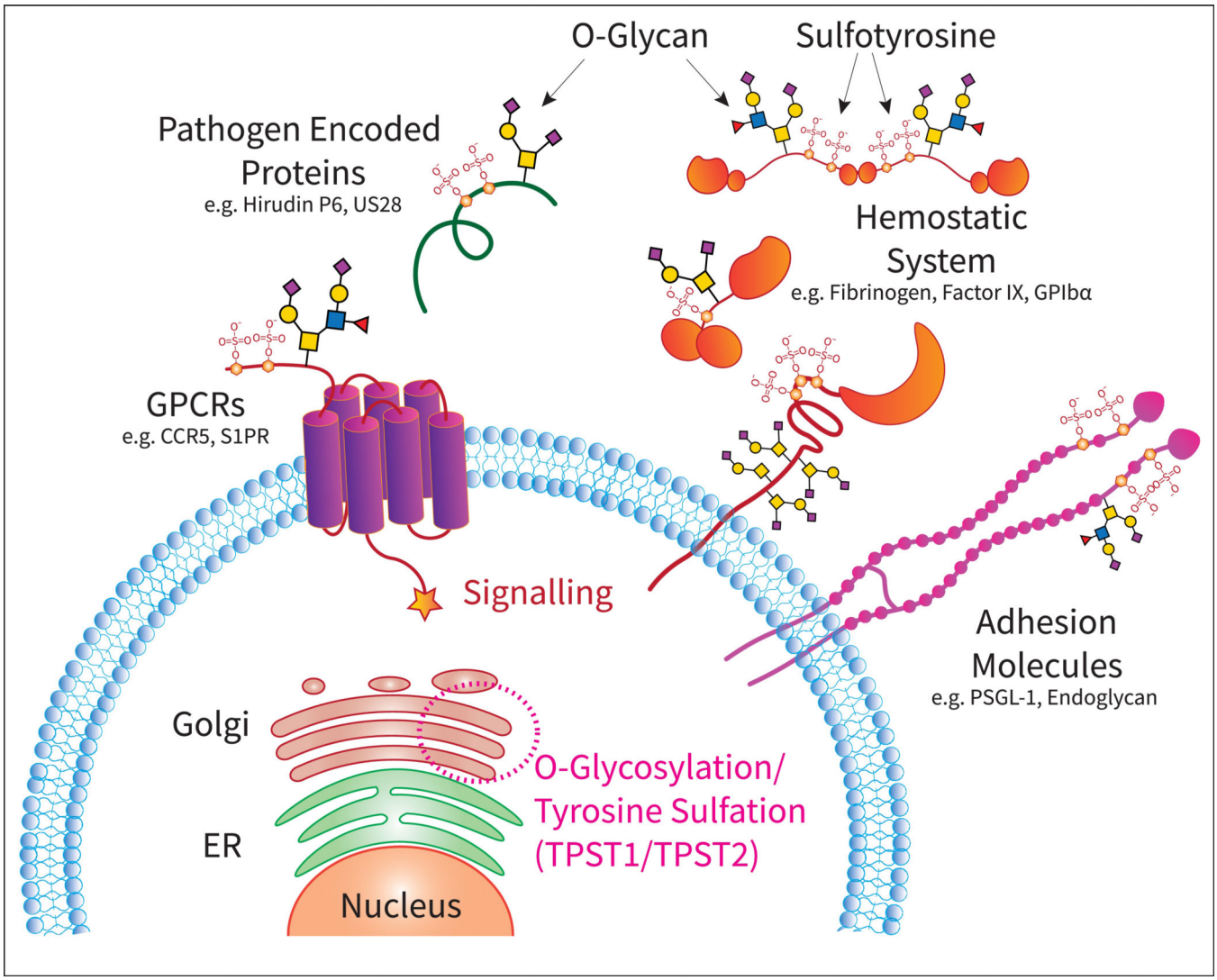


Figure 1.

Overview of cellular glycosylation and sulfation, post-translational modifications (PTMs) that take place in the ER and Golgi network, and generates glycosylated and sulfated proteins that are critical in biological processes such as adhesion, signaling through G-coupled protein receptors (GPCRs), and hemostasis.

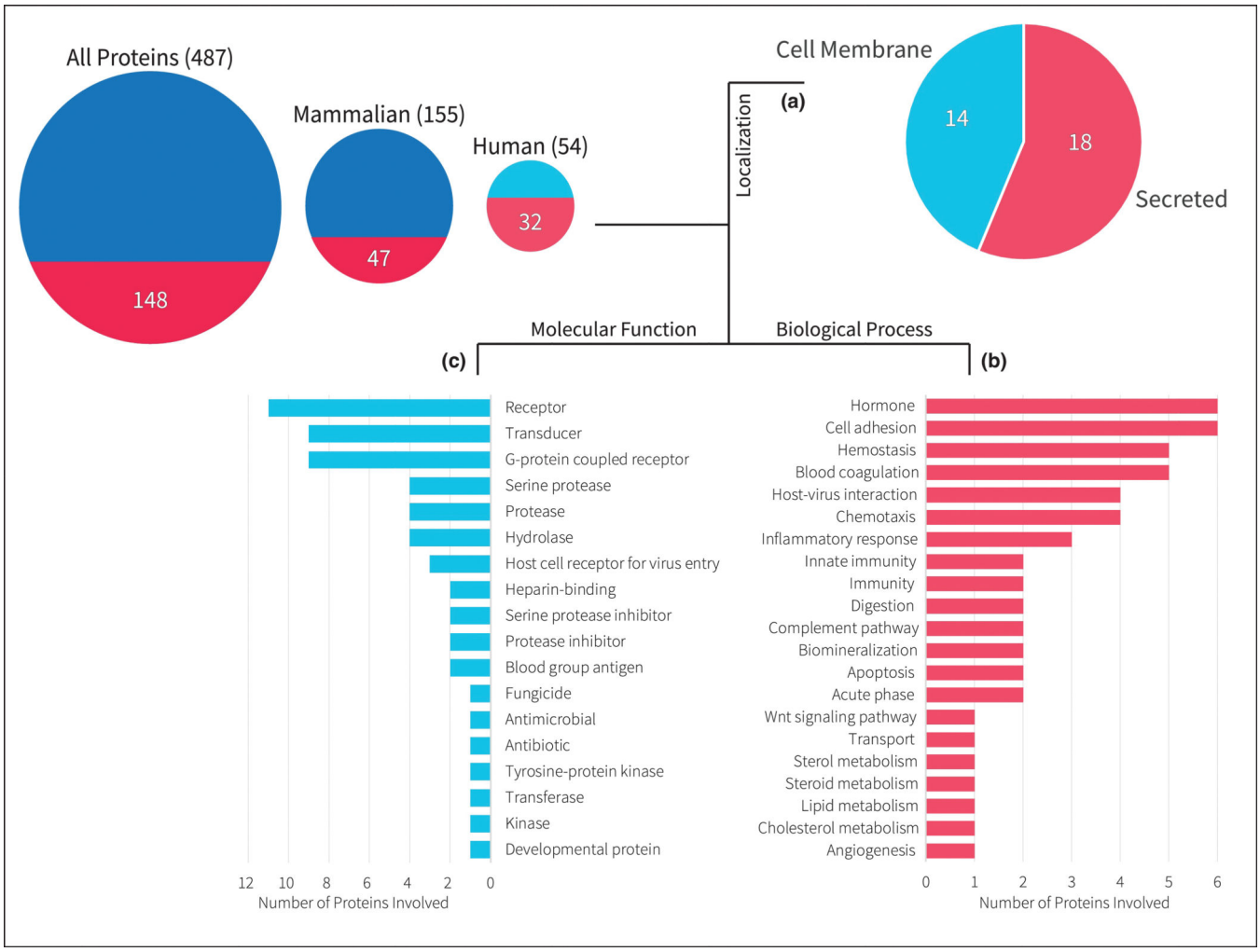


Figure 2.

Shows the number of proteins in UniProt with sulfotyrosines. The left most circle represents the data for all proteins, the center circle represents data for mammalian proteins (i.e. human, mouse, rat, bovine, rabbit), the right most circle represents data for only human proteins. The area of the circles represents the number of proteins in UniProt with known sulfotyrosine sites either through direct evidence by publication or by sequence similarity. The bottom semicircle represents the proteins with direct evidence by publications among these total proteins. Out of the 32 published proteins for humans, (A) 14 of them are localized on the cell membrane, while 18 of them are secreted; while none are found intracellularly either in the cytoplasm or nucleus. (B) Shows the biological processes these 32 proteins are involved in. (C) Shows the molecular functions these 32 proteins perform.

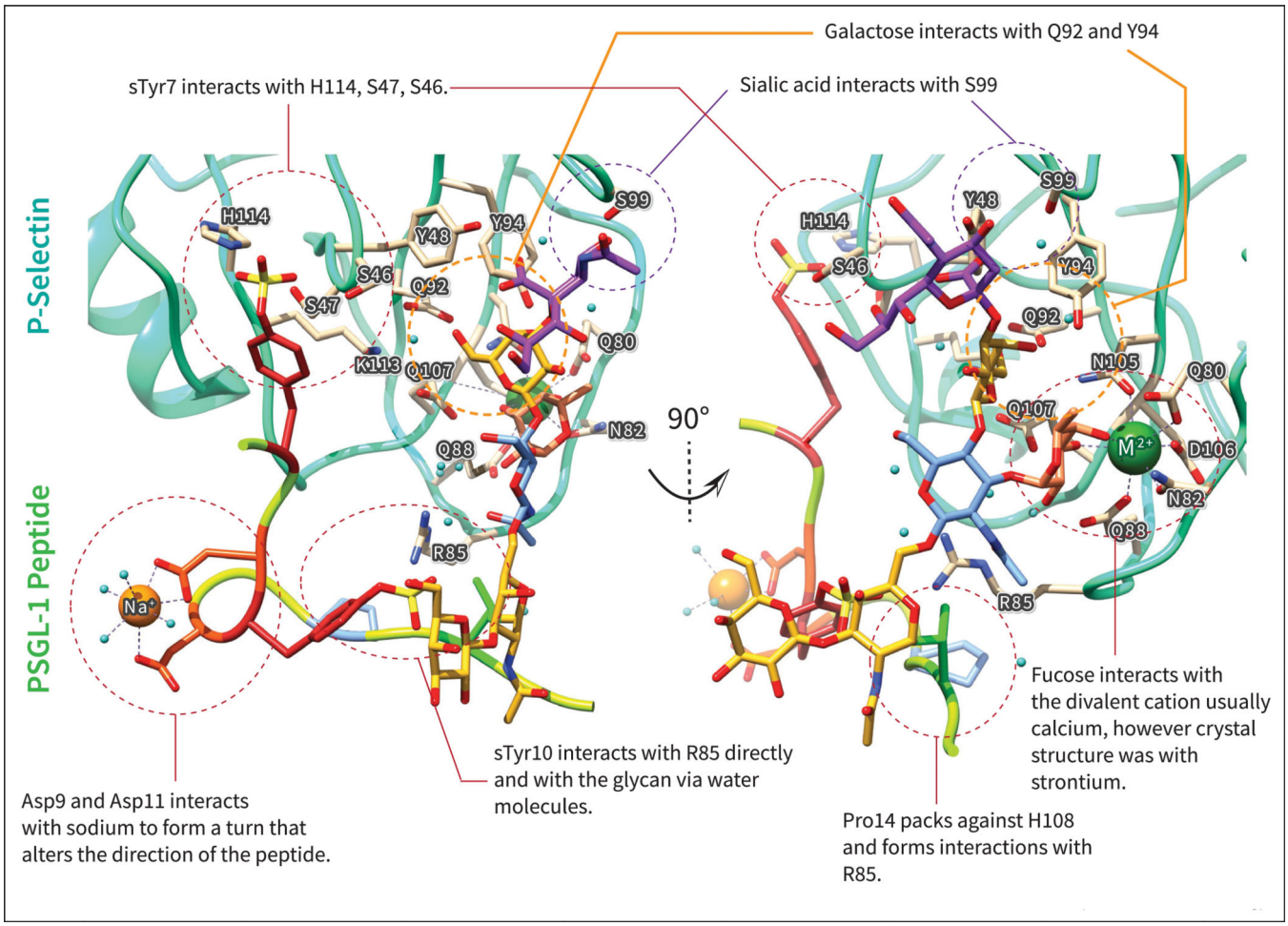


Figure 3.

Crystal structure of P-Selectin (cyan) with PSGL-1 glycosulfopeptide (multicolor) highlighting different interactions made by the sulfotyrosine (sTyr) and glycan (colored with SNFG colors to highlight different monosaccharides). Acidic residues of PSGL-1 such as Asp and sTyr are shown in shades of red. The sTyr7 corresponds to sTyr48 in the full length protein, while the sTyr10 corresponds to sTyr51 in the full length protein. The PDB ID for the structure is 1G1S from reference [65], The figures were drawn using Chimera (76).

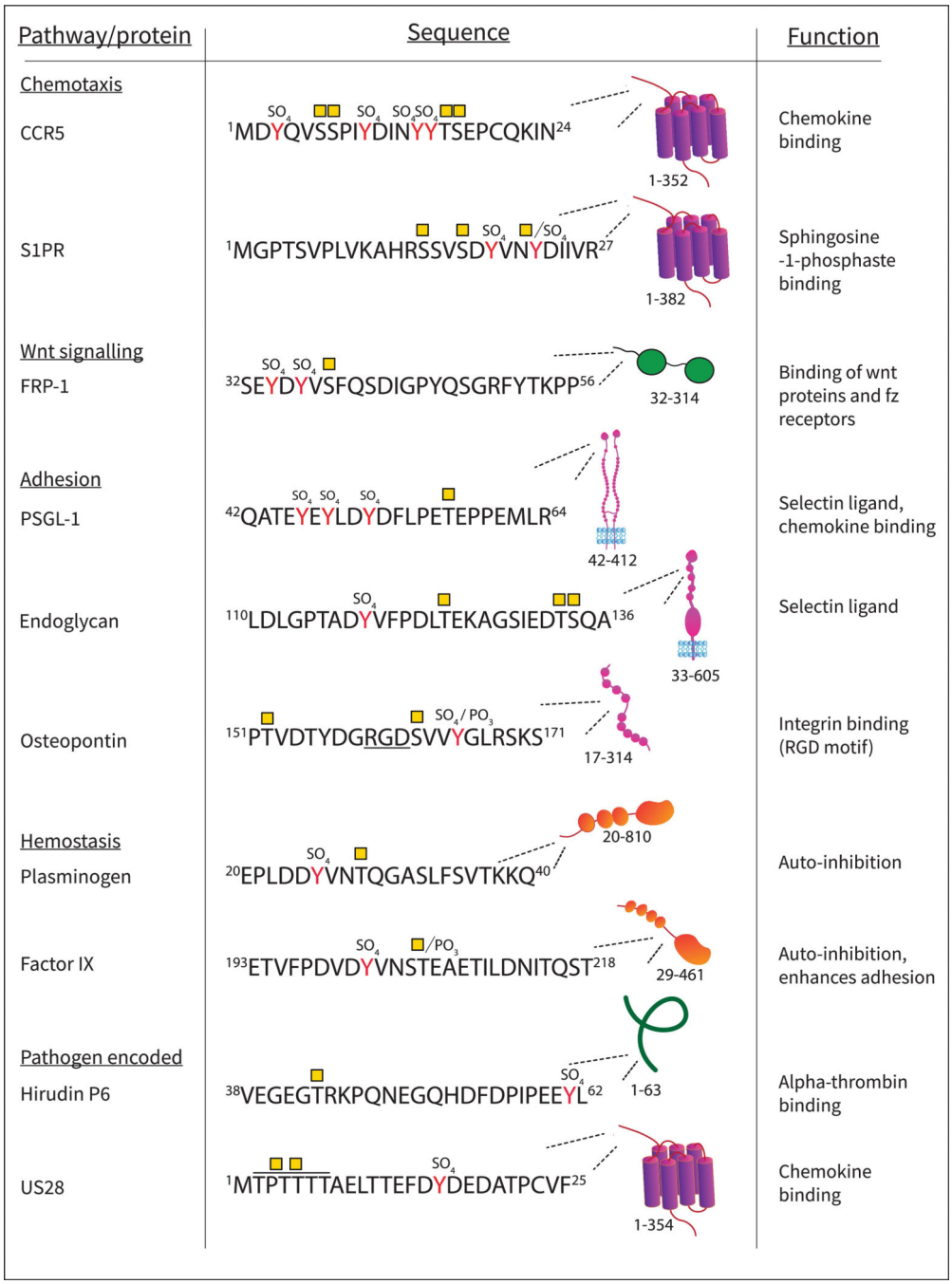


Figure 4.

Examples of proteins with multiple PTMs, showing the pathway of the protein, the sequence with sulfate groups (SO_4) attached to Tyr (Y) residues, and reported O-glycosylation sites (yellow square- GalNAc residues) in the peptide sequences. Phosphate groups (PO_3) are also indicated In OPN the integrin binding domain RGD is underlined, for US28 the two O-glycosylation sites were not unambiguously identified and can be located on any of the 5 Thr marked. The amino acids in the sequence are noted, along with a diagram of where the sequence is located in the protein. The protein function of the shown domain or sequence is

noted in the right column. Abbreviations: CCR5: CC chemokine receptor 5, S1PR1 :Sphingosine-1-phosphate receptor 1, FRP-1 : Secreted frizzled-related protein 1, PSGL-1: P-selectin glycoprotein ligand 1, US28: G-protein coupled receptor homolog US28.