BIORXIV/2021/466493

Revised

Sulfotyrosine, an interaction specificity determinant for extracellular protein-protein interactions

Running title: sTyr in extracellular protein-protein interactions

Valley Stewart [1] and Pamela C. Ronald [2]

[1] Department of Microbiology & Molecular Genetics

[2] Department of Plant Pathology

[2] Genome Center

University of California, Davis

[1] vjstewart@ucdavis.edu (corresponding author) | [2] pcronald@ucdavis.edu

Keywords: C-C chemokine receptor type 5 (CCR5); cytokine; follicle-stimulating hormone (FSH); leucine-rich repeat (LRR); peptide hormone; plant hormone; post-translational modification (PTM); protein-protein interaction; thrombin; tyrosine sulfation

Abstract

Tyrosine sulfation, a post-translational modification, can enhance and often determine protein-protein interaction specificity. Sulfotyrosyl residues (sTyr) are formed by tyrosyl-protein sulfotransferase during maturation in the golgi apparatus, and most often occur singly or as a cluster of two or three sTyr within a six-residue span. With both negative charge and aromatic character, sTyr enables numerous atomic contacts as visualized in binding interface structural models, and so there is no discernible binding site consensus. Found exclusively in secreted proteins, sTyr residues occur in four broad sequence contexts. First, a single sTyr residue is critical for diverse high-affinity interactions between peptide hormones and their receptor in both plants and animals. Second, sTyr clusters within structurally flexible anionic segments are essential for a variety of processes including coreceptor binding to the HIV-1 envelope spike protein during virus entry, chemokine interactions with many chemokine receptors, and leukocyte rolling cell adhesion. Third, a subcategory of sTyr clusters occurs in the context of conserved acidic sequences termed hirudin-like motifs that enable several proteins to interact with thrombin, central to normal blood-clotting. Consequently, many proven and potential therapeutic proteins derived from blood-consuming invertebrates depend on sTyr residues for their activity. Fourth, a few proteins that interact with collagen or other proteins contain one or more sTyr residues within an acidic residue array. Refined methods to direct sTyr incorporation in peptides synthesized both in vitro and in vivo, together with continued advances in MS and affinity detection, promise to accelerate discoveries of sTyr occurrence and function.

[248 words; limit 250]

Introduction

Post-translational modifications influence protein activity in many ways. Thus, learning how these modifications act singly and in combination is important for understanding protein function (1,2). One post-translational modification increasingly recognized as critical for diverse extracellular interactions in animals, plants and certain bacteria is tyrosine O-sulfation, forming sulfotyrosyl residues (sTyr or Tys). Sulfation is catalyzed during golgi transit by **t**yrosyl-**p**rotein **s**ulfo**t**ransferase (TPST), and therefore occurs only in secreted and membrane-spanning proteins (3).

sTyr was discovered nearly 60 years ago (4), but detecting and documenting this modification remains challenging such that the extent of sTyr occurrence has been incompletely defined (5-7). Although the sTyr sulfate linkage is generally stable in weak acid, it is cleaved by the strong acid used in the Edman degradation method used to determine protein sequences (6,8). Compounding the difficulty, routine MS methods do not reliably detect sTyr residues (9), so specialized protocols are being developed to minimize sulfate loss during peptide ionization and fragmentation (7,10-16).

Research on phosphotyrosyl (pTyr) occurrence and function relies in part on a variety of anti-pTyr antibodies, some with broad recognition and others specific for pTyr residues in defined sequence contexts (17). A commercially available anti-sTyr monoclonal antibody binds sTyr residues with high affinity regardless of flanking sequence and discriminates between peptide sTyr and pTyr residues by several hundred-fold (18). Nevertheless, this reagent has been used sparingly in sTyr research (11). More recently, two antibodies were identified that may discriminate between different chemokine receptor CCR5 sulfoforms (19). It therefore may be feasible to isolate sequence-specific anti-sTyr antibodies to monitor individual sulfation sites.

Systems-level characterization of pTyr residues often employs initial affinity enrichment with a broad-specificity anti-pTyr antibody (20). High-affinity pTyr-binding ***s****rc* **h**omology **2** (SH2) domains (approximately 100 residues) can provide an inexpensive alternative (20). Therefore, SH2 domains with high affinity and specificity for sTyr may offer a useful means for identifying and enriching sTyr-containing proteins (21,22).

Finally, early research on sTyr occurrence and function was constrained by the availability of synthetic sulfopeptides for use in binding assays. No more than two sTyr could be incorporated in a peptide (23), and some studies substituted pTyr for sTyr (24). Today, homogeneous sulfopeptides can be synthesized in vitro with up to three sTyr residues (25-27), and sulfonyl analogs of sTyr provide increased stability (28). Separately, sulfopeptides with up to five sTyr residues in all combinations have been synthesized in *Escherichia coli* strains that have an expanded genetic code, in which an amber (UAG)-reading tRNA is charged with exogenous sTyr (29-32). A similar system decodes amber as sTyr in mammalian cells (33), providing the possibility for functionally analyzing individual sTyr residues in vivo.

Why sTyr?

Consider sTyr in contrast to the familiar phosphotyrosyl (pTyr) residue. Sulfation and phosphorylation both add a functional group that is fully ionized at neutral pH and consequently results in increased side-chain polarity (34). However, sTyr makes weaker hydrogen bonds due to its lesser charge (–1 vs. –2) and smaller dipole moment (34).

Functions for pTyr have been studied extensively: its critical role in cellular growth control was discovered in large part from analyzing tumor virus-encoded oncogenic proteins (35). Tyrosyl protein kinase activity serves as "writer" for signaling, whereas separate phosphoprotein phosphatase activity serves as "eraser." For signal propagation, small pTyr binding domains such as SH2 provide "reader" function for multi-domain output complexes (36,37). By contrast, sTyr is a long-lived post-translational modification (38), and in most cases the sulfate ester likely is stable in physiological conditions (6). Neither "reader" (e.g., portable sTyr-binding domains) nor "eraser" (e.g., sulfoprotein sulfatase) components are known.

The sTyr residue potentiates protein-protein interactions through dual means: the sulfate group, which can make multiple electrostatic interactions to basic Arg or Lys residues in the binding partner; and the Tyr aromatic ring, which engages in both nonpolar and stacking interactions with diverse binding partner residues. Although pTyr can at least partially replace sTyr for certain peptide-peptide interactions (24,39,40), the sTyr sulfate makes distinct ionic contacts (41,42) and therefore in most cases provides a unique in vivo interaction specificity determinant.

sTyr residues are documented in a variety of extracellular proteins (3,9,27,43-46), and currently are known to occur predominately in four broad contexts (**Fig. 1**): (a) in peptide hormones or their receptors, almost all with a single sTyr residue; (b) in conformationally-flexible segments at the amino-termini of cell-surface proteins, most with two or three sTyr clustered within a span of six residues; (c) in conserved sequences (hirudin-like motifs) that interact with thrombin, all with two or three sTyr clustered within a span of six residues; and (d) in tracts of acidic residues at the amino-termini of certain secreted proteins. Thus, sTyr-based interactions usually involve relatively short protein segments, often within large multidomain proteins.

Recent discoveries of previously unknown sTyr occurrence (47,48) and functions (26,49-51) show that the repertoire of known sulfoproteins will continue to expand. Here we first describe the enzyme responsible for sTyr formation, and then review examples that illustrate each of these four sTyr contexts (**Fig. 1**). We finish the article by considering some challenges and opportunities for future sTyr research.

Tyrosyl-protein sulfotransferase (TPST)

Sulfotransferases share similar globular folds that bind the sulfodonor 3'-**p**hospho**a**denosine 5'-**p**hospho**s**ulfate (PAPS) (**Fig. 2A**), the activated intermediate for sulfur assimilation. PAPS synthesis and functions are described elsewhere (52,53). Two short sequence motifs, termed 5'-PSB and 3'-PB, coordinate the 5'-phosphosulfate and 3'-phosphate groups, respectively (54) (**Fig. 2B**). Sulfotransfer and kinase reactions are similar (**Fig. 2A**), and the geometry of PAPS binding in sulfotransferases is similar to that of ATP binding in nucleotide kinases (55). TPST enzymology has been reviewed elsewhere (56).

Cytoplasmic sulfotransferases termed SULT modify a variety of small molecules such as xenobiotics and hormones, generally share similar overall sequence (57,58). By contrast, tyrosyl-protein and polysaccharide sulfotransferases share little sequence similarity with one another and have different spacing between the 5'-PSB motif, catalytic base residue, and the 3'-PB motif (**Figs. 2B and 2C**), reflecting their engagement with different polymeric substrates.

Metazoan tyrosyl-protein and polysaccharide sulfotransferases are anchored to the Golgi lumen through an amino-terminal transmembrane segment, and thus are type II transmembrane proteins. These sulfotransferases modify proteins and sulfated **g**lycos**a**mino**g**lycans (GAGs) in spatial and temporal coordination with other modifications such as protein glycosylation (2,3,5,43,59,60). Two TPST isoenzymes are encoded by most genera throughout the Metazoa (5). The human TPST isoenzymes are expressed broadly, with one or the other predominating in certain tissues (43,61). Understanding these expression patterns requires defining the relationship between isoenzyme expression and the availability and relative affinity of substrates.

For example, homozygous *tpst-1* null mice have lower average body weight, which may result in part from reduced levels of the sulfated hormones **c**hole**c**ysto**k**inin (CCK) and gastrin (described below). By contrast, homozygous *tpst-2* null mice display primary hypothyroidism, likely resulting from failure to sulfate the receptor for thyroid stimulating hormone (see below). Homozygous *tpst-2* null males also are infertile. Finally, homozygous *tpst-1 tpst-2* double null mice usually die soon after birth due to cardiopulmonary insufficiency (62-66). Thus, TPST-1 and TPST-2 functions overlap only partially. The challenge remains to identify specific bases for most of these phenotypes.

Metazoan TPST displays broad substrate specificity. Most substrates with at least moderate affinity have an Asp residue just proximal to the sulfoaccepting Tyr residue (61,67-71) (**Fig. 1**). Non-conserved flanking residues, often rich in acidic Asp and Glu residues, make multiple interactions with different TPST residues near the active site (67,68,70). Overall, these broader sequence features and their roles in metazoan TPST binding affinity and sulfation efficiency remain obscure (72).

Plants encode a single TPST (53). In the model plant *Arabidopsis thaliana*, a null mutant displays developmental phenotypes such as dwarfism (73), consistent with the functions for the only known plant TPST substrates, sulfopeptide hormones involved in multicellular development (described below).

Plant TPST has a carboxyl-terminal transmembrane segment (type I transmembrane protein) in contrast to metazoan TPST and GAG sulfotransferases (73,74) (**Fig. 2C**). Sequences for plant and metazoan TPST are not obviously related, although plant TPST does share carboxyl-terminal sequence similarity with **h**eparan **s**ulfate **6**-**O**-**s**ulfo**t**ransferase (HS6OST), a GAG sulfotransferase (73). This carboxyl terminus is a prominent α-helix in the HS6OST-3 X-ray structure (75) and is not present in other characterized sulfotransferases.

Initial analysis did not detect plant TPST sequence similarity to the conserved 5'-PSB and 3'-PB motifs (73). With more sequences now available, it is evident that different polymer sulfotransferases have PAPS-binding 5'-PSB and 3'-PB motifs with varied sequences and spacings, and so plant TPST sequences do include putative PAPS-binding motifs that match well with those from other polymer sulfotransferases (**Fig. 2B**). Experimental work is necessary to test key residues in the plant TPST putative 5'-PSB and 3'-PB motifs and to examine substrate specificity determinants.
Plant TPST orthologs are conserved throughout green plants including unicellular algae like *Chlamydomonas* spp. that are not known to synthesize sulfopeptide hormones (53,73). It will be interesting to learn the algal substrate proteins for TPST, and to determine if any are conserved in land plants.

Most bacteria, archaea and non-photosynthetic eukaryotic microbes do not encode TPST. Nevertheless, TPST is made by some species of plant-pathogenic bacteria in the genus *Xanthomonas* that also synthesize the RaxX protein (**r**equired for **a**ctivation of **X**a21-mediated immunity), a molecular mimic of the PSY sulfopeptide hormone (76,77) (**Fig. 3A**). The bacterial TPST sequence is most similar to that of the Golgi-localized metazoan TPST, except that bacterial TPST acts in the cytoplasm prior to substrate secretion and therefore is not membrane-anchored (78).

sTyr increases binding affinity for peptide hormone-receptor pairs

Secreted sulfopeptide hormones of roughly 10-20 residues control growth and development in land plants (38,79-81), and digestion, neurotransmission and endocrine functions in animals (82,83). Most plant sulfopeptide hormones possess the invariant amino-terminal residue pair Asp-1 sTyr-2, where the Asp residue presumably directs plant TPST to the substrate Tyr residue (**Fig. 3A**).

Plant peptide hormones bind their cognate receptors through **l**eucine-**r**ich **r**epeat (LRR) extracellular domains (84,85). The plant sulfopeptide hormones tested display high affinities for their cognate receptor LRR domains, ranging from 1 to about 300 nM, and in each case the sTyr residue contributes tenfold or more to affinity compared with the nonsulfated peptide. (48,78,86,87).

Plant receptor binding to the sTyr residue is shown in X-ray co-crystal structures for the sulfopeptide hormones **r**oot meristem **g**rowth **f**actor (RGF) and **C**asparian strip **i**ntegrity **f**actor (CIF) with their cognate receptor LRR domains (86,87). Both structures reveal overall hydrophobic sTyr-binding interfaces, but only one Arg residue is conserved among residues that make specific contacts (**Fig. 3B**). Based on these two examples, it appears that plant LRR receptor sTyr binding interfaces have few sequence constraints and therefore might readily be formed during evolution. Further examples of LRR-sulfopeptide interactions are needed to determine the limits of these constraints.

The pentapeptide **p**hyto**s**ulfo**k**ine (PSK) is the only known plant hormone that contains two sTyr residues, which together confer about 40-fold higher affinity for the PSK receptor (38,88,89). The PSK receptor LRR domain contains a 36 residue "island" that contacts both sTyr residues through multiple interactions (88) (**Fig. 3C**). These different features illustrate how sTyr residues are versatile interaction determinants even in a broadly similar context such as sulfopeptide-LRR binding.

The animal sulfopeptide hormones CCK and gastrin are processed to generate different length bioactive peptides with a common carboxyl-terminal Phe-amide (**Fig. 3A**). The CCK sTyr residue mostly is sulfated (83), whereas the gastrin sTyr residue is heterogeneously sulfated, due perhaps to a low-affinity sequence environment for TPST recognition (67).

CCK and gastrin signal through the homologous G protein-coupled receptors CCK1R and CCK2R. CCK1R binds sulfated CCK with up to one-thousand-fold greater affinity than nonsulfated CCK, sulfated gastrin, or nonsulfated gastrin. In contrast, the CCK2R receptor makes little distinction (90). Thus, ligand sulfation state is a specificity determinant for CCK1R-CCK interaction. No X-ray structure is available for either CCK1R or CCK2R, but various studies indicate that these sulfopeptides interact across a broad region of the receptor external face (82), and that a conserved Arg residue in a CCK1R extracellular loop interacts specifically with the CCK sTyr (91).

Mirroring these interactions between nonsulfated receptors and sulfated peptide hormones, some nonsulfated ligands bind single sTyr-containing receptors (**Fig. 3D**). Glycoprotein hormones such as follicle-stimulating hormone are central to the complex endocrine system that regulates normal growth, sexual development, and reproductive function. Receptors for glycoprotein hormones comprise an amino-terminal LRR domain connected through a flexible region to a carboxyl-terminal G protein-coupled receptor (**Fig. 2A**). This interdomain region contains an sTyr residue that is indispensable for hormone recognition and signaling (92,93).

Upon binding to the LRR domain, follicle-stimulating hormone (about 200 residues) exposes a hydrophobic interface that contains two positively charged residues for electrostatic interactions with the sTyr sulfate (92,93) (**Fig. 3E**). Homologous receptors for thyroid stimulating hormone and leutenizing hormone similarly contain an essential interdomain flexible region with a single sTyr residue (94) (**Fig. 3D**).

Component C3a, a 77 residue peptide generated through the complement cascade, acts through its receptor to stimulate many aspects of the inflammatory response, sometimes leading to anaphylaxis (95). The G protein-coupled C3a receptor has an sTyr residue, essential for binding C3a, within an unusually large extracellular loop (96) (**Fig. 3D**). To date, an atomic-level view of receptor binding and activation is not available (97).

sTyr clusters mediate a variety of protein-protein interactions

In most cases sTyr residues occur in clusters of two or three within a six-residue span. This arrangement enables a wide range of interactions as illustrated by the examples below.

sTyr clusters are essential for HIV-1 entry. HIV-1 binding and entry depends on the viral envelope **g**lyco**p**rotein (gp) spike, a trimer of gp120-gp41 heterodimers (98). First, the spike binds to the cell-surface receptor CD4, thereby altering gp120 conformation to expose its bridging sheet element (99). Then, the coreceptor amino terminus binds the bridging sheet, triggering membrane fusion and viral entry (98,100,101).

The principal T cell coreceptors for HIV-1 entry are the chemokine receptors CCR5 and CXCR4 (102). These are G protein-coupled receptors for **chemo**tactic cyto**kines** (chemokines), secreted signaling proteins of approximately 75 residues that direct leukocyte movement toward injury or infection sites (103). The amino-terminal segment in many chemokine receptors includes an sTyr cluster that is essential for high-affinity chemokine binding (44,46,104-107) (**Fig. 2B & Fig. 4A**). Interest in sTyr function accelerated with discoveries that the CCR5 sTyr cluster is essential for HIV-1 entry (100) and is mimicked in some HIV-1 broadly-neutralizing antibodies (108).

CCR5 residues sTyr-10 and sTyr-14 are necessary for proper gp120-CCR5 interaction, whereas sulfation at residues Tyr-3 or Tyr-15 dispensable (109-111) (**Fig. 4A**). A cryo-EM structural model shows the CCR5 amino-terminal segment in an extended conformation across the gp120 bridging sheet. In this model, residue sTyr-14 is mostly buried in a pocket lined with hydrophobic contacts and capped by electrostatic interactions. By contrast, residue sTyr-10 is mostly exposed (98,101) (**Fig. 4B**).

Certain HIV-1-infected individuals produce antibodies that recognize the **CD4**-**i**nduced (CD4i) conformation of gp120 with the bridging sheet exposed (108,112). Most of these CD4i antibodies contain **h**eavy chains with unusually long and diverse sequences for the third **c**omplementarity-**d**etermining regions (**HCD3**) that form protruding loops to contact otherwise poorly accessible epitopes (112-114). Many of these CD4i antibody HCD3 sequences include sTyr residues required for gp120 binding (108,112) (**Fig. 4C**).

In an X-ray co-crystal structure with the gp120-CD4 complex, residues sTyr-100c^{412d} and sTyr-100^{412d} bind the same gp120 interfaces as the modeled residues sTyr-10^{CCR5} and sTyr-14^{CCR5}, respectively (101,115) (**Fig. 4D**). (The antibody numbering accounts for variable numbers of residues, labeled 100a-100w, between HCD3 residues at positions 100 and 101.) Separate studies with CCR5 sulfopeptides and antibody 412d sulfoforms expressed with the expanded genetic code reveal that the the buried residues sTyr-14^{CCR5} and sTyr-100c^{412d} are essential for binding, whereas the more exposed residues sTyr-10^{CCR5} and sTyr-100^{412d} are ancillary (111,116). Thus, the CCR5 and 412d sTyr-containing elements apparently share similar interaction with gp120.

A cryo-EM structural model shows a different binding pattern by antibody E51 (117) (**Fig. 4E**). Although residue sTyr-100i^{E51} binds to the same gp120 interface as the exposed ancillary residues sTyr-10^{CCR5} and 412d sTyr-100^{412d}, residue sTyr-100f^{E51} instead makes unique interactions (**Fig. 4E**). Interestingly, for antibody E51 sulfoforms, exposed residue sTyr-100i^{E51} is essential whereas the unique residue sTyr-100f^{E51} is ancillary (32).

Thus, antibody E51 sTyr-gp120 interactions are distinct from those for chemokine receptor CCR5 and antibody 412d. Intriguingly, E51 also is more potent than 412d (117), and forms the basis for an effective anti-HIV peptide (118). Apparently the gp120 bridging sheet is versatile enough to make high-affinity interactions with different sTyr-containing flexible segments, increasing opportunities to meet the stereochemical constraints to binding sulfate (112).

sTyr clusters contribute to combinatorial receptor-peptide interactions. Human cells express about 50 chemokines and about 20 chemokine receptors acting in different combinations in different tissues (119), so binding interface versatility is a hallmark. Roughly half of the chemokine receptors contain sTyr residues in their amino-termini (44). The sTyr cluster forms the core for binding the chemokine ligand globular domain and helps enable a given receptor to interact with multiple chemokines (46,104,107,119). The CCR5 amino-terminal segment has yet to be captured in a receptor-chemokine X-ray or cryo-EM structure (44,101,107,120,121), suggesting that it is intrinsically disordered (122).

Indeed, NMR spectroscopy and modeling suggests that CCR5-CCL5 interactions are highly dynamic (121,123). In accordance with these observations, separate sulfopeptide binding experiments revealed that any combination of two sTyr residues at CCR5 positions 10, 14 and 15 enables strong binding to chemokine CCL5 (124). Together, these results support the hypothesis that the CCR5 flexible sTyr-containing anionic segment makes a variety of dynamic yet high-affinity contacts to the chemokine, in striking contrast to the relatively fixed contacts observed in the CCR5-gp120 interactions described above (121,124).

Chemokines also bind sulfated GAGs to form chemotactic gradients (103). Notably, GAG sulfate groups make electrostatic interactions with many of the same chemokine basic residues that contact sTyr sulfates. Thus, chemokine functionality is expanded by a single versatile interface that binds different sulfated polymers for different purposes (106,123,125).

Chemokines provide signals for leukocyte movement toward sites of infection or injury. This rolling cell adhesion is mediated in part by **P**-**s**electin **g**lycoprotein **l**igand-**1** (PSGL-1), a homodimeric mucin-like cell surface glycoprotein. PSGL-1 contains three sTyr residues in its amino-terminal flexible segment (**Fig. 4A**) and engages with the cell-surface adhesion P-selectin through a mechanosensitive catch-bond that enables rapid engagement and release under force in the bloodstream flow (126,127). An X-ray co-crystal structure of a short PSGL-1 sulfoglycopeptide bound to the P-selectin amino-terminal domain shows residue sTyr-7 with several electrostatic and hydrophobic contacts to residues in P-selectin, and residue sTyr-10 with multiple hydrophobic contacts that orient the sulfate for hydrogen bonding to a critical Arg residue (2,128).

Although all three PSGL-1 sTyr residues are necessary for full P-selectin interactions, any one sTyr suffices for partial function in a variety of assays relevant to rolling cell adhesion (129). Indeed, residue sTyr-5 is not visible in the P-selectin-PSGL-1 cocrystal structure, suggesting that the observed structure may represent one of several productive conformations (128). In this hypothesis, the sTyr cluster potentially presents a variety of conformations suitable for interaction. Nevertheless, a separate study identified sTyr-7 as essential for PSGL-1 binding to P-selectin (130), congruent with the extensive interactions made by this residue (128).

Separately, PSGL-1 helps regulate aspects of T cell function, including the progressive loss of effector function in so-called exhausted T cells that accompanies persistent antigen stimulation (131). In this context, PSGL-1 is the receptor for **V**-domain **i**mmunoglobulin **s**uppressor of **T** cell **a**ctivation (VISTA), a pH-responsive T cell inhibitor (131). In a computational model, all three PGSL-1 sTyr residues make ionic interactions with His residues along one edge of VISTA. In this model, sTyr interaction depends upon His protonation, as would occur in acidic tumor microenvironments (pH $\leq$ 6) (49). Sulfation is critical for VISTA-PSGL-1 binding, but contributions of individual sTyr residues are not known (49). Together, interactions with P-selectin and VISTA illustrate how sTyr cluster structural flexibility enables impressive functional versatility by the PSGL-1 amino-terminal segment.

Complement protein C5a (74 residues), generated through the complement cascade, stimulates inflammatory mediator release and also is a potent chemoattractant (95). Similar to chemokine receptors, the G protein-coupled C5a receptor amino-terminal segment contains two sTyr residues (132) (**Fig. 4A**). **Ch**emotaxis **i**nhibitory **p**rotein of ***S****taphylococcus aureus* (CHIPS; 121 residues) also binds the amino-terminal segment, thereby inhibiting C5a-dependent inflammatory responses (97). The C5a receptor sTyr residue pair increases CHIPS binding affinity for peptides from the C5a receptor amino-terminal segment by approximately 1,000-fold (133). In the NMR structure of CHIPS bound to the C5a receptor amino-terminal peptide, the two sTyr residues are within a five residue β-strand and make several hydrophobic and electrostatic contacts to residues in CHIPS (133).

These three examples – CCR5, PSGL-1, and C5aR – illustrate a variety of sTyr interactions ranging from relatively ordered, as with the sulfoantibody-gp120 and C5aR-CHIPS interactions, to highly dynamic, as with the CCR5-CCL5 and PSGL-1-P-selectin interactions. Thus, it appears that two sTyr residues can enable a wider range of binding modes and partners than observed in the single sTyr residue hormones and receptors described above.

sTyr clusters in hirudin-like motifs bind the blood clotting enzyme thrombin

The sTyr cluster-containing anionic segments described above are versatile and dynamic, and there is no obviously conserved sequence pattern. Here, a different collection of anionic segments makes conserved interactions and contain conserved sTyr cluster sequences. Thus, sTyr clusters participate across a range of anionic segment functions.

Aberrant hemostasis – ranging from profuse bleeding to excessive clotting – underlies or complicates several human conditions, and many treatments have been developed (134-136). The serine endoprotease α-thrombin is central to controlling the balance between initiating and terminating blood clotting (thrombosis). Thrombin function depends on its interactions with molecular tethers in several different substrates and inhibitors (137,138). This number and diversity of thrombin binding partners provides an opportunity to explore unity and diversity in sTyr cluster interactions.

One clot-inhibiting therapeutic used for over one hundred years is hirudin, an sTyr-containing **d**irect **t**hrombin **i**nhibitor (DTI) made by *Hirudo* leeches (139,140). Hirudin (65 residues) comprises a carboxyl-terminal anionic segment that binds thrombin and an amino-terminal domain that occludes the thrombin active site. Similar anionic segments, termed hirudin-like motifs (141), are present in other thrombin-binding proteins (137,138) (**Fig. 5A**).

Thrombin contains two cationic surface patches, termed anion-binding exosites, that bind different protein substrates through their hirudin-like motifs (137,138). Thrombin exosite 1 aligns both endoprotease substrates and DTIs such as hirudin with respect to the adjacent active site, whereas exosite 2 tethers thrombin at the site of blood clots, binding platelet surfaces via platelet **g**lyco**p**rotein **Ibα** (GpIbα) and to fibrin clots via fibrinogen γ' (142,143). Exosite 2 also binds the DTIs madanin and **t**setse **t**hrombin **i**nhibitor (TTI) (51,144). Each exosite contains several basic Arg and Lys residues, enabling different contacts with acidic Asp, Glu and sTyr residues in different hirudin-like motif sequences.

As defined by the sequence alignments presented in **Fig. 5A**, the hirudin-like motif spans eight contiguous residues, of which most are acidic, aromatic, or both in the case of sTyr. Critically, hirudin-like motifs in all four proteins known to bind thrombin exosite 2 contain an sTyr cluster, whereas hirudin-like motifs from exosite 1-binding proteins, including hirudin itself, rarely contain sTyr residues and none at conserved positions (**Fig. 5A**).

Co-crystal X-ray structures of thrombin with exosite 2-bound sulfopeptides, illustrated here with the DTI madanin (144), reveal conserved contacts to thrombin residues across exosite 2 (24,51,145) (**Fig. 5B**). Residues sTyr-32^{madanin} and Asp-33^{madanin}, which occupy the hirudin-like motif conserved positions 2 and 3, make both ionic and nonpolar contacts to numerous exosite 2 residues (144) that form the core interaction with thrombin (24). Residue sTyr-35^{madanin} is more exposed (**Fig. 5B**). Thrombin contacts to all three residues are conserved in all four available co-crystal X-ray structures (51), indicating that the sTyr cluster-containing hirudin-like motif provides a well-defined high-affinity binding determinant for exosite 2.

Peptide binding analyses with sulfopeptides (51,144) or phosphopeptides (24,145) illuminate these structural models, again illustrated here with sulfo-madanin. The thrombin inhibition constant for doubly sulfated madanin is about 400-fold lower than for non-sulfated madanin, showing the essentiality of sTyr. Madanin with only sTyr-35 is two-fold more effective than non-sulfated madanin, whereas madanin with only sTyr-32 is about 25-fold more effective (144). Thus, like chemokine receptor CCR5 interaction with gp120, the core sTyr residue (sTyr-32^{madanin}; sTyr-14^{CCR5}) is essential whereas the ancillary sTyr residue (sTyr-35^{madanin}; sTyr-10^{CCR5}) enhances binding only in the presence of the core (111,144) (**Fig. 5B**).

Like hemostasis, the complement system is activated through an endoprotease cascade (146). Complement component C1, a serine protease homologous to thrombin, has an anion-binding exosite that interacts with an sTyr cluster-containing hirudin-like motif in the substrate C4 (147,148) (**Fig. 5A**). The X-ray co-crystal structure of complement C1 bound to gigastasin, a C1 inhibitor made by *Haementaria* leeches, reveals electrostatic contacts from C1 exosite basic residues to sulfate O atoms from gigastasin residues sTyr-117 and sTyr-119 (149). Thus, the homologous hemostasis and complement systems share the use of hirudin-like motifs to direct partner protein binding to exosites.

Segments with sTyr clusters are prevalent in bloodstream activities, including leukocyte migration and signaling, blood clotting, complement activation and triglyceride metabolism. These activities also are modulated extensively by interactions with sulfated GAGs such as heparan, anionic oligosaccharides that help control hemostasis through binding thrombin (103,150-152). Indeed, thrombin exosite 2 was identified initially as the binding site for the highly-sulfated GAG heparin (142). Thus, a cluster of two or three sTyr residues can resemble a sulfated GAG to make conserved contacts with some of the many available basic residues in exosite 2, thereby expanding its valency beyond that of heparin-binding. Notably, basic exosite residues that contact GAG substrates (142) mostly are distinct from those that contact sTyr-containing substrates, illustrating the broad binding versatility provided by exosite 2.

sTyr in acidic arrays

Finally, some sTyr residues lie within relatively long tracts of acidic Asp and Glu residues (**Fig. 2D**). Glycosyl**p**hosphatidyl**i**nositol-anchored **h**igh-density lipoprotein **b**inding **p**rotein **1** (GPIHBP1) is a membrane protein critical for triglyceride metabolism. The GPIHBP1 amino-terminal acidic tract requires a single sTyr residue for proper interaction and function with its binding partner, lipoprotein lipase (47). As with other examples above, the GPIHBP1 amino-terminal acidic tract is not visible in the lipoprotein lipase-GPIHBP1 co-crystal X-ray structure (153).
At the opposite extreme lie certain extracellular matrix proteins such as fibromodulin, whose amino terminus has an extended tract of acidic residues including several sTyr (154,155). This region binds collagen and a variety of heparin-binding proteins (156,157). The collagen-modifying enzyme lysyl oxidase has a similar sTyr-containing acidic tract through which it binds collagen (158). The fibrinogen and lysyl oxidase

acidic tracts contain several Tyr residues, but in most cases, it has not been possible to determine which of these is sulfated to form sTyr (157,158).

Challenges and prospects

The examples presented here exemplify the diverse contexts for the sTyr post-translational modification. A single sTyr residue can enhance binding affinity by several hundredfold. A cluster of two or three sTyr residues provides a wider range of binding modes, from versatile, as illustrated by the chemokine receptor CCR5 sTyr-containing amino-terminal segment, to stringent, as illustrated by hirudin-like motif interactions with thrombin exosite 2.

Post-translational modifications often work in concert, with different combinations exerting different effects (1). sTyr residues often occur with nearby *N*- or *O*-linked glycans, and differential glycosylation can affect function (2). For example, PSGL-1 residue Thr-16, near the sTyr cluster (**Fig. 4A**), is glycosylated in leukocytes to enable interaction with P-selectin, but it is not glycosylated in certain T cells wherein PSGL-1 interacts with VISTA (49,159). This encourages a parallel hypothesis, that differential tyrosyl sulfation also can help determine binding partner selection. Indeed, the chemokine receptor CXCR4 (for one example) requires residue sTyr-21 for chemokine ligand CXCL12 binding but not for HIV-1 entry (160). It is not known if or how CXCR4 tyrosine sulfation is regulated with respect to cell type or external stimulus.
The extent of sulfation at a given Tyr residue can be incomplete, such that only a fraction of proteins in a given population carry a particular sTyr residue (19,44). This is difficult to evaluate directly because MS analyses may result in sulfate loss (9). Does heterogeneous sulfation result from stochastic TPST catalysis? Or are different sulfation states (potentially with distinct ligand-binding properties) programmed in different cell types or in response to different stimuli?

Evaluating the location and extent of tyrosine sulfation also requires better understanding of how TPST engages its substrates and how its activity is regulated. Recently described TPST inhibitors may prove helpful (161). Substrates require nearby Asp or Glu residues for efficient sulfotransfer, but overall specificity determinants are not well-defined (70-72) in comparison to tyrosine kinases (162). However, animal cells make hundreds of tyrosine kinases but only two TPSTs. Therefore, TPST might catalyze some level of indiscriminate sulfation within a flexible anionic segment containing multiple Tyr residues. For example, several nearby Tyr residues within glycopeptide receptors and the C3a receptor are sulfated in vivo (**Fig. 3D**), even though sulfates at these positions are not required for receptor function (94,96).

Similarly, chemoreceptor CCR5 residue sTyr-3 (**Fig. 4A**) is considered to have minimal functional significance (109,110,163,164) and therefore usually is not studied (111,124). However, in vitro studies show that position Tyr-3 is the first to be sulfated in a CCR5 peptide that includes residue Met-1 (165,166). Does sTyr-3 result simply from "bystander" sulfation by TPST enzymes with broad substrate recognition? Or does it have a defined function (19)?

In vitro, pTyr can substitute for sTyr in some contexts (24,39,40) but not in others (164). Nevertheless, stringent specificity for sTyr in vivo is seen from the number of molecular mimics that contain sTyr, including anti-gp120 CD4i antibodies (115,117) as well as inhibitors of chemokine signaling (50), thrombin (26,51,144,167), and complement C1s (149). A striking example is the bacterial RaxX sulfopeptide, which acts as a molecular mimic of the plant PSY sulfopeptide hormone (76,77) (**Fig. 3A**). Although genes encoding TPST are ubiquitous in metazoa and plants, they are only sparsely distributed among relatively few bacterial lineages (168), and often are adjacent to genes encoding synthesis and export of extracytoplasmic proteins (V. Stewart and P. C. Ronald, unpublished observations). Thus, because RaxX requires sTyr for activity, the bacterium must synthesize TPST.

One potential research goal is to make defined sTyr residues to facilitate engineering of new interaction determinants. An initial proof-of-principle is eCD4-Ig, a chimeric protein that effectively blocks HIV-1 entry in rhesus macaques (118,169,170). eCD4-Ig contains an essential sTyr-rich peptide, derived from the sTyr-containing anti-gp120 antibody E51 variable region (**Fig. 4C**), and therefore is modeled closely after a pre-existing sTyr-containing segment.

More challenging will be to design sTyr interactions de novo. Placing sTyr residues at predetermined locations would require understanding and manipulating the substrate binding site as described above, or it could be accomplished through genetic code modification (29,32,33). Designing an effective sTyr binding site might be more feasible, given the wide range of naturally occurring sites characterized to date. An alternative approach is to use SH2 domains modified to recognize sTyr in place of pTyr (21).

A separate platform for engineering sTyr residues may come from **ri**bosomally synthesized and **p**ost-translationally modified **p**eptides (RiPPs), microbial natural products with diverse chemistry and potential applications (171). RiPPs are matured through a variety of post-translational modifications, and there is interest in engineering these modifications to create new RiPPs. The bacterial RaxST TPST, which synthesizes the sTyr residue in the RaxX RiPP, is a good candidate for strategies to introduce novel sTyr residues in engineered RiPPs (78,171). Additionally, certain polyketide biosynthesis complexes include SULT-type sulfotransferases among several modifying enzymes (172-174), so it may be possible to add TPST modules to non-ribosomal peptide synthesis complexes (175).

Finally, sTyr residues touch many topics in human medicine. For example, there are numerous sTyr residues in the homologous multidomain proteins, coagulation factors VIII (FVIII; six sTyr) and V (FV; seven sTyr) (176). Classic hemophilia results from FVIII deficiency, and one of the many causative *F8* alleles encodes a missense substitution of FVIII Phe for sTyr-1680 (176), documenting an important function for sTyr in hemostasis. Indeed, synthetic FVIII used for human therapy contains a full complement of sTyr residues (177). Nevertheless, relatively little is known about sTyr function in these critical hemostasis proteins (178,179). Research in FVIII and FV surely can benefit from the technical advances that now enable synthesis of sTyr residues at defined locations in vitro (25,27) and in vivo (29,32,33).

sTyr residues are versatile interaction determinants, essential for many critical interactions but also challenging to study. This means that sTyr residues are under-documented and -studied relative to other post-translational modifications. Nevertheless, the accumulated knowledge base makes it easier to predict and analyze newly discovered sTyr residues, and such discovery will be facilitated as MS methods for sTyr detection increasingly are refined. One expects to see accelerated progress, not only in finding new sTyr residues but also in understanding functions for currently known examples.

Supporting information — None.

Acknowledgements — We thank Drs. Anna Joe, Dee Dee Luu and Chet Price for their helpful critiques on an early draft version of this manuscript, and the anonymous reviewers who provided candid, very helpful comments and suggestions.

Author contributions — V. S. and P. C. R. wrote and edited the manuscript.

Funding and additional information — Tyrosine sulfation research in the Ronald laboratory is supported by Public Health Service grant GM122968 from the National Institute of General Medical Sciences (awarded to P.C.R.). Molecular graphics and analyses were 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.

Conflict of interest — The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations — The abbreviations used are: CCL, CC-type chemokine ligand; CCR, CC-type chemokine receptor; CD4i, CD4-induced; CIF, Casparian strip integrity factor; CXCR, CXC-type chemokine receptor; DTI, direct thrombin inhibitor; FSHR, follicle-stimulating hormone receptor; FV, coagulation factor V; FVIII, coagulation factor VIII; GAG, glycosaminoglycan; GpIbα, glycoprotein Ibα; gp120, glycoprotein spike 120 kDa subunit; HCD3, heavy-chain complementarity-determining region 3; LRR, leucine-rich repeat; PSY, plant peptide containing sulfated tyrosine; Rax, required for activation of Xa21-mediated immunity; RGF, root meristem growth factor; RiPP, ribosomally synthesized and post-translationally modified peptide; PSGL-1, P-selectin glycoprotein ligand-1; PAPS, 3'-phosphoadenosine 5'-phosphosulfate; PSK, phytosulfokine; SH2, *src* homology 2; sTyr, sulfotyrosyl residue; TPST, tyrosyl-protein sulfotransferase; TTI, tsetse thrombin inhibitor

Intentionally blank page to enable side-by-side figures and legends.

Fig. 1. Amino acid sequence contexts for representative sTyr residues

Sketch for journal artist | two-column figure (7.2" X 5.6")

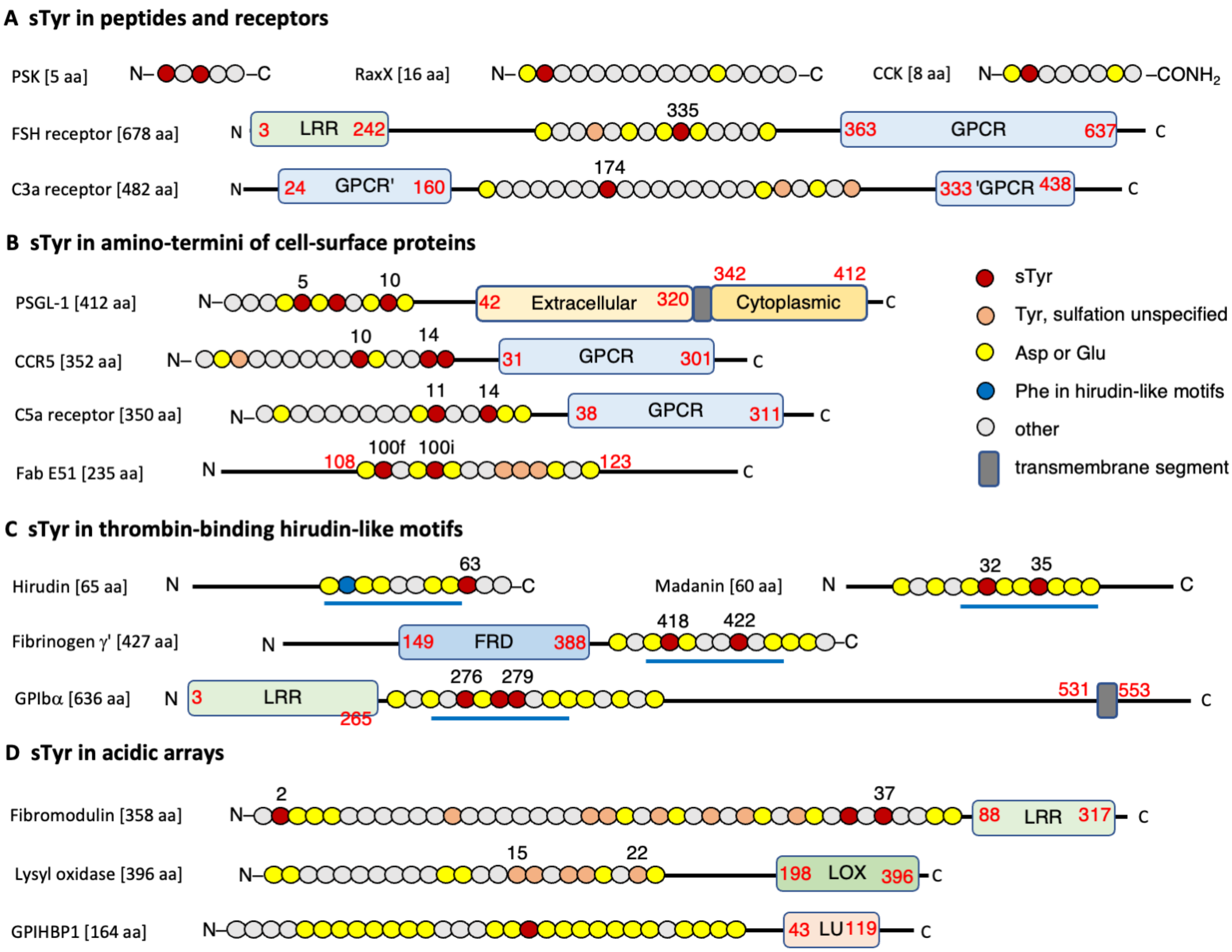

Fig. 1. Amino acid sequence contexts for representative sTyr residues

Primary structures of sTyr-containing proteins are shown schematically with major domains drawn approximately to scale. Segments with sTyr residues are depicted as circles, each corresponding to a different amino acid, and colored according to the key. Proteins and peptides are described in the text. The amino- and carboxyl-termini are denoted N and C, respectively, and numbers indicate residue position in the sequence.

A. sTyr in peptides and receptors. The CCK-8 Phe-amide terminus is denoted $CONH_2$. The sTyr-containing sequence in the C3a receptor is in a loop within the seven transmembrane GPCR domain.

B. sTyr in amino-termini of cell-surface proteins.

C. sTyr in hirudin-like motifs. The blue lines denote the eight-residue hirudin-like motif.

D. sTyr in acidic arrays.

Fig. 2. Tyrosyl-protein sulfotransferase (TPST)

Sketch for journal artist | two-column figure (7.2" X 4.6")

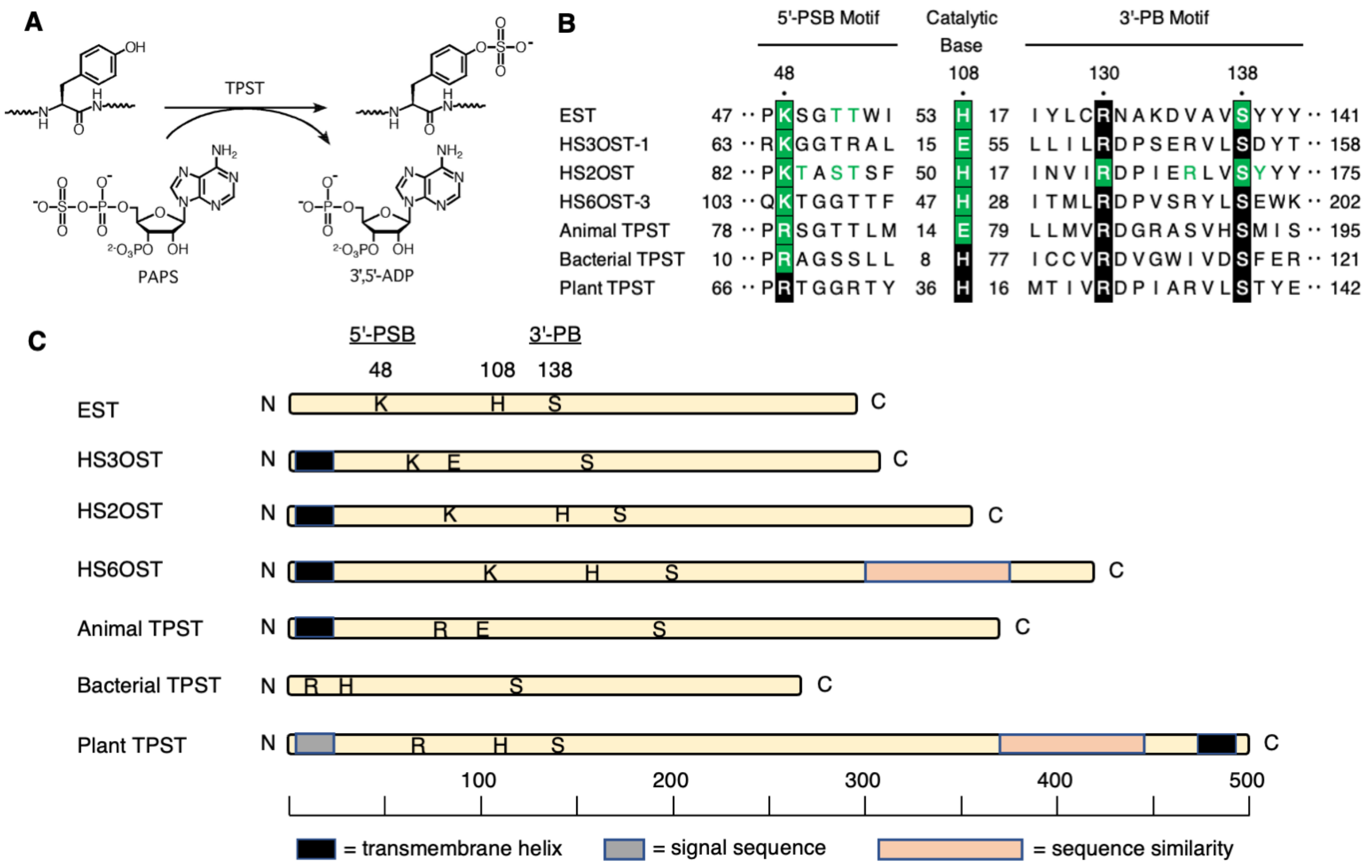

Fig. 2. Tyrosyl-protein sulfotransferase (TPST)

A. Tyrosyl-protein sulfotransferase catalyzes the transfer of sulfate from the universal sulfate donor PAPS (3'-**p**hospho**a**denosine 5'-**p**hospho**s**ulfate) to the hydroxyl group of a peptidyl-tyrosine residue to form a tyrosyl O4-sulfate ester and 3',5'-ADP (PAP; 3'-**p**hospho**a**denosine 5'-**p**hosphate). Reproduced from reference (5).

B. Sequence motifs from representative sulfotransferases. Substitutions at residues highlighted in green strongly decrease enzyme activity. Numbers at the left and right ends show positions within the sequence; internal numbers show the number of residues between elements.

C. Primary structures of representative sulfotransferases, drawn approximately to scale. Letters show the approximate positions of the 5'-PSB (K or R), catalytic base (H or E) and 3'PB (S) motifs shown in panel B. Enzymes are: EST, estrogen 17-β sulfotransferase (*Mus musculus*; PDB code 1AQU); HS3OST-1, heparan sulfate (3-O) sulfotransferase (*M. musculus*; 1S6T); HS2OST, heparan sulfate (2-O) sulfotransferase (*Homo sapiens*; 3F5F); HS6OST-3, heparan sulfate (6-O) sulfotransferase (*Danio rerio*; 5T03); animal TPST (*H. sapiens*; 5WRI); bacterial TPST (*Xanthomonas* spp.; GenBank accession WP_027703307); plant TPST (*Arabidopsis thaliana*; BAI22702).

Fig. 3. sTyr in peptides and receptors

two-column figure (7.2" X 5.3")

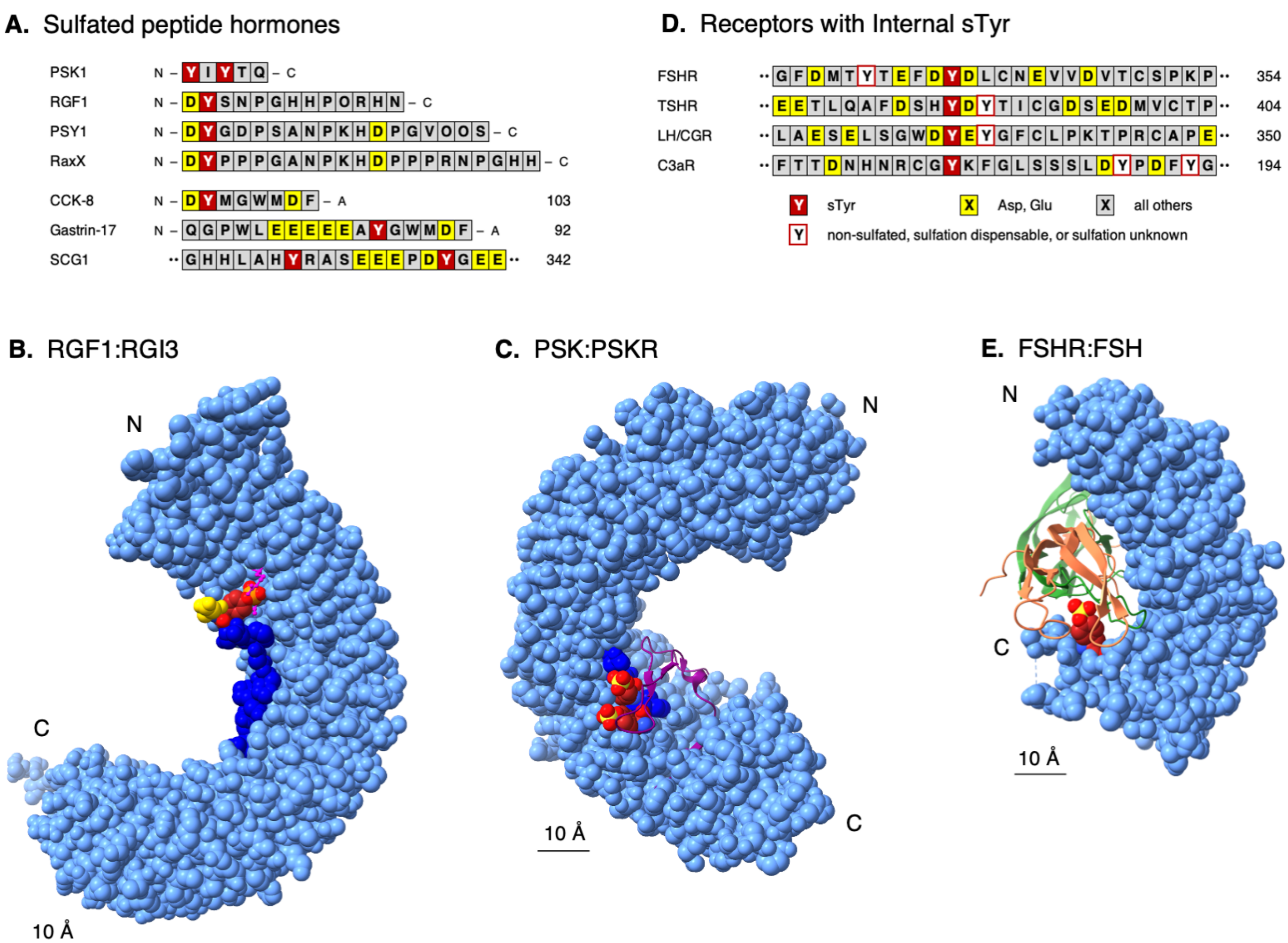

Fig. 3. sTyr in peptides and receptors

A. sTyr residue contexts are shown for the mature forms of the plant sulfopeptide hormones PSK1 (phytosulfokine), RGF1 (root meristem growth factor), PSY1 (peptide sulfated on tyrosine), the hormone mimic RaxX (required for activating XA21 protein X), and the animal sulfopeptide hormones CCK-8 (cholecystokinin), gastrin-17, and one of two sulfopeptides derived from the secretogranin-1 (SGC-1) precursor. Residues are colored according to the key in panel D. O denotes hydr**o**xyproline.

B. The co-crystal X-ray model of sulfated RGF (dark blue) bound to the RGI3 (RGF-insensitive-3) receptor ectodomain (light blue) is depicted at 2.6 Å resolution (86) (PDB code 5HYX). Atoms are shown as Van der Waals spheres. The RGI3 ectodomain is built from 23 LRRs (leucine-rich repeats) stacked one upon the next. RGF residue sTyr-2 (dark red; S atom yellow) contacts RGI3 residues Arg-195 and Ala-222, depicted as balls-and-sticks (magenta). RGF residue Asp-1 (yellow) is mostly exposed.

C. The co-crystal X-ray model of sulfated PSK (dark blue) bound to the PSK receptor ectodomain (light blue) is depicted at 2.5 Å resolution (88) (PDB code 4Z63). Atoms are shown as Van der Waals spheres. The RGI3 ectodomain is built from 21 LRRs (leucine-rich repeats) stacked one upon the next, with a 36-residue island domain, shown as a ribbon (magenta), inserted into LRR number 18. PSK residues sTyr-1 and sTyr-3 (dark red; S atoms yellow) mostly contact residues in the island domain.

D. sTyr residue contexts are shown for internal segments (**Fig. 1A**) of the receptors for FSH (follicle-stimulating hormone), TSH (thyroid stimulating hormone), LH/CGR (luteinizing hormone/ choriogonadotropin) and complement component C3a.

E. The co-crystal X-ray model of FSH (follicle-stimulating hormone; α subunit salmon; β subunit light green) bound to the sulfated FSH receptor ectodomain (light blue) is depicted at 2.5 Å resolution (93) (PDB code 4AY9). The FSHR ectodomain is built from 12 LRRs stacked one upon the next. FSHR residue sTyr-335 (dark red; S atom yellow) contacts several residues each in both FSH subunits.

Fig. 4. sTyr in receptor amino-termini and in anti-gp120 antibodies

two-column figure (7.2" X 5.3")

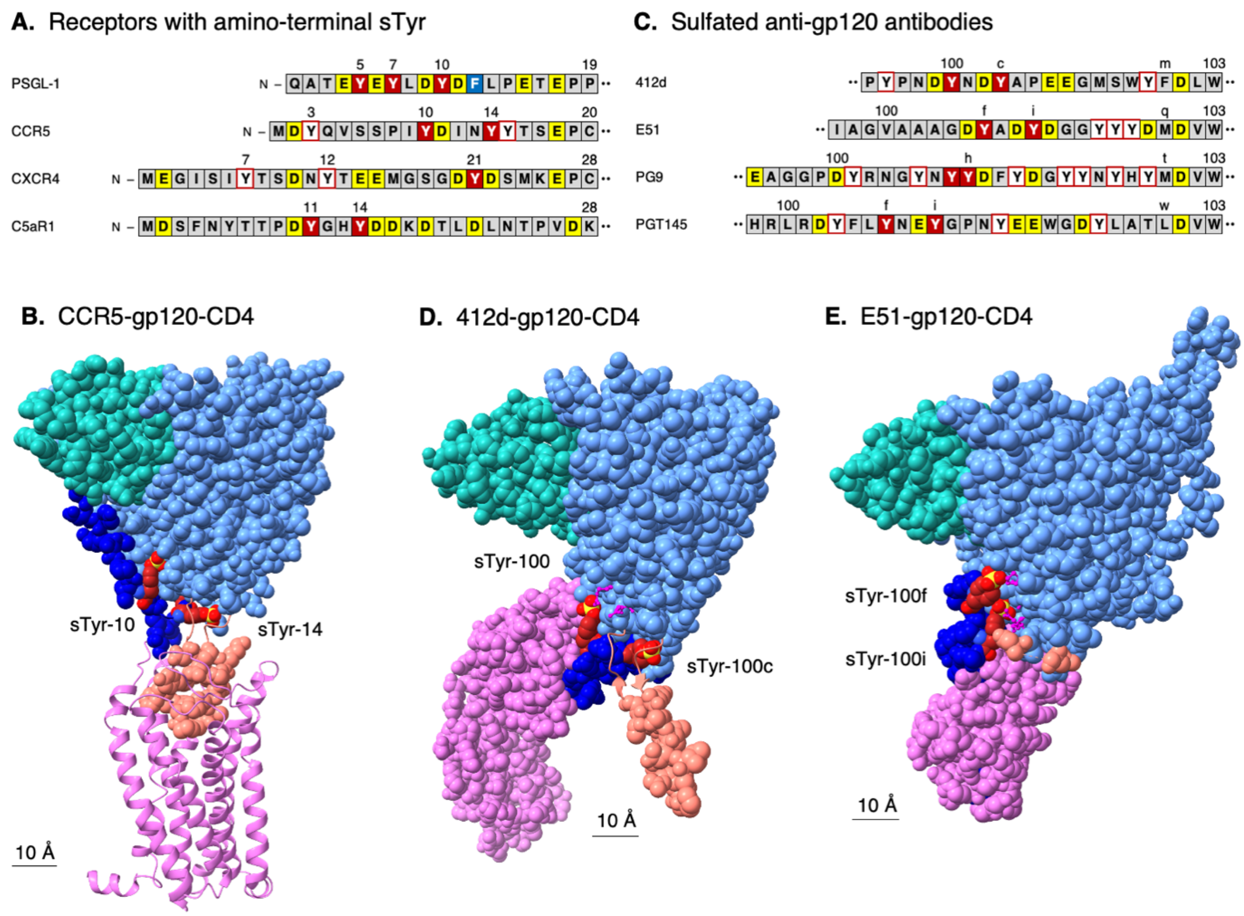

Fig. 4. sTyr in receptor amino-termini and in anti-gp120 antibodies

A. sTyr cluster contexts are shown for the amino-termini of the cell-surface proteins PSGL-1 (P-selectin glycoprotein ligand-1), CCR5 (CC-type chemokine receptor 5), CXCR4 (CXC-type chemokine receptor 4) and C5aR (complement component C5a receptor). Residues are colored according to the key in **Fig. 3D**.

B. The cryo-EM model of HIV-1 gp120 (glycoprotein 120; light blue) bound to host CCR5 (light purple) and one subdomain of host CD4 (cluster of differentiation 4 receptor; light green) is depicted at 3.9 Å resolution (101) (PDB code 6MEO). Atoms are shown as Van der Waals spheres except for CCR5, depicted as ribbons the better to display the seven transmembrane helix structure and to visualize interaction with the gp120 V3 loop residues (salmon). The CCR5 amino-terminal segment (dark blue) includes residues sTyr-10 and sTyr-14 (dark red; S atoms yellow) that interact with residues in the gp120 bridging sheet element including the base of the V3 loop. Certain residues at the base of the V3 loop are shown as ribbons the better to visualize the sTyr residues.

C. sTyr cluster contexts are shown for the HCD3 regions from representative CD4i anti-gp120 antibodies: 412d (PDB code 2QAD); E51 (6U0L); PG9 (3U2S); PGT145 (3U1S).

D. The co-crystal X-ray model of HIV-1 gp120 (glycoprotein 120; light blue) bound to the antibody 412d heavy chain (light purple) and one subdomain of host CD4 (light green) is depicted at 3.3 Å resolution (112) (PDB code 2QAD). Certain residues at the base of the V3 loop are shown as ribbons the better to visualize the sTyr residues. The HCD3 region (dark blue) includes residues sTyr-100 and sTyr-100c (dark red; S atoms yellow) that interact with residues in the gp120 bridging sheet element including the base of the V3 loop. Certain residues at the base of the V3 loop are shown as ribbons the better to visualize the sTyr residues. Only a portion of the 412d light chain is shown.

E. The cryo-EM model of HIV-1 gp120 (light blue) bound to the antibody E51 heavy chain (light purple) and one subdomain of host CD4 (light green) is depicted at 3.3 Å resolution (117) (PDB code 6U0L). The HCD3 region (dark blue) includes residues sTyr-100f and sTyr-100i (dark red; S atoms yellow) that interact with residues in the gp120 bridging sheet element. The V3 loop is not resolved in this structure; gp120 residues at the V3 base (salmon) indicate its position. Only a portion of the E51 light chain is shown.

Fig. 5. sTyr in hirudin-like motifs

one and one-half column figure (5" X 2.6")

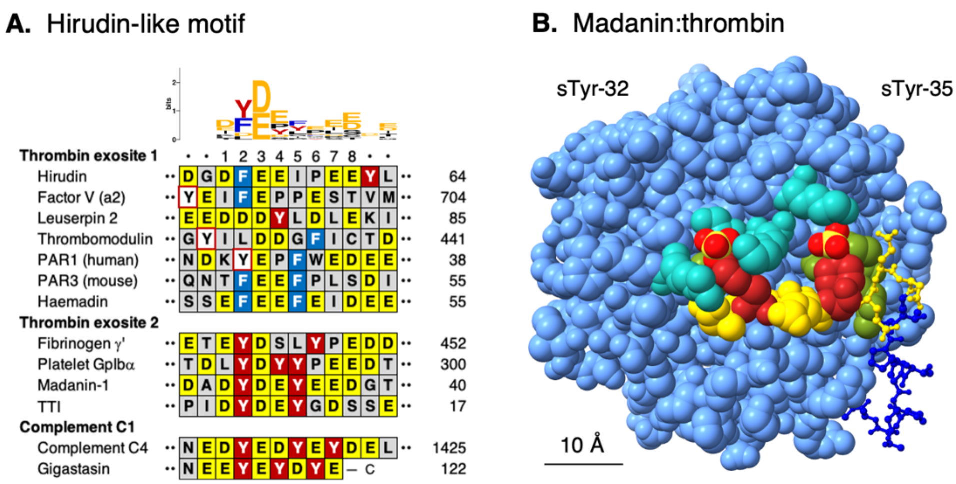

Fig. 5. sTyr in hirudin-like motifs

A. sTyr cluster contexts are shown for thrombin-binding hirudin-like motifs from hirudin (PDB code 2PW8); factor V (a2 segment; 3P6Z); leuserpin 2 (1JMO); thrombomodulin (1DX5); peptide-activated receptor-1 (PAR1; human; 3LU9); peptide-activated receptor-3 (PAR3; mouse;2PUX); haemadin (1E0F); madanin-1 (5L6N); fibrinogen γ' (2HWL); platelet GpIbα (4CH2); tsetse thrombin inhibitor (TTI; 6TKG); complement C4 (GenBank accession AAA51855); and gigastasin (5UBM). Residues are colored according to the key in **Fig. 3D**. The sequence logo (180,181) was generated from the alignment shown. The hirudin-like motif is defined here as positions 1-8. Positions 2 and 3 comprise a conserved aromatic-acidic residue pair, consisting of Phe, Tyr or sTyr at position 2 and Asp or Glu at position 3. The remaining hirudin-like motif residues are hydrophobic, negatively charged, or both in the case of sTyr, and binding to thrombin exosites involves both electrostatic and hydrophobic contacts (142).

B. The co-crystal X-ray model of sulfated madanin bound to exosite 2 on thrombin (light blue) is depicted at 1.6 Å resolution (144) (PDB code 5L6N). Atoms are shown as Van der Waals spheres, except for madanin residues 36-60 shown as balls-and-sticks. Madanin residues sTyr-32 and sTyr-35 (dark red; S atoms yellow) make electrostatic interactions with exosite 2 Arg and Lys residues (light green) and hydrophobic interactions with other residues (olive green). Madanin Asp and Glu residues (yellow) are highlighted for reference.

References

1. Nussinov, R., Ma, B., Tsai, C. J., and Csermely, P. (2013) Allosteric conformational barcodes direct signaling in the cell. *Structure* **21**, 1509-1521

2. Mehta, A. Y., Heimburg-Molinaro, J., Cummings, R. D., and Goth, C. K. (2020) Emerging patterns of tyrosine sulfation and O-glycosylation cross-talk and co-localization. *Curr Opin Struct Biol* **62**, 102-111

3. Huttner, W. B. (1988) Tyrosine sulfation and the secretory pathway. *Annu Rev Physiol* **50**, 363-376

4. Jevons, F. R. (1963) Tyrosine O-sulphate in fibrinogen and fibrin. *Biochem J* **89**, 621-624

5. Moore, K. L. (2003) The biology and enzymology of protein tyrosine O-sulfation. *J Biol Chem* **278**, 24243-24246

6. Balsved, D., Bundgaard, J. R., and Sen, J. W. (2007) Stability of tyrosine sulfate in acidic solutions. *Anal Biochem* **363**, 70-76

7. Chen, G., Zhang, Y., Trinidad, J. C., and Dann, C., 3rd. (2018) Distinguishing sulfotyrosine containing peptides from their phosphotyrosine counterparts using mass spectrometry. *J Am Soc Mass Spectrom* **29**, 455-462

8. Monigatti, F., Hekking, B., and Steen, H. (2006) Protein sulfation analysis--A primer. *Biochim Biophys Acta* **1764**, 1904-1913

9. Seibert, C., Sanfiz, A., Sakmar, T. P., and Veldkamp, C. T. (2016) Preparation and analysis of N-terminal chemokine receptor sulfopeptides using tyrosylprotein sulfotransferase enzymes. *Methods Enzymol* **570**, 357-388

10. Robinson, M. R., Moore, K. L., and Brodbelt, J. S. (2014) Direct identification of tyrosine sulfation by using ultraviolet photodissociation mass spectrometry. *J Am Soc Mass Spectrom* **25**, 1461-1471

11. Robinson, M. R., and Brodbelt, J. S. (2016) Integrating weak anion exchange and ultraviolet photodissociation mass spectrometry with strategic modulation of peptide basicity for the enrichment of sulfopeptides. *Anal Chem* **88**, 11037-11045

12. Borotto, N. B., Ileka, K. M., Tom, C., Martin, B. R., and Hakansson, K. (2018) Free radical initiated peptide sequencing for direct site localization of sulfation and phosphorylation with negative ion mode mass spectrometry. *Anal Chem* **90**, 9682-9686

13. Halim, M. A., MacAleese, L., Lemoine, J., Antoine, R., Dugourd, P., and Girod, M. (2018) Ultraviolet, infrared, and high-low energy photodissociation of post-translationally modified peptides. *J Am Soc Mass Spectrom* **29**, 270-283

14. Yang, N., Anapindi, K. D. B., Rubakhin, S. S., Wei, P., Yu, Q., Li, L., Kenny, P. J., and Sweedler, J. V. (2018) Neuropeptidomics of the Rat Habenular Nuclei. *J Proteome Res* **17**, 1463-1473

15. Tyshchuk, O., Gstottner, C., Funk, D., Nicolardi, S., Frost, S., Klostermann, S., Becker, T., Jolkver, E., Schumacher, F., Koller, C. F., Volger, H. R., Wuhrer, M., Bulau, P., and Molhoj, M. (2019) Characterization and prediction of positional 4-hydroxyproline and sulfotyrosine, two post-translational modifications that can occur at substantial levels in CHO cells-expressed biotherapeutics. *MAbs* **11**, 1219-1232

16. Brodbelt, J. S., Morrison, L. J., and Santos, I. (2020) Ultraviolet photodissociation mass spectrometry for analysis of biological molecules. *Chem Rev* **120**, 3328-3380

17. Hunter, T. (2014) The genesis of tyrosine phosphorylation. *Cold Spring Harb Perspect Biol* **6**, a020644

18. Hoffhines, A. J., Damoc, E., Bridges, K. G., Leary, J. A., and Moore, K. L. (2006) Detection and purification of tyrosine-sulfated proteins using a novel anti-sulfotyrosine monoclonal antibody. *J Biol Chem* **281**, 37877-37887

19. Scurci, I., Akondi, K. B., Pinheiro, I., Paolini-Bertrand, M., Borgeat, A., Cerini, F., and Hartley, O. (2021) CCR5 tyrosine sulfation heterogeneity generates cell surface receptor subpopulations with different ligand binding properties. *Biochim Biophys Acta Gen Subj* **1865**, 129753

20. Bian, Y., Li, L., Dong, M., Liu, X., Kaneko, T., Cheng, K., Liu, H., Voss, C., Cao, X., Wang, Y., Litchfield, D., Ye, M., Li, S. S., and Zou, H. (2016) Ultra-deep tyrosine phosphoproteomics enabled by a phosphotyrosine superbinder. *Nat Chem Biol* **12**, 959-966

21. Ju, T., Niu, W., and Guo, J. (2016) Evolution of Src Homology 2 (SH2) domain to recognize sulfotyrosine. *ACS Chem Biol* **11**, 2551-2557

22. Lawrie, J., Waldrop, S., Morozov, A., Niu, W., and Guo, J. (2021) Engineering of a small protein scaffold to recognize sulfotyrosine with high specificity. *ACS Chem Biol* **16**, 1508-1517

23. Farzan, M., Chung, S., Li, W., Vasilieva, N., Wright, P. L., Schnitzler, C. E., Marchione, R. J., Gerard, C., Gerard, N. P., Sodroski, J., and Choe, H. (2002) Tyrosine-sulfated peptides functionally reconstitute a CCR5 variant lacking a critical amino-terminal region. *J Biol Chem* **277**, 40397-40402

24. Lechtenberg, B. C., Freund, S. M., and Huntington, J. A. (2014) GpIb alpha interacts exclusively with exosite II of thrombin. *J Mol Biol* **426**, 881-893

25. Stone, M. J., and Payne, R. J. (2015) Homogeneous sulfopeptides and sulfoproteins: synthetic approaches and applications to characterize the effects of tyrosine sulfation on biochemical function. *Acc Chem Res* **48**, 2251-2261

26. Watson, E. E., Ripoll-Rozada, J., Lee, A. C., Wu, M. C. L., Franck, C., Pasch, T., Premdjee, B., Sayers, J., Pinto, M. F., Martins, P. M., Jackson, S. P., Pereira, P. J. B., and Payne, R. J. (2019) Rapid assembly and profiling of an anticoagulant sulfoprotein library. *Proc Natl Acad Sci U S A* **116**, 13873-13878

27. Maxwell, J. W. C., and Payne, R. J. (2020) Revealing the functional roles of tyrosine sulfation using synthetic sulfopeptides and sulfoproteins. *Curr Opin Chem Biol* **58**, 72-85

28. Dowman, L. J., Agten, S. M., Ripoll-Rozada, J., Calisto, B. M., Pereira, P. J. B., and Payne, R. J. (2021) Synthesis and evaluation of peptidic thrombin inhibitors bearing acid-stable sulfotyrosine analogues. *Chem Commun (Camb)* **57**, 10923-10926

29. Liu, C. C., and Schultz, P. G. (2006) Recombinant expression of selectively sulfated proteins in *Escherichia coli*. *Nat Biotechnol* **24**, 1436-1440

30. Liu, C. C., Brustad, E., Liu, W., and Schultz, P. G. (2007) Crystal structure of a biosynthetic sulfo-hirudin complexed to thrombin. *J Am Chem Soc* **129**, 10648-10649

31. Schwessinger, B., Li, X., Ellinghaus, T. L., Chan, L. J., Wei, T., Joe, A., Thomas, N., Pruitt, R., Adams, P. D., Chern, M. S., Petzold, C. J., Liu, C. C., and Ronald, P. C. (2016) A second-generation expression system for tyrosine-sulfated proteins and its application in crop protection. *Integr Biol (Camb)* **8**, 542-545

32. Li, X., Hitomi, J., and Liu, C. C. (2018) Characterization of a sulfated anti-HIV antibody using an expanded genetic code. *Biochemistry* **57**, 2903-2907

33. Italia, J. S., Peeler, J. C., Hillenbrand, C. M., Latour, C., Weerapana, E., and Chatterjee, A. (2020) Genetically encoded protein sulfation in mammalian cells. *Nat Chem Biol* **16**, 379-382

34. Rapp, C., Klerman, H., Levine, E., and McClendon, C. L. (2013) Hydrogen bond strengths in phosphorylated and sulfated amino acid residues. *PLoS One* **8**, e57804

35. Hunter, T. (2015) Discovering the first tyrosine kinase. *Proc Natl Acad Sci U S A* **112**, 7877-7882

36. Lim, W. A., and Pawson, T. (2010) Phosphotyrosine signaling: evolving a new cellular communication system. *Cell* **142**, 661-667

37. Jin, J., and Pawson, T. (2012) Modular evolution of phosphorylation-based signalling systems. *Philos Trans R Soc Lond B Biol Sci* **367**, 2540-2555

38. Kaufmann, C., and Sauter, M. (2019) Sulfated plant peptide hormones. *J Exp Bot* **70**, 4267-4277

39. Hofsteenge, J., Stone, S. R., Donella-Deana, A., and Pinna, L. A. (1990) The effect of substituting phosphotyrosine for sulphotyrosine on the activity of hirudin. *Eur J Biochem* **188**, 55-59

40. Phillips, A. J., Taleski, D., Koplinski, C. A., Getschman, A. E., Moussouras, N. A., Richard, A. M., Peterson, F. C., Dwinell, M. B., Volkman, B. F., Payne, R. J., and Veldkamp, C. T. (2017) CCR7 sulfotyrosine enhances CCL21 binding. *Int J Mol Sci* **18**, 1857

41. Kanyo, Z. F., and Christianson, D. W. (1991) Biological recognition of phosphate and sulfate. *J Biol Chem* **266**, 4264-4268

42. Quiocho, F. A. (1996) Atomic basis of the exquisite specificity of phosphate and sulfate transport receptors. *Kidney Int* **49**, 943-946

43. Stone, M. J., Chuang, S., Hou, X., Shoham, M., and Zhu, J. Z. (2009) Tyrosine sulfation: an increasingly recognised post-translational modification of secreted proteins. *N Biotechnol* **25**, 299-317

44. Ludeman, J. P., and Stone, M. J. (2014) The structural role of receptor tyrosine sulfation in chemokine recognition. *Br J Pharmacol* **171**, 1167-1179

45. Matsubayashi, Y. (2014) Posttranslationally modified small-peptide signals in plants. *Annu Rev Plant Biol* **65**, 385-413

46. Stone, M. J., Hayward, J. A., Huang, C., Z, E. H., and Sanchez, J. (2017) Mechanisms of regulation of the chemokine-receptor network. *Int J Mol Sci* **18**

47. Kristensen, K. K., Midtgaard, S. R., Mysling, S., Kovrov, O., Hansen, L. B., Skar-Gislinge, N., Beigneux, A. P., Kragelund, B. B., Olivecrona, G., Young, S. G., Jorgensen, T. J. D., Fong, L. G., and Ploug, M. (2018) A disordered acidic domain in GPIHBP1 harboring a sulfated tyrosine regulates lipoprotein lipase. *Proc Natl Acad Sci U S A* **115**, E6020-E6029

48. Doll, N. M., Royek, S., Fujita, S., Okuda, S., Chamot, S., Stintzi, A., Widiez, T., Hothorn, M., Schaller, A., Geldner, N., and Ingram, G. (2020) A two-way molecular dialogue between embryo and endosperm is required for seed development. *Science* **367**, 431-435

49. Johnston, R. J., Su, L. J., Pinckney, J., Critton, D., Boyer, E., Krishnakumar, A., Corbett, M., Rankin, A. L., Dibella, R., Campbell, L., Martin, G. H., Lemar, H., Cayton, T., Huang, R. Y., Deng, X., Nayeem, A., Chen, H., Ergel, B., Rizzo, J. M., Yamniuk, A. P., Dutta, S., Ngo, J., Shorts, A. O., Ramakrishnan, R., Kozhich, A., Holloway, J., Fang, H., Wang, Y. K., Yang, Z., Thiam, K., Rakestraw, G., Rajpal, A., Sheppard, P., Quigley, M., Bahjat, K. S., and Korman, A. J. (2019) VISTA is an acidic pH-selective ligand for PSGL-1. *Nature* **574**, 565-570

50. Franck, C., Foster, S. R., Johansen-Leete, J., Chowdhury, S., Cielesh, M., Bhusal, R. P., Mackay, J. P., Larance, M., Stone, M. J., and Payne, R. J. (2020) Semisynthesis of an evasin from tick saliva reveals a critical role of tyrosine sulfation for chemokine binding and inhibition. *Proc Natl Acad Sci U S A* **117**, 12657-12664

51. Calisto, B. M., Ripoll-Rozada, J., Dowman, L. J., Franck, C., Agten, S. M., Parker, B. L., Veloso, R. C., Vale, N., Gomes, P., de Sanctis, D., Payne, R. J., and Pereira, P. J. B. (2021) Sulfotyrosine-mediated recognition of human thrombin by a tsetse fly anticoagulant mimics physiological substrates. *Cell Chem Biol* **28**, 26-33 e28

52. Günal, S., Hardman, R., Kopriva, S., and Mueller, J. W. (2019) Sulfation pathways from red to green. *J Biol Chem* **294**, 12293-12312

53. Zhao, C., Wang, Y., Chan, K. X., Marchant, D. B., Franks, P. J., Randall, D., Tee, E. E., Chen, G., Ramesh, S., Phua, S. Y., Zhang, B., Hills, A., Dai, F., Xue, D., Gilliham, M., Tyerman, S., Nevo, E., Wu, F., Zhang, G., Wong, G. K., Leebens-Mack, J. H., Melkonian, M., Blatt, M. R., Soltis, P. S., Soltis, D. E., Pogson, B. J., and Chen, Z. H. (2019) Evolution of chloroplast retrograde signaling facilitates green plant adaptation to land. *Proc Natl Acad Sci U S A* **116**, 5015-5020

54. Kakuta, Y., Pedersen, L. G., Pedersen, L. C., and Negishi, M. (1998) Conserved structural motifs in the sulfotransferase family. *Trends Biochem Sci* **23**, 129-130

55. Negishi, M., Pedersen, L. G., Petrotchenko, E., Shevtsov, S., Gorokhov, A., Kakuta, Y., and Pedersen, L. C. (2001) Structure and function of sulfotransferases. *Arch Biochem Biophys* **390**, 149-157

56. Yang, Y. S., Wang, C. C., Chen, B. H., Hou, Y. H., Hung, K. S., and Mao, Y. C. (2015) Tyrosine sulfation as a protein post-translational modification. *Molecules* **20**, 2138-2164

57. Dong, D., Ako, R., and Wu, B. (2012) Crystal structures of human sulfotransferases: insights into the mechanisms of action and substrate selectivity. *Expert Opin Drug Metab Toxicol* **8**, 635-646

58. Rath, V. L., Verdugo, D., and Hemmerich, S. (2004) Sulfotransferase structural biology and inhibitor discovery. *Drug Discov Today* **9**, 1003-1011

59. Esko, J. D., and Selleck, S. B. (2002) Order out of chaos: assembly of ligand binding sites in heparan sulfate. *Annu Rev Biochem* **71**, 435-471

60. Xu, D., and Esko, J. D. (2014) Demystifying heparan sulfate-protein interactions. *Annu Rev Biochem* **83**, 129-157

61. Mishiro, E., Sakakibara, Y., Liu, M. C., and Suiko, M. (2006) Differential enzymatic characteristics and tissue-specific expression of human TPST-1 and TPST-2. *J Biochem* **140**, 731-737

62. Ouyang, Y. B., Crawley, J. T., Aston, C. E., and Moore, K. L. (2002) Reduced body weight and increased postimplantation fetal death in tyrosylprotein sulfotransferase-1-deficient mice. *J Biol Chem* **277**, 23781-23787

63. Sasaki, N., Hosoda, Y., Nagata, A., Ding, M., Cheng, J. M., Miyamoto, T., Okano, S., Asano, A., Miyoshi, I., and Agui, T. (2007) A mutation in *Tpst2* encoding tyrosylprotein sulfotransferase causes dwarfism associated with hypothyroidism. *Mol Endocrinol* **21**, 1713-1721

64. Westmuckett, A. D., Hoffhines, A. J., Borghei, A., and Moore, K. L. (2008) Early postnatal pulmonary failure and primary hypothyroidism in mice with combined TPST-1 and TPST-2 deficiency. *Gen Comp Endocrinol* **156**, 145-153

65. Hoffhines, A. J., Jen, C. H., Leary, J. A., and Moore, K. L. (2009) Tyrosylprotein sulfotransferase-2 expression is required for sulfation of RNase 9 and Mfge8 in vivo. *J Biol Chem* **284**, 3096-3105

66. Westmuckett, A. D., Siefert, J. C., Tesiram, Y. A., Pinson, D. M., and Moore, K. L. (2013) Salivary gland hypofunction in tyrosylprotein sulfotransferase-2 knockout mice is due to primary hypothyroidism. *PLoS One* **8**, e71822

67. Bundgaard, J. R., Vuust, J., and Rehfeld, J. F. (1997) New consensus features for tyrosine O-sulfation determined by mutational analysis. *J Biol Chem* **272**, 21700-21705

68. Lin, W. H., Larsen, K., Hortin, G. L., and Roth, J. A. (1992) Recognition of substrates by tyrosylprotein sulfotransferase. Determination of affinity by acidic amino acids near the target sites. *J Biol Chem* **267**, 2876-2879

69. Niehrs, C., Kraft, M., Lee, R. W., and Huttner, W. B. (1990) Analysis of the substrate specificity of tyrosylprotein sulfotransferase using synthetic peptides. *J Biol Chem* **265**, 8525-8532

70. Tanaka, S., Nishiyori, T., Kojo, H., Otsubo, R., Tsuruta, M., Kurogi, K., Liu, M. C., Suiko, M., Sakakibara, Y., and Kakuta, Y. (2017) Structural basis for the broad substrate specificity of the human tyrosylprotein sulfotransferase-1. *Sci Rep* **7**, 8776

71. Teramoto, T., Fujikawa, Y., Kawaguchi, Y., Kurogi, K., Soejima, M., Adachi, R., Nakanishi, Y., Mishiro-Sato, E., Liu, M. C., Sakakibara, Y., Suiko, M., Kimura, M., and Kakuta, Y. (2013) Crystal structure of human tyrosylprotein sulfotransferase-2 reveals the mechanism of protein tyrosine sulfation reaction. *Nat Commun* **4**, 1572

72. Nedumpully-Govindan, P., Li, L., Alexov, E. G., Blenner, M. A., and Ding, F. (2014) Structural and energetic determinants of tyrosylprotein sulfotransferase sulfation specificity. *Bioinformatics* **30**, 2302-2309

73. Komori, R., Amano, Y., Ogawa-Ohnishi, M., and Matsubayashi, Y. (2009) Identification of tyrosylprotein sulfotransferase in *Arabidopsis*. *Proc Natl Acad Sci U S A* **106**, 15067-15072

74. Moore, K. L. (2009) Protein tyrosine sulfation: a critical posttranslation modification in plants and animals. *Proc Natl Acad Sci U S A* **106**, 14741-14742

75. Xu, Y., Moon, A. F., Xu, S., Krahn, J. M., Liu, J., and Pedersen, L. C. (2017) Structure based substrate specificity analysis of heparan sulfate 6-O-sulfotransferases. *ACS Chem Biol* **12**, 73-82

76. Pruitt, R. N., Joe, A., Zhang, W., Feng, W., Stewart, V., Schwessinger, B., Dinneny, J. R., and Ronald, P. C. (2017) A microbially derived tyrosine-sulfated peptide mimics a plant peptide hormone. *New Phytol* **215**, 725-736

77. Pruitt, R. N., Schwessinger, B., Joe, A., Thomas, N., Liu, F., Albert, M., Robinson, M. R., Chan, L. J., Luu, D. D., Chen, H., Bahar, O., Daudi, A., De Vleesschauwer, D., Caddell, D., Zhang, W., Zhao, X., Li, X., Heazlewood, J. L., Ruan, D., Majumder, D., Chern, M., Kalbacher, H., Midha, S., Patil, P. B., Sonti, R. V., Petzold, C. J., Liu, C. C., Brodbelt, J. S., Felix, G., and Ronald, P. C. (2015) The rice immune receptor XA21 recognizes a tyrosine-sulfated protein from a Gram-negative bacterium. *Sci Adv* **1**, e1500245

78. Luu, D. D., Joe, A., Chen, Y., Parys, K., Bahar, O., Pruitt, R. N., Chan, L. J. G., Petzold, C. J., Long, K., Adamchak, C., Stewart, V., Belkhadir, Y., and Ronald, P. C. (2019) Biosynthesis and secretion of the microbial sulfated peptide RaxX and binding to the rice XA21 immune receptor. *Proc Natl Acad Sci U S A* **116**, in press

79. Okamoto, S., Tabata, R., and Matsubayashi, Y. (2016) Long-distance peptide signaling essential for nutrient homeostasis in plants. *Curr Opin Plant Biol* **34**, 35-40

80. Matsubayashi, Y. (2018) Exploring peptide hormones in plants: identification of four peptide hormone-receptor pairs and two post-translational modification enzymes. *Proc Jpn Acad Ser B Phys Biol Sci* **94**, 59-74

81. Stührwohldt, N., and Schaller, A. (2019) Regulation of plant peptide hormones and growth factors by post-translational modification. *Plant Biol (Stuttg)* **21 Suppl 1**, 49-63

82. Miller, L. J., and Desai, A. J. (2016) Metabolic Actions of the Type 1 Cholecystokinin Receptor: Its Potential as a Therapeutic Target. *Trends Endocrinol Metab* **27**, 609-619

83. Rehfeld, J. F. (2017) Cholecystokinin-From Local Gut Hormone to Ubiquitous Messenger. *Front Endocrinol (Lausanne)* **8**, 47

84. Breiden, M., and Simon, R. (2016) Q&A: How does peptide signaling direct plant development? *BMC Biol* **14**, 58

85. Tang, D., Wang, G., and Zhou, J. M. (2017) Receptor kinases in plant-pathogen interactions: more than pattern recognition. *Plant Cell* **29**, 618-637

86. Song, W., Liu, L., Wang, J., Wu, Z., Zhang, H., Tang, J., Lin, G., Wang, Y., Wen, X., Li, W., Han, Z., Guo, H., and Chai, J. (2016) Signature motif-guided identification of receptors for peptide hormones essential for root meristem growth. *Cell Res* **26**, 674-685

87. Okuda, S., Fujita, S., Moretti, A., Hohmann, U., Doblas, V. G., Ma, Y., Pfister, A., Brandt, B., Geldner, N., and Hothorn, M. (2020) Molecular mechanism for the recognition of sequence-divergent CIF peptides by the plant receptor kinases GSO1/SGN3 and GSO2. *Proc Natl Acad Sci U S A* **117**, 2693-2703

88. Wang, J., Li, H., Han, Z., Zhang, H., Wang, T., Lin, G., Chang, J., Yang, W., and Chai, J. (2015) Allosteric receptor activation by the plant peptide hormone phytosulfokine. *Nature* **525**, 265-268

89. Sauter, M. (2015) Phytosulfokine peptide signalling. *J Exp Bot* **66**, 5161-5169

90. Dufresne, M., Seva, C., and Fourmy, D. (2006) Cholecystokinin and gastrin receptors. *Physiol Rev* **86**, 805-847

91. Miller, L. J., and Gao, F. (2008) Structural basis of cholecystokinin receptor binding and regulation. *Pharmacol Ther* **119**, 83-95

92. Jiang, X., Dias, J. A., and He, X. (2014) Structural biology of glycoprotein hormones and their receptors: insights to signaling. *Mol Cell Endocrinol* **382**, 424-451

93. Jiang, X., Liu, H., Chen, X., Chen, P. H., Fischer, D., Sriraman, V., Yu, H. N., Arkinstall, S., and He, X. (2012) Structure of follicle-stimulating hormone in complex with the entire ectodomain of its receptor. *Proc Natl Acad Sci U S A* **109**, 12491-12496

94. Bonomi, M., Busnelli, M., Persani, L., Vassart, G., and Costagliola, S. (2006) Structural differences in the hinge region of the glycoprotein hormone receptors: evidence from the sulfated tyrosine residues. *Mol Endocrinol* **20**, 3351-3363

95. Barnum, S. R. (2017) Complement: A primer for the coming therapeutic revolution. *Pharmacol Ther* **172**, 63-72

96. Gao, J., Choe, H., Bota, D., Wright, P. L., Gerard, C., and Gerard, N. P. (2003) Sulfation of tyrosine 174 in the human C3a receptor is essential for binding of C3a anaphylatoxin. *J Biol Chem* **278**, 37902-37908

97. Klos, A., Wende, E., Wareham, K. J., and Monk, P. N. (2013) International Union of Basic and Clinical Pharmacology. [corrected]. LXXXVII. Complement peptide C5a, C4a, and C3a receptors. *Pharmacol Rev* **65**, 500-543

98. Chen, B. (2019) Molecular mechanism of HIV-1 entry. *Trends Microbiol* **27**, 878-891

99. Rizzuto, C. D., Wyatt, R., Hernandez-Ramos, N., Sun, Y., Kwong, P. D., Hendrickson, W. A., and Sodroski, J. (1998) A conserved HIV gp120 glycoprotein structure involved in chemokine receptor binding. *Science* **280**, 1949-1953

100. Farzan, M., Mirzabekov, T., Kolchinsky, P., Wyatt, R., Cayabyab, M., Gerard, N. P., Gerard, C., Sodroski, J., and Choe, H. (1999) Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. *Cell* **96**, 667-676

101. Shaik, M. M., Peng, H., Lu, J., Rits-Volloch, S., Xu, C., Liao, M., and Chen, B. (2019) Structural basis of coreceptor recognition by HIV-1 envelope spike. *Nature* **565**, 318-323

102. Fauci, A. S. (1996) Host factors and the pathogenesis of HIV-induced disease. *Nature* **384**, 529-534

103. Graham, G. J., Handel, T. M., and Proudfoot, A. E. I. (2019) Leukocyte adhesion: reconceptualizing chemokine presentation by glycosaminoglycans. *Trends Immunol* **40**, 472-481

104. Kufareva, I. (2016) Chemokines and their receptors: insights from molecular modeling and crystallography. *Curr Opin Pharmacol* **30**, 27-37

105. Arimont, M., Sun, S. L., Leurs, R., Smit, M., de Esch, I. J. P., and de Graaf, C. (2017) Structural analysis of chemokine receptor-ligand iInteractions. *J Med Chem* **60**, 4735-4779

106. Kufareva, I., Gustavsson, M., Zheng, Y., Stephens, B. S., and Handel, T. M. (2017) What do structures tell us about chemokine receptor function and antagonism? *Annu Rev Biophys* **46**, 175-198

107. Bhusal, R. P., Foster, S. R., and Stone, M. J. (2020) Structural basis of chemokine and receptor interactions: Key regulators of leukocyte recruitment in inflammatory responses. *Protein Sci* **29**, 420-432

108. Choe, H., Li, W., Wright, P. L., Vasilieva, N., Venturi, M., Huang, C. C., Grundner, C., Dorfman, T., Zwick, M. B., Wang, L., Rosenberg, E. S., Kwong, P. D., Burton, D. R., Robinson, J. E., Sodroski, J. G., and Farzan, M. (2003) Tyrosine sulfation of human antibodies contributes to recognition of the CCR5 binding region of HIV-1 gp120. *Cell* **114**, 161-170

109. Farzan, M., Choe, H., Vaca, L., Martin, K., Sun, Y., Desjardins, E., Ruffing, N., Wu, L., Wyatt, R., Gerard, N., Gerard, C., and Sodroski, J. (1998) A tyrosine-rich region in the N terminus of CCR5 is important for human immunodeficiency virus type 1 entry and mediates an association between gp120 and CCR5. *J Virol* **72**, 1160-1164

110. Rabut, G. E., Konner, J. A., Kajumo, F., Moore, J. P., and Dragic, T. (1998) Alanine substitutions of polar and nonpolar residues in the amino-terminal domain of CCR5 differently impair entry of macrophage- and dualtropic isolates of human immunodeficiency virus type 1. *J Virol* **72**, 3464-3468

111. Liu, X., Malins, L. R., Roche, M., Sterjovski, J., Duncan, R., Garcia, M. L., Barnes, N. C., Anderson, D. A., Stone, M. J., Gorry, P. R., and Payne, R. J. (2014) Site-selective solid-phase synthesis of a CCR5 sulfopeptide library to interrogate HIV binding and entry. *ACS Chem Biol* **9**, 2074-2081

112. Huang, C. C., Venturi, M., Majeed, S., Moore, M. J., Phogat, S., Zhang, M. Y., Dimitrov, D. S., Hendrickson, W. A., Robinson, J., Sodroski, J., Wyatt, R., Choe, H., Farzan, M., and Kwong, P. D. (2004) Structural basis of tyrosine sulfation and V_H-gene usage in antibodies that recognize the HIV type 1 coreceptor-binding site on gp120. *Proc Natl Acad Sci U S A* **101**, 2706-2711

113. Stanfield, R. L., and Wilson, I. A. (2014) Antibody structure. *Microbiol Spectr* **2**

114. Chen, F., Tzarum, N., Wilson, I. A., and Law, M. (2019) VH1-69 antiviral broadly neutralizing antibodies: genetics, structures, and relevance to rational vaccine design. *Curr Opin Virol* **34**, 149-159

115. Huang, C. C., Lam, S. N., Acharya, P., Tang, M., Xiang, S. H., Hussan, S. S., Stanfield, R. L., Robinson, J., Sodroski, J., Wilson, I. A., Wyatt, R., Bewley, C. A., and Kwong, P. D. (2007) Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science* **317**, 1930-1934

116. Liu, C. C., Choe, H., Farzan, M., Smider, V. V., and Schultz, P. G. (2009) Mutagenesis and evolution of sulfated antibodies using an expanded genetic code. *Biochemistry* **48**, 8891-8898

117. Yang, Z., Wang, H., Liu, A. Z., Gristick, H. B., and Bjorkman, P. J. (2019) Asymmetric opening of HIV-1 Env bound to CD4 and a coreceptor-mimicking antibody. *Nat Struct Mol Biol* **26**, 1167-1175

118. Gardner, M. R., Kattenhorn, L. M., Kondur, H. R., von Schaewen, M., Dorfman, T., Chiang, J. J., Haworth, K. G., Decker, J. M., Alpert, M. D., Bailey, C. C., Neale, E. S., Jr., Fellinger, C. H., Joshi, V. R., Fuchs, S. P., Martinez-Navio, J. M., Quinlan, B. D., Yao, A. Y., Mouquet, H., Gorman, J., Zhang, B., Poignard, P., Nussenzweig, M. C., Burton, D. R., Kwong, P. D., Piatak, M., Jr., Lifson, J. D., Gao, G., Desrosiers, R. C., Evans, D. T., Hahn, B. H., Ploss, A., Cannon, P. M., Seaman, M. S., and Farzan, M. (2015) AAV-expressed eCD4-Ig provides durable protection from multiple SHIV challenges. *Nature* **519**, 87-91

119. Hughes, C. E., and Nibbs, R. J. B. (2018) A guide to chemokines and their receptors. *FEBS J* **285**, 2944-2971

120. Zheng, Y., Han, G. W., Abagyan, R., Wu, B., Stevens, R. C., Cherezov, V., Kufareva, I., and Handel, T. M. (2017) Structure of CC chemokine receptor 5 with a potent chemokine antagonist reveals mechanisms of chemokine recognition and molecular mimicry by HIV. *Immunity* **46**, 1005-1017 e1005

121. Wedemeyer, M. J., Mueller, B. K., Bender, B. J., Meiler, J., and Volkman, B. F. (2020) Comparative modeling and docking of chemokine-receptor interactions with Rosetta. *Biochem Biophys Res Commun* **528**, 389-397

122. Olsen, J. G., Teilum, K., and Kragelund, B. B. (2017) Behaviour of intrinsically disordered proteins in protein-protein complexes with an emphasis on fuzziness. *Cell Mol Life Sci* **74**, 3175-3183

123. Abayev, M., Rodrigues, J., Srivastava, G., Arshava, B., Jaremko, L., Jaremko, M., Naider, F., Levitt, M., and Anglister, J. (2018) The solution structure of monomeric CCL5 in complex with a doubly sulfated N-terminal segment of CCR5. *FEBS J* **285**, 1988-2003

124. Kessler, N., Akabayov, S. R., Moseri, A., Cohen, L. S., Sakhapov, D., Bolton, D., Fridman, B., Kay, L. E., Naider, F., and Anglister, J. (2021) Allovalency observed by transferred NOE: interactions of sulfated tyrosine residues in the N-terminal segment of CCR5 with the CCL5 chemokine. *FEBS J* **288**, 1648-1663

125. Deshauer, C., Morgan, A. M., Ryan, E. O., Handel, T. M., Prestegard, J. H., and Wang, X. (2015) Interactions of the chemokine CCL5/RANTES with medium-sized chondroitin sulfate ligands. *Structure* **23**, 1066-1077

126. Chakrabarti, S., Hinczewski, M., and Thirumalai, D. (2017) Phenomenological and microscopic theories for catch bonds. *J Struct Biol* **197**, 50-56

127. McEver, R. P., and Zhu, C. (2010) Rolling cell adhesion. *Annu Rev Cell Dev Biol* **26**, 363-396

128. Somers, W. S., Tang, J., Shaw, G. D., and Camphausen, R. T. (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

129. Xiao, B., Tong, C., Jia, X., Guo, R., Lu, S., Zhang, Y., McEver, R. P., Zhu, C., and Long, M. (2012) Tyrosine replacement of PSGL-1 reduces association kinetics with P- and L-selectin on the cell membrane. *Biophys J* **103**, 777-785

130. Hirata, T., Furukawa, Y., Yang, B. G., Hieshima, K., Fukuda, M., Kannagi, R., Yoshie, O., and Miyasaka, M. (2004) Human P-selectin glycoprotein ligand-1 (PSGL-1) interacts with the skin-associated chemokine CCL27 via sulfated tyrosines at the PSGL-1 amino terminus. *J Biol Chem* **279**, 51775-51782

131. Tinoco, R., Otero, D. C., Takahashi, A. A., and Bradley, L. M. (2017) PSGL-1: a new player in the immune checkpoint landscape. *Trends Immunol* **38**, 323-335

132. Farzan, M., Schnitzler, C. E., Vasilieva, N., Leung, D., Kuhn, J., Gerard, C., Gerard, N. P., and Choe, H. (2001) Sulfated tyrosines contribute to the formation of the C5a docking site of the human C5a anaphylatoxin receptor. *J Exp Med* **193**, 1059-1066

133. Ippel, J. H., de Haas, C. J., Bunschoten, A., van Strijp, J. A., Kruijtzer, J. A., Liskamp, R. M., and Kemmink, J. (2009) Structure of the tyrosine-sulfated C5a receptor N terminus in complex with chemotaxis inhibitory protein of Staphylococcus aureus. *J Biol Chem* **284**, 12363-12372

134. Nellenbach, K., and Brown, A. C. (2017) Peptide mimetic drugs for modulating thrombosis and hemostasis. *Drug Dev Res* **78**, 236-244

135. Plow, E. F., Wang, Y., and Simon, D. I. (2018) The search for new antithrombotic mechanisms and therapies that may spare hemostasis. *Blood* **131**, 1899-1902

136. Mackman, N., Bergmeier, W., Stouffer, G. A., and Weitz, J. I. (2020) Therapeutic strategies for thrombosis: new targets and approaches. *Nat Rev Drug Discov* **19**, 333-352

137. Bode, W. (2006) Structure and interaction modes of thrombin. *Blood Cells Mol Dis* **36**, 122-130

138. Huntington, J. A. (2013) Thrombin inhibition by the serpins. *J Thromb Haemost* **11 Suppl 1**, 254-264

139. Nowak, G., and Schror, K. (2007) Hirudin--the long and stony way from an anticoagulant peptide in the saliva of medicinal leech to a recombinant drug and beyond. A historical piece. *Thromb Haemost* **98**, 116-119

140. Greinacher, A., and Warkentin, T. E. (2008) The direct thrombin inhibitor hirudin. *Thromb Haemost* **99**, 819-829

141. Vu, T. K., Wheaton, V. I., Hung, D. T., Charo, I., and Coughlin, S. R. (1991) Domains specifying thrombin-receptor interaction. *Nature* **353**, 674-677

142. Huntington, J. A. (2005) Molecular recognition mechanisms of thrombin. *J Thromb Haemost* **3**, 1861-1872

143. Bock, P. E., Panizzi, P., and Verhamme, I. M. (2007) Exosites in the substrate specificity of blood coagulation reactions. *J Thromb Haemost* **5 Suppl 1**, 81-94

144. Thompson, R. E., Liu, X., Ripoll-Rozada, J., Alonso-Garcia, N., Parker, B. L., Pereira, P. J. B., and Payne, R. J. (2017) Tyrosine sulfation modulates activity of tick-derived thrombin inhibitors. *Nat Chem* **9**, 909-917

145. Pineda, A. O., Chen, Z. W., Marino, F., Mathews, F. S., Mosesson, M. W., and Di Cera, E. (2007) Crystal structure of thrombin in complex with fibrinogen gamma' peptide. *Biophys Chem* **125**, 556-559

146. Keragala, C. B., Draxler, D. F., McQuilten, Z. K., and Medcalf, R. L. (2018) Haemostasis and innate immunity - a complementary relationship: A review of the intricate relationship between coagulation and complement pathways. *Br J Haematol* **180**, 782-798

147. Kidmose, R. T., Laursen, N. S., Dobo, J., Kjaer, T. R., Sirotkina, S., Yatime, L., Sottrup-Jensen, L., Thiel, S., Gal, P., and Andersen, G. R. (2012) Structural basis for activation of the complement system by component C4 cleavage. *Proc Natl Acad Sci U S A* **109**, 15425-15430

148. Perry, A. J., Wijeyewickrema, L. C., Wilmann, P. G., Gunzburg, M. J., D'Andrea, L., Irving, J. A., Pang, S. S., Duncan, R. C., Wilce, J. A., Whisstock, J. C., and Pike, R. N. (2013) A molecular switch governs the interaction between the human complement protease C1s and its substrate, complement C4. *J Biol Chem* **288**, 15821-15829

149. Pang, S. S., Wijeyewickrema, L. C., Hor, L., Tan, S., Lameignere, E., Conway, E. M., Blom, A. M., Mohlin, F. C., Liu, X., Payne, R. J., Whisstock, J. C., and Pike, R. N. (2017) The structural basis for complement inhibition by gigastasin, a protease inhibitor from the giant Amazon leech. *J Immunol* **199**, 3883-3891

150. Li, W., Johnson, D. J., Esmon, C. T., and Huntington, J. A. (2004) Structure of the antithrombin-thrombin-heparin ternary complex reveals the antithrombotic mechanism of heparin. *Nat Struct Mol Biol* **11**, 857-862

151. Carter, W. J., Cama, E., and Huntington, J. A. (2005) Crystal structure of thrombin bound to heparin. *J Biol Chem* **280**, 2745-2749

152. Soares da Costa, D., Reis, R. L., and Pashkuleva, I. (2017) Sulfation of glycosaminoglycans and its implications in human health and disorders. *Annu Rev Biomed Eng* **19**, 1-26

153. Arora, R., Nimonkar, A. V., Baird, D., Wang, C., Chiu, C. H., Horton, P. A., Hanrahan, S., Cubbon, R., Weldon, S., Tschantz, W. R., Mueller, S., Brunner, R., Lehr, P., Meier, P., Ottl, J., Voznesensky, A., Pandey, P., Smith, T. M., Stojanovic, A., Flyer, A., Benson, T. E., Romanowski, M. J., and Trauger, J. W. (2019) Structure of lipoprotein lipase in complex with GPIHBP1. *Proc Natl Acad Sci U S A* **116**, 10360-10365

154. Önnerfjord, P., Heathfield, T. F., and Heinegård, D. (2004) Identification of tyrosine sulfation in extracellular leucine-rich repeat proteins using mass spectrometry. *J Biol Chem* **279**, 26-33

155. Jensen, M. M., and Karring, H. (2020) The origins and developments of sulfation-prone tyrosine-rich and acidic N- and C-terminal extensions of class II and III small leucine-rich repeat proteins shed light on connective tissue evolution in vertebrates. *BMC Evol Biol* **20**, 73

156. Tillgren, V., Önnerfjord, P., Haglund, L., and Heinegård, D. (2009) The tyrosine sulfate-rich domains of the LRR proteins fibromodulin and osteoadherin bind motifs of basic clusters in a variety of heparin-binding proteins, including bioactive factors. *J Biol Chem* **284**, 28543-28553

157. Tillgren, V., Mörgelin, M., Önnerfjord, P., Kalamajski, S., and Aspberg, A. (2016) The tyrosine sulfate domain of fibromodulin binds collagen and enhances fibril formation. *J Biol Chem* **291**, 23744-23755

158. Rosell-García, T., Paradela, A., Bravo, G., Dupont, L., Bekhouche, M., Colige, A., and Rodriguez-Pascual, F. (2019) Differential cleavage of lysyl oxidase by the metalloproteinases BMP1 and ADAMTS2/14 regulates collagen binding through a tyrosine sulfate domain. *J Biol Chem* **294**, 11087-11100

159. Sako, D., Comess, K. M., Barone, K. M., Camphausen, R. T., Cumming, D. A., and Shaw, G. D. (1995) A sulfated peptide segment at the amino terminus of PSGL-1 is critical for P-selectin binding. *Cell* **83**, 323-331

160. Farzan, M., Babcock, G. J., Vasilieva, N., Wright, P. L., Kiprilov, E., Mirzabekov, T., and Choe, H. (2002) The role of post-translational modifications of the CXCR4 amino terminus in stromal-derived factor 1 alpha association and HIV-1 entry. *J Biol Chem* **277**, 29484-29489

161. Byrne, D. P., Li, Y., Ngamlert, P., Ramakrishnan, K., Eyers, C. E., Wells, C., Drewry, D. H., Zuercher, W. J., Berry, N. G., Fernig, D. G., and Eyers, P. A. (2018) New tools for evaluating protein tyrosine sulfation: tyrosylprotein sulfotransferases (TPSTs) are novel targets for RAF protein kinase inhibitors. *Biochem J* **475**, 2435-2455

162. Shah, N. H., and Kuriyan, J. (2019) Understanding molecular mechanisms in cell signaling through natural and artificial sequence variation. *Nat Struct Mol Biol* **26**, 25-34

163. Bannert, N., Craig, S., Farzan, M., Sogah, D., Santo, N. V., 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

164. Cormier, E. G., Persuh, M., Thompson, D. A., Lin, S. W., Sakmar, T. P., Olson, W. C., and Dragic, T. (2000) Specific interaction of CCR5 amino-terminal domain peptides containing sulfotyrosines with HIV-1 envelope glycoprotein gp120. *Proc Natl Acad Sci U S A* **97**, 5762-5767

165. Seibert, C., Cadene, M., Sanfiz, A., Chait, B. T., and Sakmar, T. P. (2002) Tyrosine sulfation of CCR5 N-terminal peptide by tyrosylprotein sulfotransferases 1 and 2 follows a discrete pattern and temporal sequence. *Proc Natl Acad Sci U S A* **99**, 11031-11036

166. Jen, C. H., Moore, K. L., and Leary, J. A. (2009) Pattern and temporal sequence of sulfation of CCR5 N-terminal peptides by tyrosylprotein sulfotransferase-2: an assessment of the effects of N-terminal residues. *Biochemistry* **48**, 5332-5338

167. Watson, E. E., Liu, X., Thompson, R. E., Ripoll-Rozada, J., Wu, M., Alwis, I., Gori, A., Loh, C. T., Parker, B. L., Otting, G., Jackson, S., Pereira, P. J. B., and Payne, R. J. (2018) Mosquito-derived anophelin sulfoproteins are potent antithrombotics. *ACS Cent Sci* **4**, 468-476

168. Liu, F., McDonald, M., Schwessinger, B., Joe, A., Pruitt, R., Erickson, T., Zhao, X., Stewart, V., and Ronald, P. C. (2019) Variation and inheritance of the *Xanthomonas raxX-raxSTAB* gene cluster required for activation of XA21-mediated immunity. *Mol Plant Pathol* **20**, 656-672

169. Dorfman, T., Moore, M. J., Guth, A. C., Choe, H., and Farzan, M. (2006) A tyrosine-sulfated peptide derived from the heavy-chain CDR3 region of an HIV-1-neutralizing antibody binds gp120 and inhibits HIV-1 infection. *J Biol Chem* **281**, 28529-28535

170. Gardner, M. R., Fellinger, C. H., Kattenhorn, L. M., Davis-Gardner, M. E., Weber, J. A., Alfant, B., Zhou, A. S., Prasad, N. R., Kondur, H. R., Newton, W. A., Weisgrau, K. L., Rakasz, E. G., Lifson, J. D., Gao, G., Schultz-Darken, N., and Farzan, M. (2019) AAV-delivered eCD4-Ig protects rhesus macaques from high-dose SIVmac239 challenges. *Sci Transl Med* **11**, eaau5409

171. Li, Y., and Rebuffat, S. (2020) The manifold roles of microbial ribosomal peptide-based natural products in physiology and ecology. *J Biol Chem* **295**, 34-54

172. Bick, M. J., Banik, J. J., Darst, S. A., and Brady, S. F. (2010) Crystal structures of the glycopeptide sulfotransferase Teg12 in a complex with the teicoplanin aglycone. *Biochemistry* **49**, 4159-4168

173. McCarthy, J. G., Eisman, E. B., Kulkarni, S., Gerwick, L., Gerwick, W. H., Wipf, P., Sherman, D. H., and Smith, J. L. (2012) Structural basis of functional group activation by sulfotransferases in complex metabolic pathways. *ACS Chem Biol* **7**, 1994-2003

174. Shi, R., Lamb, S. S., Bhat, S., Sulea, T., Wright, G. D., Matte, A., and Cygler, M. (2007) Crystal structure of StaL, a glycopeptide antibiotic sulfotransferase from *Streptomyces toyocaensis*. *J Biol Chem* **282**, 13073-13086

175. Reimer, J. M., Haque, A. S., Tarry, M. J., and Schmeing, T. M. (2018) Piecing together nonribosomal peptide synthesis. *Curr Opin Struct Biol* **49**, 104-113

176. Hansson, K., and Stenflo, J. (2005) Post-translational modifications in proteins involved in blood coagulation. *J Thromb Haemost* **3**, 2633-2648

177. Ezban, M., Vad, K., and Kjalke, M. (2014) Turoctocog alfa (NovoEight(R))--from design to clinical proof of concept. *Eur J Haematol* **93**, 369-376

178. Bradford, H. N., and Krishnaswamy, S. (2019) Occlusion of anion-binding exosite 2 in meizothrombin explains its impaired ability to activate factor V. *J Biol Chem* **294**, 2422-2435

179. Schreuder, M., Reitsma, P. H., and Bos, M. H. A. (2019) Blood coagulation factor Va's key interactive residues and regions for prothrombinase assembly and prothrombin binding. *J Thromb Haemost* **17**, 1229-1239

180. Schneider, T. D., and Stephens, R. M. (1990) Sequence logos: a new way to display consensus sequences. *Nucleic Acids Res* **18**, 6097-6100

181. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res* **14**, 1188-1190