Journal of Biological Chemistry

JBC/2019/010277R1

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

Running title: Sulfotyrosine in 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 sulfotransferases (TPSTs) during maturation of certain secreted proteins. Here we consider three contexts for sTyr function. First, a single sTyr residue is critical for high-affinity peptide-receptor interactions in plant peptide hormones and animal receptors for glycopeptide hormones. Second, structurally flexible anionic segments often contain a cluster of two or three sTyr residues within a six-residue span. These sTyr residues are essential for coreceptor binding of the HIV-1 envelope spike protein during virus entry and for chemokine interactions with many chemokine receptors. Third, several proteins that interact with thrombin, central to normal blood-clotting, require the presence of sTyr residues in the context of acidic sequences termed hirudin-like motifs. Consequently, many proven and potential therapeutic proteins derived from blood-consuming invertebrates depend on sTyr residues for their activity. Technical advances in generating and documenting site-specific sTyr substitutions facilitate discovery and analysis, and promise to enable engineering of defined interaction determinants.

[167 words; limit 250]

Post-translational modifications influence protein activity in many ways. For this reason, 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, resulting in sulfotyrosyl (sTyr) residues. Sulfation is catalyzed in metazoa and plants by Golgi-localized **t**yrosyl-**p**rotein **s**ulfo**t**ransferase (TPST; **Fig. 1**), and therefore occurs only in extracellular proteins (3).

Although sTyr was discovered nearly 60 years ago (4), detecting and documenting this modification is challenging. The sTyr sulfate linkage is cleaved by the strong acid used in the Edman degradation method used to determine protein sequences, although it is generally stable in weak acid (5,6). Routine mass spectrometry methods do not reliably detect sTyr residues (7), and more recent protocols to minimize sulfate loss during peptide ionization and fragmentation have not yet been applied to large-scale analyses (8-10). Thus, systems-level technologies for detecting and validating sTyr have lagged behind those developed for some other post-translational modifications.

Research on sTyr occurrence and function previously also has been hampered by challenges to synthesizing sulfopeptides for use in binding assays. Earlier studies were limited by the maximum of two sTyr in a single peptide (11) or by the substitution of pTyr for sTyr (12). Today, homogeneous sulfopeptides can be synthesized in vitro with up to three sTyr residues (13-15). 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 a nonsense-reading tRNA is charged with exogenous sTyr (16-19). This sTyr-specific expanded genetic code functions also in mammalian cells (20). Given these earlier technical challenges, the extent of sTyr occurrence has been incompletely defined (6,9,21). Nevertheless, recent discoveries of previously-unknown

sTyr sequence contexts (22,23) and functions (14,24-26) are helping to close this knowledge gap. Here we consider a few well-studied examples that illustrate sTyr functions in diverse settings.

Why sulfotyrosine?

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 (27). However, sTyr makes weaker hydrogen bonds due to its lesser charge (–1 vs. –2) and smaller dipole moment (27).

Functions for pTyr have been studied extensively: its critical role in cellular growth control was first demonstrated in tumor virus-encoded oncogenic proteins (28). Tyrosyl protein kinase activity serves as "writer" for reversible signaling, whereas separate phosphoprotein phosphatase activity serves as "eraser." For signal propagation, small pTyr binding domains such as ***s****rc* **h**omology **2** (SH2) provide "reader" function for multi-domain output complexes (29,30).

By contrast, sTyr formation has no known response to cellular signaling, and neither "reader" (e.g., portable sTyr-binding domains) nor "eraser" (e.g., sulfoprotein sulfatase) components are known, save for SH2 domain sequences engineered to recognize sTyr (31). Instead, sTyr is a long-lived post-translational modification (32), and in most cases the sulfate ester likely is stable in physiological conditions (6). This stability means that sTyr residues can function in macromolecular tethers, as described below for thrombin binding-partners. 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 (12,33), the sTyr sulfate makes distinct ionic contacts (34,35) and therefore in many cases provides a unique interaction specificity determinant. This stringency is illustrated by cases wherein sTyr binding is conserved in proteins that mimic sTyr-dependent protein-protein interactions (described below).

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. 1**), the activated intermediate for sulfur assimilation. PAPS synthesis and function are described elsewhere (36). TPST and polysaccharide sulfotransferases, anchored to the Golgi lumen through an amino-terminal transmembrane sequence, modify proteins and sulfated **g**lycos**a**mino**g**lycans (GAGs) in spatial and temporal coordination with other modifications such as protein glycosylation (2,37).

Metazoan TPST displays broad substrate specificity. Most substrates with at least moderate affinity have an Asp residue just proximal to the sulfoaccepting Tyr residue (38-43) (**Fig. 2**). Non-conserved flanking residues, often rich in acidic Asp and Glu residues, make multiple interactions with different TPST residues (38,39,42). Overall, these broader sequence features and their roles in metazoan TPST binding affinity and sulfation efficiency remain obscure (44).

Plants encode a single TPST (45). In the model plant *Arabidopsis thaliana*, a null mutant displays several developmental phenotypes including dwarfism (46). These

phenotypes are consistent with the functions for the only known plant TPST substrates, sulfopeptide hormones involved in multicellular development (described below).

Unlike metazoans and plants (including chlorophyte algae; reference (45), most bacteria, archaea and eukaryotic microbes do not encode TPST. Nevertheless, TPST is made by some species of plant-pathogenic bacteria in the genus *Xanthomonas* that also synthesize RaxX (**r**equired for **a**ctivation of **X**a21-mediated immunity), a molecular mimic of the plant sulfopeptide hormone PSY (**p**eptide **s**ulfated on t**y**rosine; see below) (47,48) (**Fig. 2A**). Bacterial TPST is similar to the Golgi-localized metazoan TPST, except that it acts in the bacterial cytoplasm prior to substrate secretion (49).

Three sequence contexts for sTyr function

Over the years, sTyr residues have been documented in a variety of proteins and peptides (3,50-52). Most sulfated proteins have three or fewer sTyr residues, usually clustered within a span of six residues. Exceptions include fibromodulin and related proteins, in which sTyr residue tracts are hypothesized to mimic sulfated GAGs and promote extracellular matrix formation (53,54). Thus, sTyr-based interactions usually involve relatively short protein segments, often within large multidomain proteins.

We sorted sTyr-containing sequences into three broad groups as a strategy to search for functional similarities and differences: (i) sequences with a single sTyr residue, illustrated by plant sulfopeptide hormones (**Fig. 2A**); (ii) sequences with a cluster of two or three sTyr residues, illustrated by chemokine receptor interactions with HIV and with chemokines (**Fig. 2B**); and (iii) sTyr residues in acidic sequence motifs that interact with thrombin (**Fig. 2C**). Many additional examples of sTyr occurrence and function are considered elsewhere (7,15,51,55).

A single sTyr residue increases binding affinity for peptide hormone-receptor pairs

The first sequence group (**Fig. 2A; Fig. S1**) contains one sTyr residue, primarily in certain peptide hormones or their receptors. In these cases, the sTyr-containing protein binds a unique partner, either (i) a sulfopeptide ligand binding its receptor, illustrated by plant sulfopeptide hormones; or (ii) a peptide ligand binding its sulfated receptor, illustrated by the sulfated **f**ollicle-**s**timulating **h**ormone **r**eceptor (FSHR) (**Fig. 3A**). Additional examples of single sTyr residue functions in animals include sulfopeptide hormones (56,57) and the sulfated receptor for complement C3a (58) (supporting information 1.1 and 1.2, respectively).

Land plants use secreted peptide hormones of roughly 10-20 residues to control growth and development through both short- and long-range intercellular signaling. Most peptide hormones are encoded by multigene families (59,60). Mature sulfopeptide hormones result from post-translational modifications during Golgi transit and possess the invariant amino-terminal residue pair Asp-1 sTyr-2 (32,52,61,62) (**Fig. 2A**). The Asp residue presumably is important for directing plant TPST to the substrate Tyr residue.

Plant peptide hormones bind their cognate receptors through **l**eucine-**r**ich **r**epeat (LRR) extracellular domains (59,63). X-ray co-crystal structures show both nonsulfated and sulfated peptide hormones as making similar main-chain contacts throughout their length to the LRR solenoid concave surface (64-66) (**Fig. 4A**). Receptor binding to the sulfopeptide sTyr residue is inferred from 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 (67,68). Both structures reveal multiple contacts to the sTyr residue sulfate and C atoms (**Fig. 4B**). Although the contacting

residues are at the same relative positions in the two receptor LRR sequences, several specific residues and contacts are not conserved. Thus, these geometrically-similar sTyr binding sites do not require obvious consensus sequence features, presumably facilitating their formation (and loss) through evolution.

The plant sulfopeptide hormones tested display high affinities for their cognate receptor LRR domains, ranging from 1 to about 300 nM (23,49,67,68). In each case, the sTyr residue is critical for binding, contributing tenfold or more to affinity compared with the nonsulfated peptide. For example, the steady-state binding affinity for CIF2 hormone (**Fig. 2A**) binding decreases 200-fold or more upon changing either (i) the CIF2 peptide sTyr residue (to Tyr or Phe); or (ii) a pair of sTyr-interacting residues in the cognate receptor LRR domain. By contrast, substitutions in other peptide or receptor residues (including peptide hydroxy-Pro) decrease binding affinity only about tenfold (68). These results demonstrate the substantial contribution by the sTyr residue to sulfopeptide-receptor binding affinity.

The specific requirement of sTyr for functional interactions is illustrated by the RaxX peptide, a bacterial mimic of the plant PSY hormone (47,48) (**Fig. 2A**) (see above). Because RaxX requires sulfation for activity, the bacterium must synthesize TPST, nominally a eukaryotic function (48,49). This is the first of several examples in which a molecular mimic contains one or more critical sTyr residues.

Mirroring these interactions between nonsulfated receptors and sulfated peptide hormones, some nonsulfated ligands bind single sTyr-containing receptors (**Fig. 2A; Fig. S1**). 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. 4C**). This interdomain region contains an sTyr residue that is indispensable for hormone recognition and signaling (69,70).

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 (69,70) (**Fig. 4D**). Homologous receptors for thyroid stimulating hormone and leutenizing hormone similarly contain an essential interdomain flexible region with a single sTyr residue (71) (**Fig. S1**).

The plant sulfopeptide hormone **p**hyto**s**ulfo**k**ine (PSK) has unique sequence and receptor-binding mode compared to the sulfopeptides described above (32,72). The PSK pentapeptide contains two sTyr residues and binds its cognate receptor with about 40-fold higher affinity than the corresponding nonsulfated peptide (73). The PSK receptor LRR repeats contain a 36 residue interruption, termed an island domain, that contacts both sTyr residues through multiple interactions (73). These different features illustrate how sTyr residues are versatile interaction determinants even in a broadly similar context such as sulfopeptide-LRR binding.

A cluster of two or three sTyr residues contributes to combinatorial receptor-peptide interactions

The second sequence group (**Fig. 2B; Fig. S2**) has clusters of two or three sTyr residues, often interspersed with several acidic Asp and Glu residues. These intrinsically-disordered segments, illustrated by the amino-termini of chemokine receptors, enable specific binding to different partners (74) (**Fig. 3B**). Additional examples of sTyr cluster functions include **P**-**s**electin **g**lycoprotein **l**igand-**1** (PSGL-1)

(2,75) and the sulfated receptor for complement C5a (76,77) (supporting information 2.1 and 2.2, respectively).

Chemotactic cyto**kines** (chemokines) are secreted signaling proteins of approximately 75 residues that form chemical gradients to direct leukocyte movement toward injury or infection sites. Chemokines are classified according to sequence: **CC**-type chemokine **l**igands (CCL) such as CCL5 have adjacent Cys residues, whereas **CXC**-type chemokine **l**igands (CXCL) have an intervening non-Cys residue. Human cells express about 50 chemokines and about 20 chemokine receptors acting in different combinations in different tissues (55,78-80).

Chemokine receptors are G protein-coupled receptors in which the flexible amino-terminal anionic sequence segment often includes an sTyr cluster that is essential for high-affinity chemokine binding (81-83) (**Fig. 2B; Fig. S2**). Chemokine receptors have other functions; for example, CCR5 is a major T cell coreceptor for HIV-1 infection (84). Interest in sTyr function accelerated with discoveries that the CCR5 sTyr cluster is essential for virus entry (81) and is mimicked in some HIV-1 broadly-neutralizing antibodies (85). Several helpful reviews organize the literature on sTyr-containing chemokine receptors (33,51,55,83,86-88). Here we focus on chemokine receptor CCR5 interactions with HIV-1 and with the chemokine CCL5.

HIV-1 gp120 interaction with chemokine receptor CCR5

HIV-1 binding and entry depends on the viral envelope **g**lyco**p**rotein (gp) spike, a trimer of gp120-gp41 heterodimers (89). The spike initially binds to the cell-surface receptor CD4, thereby altering gp120 conformation to expose its bridging sheet element including coreceptor binding determinants (90). The CCR5 amino-terminal segment,

including sTyr residues 10 and 14 (**Fig. 2B**), binds to residues near the bridging sheet base to trigger membrane fusion and viral entry (81,89,91).

CCR5 interactions with gp120 are revealed through a cryo-EM structure that shows a three-member complex containing nonsulfated CCR5 coreceptor, gp120, and part of the CD4 receptor (89,91) (**Fig. 5A**). In this model, the CCR5 amino-terminal segment adopts an extended conformation across the gp120 bridging sheet element, with residues Tyr-10, Tyr-14 and Tyr-15 each occupying separate binding interfaces through multiple contacts to different gp120 residues (89,91) (**Fig. 5B**). Although residues Tyr-10 and Tyr-14 can be modeled as sTyr, the Tyr-15 binding interface cannot accommodate sulfate (91).

With the modeled sulfates added, residue sTyr-14 occupies a deep pocket lined with hydrophobic contacts and capped by electrostatic interactions, the latter for example with gp120 residue Asn-302 (**Fig. 5B**). By contrast, residue sTyr-10 has more peripheral intreactions with a distinct set of residues, for example with gp120 residue Arg-414 (**Fig. 5B**). Although the contacting residues are different, these sTyr-binding interfaces resemble those provided by FSH (**Fig. 4D**) and RGFR (**Fig. 4B**), respectively.

Three sets of experiments support and extend these predicted sTyr interactions (91). First, Ala and Phe missense substitutions in CCR5 identify residues Tyr-10, Tyr-14 and Tyr-15 as necessary for proper gp120-CCR5 interaction (92,93). Importantly, Phe can substitute for Tyr-15 but not for Tyr-10 or Tyr-14 (93). This suggests that sulfation is not necessary at position Tyr-15, because Phe resembles Tyr but cannot be sulfated. Second, sulfation at CCR5 residue Tyr-14 is essential for sulfopeptide binding to gp120 in vitro, whereas sulfation at residue Tyr-10 enhances binding only in the presence of sTyr-14 (94) (**Fig. 6A**). This result might have been be predicted from observing the

more extensive binding interface for residue sTyr-14 (**Fig. 4B**). Separately, sulfation at residue Tyr-15 has little effect on peptide binding, congruent with the absence of modeled sulfate at Tyr-15 in the cryo-EM model (94) and the permissivity of Phe at this position (93) (**Fig. 4B**). Thus, results obtained with both mutants and peptides support conclusions drawn from the structures.

HIV-1 gp120 interaction with sulfated anti-gp120 antibodies

The third set of experiments supporting the modeled sTyr interactions in the gp-120-CCR5 cryo-EM structure comes from studies with sTyr-containing anti-gp120 antibodies. Certain HIV-1-infected individuals produce antibodies that recognize the **CD4**-**i**nduced (CD4i) conformation of gp120 that exposes the bridging sheet element essential for CCR5 binding (85,95). Most of these CD4i antibodies are derived from the *Vh1-69* gene, which encodes **h**eavy chains with unusually long and diverse sequences for the third **c**omplementarity-**d**etermining region (**HCD3**). These protruding HCD3 loops can contact otherwise poorly-accessible epitopes (95-97).

The variable HCD3 sequence for several of these CD4i antibodies includes sTyr residues required for gp120 binding (85,95) (**Fig. 2B**). Two CDi antibodies, termed 412d and E51, have been studied in X-ray co-crystal structures with the gp120-CD4 complex (**Fig. 5C**). For antibody 412d, residues sTyr-100C and sTyr-100 bind the same gp120 interfaces as CCR5 residues sTyr-10 and sTyr-14, respectively (91,98). (The numbering accounts for variable numbers of residues, labeled 100A-100Q, between HCD3 residues at positions 100 and 101.)

By contrast, there is no sTyr-14 structural mimic in antibody E51. Although antibody residue sTyr-100I binds to the same gp120 interface as CCR5 residue sTyr-10,

antibody residue sTyr-100F instead makes distinct interactions that include contacts with two residues in the sTyr-10/sTyr-100I interface (**Fig. 5D**).

Antibody heavy chain variable region sulfoforms, expressed with the expanded genetic code, reveal distinct functions for the two sTyr residues. For antibody 412d, the sTyr-14 mimic (residue sTyr-100C) is essential for binding and the sTyr-10 mimic (residue sTyr-100) is ancillary (99), fully congruent with results for the CCR5 peptide sulfoforms (94). For antibody E51, the sTyr-10 mimicking residue sTyr-100I and the non-mimicking residue sTyr-100F likewise are essential and ancillary, respectively, for gp120 binding (19).

Antibody E51, more potent than 412d (100), therefore displays a different sTyr-gp120 interaction. Although the sTyr-14/sTyr-100C binding interface is essential for binding CCR5 and antibody 412d (**Fig. 4B**), E51 does not engage that interface, at least in the co-crystal (**Fig. 4D**). Additionally, 412d residue sTyr-100C helps to stabilize the gp120 V3 (variable loop 3) element, which mimicks chemokine binding (83) and inserts into the chemoreceptor CCR5 transmembrane helix bundle (98) (**Fig. 4B**). By contrast, the V3 loop is disordered in the E51-gp120 structure (100), which lacks interactions at the sTyr-14/sTyr-100C binding interface (**Fig. 4D**).

Thus, this broad binding region in gp120 can accomodate different sTyr-containing flexible loops, perhaps enabling interactions with sTyr-containing amino-terminal segments from different chemokine receptors (**Fig. S2**). Morever, a versatile sTyr-interacting region increases opportunities to meet the stereochemical constraints required specifically for binding sulfate (95).

Chemokine CCL5 interaction with chemokine receptor CCR5

The chemokine receptor CCR5 amino-terminal segment interacts differently with partner chemokines such as CCL5. The CCR5 sTyr cluster forms the core for binding the chemokine globular domain, and helps enable a given chemokine receptor to interact with multiple chemokines (55,78,83,88). Nevertheless, the amino-terminal segment has never been captured in a receptor-chemokine co-crystal X-ray structure (51,88,91,101,102), indicating that it is an intrinsicially disordered protein (74).

NMR spectroscopy of chemokine CCL5 in complex with a CCR5 amino-terminal synthetic sulfopeptide reveals two sTyr binding regions (103). NMR structures of different chemokines complexed with receptor CRS1 peptides display different relative orientations, creating uncertainty in their interpretation (88,102). Nevertheless, the CCR5-CCL5 overall interaction is highly dynamic, with the flexible anionic segment making different contacts to chemokine residues in different models (102,103).
This dynamic interaction was explored recently with synthetic sulfopeptide binding experiments, which found that any combination of two sTyr residues at positions 10, 14 and 15 (**Fig. 2B**) enables strong binding to chemokine CCL5 (104) (**Fig. 6C**). Together, these results support the hypothesis that the CCR5 flexible sTyr-containing anionic segment can make a variety of dynamic but high-affinity contacts to the chemokine. This contrasts with the relatively fixed contacts observed in the CCR5-gp120 interactions described above (102,104).

Chemokines also bind sulfated GAGs to form chemotactic gradients (80). 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 (87,103,105).

The blood clotting enzyme thrombin binds proteins with hirudin-like sequence motifs, many of which contain a cluster of two or three sTyr residues.

The third sequence group (**Fig. 2C; Fig. S3; Fig. S4**) also has clusters of two or three sTyr residues. However, these are in hirudin-like sequence motifs that bind the serine endoprotease α-thrombin (**Fig. 3C**), which regulates the balance between initiating and terminating blood clotting (thrombosis). Additional examples of sTyr function in thrombin-binding proteins are in the supporting information (3.1-3.5).

Aberrant hemostasis – ranging from profuse bleeding to excessive clotting – underlies or complicates several medical conditions, and many treatments have been developed (106-108). Hirudin, a clot-inhibiting therapeutic used for over one hundred years, is an sTyr-containing **d**irect **t**hrombin **i**nhibitor (DTI) made by *Hirudo* leeches (17,109,110). This 65 residue protein comprises a carboxyl-terminal short flexible anionic segment – the hirudin-like motif (111) (**Fig. S3**) – that binds thrombin, and an amino-terminal domain that occludes the thrombin active site (112,113).

Thrombin contains two cationic surface patches, termed anion-binding exosites, that bind different protein substrates through hirudin-like motifs (112,113) (**Fig. 7**). Thrombin exosite 1 positions both endoprotease substrates and DTIs such as hirudin with respect to the adjacent serine protease active site, whereas exosite 2 tethers thrombin to platelet surfaces via platelet **g**lyco**p**rotein **Ib**α (GpIbα) and to fibrin clots via fibrinogen γ' (114,115). Exosite 2 also binds the DTIs madanin and **t**setse **t**hrombin **i**nhibitor (TTI) (26,116). 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 here (**Fig. 2C and Fig. S3A**), the hirudin-like motif spans seven contiguous aromatic or acidic residues (**Fig. S3B**). Hirudin-like motifs in all four proteins that 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. S3A**) (supporting information 3.1).

Co-crystal X-ray structures of thrombin with exosite 2-bound sulfopeptides (**Fig. 2C**), illustrated here with the DTI madanin (116), reveal conserved contacts to thrombin residues across exosite 2 (12,26,117) (**Fig. 7A**). As with the structure for CCR5-gp120 described above (**Fig. 5B**), both madanin sTyr residues occupy extensive binding interfaces made from distinct groups of exosite 2 residues. Residues sTyr-32 and Asp-33, which occupy the hirudin-like motif conserved positions 1 and 2 (**Fig. 2C**), make both ionic and nonpolar contacts to numerous exosite 2 residues (116) (**Fig. 7B**). These contacts form the core interaction with thrombin (12). Madanin residue sTyr-35 contacts a separate set of exosite 2 residue (**Fig. 7B**). Thrombin contacts to all three residues are conserved in all four available co-crystal X-ray structures (26), indicating that the sTyr cluster-containing hirudin-like motif provides a defined high-affinity binding determinant for exosite 2. Thus, these sTyr contacts overall resemble the CCR5- and antibody 412d-gp120 contacts described above.

Peptide binding analyses with sulfopeptides (26,116) or phosphopeptides (12,117) support and illuminate these structural models, again illustrated here with sulfo-madanin (**Fig. 6D**). 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 (116). Thus, again similar to

chemokine receptor CCR5, the core sTyr residue (madanin sTyr-32; CCR5 sTyr-14) is essential whereas the ancillary sTyr residue (madanin sTyr-35; CCR5 sTyr-10) enhances binding only in the presence of the core sTyr residue (94,116) (**Fig. 6A and Fig. 6D**).

Segments with sTyr clusters are prevalent in bloodstream activities, including leukocyte migration and signaling, blood clotting and complement activation. These activities also are modulated extensively by interactions with sulfated GAGs, anionic oligosaccharides that help control hemostasis through binding thrombin (80,118). Indeed, thrombin exosite 2 was identified initially as the binding site for the highly-sulfated GAG heparin (114). Thus, a cluster of two 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 far beyond merely "heparin-binding." Notably, basic exosite residues that contact GAG substrates (114) mostly are distinct from those that contact sTyr-containing substrates, illustrating the broad binding versatility provided by exosite 2.

Prospects

The examples presented here illustrate diverse contexts for the sTyr post-translational modification. A single sTyr residue can be as effective as a cluster of two or three, in terms of binding affinity enhancement. However, an sTyr cluster potentially confers greater versatility, as demonstrated by the chemokine receptor CCR5 sTyr-containing amino-terminal segment, or greater stringency, as demonstrated by conserved interactions with thrombin exosite 2 to enforce flanking sequence geometries.

It is striking that mimics of sTyr-containing proteins require sTyr residues for function. This is observed with the RaxX sulfopeptide hormone mimic (48,49), anti-gp120 antibodies (98,100), and inhibitors of chemokine signaling (25), thrombin (14,26,116,119), and complement C1s (120). These examples show that its unique blend of hydrophobic and ionic contacts renders the sTyr residue irreplaceable.

Better understanding of TPST substrate specificity is necessary to evaluate residues such as chemoreceptor CCR5 sTyr-3 (**Fig. 2B**), which supposedly has minimal functional significance (92,93,121,122) and therefore usually is not studied (94,104). However, position Tyr-3 is the first to be sulfated in a CCR5 peptide that includes residue Met-1 (123,124). Does sTyr-3 result simply from "bystander" sulfation by TPST enzymes with broad substrate recognition? Or does it have a defined function?

A related topic is the hypothesis that different amounts and types of incompletely-sulfated proteins are produced in vivo (51). This is supported by a recent cell culture study that used antibodies to distinguish differently-sulfated chemokine receptor CCR5 molecules with different ligand-binding properties (125). 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?

Post-translational modifications often work in concert, with different combinations exerting different effects (1). A recent analysis (2) highlights instances of sTyr residues functioning with nearby *N*- or *O*-linked glycans. For example, PSGL-1 is glycosylated in leukocytes, in which the interaction with one binding partner requires both the sTyr residues and the glycan (126), whereas PSGL-1 is not glycosylated in certain T cells, where interaction with a different binding partner requries only the sTyr residues (24).

Does differential tyrosyl sulfation in different cell types or under different conditions similarly help determine binding partner selection?

The technical advances described above mean that it now is possible to engineer and document defined sTyr residues to add novel interaction determinants. One proof-of-principle is eCD4-Ig, a chimeric protein that effectively blocks HIV-1 entry in rhesus macaques (127-129). eCD4-Ig contains an essential sTyr-rich peptide, derived from the sTyr-containing anti-gp120 antibody E51 variable region (**Fig. 2B**), that mimics HIV-1 gp120 interaction with the CCR5 coreceptor.

Specific sTyr residues can be engineered through peptide synthesis (13-15) or through genetic code modification (16-19). 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 (130). 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 (49,130).

In biomedical research, it will be important to examine further the large number of sTyr residues involved in hemostasis, particularly those in the homologous multidomain proteins, coagulation factors VIII (FVIII) and V (FV) (131) (**Fig. S4**) (supporting information 4.1 and 4.2, respectively). In humans, classic hemophilia results from FVIII deficiency. One of the many causative *F8* alleles encodes a missense substitution of Phe for sTyr-1680 in FVIII (131), documenting an important function for sTyr in hemostasis. Indeed, synthetic FVIII used for human therapy contains a full complement

of sTyr residues (132). Nevertheless, relatively little is known about sTyr function in these critical hemostasis proteins (133,134).

Supporting information — This article contains supporting information. References are listed after the main text references.

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.

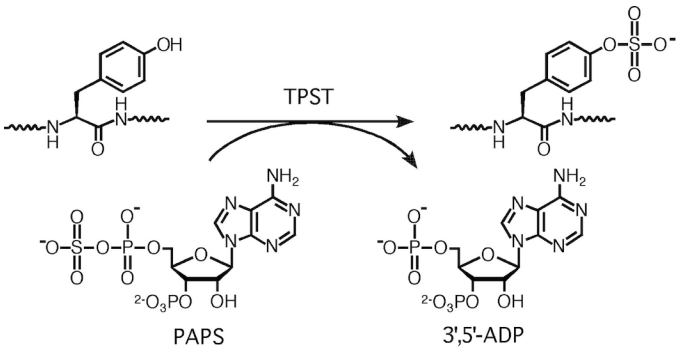
Funding and additional information — Tyrosine sulfation research in the Ronald laboratory is supported by Public Health Service grant GM122968 from the National Institute of General Medical Sciences (awarded to P.C.R.). Molecular graphics and analyses were performed with UCSF Chimera, developed by the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

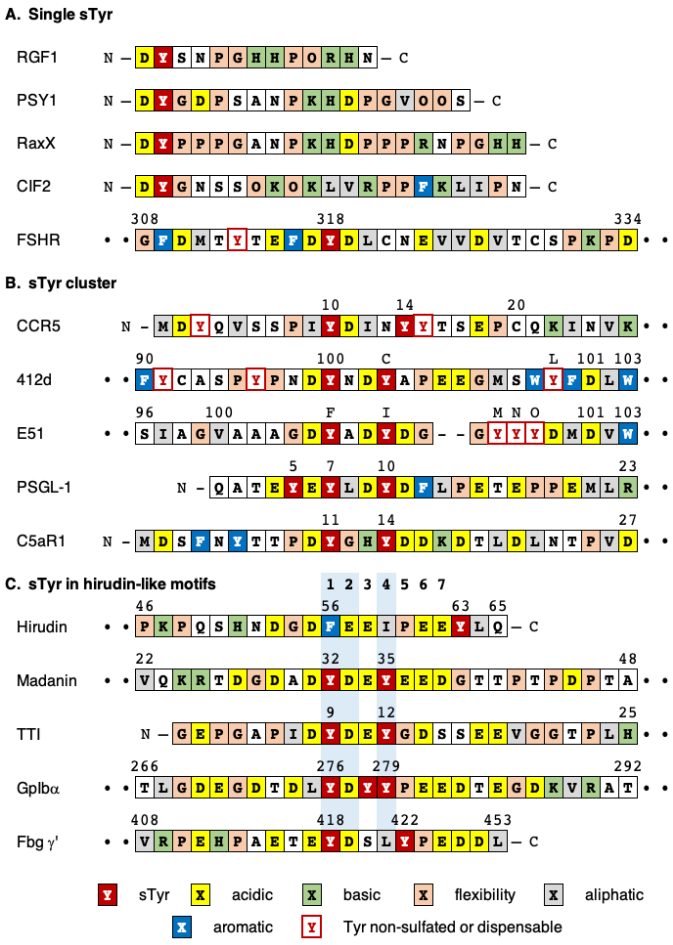
Fig. 1. Tyrosyl-protein sulfotransferase (TPST).



One column figure (3.5″ wide)

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. Reproduced from reference (21).

Fig. 2. Amino acid sequence contexts for representative sTyr residues.



One column figure (3.5" wide)

Sequences are denoted in single-letter code and colored according to sidechain as shown in the key. The amino- and carboxyl-termini are denoted N and C, respectively. Numbers indicate residue position in the sequence. Dashes in the E51 sequence denote spaces introduced for alignment. O denotes hydr**o**xyproline.

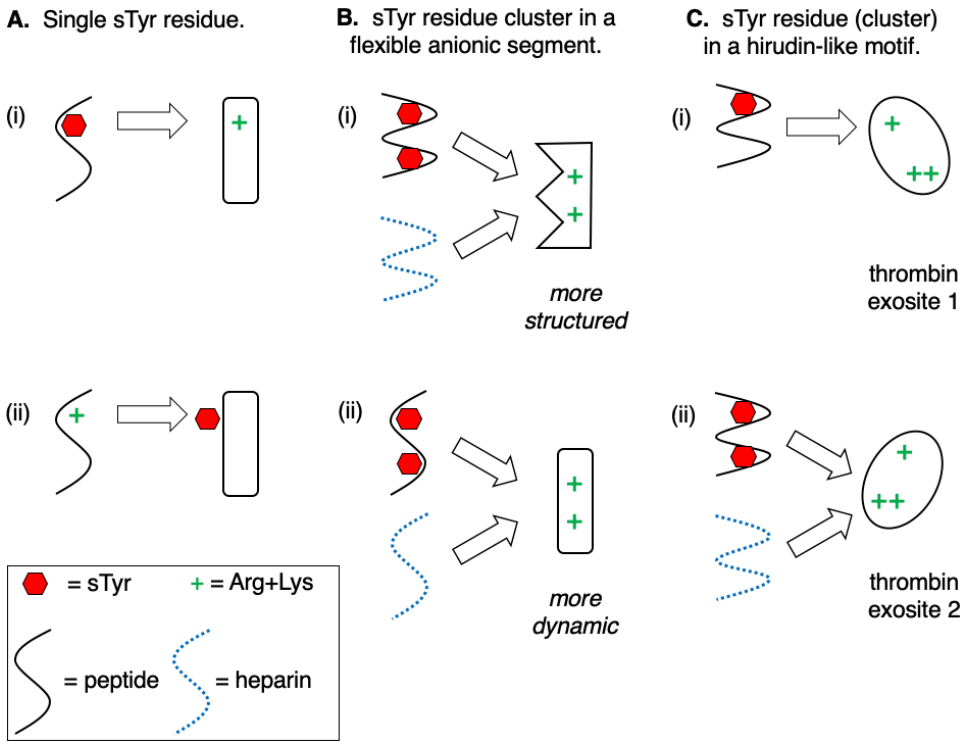
Fig. 2. (cotinued) **Amino acid sequence contexts for representative sTyr residues.**

A. Single sTyr residue contexts are shown for the mature forms of RGF1 (root meristem growth factor), PSY1 (peptide sulfated on tyrosine), RaxX (required for activating XA21 protein X), and CIF2 (Casperian strip integrity factor). The sequence shown for FSHR (follicle-stimulating hormone receptor) corresponds to a portion of the flexible segment distal to the LRR hormone-binding domain (**Fig. 4C**). Additional sequences and accession numbers are in **Fig. S1**.

B. sTyr clusters that bind gp120 include the chemokine receptor CCR5 (CC-type chemokine receptor 5) amino-terminal segment and the heavy chain CD3 regions from the anti-gp120 antibodies 412d and E51. PSGL-1 (P-selectin glycoprotein ligand-1) and C5aR (complement C5a receptor) also have sTyr cluster-containing amino-terminal segments. Additional sequences and accession numbers are in **Fig. S2**.

C. sTyr clusters in hirudin-like motifs that bind thrombin exosite 2 include madanin, TTI (tsetse thrombin inhibitor), GpIbα (platelet glycoprotein Ibα) and Fbg γ' (fibrinogen γ'). The sequence for hirudin, which binds exosite 1, is shown for reference. Hirudin-like motif positions are indicated with numbers and conserved positions (**Fig. S3B**) are highlighted in light blue. Additional sequences and accession numbers are in **Figs. S3 and S5**.

Fig. 3. sTyr sequence contexts represent distinct binding modes. (sketch for JBC artist)



One and one-half column figure (5" wide)

A. A single sTyr residue, as part of a multi-residue interaction determinant, can be on either the ligand (i) or the receptor (ii) depending on the example considered.

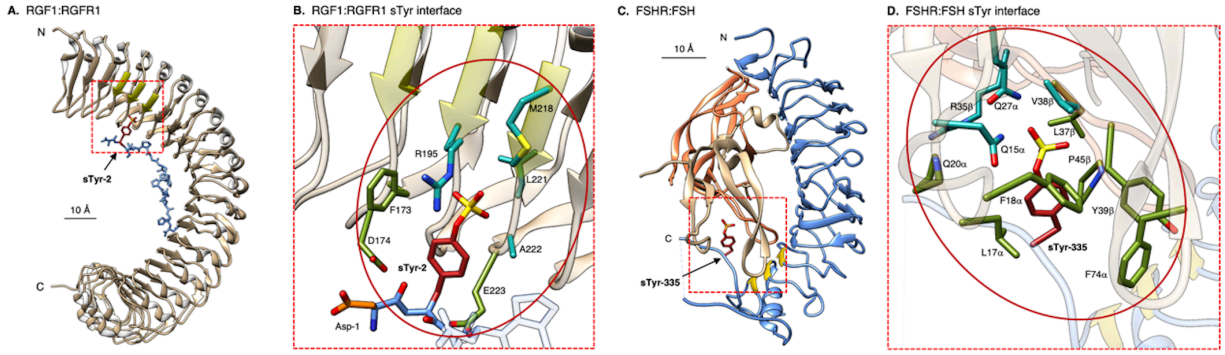
B. An sTyr cluster may be part of a more ordered binding determinant (i) or part of a highly flexible segment (ii). In either case, sulfoprotein binding may compete with binding of sulfated GAGs such as heparan sulfate.

Fig. 3. (cotinued) **sTyr sequence contexts represent distinct binding modes.**

(Sketch for JBC artist)

C. Hirudin-like motif proteins with 0 or 1 sTyr residue bind thrombin exosite 1 (i), whereas hirudin-like motif proteins with 2 or 3 sTyr residues bind thrombin exosite 2 (ii). Exosite 2 also is the heparin-binding site.

Fig. 4. Single sTyr residue interactions.



Two column figure (7.2" wide)

A. The co-crystal X-ray model of sulfated RGF (root meristem growth factor; blue) bound to the RGF receptor ectodomain (tan) is depicted at 2.6 Å resolution **(67)** (PDB code 5HYX). Backbone atoms are shown as ribbons with helices and arrows indicating α-helices and β-strands, respectively. The RGFR ectodomain is built from 23 LRRs (leucine-rich repeats) stacked one upon the next. β-strands in LRRs 3-5 are colored yellow-green for reference to panel B.

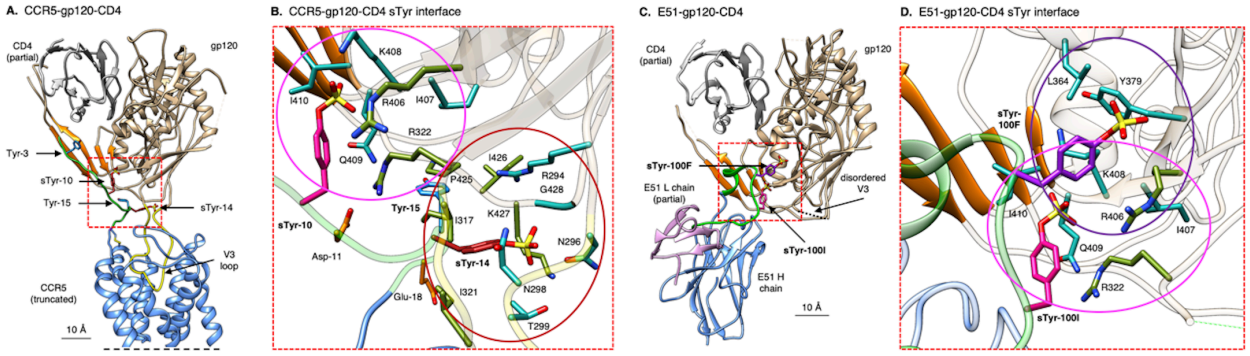
B. The sTyr binding interface. RGF residue sTyr-2 (dark red) includes the sulfate (S, yellow; O, red). RGFR residues R195, A196, M218, L221 and A222 (light green) are predicted to interact with the sulfate, whereas RGFR residues F73, D74 and E223 (dark green) are predicted to interact with sTyr C atoms. Together these form the sTyr binding interface encircled in dark red.

Fig. 4. (continued) **Single sTyr residue interactions.**

C. The co-crystal X-ray model of FSH (follicle-stimulating hormone; α subunit tan and β subunit coral) bound to the sulfated FSH receptor ectodomain (blue) is depicted at 2.5 Å resolution (70) (PDB code 4AY9). The FSHR ectodomain is built from 12 LRRs stacked one upon the next. β-strands in LRRs 10-12 are colored yellow for reference to panel D.

D. The sTyr binding interface. FSHR residue sTyr-335 (dark red), including the sulfate (S, yellow; O, red). Residues from both FSH subunits make contacts to sTyr: residues Q15α, Q27α and V38β (light green) are predicted to interact with the sulfate, whereas FSH residues L17α, F18α, Q20α, L37β, Y39β, P45β and F74α (dark green) are predicted to interact with sTyr C atoms. Together these form the sTyr binding interface encircled in dark red.

Fig. 5. sTyr cluster interactions with HIV-1 gp120.



Two column figure (7.2" wide)

A. The cryo-EM model of HIV-1 gp120 (glycoprotein 120; tan) bound to host CCR5 (C-C-type chemokine receptor 5; blue) and host CD4 (cluster of differentiation 4 receptor; gray) is depicted at 3.9 Å resolution (91) (PDB code 6MEO). Backbone atoms are depicted as ribbons with helices and arrows indicating α-helices and β-strands, respectively. CCR5 residues 1-20 are green, and residues in the gp120 bridging sheet and V3 loop are orange and yellow, respectively. Only one CD4 domain is shown. Only the top of the CCR5 transmembrane helix bundle is shown.

B. The sTyr binding interface. CCR5 residues Tyr-14 (dark red) and Tyr-10 (magenta) are modeled to include the sulfate (S, yellow; O, red). gp120 residues R294, N296, N298, T299 and G428 (light green) are predicted to interact with the sTyr-14 sulfate, whereas gp120 residues I317, I321, P425, I426 and K427 (dark green) are predicted to interact with sTyr-14 C atoms. Together these form the sTyr-14 binding interface

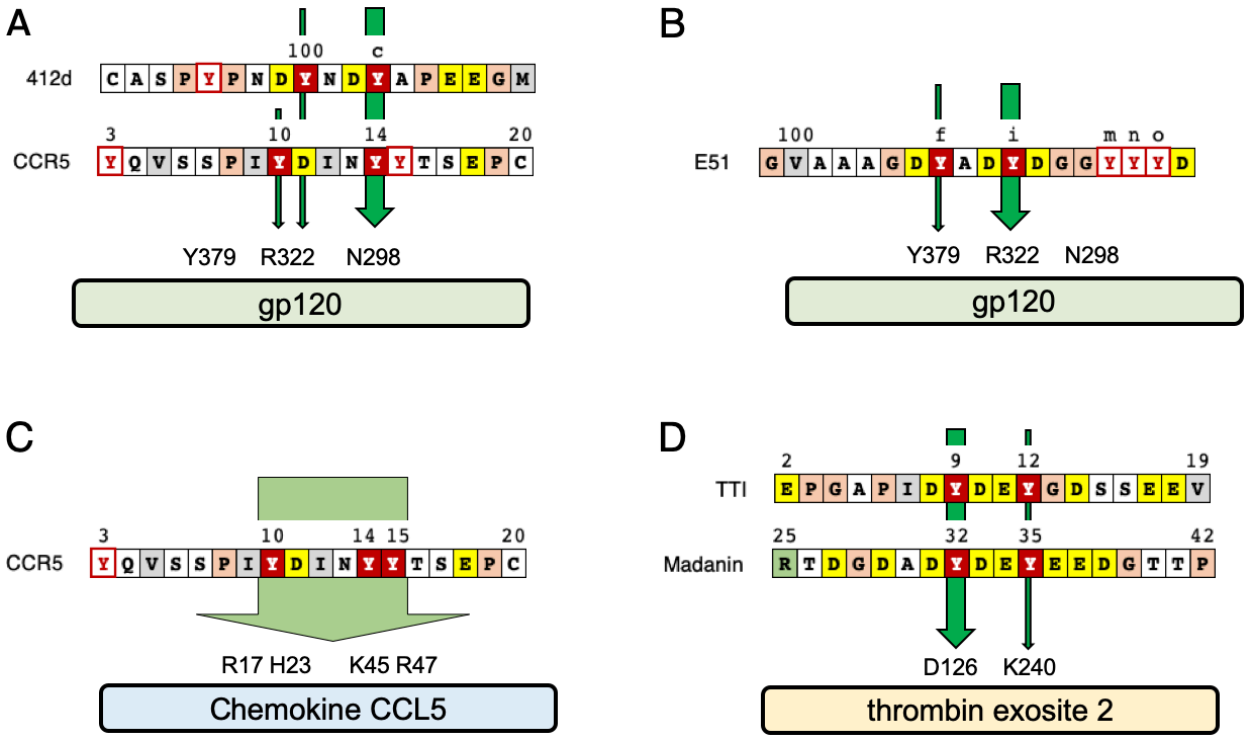
Fig. 5.(continued) **sTyr cluster interactions with HIV-1 gp120.**

encircled in dark red. Separately, gp120 residues K408, Q409 and I410 (light green) are predicted to interact with the sTyr-10 sulfate, whereas gp120 residues R322 and R406 (dark green) are predicted to interact with sTyr-10 C atoms. Together these form the sTyr-10 binding interface encircled in magenta. Virtually all of these gp120 residues predicted to contact CCR5 were identified previously in directed screens for mutants with diminished CCR5 binding (90,135). (gp120 residues are numbered according to the HIV-1 isolate YU2; accession number P35961.)

C. The cryo-EM model of HIV-1 gp120 (tan) bound to the sulfated antibody E51 Fab domain (heavy chain, blue; light chain, light purple) and host CD4 (gray) is depicted at 3.3 Å resolution (100) (PDB code 6U0L). Backbone atoms are depicted as ribbons with helices and arrows indicating α-helices and β-strands, respectively. gp120 residues in the bridging sheet are orange. E51 residues in the HCD3 (heavy chain complementary determining region 3) are green. Only one CD4 domain is shown. Only a portion of the E51 light chain is shown.

D. The sTyr binding interface. E51 residues sTyr-100I (magenta) and sTyr-100F (purple) include the sulfate (S, yellow; O, red). gp120 residues R414, K416, Q417 and I418 (light green) are predicted to interact with the sTyr-100I sulfate, whereas gp120 residue R326 (dark green) is predicted to interact with an sTyr-100I C atom. Together these form the sTyr-10-binding interface encircled in magenta. Separately, gp120 residues L364 and Y379 (light green) are predicted to interact with the sTyr-100F sulfate, whereas gp120 residues R414 and K416 (dark green), shared with the sTyr-100I interface, are predicted to interact with sTyr-100F C atoms.

Fig. 6. sTyr binding interfaces display differential affinities. (sketch for JBC artist**).**



Sketch for JBC artist

A. Sulfopeptides derived from chemokine receptor CCR5 require residue sTyr-14 interaction with the interface defined by gp120 residue Asn-298 for effective binding (**Fig. 5B**). Residue sTyr-10 interaction with the interface defined by gp120 residue Arg-322 stimulates binding only in the presence of sTyr-14 (91,94). Antibody 412d residues sTyr-100C and sTyr-100 mimic these interactions (98,99).

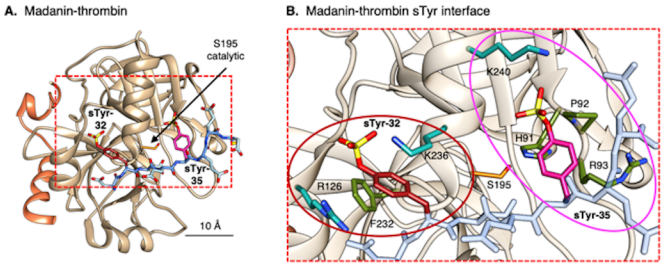
Fig. 6. (continued) **sTyr binding interfaces display differential affinities.** (sketch for JBC artist**).**

B. Sulfopeptides derived from antibody E51 require residue sTyr-100I interaction with the interface defined by gp120 residue Arg-322 for effective binding. Residue sTyr-100F interaction with the distinct interface defined by gp120 residue Tyr-379 stimulates binding only in the presence of sTyr-100I (19,100). Thus, antibody E51 sTyr residues do not occupy the Asn-298 interface critical for CCR5 and antibody 412d binding.

C. Sulfopeptides derived from chemokine receptor CCR5 require any two of the residues sTyr-10, sTyr-14 and sTyr-15 for effective chemokine CCL5 binding (82). These sTyr residues interact dynamically with interfaces defined by CCL5 residues Arg-17 plus His-23 and by Lys-45 plus Arg-47 (102,103).

D. Sulfopeptides derived from the DTI madanin require residue sTyr-32 interaction with the interface defined by thrombin exosite 2 residue Asp-126 for effective binding. Residue sTyr-35 interaction with the interface defined by exosite 2 residue Lys-240 stimulates binding only in the presence of sTyr-32 (116). Sulfopeptides derived from TTI display similar binding properties (26).

Fig. 7. Madanin sTyr cluster interaction with thrombin exosite 2.



One column figure (3.5" wide)

A. The co-crystal X-ray model of sulfated madanin (blue) bound to exosite 2 on α-thrombin (A-chain, coral; B-chain, tan) is depicted at 1.6 Å resolution (116) (PDB code 5L6N). Backbone atoms are shown as ribbons with helices and arrows indicating α-helices and β-strands, respectively. Active site residue Ser-195 (orange) is shown for reference.

B. The sTyr binding interface. Madanin residues sTyr-32 (dark red) and sTyr-35 (magenta) include the sulfate (S, yellow; O, red). Thrombin exosite 2 residues R126 and K236 (light green) are predicted to interact with the sTyr-32 sulfate, whereas exosite 2 residue F232 (dark green) is predicted to interact with sTyr-32 C atoms. Together these form the sTyr-32 binding interface encircled in dark red. Separately, exosite 2 residues K240 (light green) is predicted to interact with the sTyr-35 sulfate, whereas exosite 2 residues H91, P92 and H93 and R406 (dark green) are predicted to interact with sTyr-35 C atoms. Together these form the sTyr-35 binding interface encircled in magenta.

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

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

7. 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

8. 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

9. 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

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

11. 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

12. 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

13. 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

14. 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

15. 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

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

17. 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

18. 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

19. 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

20. 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

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

22. 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

23. 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

24. 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

25. 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

26. 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

27. 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

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

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

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

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

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

33. 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

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

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

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

37. 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

38. 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

39. 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

40. 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

41. 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

42. 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

43. 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

44. 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

45. 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

46. 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

47. 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

48. 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

49. 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

50. 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

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

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

53. 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

54. 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

55. 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**

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

57. 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

58. 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

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

60. 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

61. 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

62. 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

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

64. Zhang, H., Han, Z., Song, W., and Chai, J. (2016) Structural insight into recognition of plant peptide hormones by receptors. *Mol Plant* **9**, 1454-1463

65. Hohmann, U., Lau, K., and Hothorn, M. (2017) The structural basis of ligand perception and signal activation by receptor kinases. *Annu Rev Plant Biol* **68**, 109-137

66. Gou, X., and Li, J. (2020) Paired receptor and coreceptor kinases perceive extracellular signals to control plant development. *Plant Physiol* **182**, 1667-1681

67. 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

68. 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

69. 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

70. 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

71. 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

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

73. 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

74. 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

75. 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

76. 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

77. 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

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

79. Thomas, M. A., Kleist, A. B., and Volkman, B. F. (2018) Decoding the chemotactic signal. *J Leukoc Biol* **104**, 359-374

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

81. 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

82. Kleist, A. B., Getschman, A. E., Ziarek, J. J., Nevins, A. M., Gauthier, P. A., Chevigne, A., Szpakowska, M., and Volkman, B. F. (2016) New paradigms in chemokine receptor signal transduction: Moving beyond the two-site model. *Biochem Pharmacol* **114**, 53-68

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

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

85. 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

86. 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

87. 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

88. 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

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

90. 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

91. 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

92. 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

93. 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

94. 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

95. 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

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

97. 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

98. 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

99. 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

100. 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

101. 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

102. 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

103. 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

104. 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

105. 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

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

107. 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

108. 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

109. 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

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

111. 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

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

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

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

115. 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

116. 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

117. 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

118. 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

119. 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

120. 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

121. 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

122. 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

123. 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

124. 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

125. 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

126. 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

127. 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

128. 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

129. 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

130. 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

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

132. 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

133. 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

134. 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

135. Cormier, E. G., Tran, D. N., Yukhayeva, L., Olson, W. C., and Dragic, T. (2001) Mapping the determinants of the CCR5 amino-terminal sulfopeptide interaction with soluble human immunodeficiency virus type 1 gp120-CD4 complexes. *J Virol* **75**, 5541-5549

References for supporting information

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

137. Vishnuvardhan, D., and Beinfeld, M. C. (2000) Role of tyrosine sulfation and serine phosphorylation in the processing of procholecystokinin to amidated cholecystokinin and its secretion in transfected AtT-20 cells. *Biochemistry* **39**, 13825-13830

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

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

140. 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

141. Matsuzaki, Y., Ogawa-Ohnishi, M., Mori, A., and Matsubayashi, Y. (2010) Secreted peptide signals required for maintenance of root stem cell niche in *Arabidopsis*. *Science* **329**, 1065-1067

142. Nakayama, T., Shinohara, H., Tanaka, M., Baba, K., Ogawa-Ohnishi, M., and Matsubayashi, Y. (2017) A peptide hormone required for Casparian strip diffusion barrier formation in Arabidopsis roots. *Science* **355**, 284-286

143. Amano, Y., Tsubouchi, H., Shinohara, H., Ogawa, M., and Matsubayashi, Y. (2007) Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in *Arabidopsis*. *Proc Natl Acad Sci U S A* **104**, 18333-18338

144. Okamoto, S., Suzuki, T., Kawaguchi, M., Higashiyama, T., and Matsubayashi, Y. (2015) A comprehensive strategy for identifying long-distance mobile peptides in xylem sap. *Plant J* **84**, 611-620

145. Patel, N., Mohd-Radzman, N. A., Corcilius, L., Crossett, B., Connolly, A., Cordwell, S. J., Ivanovici, A., Taylor, K., Williams, J., Binos, S., Mariani, M., Payne, R. J., and Djordjevic, M. A. (2018) Diverse peptide hormones affecting root growth identified in the *Medicago truncatula* secreted peptidome. *Mol Cell Proteomics* **17**, 160-174

146. Bundgaard, J. R., Vuust, J., and Rehfeld, J. F. (1995) Tyrosine O-sulfation promotes proteolytic processing of progastrin. *EMBO J* **14**, 3073-3079

147. 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

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

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

150. 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

151. 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

152. 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

153. Sanchez, J., Z, E. H., Lane, J. R., Liu, X., Bridgford, J. L., Payne, R. J., Canals, M., and Stone, M. J. (2019) Evaluation and extension of the two-site, two-step model for binding and activation of the chemokine receptor CCR1. *J Biol Chem* **294**, 3464-3475

154. Preobrazhensky, A. A., Dragan, S., Kawano, T., Gavrilin, M. A., Gulina, I. V., Chakravarty, L., and Kolattukudy, P. E. (2000) Monocyte chemotactic protein-1 receptor CCR2B is a glycoprotein that has tyrosine sulfation in a conserved extracellular N-terminal region. *J Immunol* **165**, 5295-5303

155. Zhu, J. Z., Millard, C. J., Ludeman, J. P., Simpson, L. S., Clayton, D. J., Payne, R. J., Widlanski, T. S., and Stone, M. J. (2011) Tyrosine sulfation influences the chemokine binding selectivity of peptides derived from chemokine receptor CCR3. *Biochemistry* **50**, 1524-1534

156. Gutierrez, J., Kremer, L., Zaballos, A., Goya, I., Martinez, A. C., and Marquez, G. (2004) Analysis of post-translational CCR8 modifications and their influence on receptor activity. *J Biol Chem* **279**, 14726-14733

157. Moussouras, N. A., Getschman, A. E., Lackner, E. R., Veldkamp, C. T., Dwinell, M. B., and Volkman, B. F. (2017) Differences in sulfotyrosine binding amongst CXCR1 and CXCR2 chemokine ligands. *Int J Mol Sci* **18,** 1894

158. Colvin, R. A., Campanella, G. S., Manice, L. A., and Luster, A. D. (2006) CXCR3 requires tyrosine sulfation for ligand binding and a second extracellular loop arginine residue for ligand-induced chemotaxis. *Mol Cell Biol* **26**, 5838-5849

159. 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

160. Fong, A. M., Alam, S. M., Imai, T., Haribabu, B., and Patel, D. D. (2002) CX3CR1 tyrosine sulfation enhances fractalkine-induced cell adhesion. *J Biol Chem* **277**, 19418-19423

161. Choe, H., Moore, M. J., Owens, C. M., Wright, P. L., Vasilieva, N., Li, W., Singh, A. P., Shakri, R., Chitnis, C. E., and Farzan, M. (2005) Sulphated tyrosines mediate association of chemokines and Plasmodium vivax Duffy binding protein with the Duffy antigen/receptor for chemokines (DARC). *Mol Microbiol* **55**, 1413-1422

162. Hewit, K. D., Fraser, A., Nibbs, R. J., and Graham, G. J. (2014) The N-terminal region of the atypical chemokine receptor ACKR2 is a key determinant of ligand binding. *J Biol Chem* **289**, 12330-12342

163. Pejchal, R., Walker, L. M., Stanfield, R. L., Phogat, S. K., Koff, W. C., Poignard, P., Burton, D. R., and Wilson, I. A. (2010) Structure and function of broadly reactive antibody PG16 reveal an H3 subdomain that mediates potent neutralization of HIV-1. *Proc Natl Acad Sci U S A* **107**, 11483-11488

164. McLellan, J. S., Pancera, M., Carrico, C., Gorman, J., Julien, J. P., Khayat, R., Louder, R., Pejchal, R., Sastry, M., Dai, K., O'Dell, S., Patel, N., Shahzad-ul-Hussan, S., Yang, Y., Zhang, B., Zhou, T., Zhu, J., Boyington, J. C., Chuang, G. Y., Diwanji, D., Georgiev, I., Kwon, Y. D., Lee, D., Louder, M. K., Moquin, S., Schmidt, S. D., Yang, Z. Y., Bonsignori, M., Crump, J. A., Kapiga, S. H., Sam, N. E., Haynes, B. F., Burton, D. R., Koff, W. C., Walker, L. M., Phogat, S., Wyatt, R., Orwenyo, J., Wang, L. X., Arthos, J., Bewley, C. A., Mascola, J. R., Nabel, G. J., Schief, W. R., Ward, A. B., Wilson, I. A., and Kwong, P. D. (2011) Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. *Nature* **480**, 336-343

165. Changela, A., Wu, X., Yang, Y., Zhang, B., Zhu, J., Nardone, G. A., O'Dell, S., Pancera, M., Gorny, M. K., Phogat, S., Robinson, J. E., Stamatatos, L., Zolla-Pazner, S., Mascola, J. R., and Kwong, P. D. (2011) Crystal structure of human antibody 2909 reveals conserved features of quaternary structure-specific antibodies that potently neutralize HIV-1. *J Virol* **85**, 2524-2535

166. Chiang, J. J., Gardner, M. R., Quinlan, B. D., Dorfman, T., Choe, H., and Farzan, M. (2012) Enhanced recognition and neutralization of HIV-1 by antibody-derived CCR5-mimetic peptide variants. *J Virol* **86**, 12417-12421

167. Fetzer, I., Gardner, M. R., Davis-Gardner, M. E., Prasad, N. R., Alfant, B., Weber, J. A., and Farzan, M. (2018) eCD4-Ig variants that more potently neutralize HIV-1. *J Virol* **92,** e02011-17

168. Rydel, T. J., Tulinsky, A., Bode, W., and Huber, R. (1991) Refined structure of the hirudin-thrombin complex. *J Mol Biol* **221**, 583-601

169. Corral-Rodriguez, M. A., Macedo-Ribeiro, S., Pereira, P. J., and Fuentes-Prior, P. (2010) Leech-derived thrombin inhibitors: from structures to mechanisms to clinical applications. *J Med Chem* **53**, 3847-3861

170. Gandhi, P. S., Chen, Z., Appelbaum, E., Zapata, F., and Di Cera, E. (2011) Structural basis of thrombin-protease-activated receptor interactions. *IUBMB Life* **63**, 375-382

171. Figueiredo, A. C., de Sanctis, D., Gutierrez-Gallego, R., Cereija, T. B., Macedo-Ribeiro, S., Fuentes-Prior, P., and Pereira, P. J. (2012) Unique thrombin inhibition mechanism by anophelin, an anticoagulant from the malaria vector. *Proc Natl Acad Sci U S A* **109**, E3649-3658

172. Krishnaswamy, S. (2013) The transition of prothrombin to thrombin. *J Thromb Haemost* **11 Suppl 1**, 265-276

173. Arni, R. K., Padmanabhan, K., Padmanabhan, K. P., Wu, T. P., and Tulinsky, A. (1993) Structures of the noncovalent complexes of human and bovine prothrombin fragment 2 with human PPACK-thrombin. *Biochemistry* **32**, 4727-4737

174. Richardson, J. L., Kroger, B., Hoeffken, W., Sadler, J. E., Pereira, P., Huber, R., Bode, W., and Fuentes-Prior, P. (2000) Crystal structure of the human alpha-thrombin-haemadin complex: an exosite II-binding inhibitor. *EMBO J* **19**, 5650-5660

175. 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

176. 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

177. 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

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

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

180. Camire, R. M., and Bos, M. H. (2009) The molecular basis of factor V and VIII procofactor activation. *J Thromb Haemost* **7**, 1951-1961

181. Leyte, A., van Schijndel, H. B., Niehrs, C., Huttner, W. B., Verbeet, M. P., Mertens, K., and van Mourik, J. A. (1991) Sulfation of Tyr1680 of human blood coagulation factor VIII is essential for the interaction of factor VIII with von Willebrand factor. *J Biol Chem* **266**, 740-746

182. Michnick, D. A., Pittman, D. D., Wise, R. J., and Kaufman, R. J. (1994) Identification of individual tyrosine sulfation sites within factor VIII required for optimal activity and efficient thrombin cleavage. *J Biol Chem* **269**, 20095-20102

183. Dagil, L., Troelsen, K. S., Bolt, G., Thim, L., Wu, B., Zhao, X., Tuddenham, E. G. D., Nielsen, T. E., Tanner, D. A., Faber, J. H., Breinholt, J., Rasmussen, J. E., and Hansen, D. F. (2019) Interaction Between the a3 Region of Factor VIII and the TIL'E' Domains of the von Willebrand Factor. *Biophys J* **117**, 479-489

184. Newell, J. L., and Fay, P. J. (2008) Acidic residues C-terminal to the A2 domain facilitate thrombin-catalyzed activation of factor VIII. *Biochemistry* **47**, 8786-8795

185. Pittman, D. D., Tomkinson, K. N., Michnick, D., Selighsohn, U., and Kaufman, R. J. (1994) Posttranslational sulfation of factor V is required for efficient thrombin cleavage and activation and for full procoagulant activity. *Biochemistry* **33**, 6952-6959

186. Hirbawi, J., Vaughn, J. L., Bukys, M. A., Vos, H. L., and Kalafatis, M. (2010) Contribution of amino acid region 659-663 of Factor Va heavy chain to the activity of factor Xa within prothrombinase. *Biochemistry* **49**, 8520-8534

187. Corral-Rodriguez, M. A., Bock, P. E., Hernandez-Carvajal, E., Gutierrez-Gallego, R., and Fuentes-Prior, P. (2011) Structural basis of thrombin-mediated factor V activation: the Glu666-Glu672 sequence is critical for processing at the heavy chain-B domain junction. *Blood* **117**, 7164-7173

188. Shim, J. Y., Lee, C. J., Wu, S., and Pedersen, L. G. (2015) A model for the unique role of factor Va A2 domain extension in the human ternary thrombin-generating complex. *Biophys Chem* **199**, 46-50

Supporting Information

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

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

[1] Department of Microbiology & Molecular Genetics
[2] Department of Plant Pathology and Genome Center
University of California, Davis

---

1. A single sTyr residue increases binding affinity for peptide hormone-receptor pairs

1.1 An sTyr residue determines receptor selectivity for an animal sulfopeptide hormone .....S-X

1.2 An sTyr residue in the receptor for complement C3a is essential for binding.. .....S-X

Fig. S1 Single sTyr sequence contexts..... S-13

2. A cluster of two or three sTyr residues contributes to combinatorial receptor-peptide interactions.

2.1 PSGL-1 illustrates sTyr cluster functional versatility..... S-

2.2 The complement C5a receptor sTyr residue pair is in a five-residue β-strand when bound to a bacterial inhibitor ..... S-

Fig. S2. sTyr cluster sequence contexts S-15

3. The blood clotting enzyme thrombin binds proteins with hirudin-like motifs, many of which contain a cluster of two or three sTyr residues.

3.1 Hirudin is a relatively rare example of an sTyr-containing protein that binds thrombin exosite 1 .....

3.2 Anophelin binds exosite 1 in reverse.....

3.3 Exosite 1 contacts are not universally conserved .....

3.4 Haemidin-binding exosite is uncertain.....

Fig. S3a. Hiruden-like motif sequences .....S-X
Fig. S3b. Hiruden-like motif sequences (continued) .....S-X
Fig. S3c. Hiruden-like motif sequence logo .....S-X

4. The blood clotting factors FVIII and FV contain six and seven sTyr residues, respectively.

4.1 sTyr residues in coagulation factor FVIII

4.2 sTyr residues in coagulation factor FV

Fig. S6a. FVIII and FV domains.....S-X
Fig. S6b. FVIII and FV sTyr sequence contexts .....S-X

1. A single sTyr residue increases binding affinity for peptide hormone-receptor pairs

Fig. S1 shows sequences with a single sTyr residue.

1.1 An sTyr residue determines receptor selectivity for an animal sulfopeptide hormone

The secreted peptide hormones **c**hole**c**ysto**k**inin (CCK) and gastrin are involved in many processes including digestion, neurotransmission and endocrine functions (57,136). Proteolytic maturation generates different length bioactive peptides with a common carboxyl-terminal Phe-amide (the shortest of these is depicted in **Fig. 2**). In the cases examined, the CCK sTyr residue mostly is sulfated (136), whereas the gastrin sTyr residue is heterogeneously sulfated, due perhaps to a low-affinity sequence environment for TPST recognition (38). Separately, the CCK carboxyl-terminal peptide, removed during Phe-amide biogenesis, contains two sTyr (**Fig. S1**) whose functions are not known (137).

The homologous G protein-coupled receptors CCK1R and CCK2R initiate signaling in response to CCK and/or gastrin. 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 (56). Thus, ligand sulfation state is a critical specificity determinant for CCK1R-CCK interaction.

No X-ray structure is available for either CCK1R or CCK2R. Nevertheless, various studies indicate that the peptides interact across a broad region of the receptor external face (57), and that a conserved Arg residue in a CCK1R extracellular loop interacts specifically with the CCK sTyr (138).

1.2 An sTyr residue in the receptor for complement C3a is essential for binding

A separate example of a single sTyr residue is the receptor for complement component C3a. This 77 residue peptide, generated through the complement cascade, acts through its receptor to stimulate many aspects of the inflammatory response, sometimes leading to anaphylaxis (139). The G protein-coupled C3a receptor has an essential sTyr residue within an unusually large extracellular loop (58) (**Fig. S1**). Although this sTyr residue is essential for binding C3a, an atomic-level view of receptor binding and activation is not available (140).

As with the glycopeptide receptors, other Tyr residues within the C3a receptor also are sulfated in vivo. However, in these cases missense changes from (s)Tyr to Phe do not affect receptor function, indicating that sulfates at these positions are not important for receptor function (58,71). Rather, these may be examples of non-specific sulfation resulting from broad substrate recognition by TPST.

Legend for Fig. S1 (single sTyr sequence contexts, following page).

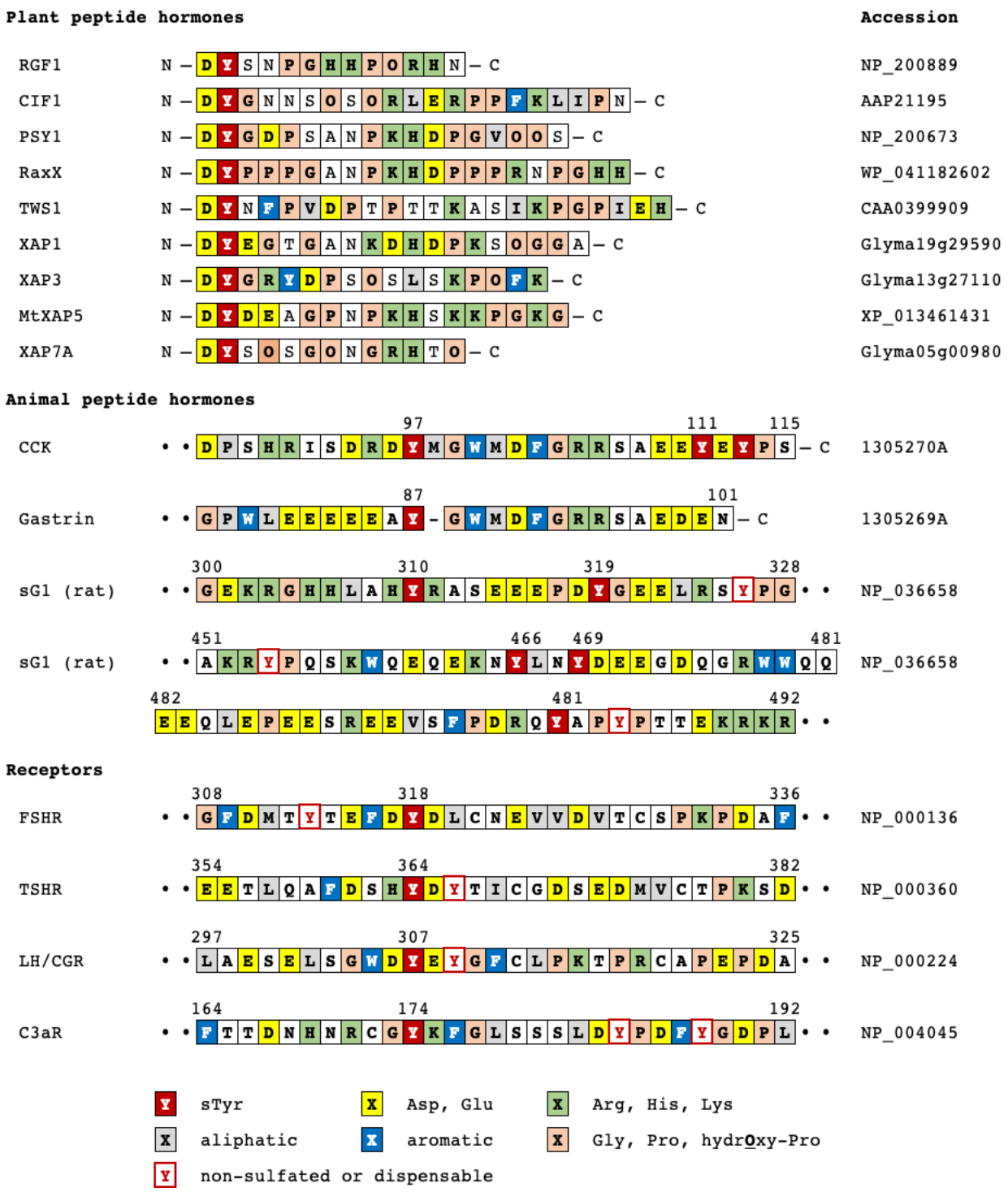
Sulfopeptide hormone sequences are shown as mature forms for plants and as precursor forms for animals. The amino- and carboxyl-termini are denoted N and C, respectively, and residues are colored according to sidechain. The sequences shown for FSHR, TSHR and LH/CGR corresponds to a portion of the flexible segment distal to the LRR hormone-binding domain. The sequence shown for C3aR corresponds to a portion of the large extracellular loop.

Notes for Fig. S1 (single sTyr sequence contexts, following page).

Symbol [a]	Function	Ref. for Sulfation
RGF	root meristem growth factor	(141)
CIF	Casparian strip-inducing factor	(142)
PSY	peptide sulfated on tyrosine	(143)
RaxX	PSY mimic (*Xanthomonas*)	(48)
TWS	twisted seed phenotype	(23)
XAP	xylem-associated peptite (soybean)	(144)
MtXAP	xylem-associated peptide (alfalfa)	(145)
CCK	cholecystokinin	(137)
gastrin	gastrin	(146)
sG	secretogranin (rat)	(147)
FSHR	follicle-stimulating hormone receptor	(71)
TSHR	thyroid-stimulating hormone receptor	(71)
LH/CGR	leutenizing hormone/chorionic gonadotropin receptor	(71)
C3aR	complement C3a receptor	(58)

[a] Plant sequences are from *Arabidopsis thaliana* except as noted. Animal sequences are from *Homo sapiens* except as noted.

Fig. S1. Amino acid sequence contexts for proteins that contain a single sTyr residue.



See p. S-4 for legend and notes.

2. A cluster of two or three sTyr residues contributes to combinatorial receptor-peptide interactions.

Fig. S2 shows sequences with sTyr residue clusters.

2.1 P-selectin glycoprotein ligand-1 illustrates sTyr cluster functional versatility

Leukocytes move toward sites of infection or injury through rolling cell adhesion, mediated in part by **P**-**s**electin **g**lycoprotein **l**igand-**1** (PSGL-1), a homodimeric mucin-like glycoprotein on lymphoid and myeloid cell surfaces. PSGL-1 engages with cell-surface adhesion molecules (selectins) through a mechanosensitive catch-bond that enables rapid engagement and release under force in the bloodstream flow (148,149). 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 (150). 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 (150).

Structural analyses reveal distinct binding modes for the PSGL-1 sTyr-containing anionic segment. The first binding mode is from an X-ray co-crystal structure of a 19 residue PSGL-1 sulfoglycopeptide bound to the P-selectin amino-terminal domain (75). In this structure, residue sTyr-7 makes several electrostatic and hydrophobic contacts to residues in P-selectin, again illustrating the multivalent binding potential for sTyr. Likewise, residue sTyr-10 makes multiple hydrophobic contacts that orient the sulfate for hydrogen bonding to a critical Arg residue (2,75).

The second binding mode is based on a computational model wherein all three PGSL-1 sTyr residues plus the adjacent residue Glu-15 make ionic interactions with four His residues along one edge of VISTA. In this model, sTyr interaction depends upon His protonation, as would occur in acidic tumor microenvironments (pH ≤ 6) (24). Together, interactions with P-selectin and VISTA illustrate how sTyr cluster structural flexibility enables impressive functional versatility by the PSGL-1 amino-terminal segment.

All three sTyr are necessary for full P-selectin interactions, but any one sTyr suffices for partial function in a variety of assays relevant to cell rolling adhesion (151). Indeed, 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 (75). 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 (152), congruent with the extensive interactions made by this residue (75). Sulfation is critical for VISTA-PSGL-1 binding, but contributions of individual sTyr residues are not known (24).

The PSGL-1 acidic amino-terminus contains an an *O*-glycan at residue Thr-16 in addition to the sTyr cluster (**Fig. S2**), illustrating a sTyr-glycan multi-motif (2). The glycan is required for binding P-selectin (2,75) but not VISTA, and PSGL-1 in exhausted T cells is synthesized without the O-glycan (150). Thus, the PSGL-1 sTyr cluster functions in an example of multiple post-translational modifications acting in different combinations for different purposes (1,2).

2.2 The complement C5a receptor sTyr residue pair is in a five-residue β-strand when bound to a bacterial inhibitor

Complement protein C5a (74 residues), generated through the complement cascade, stimulates inflammatory mediator release and also is a potent chemoattractant (139). Similar to chemokine receptors (section 3b), C5a binds the G protein-coupled C5a receptor amino-terminal flexible acidic segment, which contains an sTyr residue pair (76) (**Fig. S2**). **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 (140). The C5a receptor sTyr residue pair increases CHIPS binding affinity for peptides representing the C5a receptor amino-terminal segment by approximately 1,000 fold (77).

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 a number of hydrophobic and electrostatic contacts to residues in CHIPS (77). This conformation again is distinct from those adopted by PSGL-1 and CCR5.

Fig. S2. Amino acid sequence contexts for proteins that contain an sTyr residue cluster.

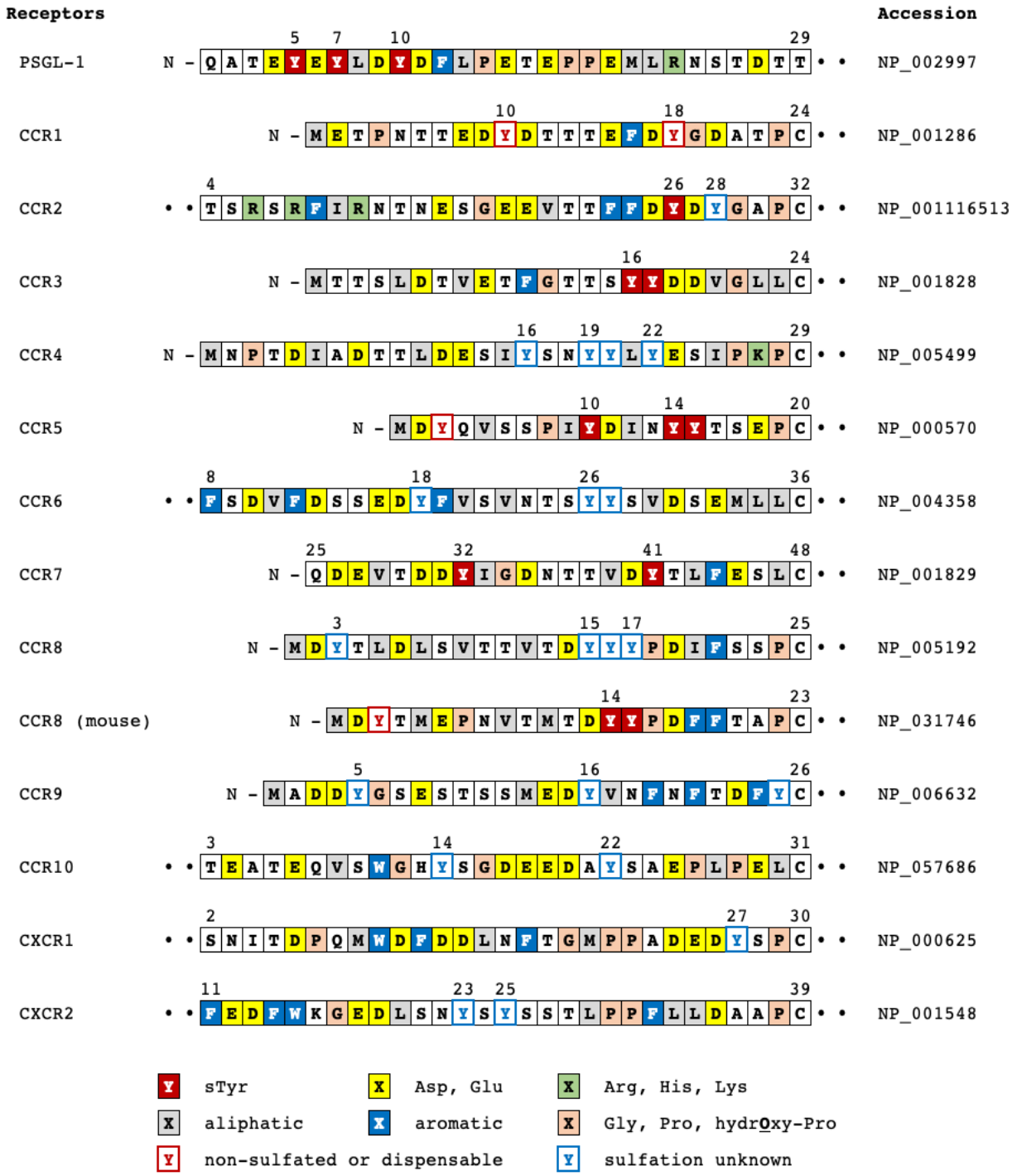


Fig. S2 (continued). Amino acid sequence contexts for proteins that contain an sTyr residue cluster.

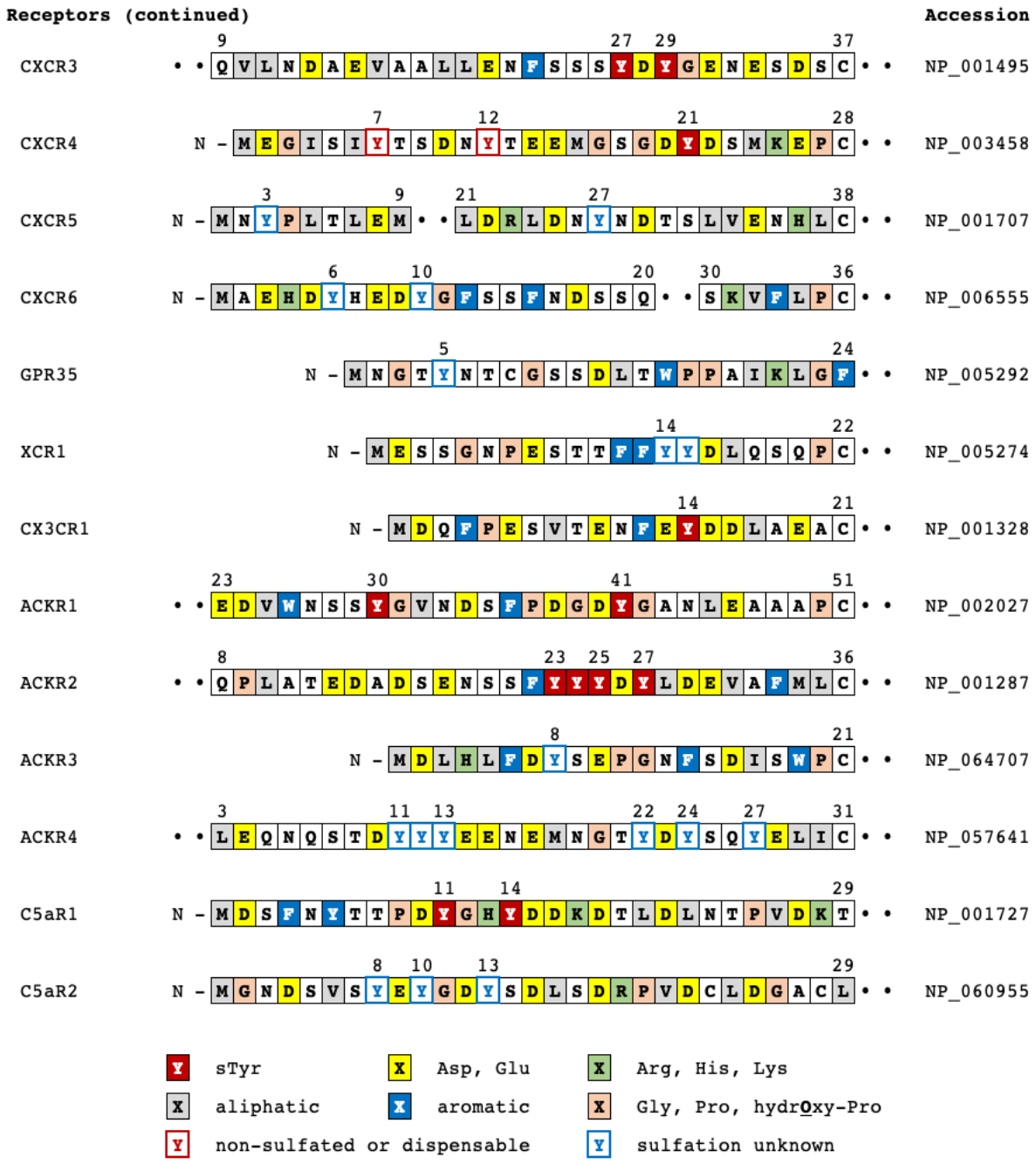
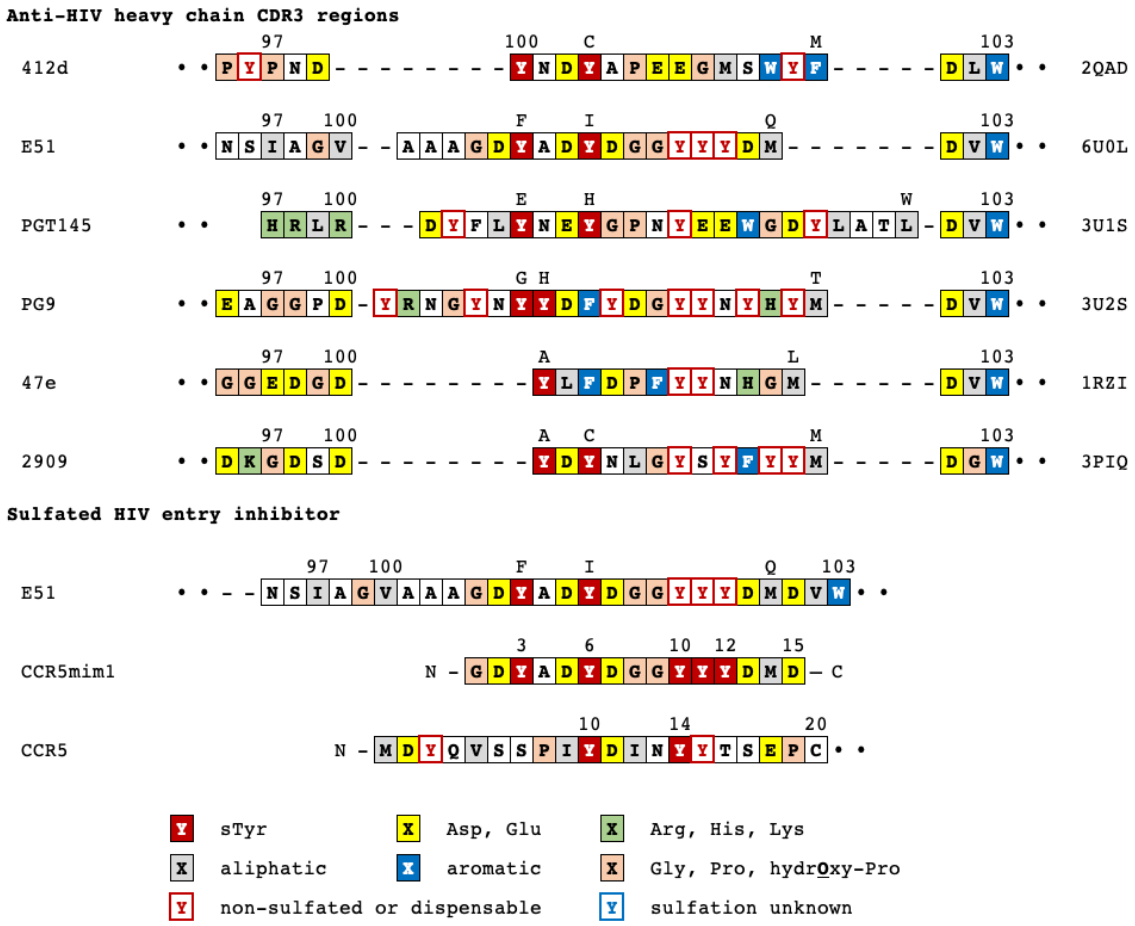


Fig. S2 (continued). Amino acid sequence contexts for proteins that contain an sTyr residue cluster.



Receptor amino-terminal sequences are shown as mature forms. The amino- and carboxyl-termini are denoted N and C, respectively, and residues are colored according to sidechain.

Notes for Fig. S2 (sTyr cluster sequence contexts, previous pages).

Symbol	Function	Ref. for Sulfation
PSGL-1	P-selectin glycoprotein ligand 1	(126)
CCR1	C-C chemokine receptor type 1	(153)
CCR2	C-C chemokine receptor type 2	(154)
CCR3	C-C chemokine receptor type 3	(155)
CCR4	C-C chemokine receptor type 4	—
CCR5	C-C chemokine receptor type 5	(81)
CCR6	C-C chemokine receptor type 6	—
CCR7	C-C chemokine receptor type 7	(33)
CCR8	C-C chemokine receptor type 8 (mouse)	(156)
CCR9	C-C chemokine receptor type 9	—
CCR10	C-C chemokine receptor type 10	—
CXCR1	C-X-C chemokine receptor type 1	(157)
CXCR2	C-X-C chemokine receptor type 2	(157)
CXCR3	C-X-C chemokine receptor type 3	(158)
CXCR4	C-X-C chemokine receptor type 4	(159)
CXCR5	C-X-C chemokine receptor type 5	—
CXCR6	C-X-C chemokine receptor type 6	—
GPR35	G-protein coupled receptor 35 (CXCR7)	—
XCR1	XC chemokine receptor 1	—
CX3CR1	CX3C chemokine receptor 1	(160)
ACKR1	atypical chemokine receptor 1 (DARC)	(161)
ACKR2	atypical chemokine receptor 2	(162)
ACKR3	atypical chemokine receptor 3	—
ACKR4	atypical chemokine receptor 4	—
C5aR1	C5a anaphylatoxin chemotactic receptor 1	(76)
C5aR2	C5a anaphylatoxin chemotactic receptor 2	—
412d	anti-HIV-1 Gp120-reactive antibody Fab heavy chain	(85)
E51	anti-HIV-1 Gp120-reactive antibody Fab heavy chain	(100)
PG9	anti-HIV-1 Gp120-reactive antibody Fab heavy chain	(163)
PG16	anti-HIV-1 Gp120-reactive antibody Fab heavy chain	(163)
47e	anti-HIV-1 Gp120-reactive antibody Fab heavy chain	(85)
PGT145	anti-HIV-1 Gp120-reactive antibody Fab heavy chain	(164)
2909	anti-HIV-1 Gp120-reactive antibody Fab heavy chain	(165)
CCR5mim1	CD4-Ig fused to a coreceptor-mimetic peptide	(166)
CCR5mim2	CD4-Ig fused to a coreceptor-mimetic peptide	(166)
CCR5mim6	CD4-Ig fused to a coreceptor-mimetic peptide	(167)

3. The blood clotting enzyme thrombin binds proteins with hirudin-like motifs, many of which contain a cluster of two or three sTyr residues.

Fig. S3A shows Hirudin-like motif sequences for proteins that bind thrombin exosite 1 (top panel) or thrombin exosite 2 (bottom panel). All sequences are from co-crystal structures, and contacts to residues in thrombin are indicated.

3.1 Hirudin is a relatively rare example of an sTyr-containing protein that binds thrombin exosite 1

Hirudin-like motif residues make conserved contacts to thrombin exosite 1 residues as revealed in co-crystal X-ray structures of thrombin bound to hirudin (17,168), to peptides from PAR1 and PAR3 (169), and to a peptide from the FV a2 region (170) (**Fig. S3A**). In each case, the conserved aromatic residue (illustrated by hirudin Phe-56) fits into a hydrophobic pocket and the conserved acidic residue (hirudin Glu-57) makes numerous electrostatic and hydrophobic contacts. This binding geometry positions the amino-terminal hirudin inhibitory domain near the active site.

The interaction of thrombin exosite 1 with hirudin is enhanced by the non-conserved residue sTyr-63, just past the hirudin-like motif (**Fig. S3A**). Thrombin affinity for sulfo-hirudin is tenfold greater than that for desulfo-hirudin, due to multiple electrostatic contacts between sulfate O atoms and exosite 1 residues (17). However, this role for sTyr is not a conserved feature of binding to thrombin exosite 1, because other known exosite 1-binding proteins do not have sTyr in a position analogous to hirudin sTyr-63. Instead, the non-conserved contacts to hirudin sTyr-63 illustrate the broad versatility of exosite 1 in presenting numerous potential electrostatic and hydrophobic contacts that enable diverse binding interactions with different partners. Indeed, both exosites accommodate sTyr residues at a variety of non-conserved positions.

3.2 Anophelin binds exosite 1 in reverse

Anophelin, a DTI from *Anopheles* mosquitos, binds exosite 1 in opposite orientation from the others (171). Remarkably, anophelin Phe-44 and Glu-42 contacts are similar to those made by residues at hirudin-like motif positions 2 and 3, respectively, reinforcing the overall importance of these interactions (**Fig. S3A**, line 3).

3.3 Contacts are not universally conserved.

Hirudin-like motif contacts to exosite 1 residues are conserved in most cases. Two partial exceptions are leuserpin-2 (heparin cofactor II) and thrombomodulin (**Fig. S3A**, lines 4 and 5). Although interactions with the acidic residues at hirudin-like motif position 2 are conserved, the aromatic residue at hirudin-like motif position 1 is absent. Perhaps to compensate, non-conserved contacts include a broader span of residues (16 for leuserpin-2 and 27 for thrombomodulin) than the others (11 for sulfo-hirudin, for example; **Fig. S3A**, line 1).

The non-sulfated prothrombin kringle domain F2 (**Fig. S3B**, line 14) binds exosite 2 transiently through a hirudin-like motif during thrombin formation (172). However, F2 binds through surface patches rather than an anionic segment (173), and therefore represents a distinct binding mode.

3.4 Haemidin-binding exosite is uncertain

Evidence indicates that haemidin, a potent DTI from *Haemadipsa* leeches, interacts with exosite 2 (169). Thrombin-haemidin co-crystals display a complex stoichiometry that complicates interpretation. Nevertheless, the X-ray structure shows haemidin hirudin-like motif contacts to exosite 1 that are similar to those made by hirudin, whereas haemidin contacts to exosite 2 are scant (174) (**Fig. S3A**, line 8).

3.5 Complement C1 has a thrombin-like exosite

Like hemostasis, the complement system is activated through an endoprotease cascade (175). 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 (176,177) (**Fig. S3B**, line 12). The X-ray co-crystal structure of complement C1 bound to gigastasin, a C1 inhibitor made by *Haementaria* leeches, reveals electrostatic contacts between sulfate O atoms from a pair of sTyr residues to C1 exosite basic residues (120). Thus, the homologous hemostasis and complement systems share the use of hirudin-like motifs to direct partner protein binding to exosites.

Fig. S3A. Amino acid sequence contexts for proteins that contain a hirudin-like motif.

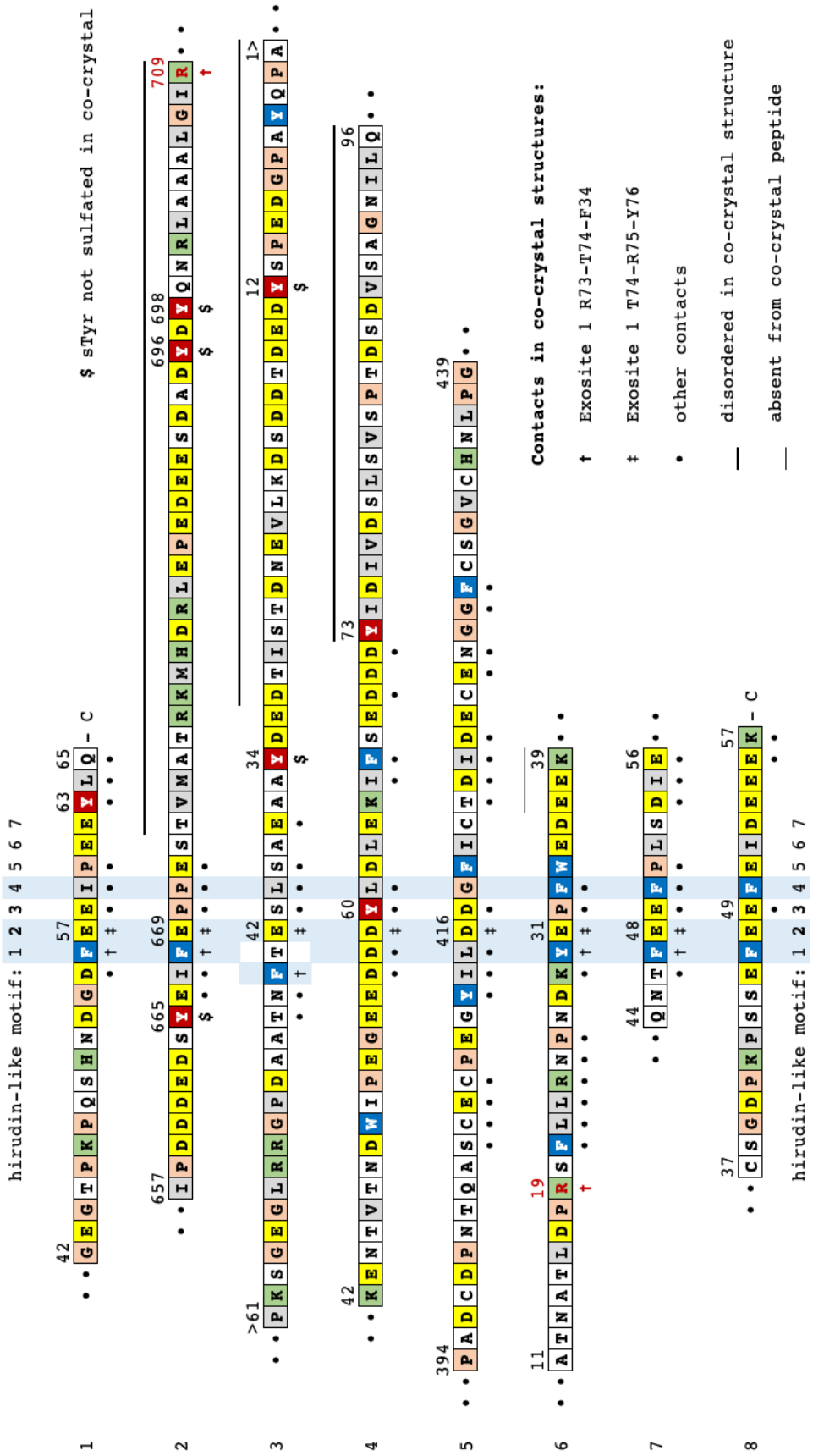
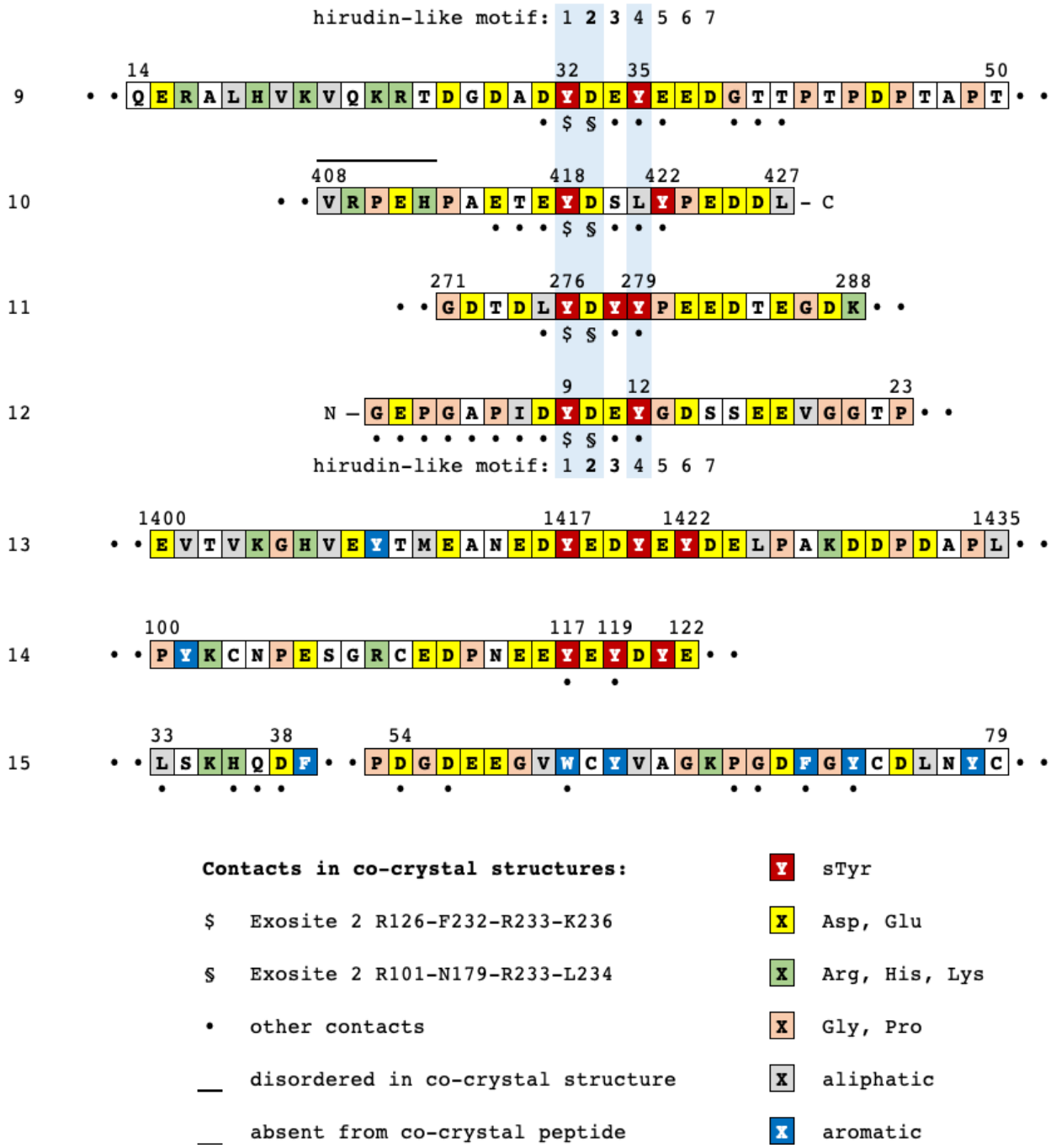


Fig. S3A (continued). Amino acid sequence contexts for proteins that contain a hirudin-like motif.



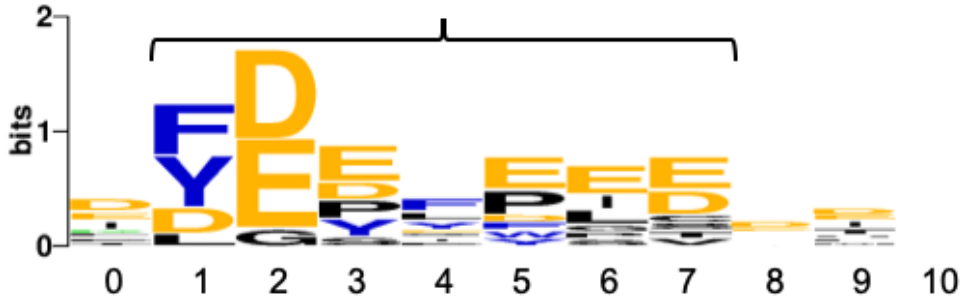
The indicated internal sequence segments are shown, and residues are colored according to sidechain. Hirudin-like motif conserved positions 1, 2 and 4 are highlighted.

Notes for Fig. S3A (hirudin-like motif sequences, previous pages).

Ln.	Protein	Size (aa)	Structure[a]	Size (aa)	Accession
1	Hirudin (variant 1)	65	full-length sulfo-	65	2PW8
2	Factor V (a2 segment)	2,224	peptide I657-R709	53	3P6Z
3	Anophelin (Aa)	61	full-length desulfo-	61	4E05
4	Leuserpin 2 (HCII)	480	full-length sulfo-	480	1JMO
5	Thrombomodulin	554	V345-C462 (TME456)	118	1DX5
6	PAR1 (human)	404	peptide 11-36	25	3LU9
7	PAR3 (mouse)	369	peptide 44-56	13	2PUX
8	Haemadin	57	full-length sulfo-	57	1E0F
9	Madanin-1	60	full-length sulfo-	60	5L6N
10	Fibrinogen γ'	427	peptide V408-L427	20	2HWL
11	Platelet GpIbα	636	peptide G271-T284	14	4CH2
12	Tsetse thrombin inhibitor	32	full-length sulfo-	32	6TKG
13	Complement C4	1,744	—	—	AAA51855
14	Gigastasin	122	full-length sulfo-	122	5UBM
15	Prothrombin F2	—	Kringle domain	79	2HPQ

[a] portion of indicated thrombin-binding protein present in the co-crystal.

Fig. S3B. Hirudin-like motif sequence logo

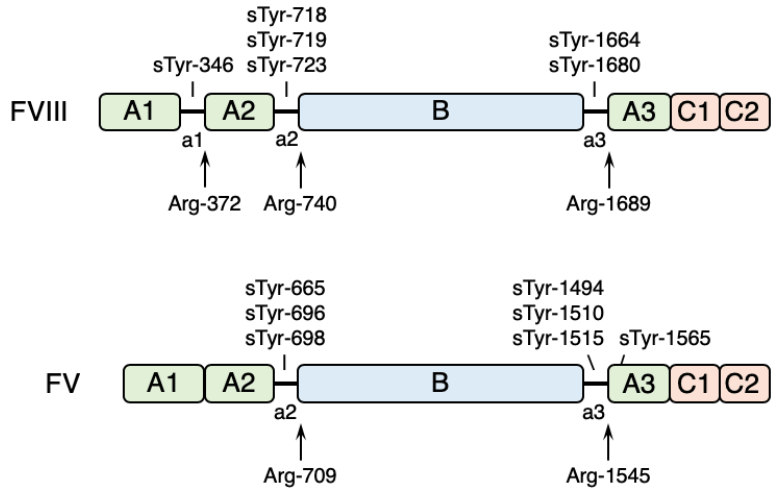


The sequence logo (178,179) was generated from the sequence alignment shown in **Fig. S3A**, and includes eleven sequences (lines 1, 2, 4-11, 14). Position 4 contains several different residues, but most are hydrophobic and relatively large (Tyr, sTyr, Leu). The hirudin-like motif, defined here as positions 1-7, is indicated with a bracket. Positions 1 and 2 comprise a conserved aromatic-acidic residue pair, consisting of Phe, Tyr or sTyr at position 1 and Asp or Glu at position 2. 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 (114).

4. The blood clotting factors FVII and FV contain six and seven sTyr residues, respectively.

Coagulation factors VIII (FVIII) and V (FV) are homologous multidomain proteins each with several sTyr residues (131) (**Fig. S4A**). Thrombin catalyzes multiple endoproteolytic cleavages to activate FVIII and FV, which in turn stimulate the rapid increase in thrombin activity necessary for clot formation.

Fig. S4A. Factor VIII and Factor V domain architectures.

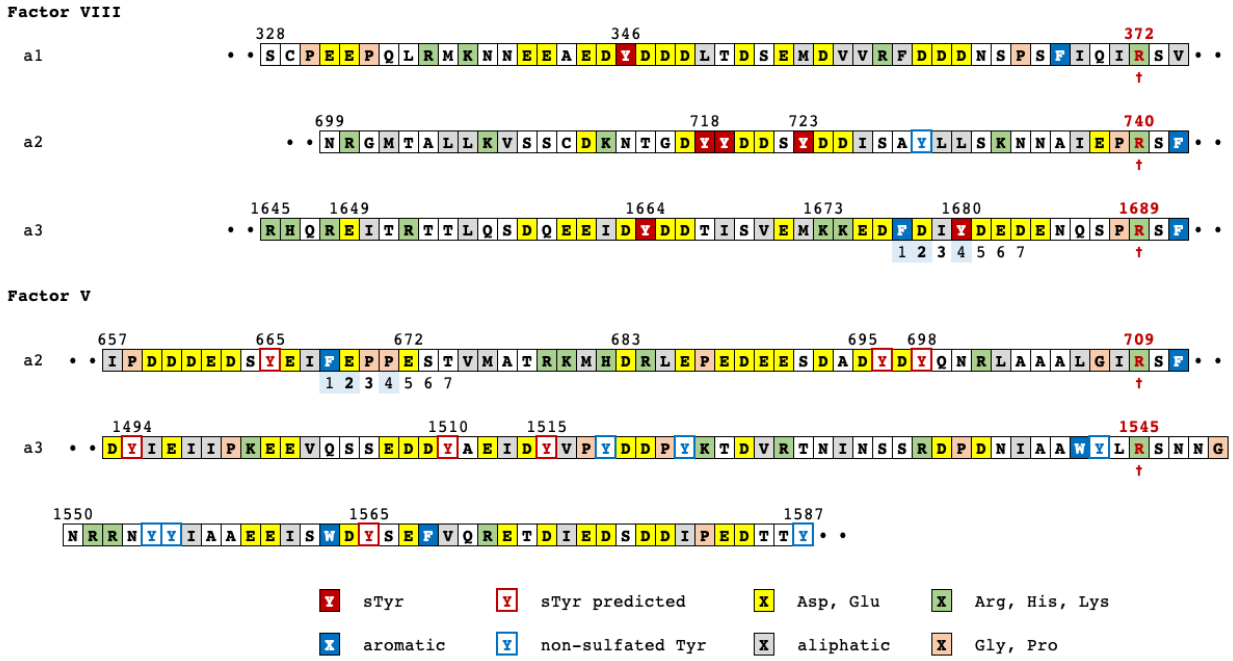


Domains are shown schematically, approximately to scale. Thrombin cleavage at the indicated Arg residues generates the active heteromeric forms. Redrawn from reference (131).

4.1 sTyr residues in coagulation factor FVIII.

FVIII residue sTyr-1680 lies in a hirudin-like motif of uncertain function (**Fig. S4B**). Although evidence suggests that FVIII binds both exosite 1 and exosite 2 (180), specific interactions are not defined at the atomic level. Separately, FVIII residue sTyr-1680 is important for the noncovalent complex between FVIII and vonWillebrand factor (vWF), a multidomain protein whose conformation depends on liquid shear flow. Tethered vWF protects circulating FVIII in part by blocking proteolysis (181,182). Again, atomic details for interactions between FVIII and vWF are not known (183).

Fig. S4B. Factor VIII and Factor V sTyr sequence contexts.



Sequence segment contexts are depicted in **Fig. S4A**. Red numbers and † symbol show thrombin cleavage sites. Hirudin-like motifs are indicated for FVIII segment a3 and for FV segment a2. Accession numbers are NP_000121 (human factor V) and NP_000123 (human factor VIII).

FVIII residues sTyr-346 and sTyr-1664 both are important for efficient prothrombin conversion (182). These are roughly 25 residues proximal to the respective endoproteolysis sites (**Fig. S4B**) and likely bind exosite 1 (114,115). Separately, the sTyr cluster of residues at positions 718, 719 and 721 was identified as "important for intrinsic activity" (182), but reportedly there is little effect of substituting Phe in place of all three sTyr residues in this cluster (184).

4.2 sTyr residues in coagulation factor FV.

Specific sTyr residues in FV were not determined (185), but rather are predicted from sequence analysis (131). Analyses with multiply-substituted mutant proteins have implicated these anionic segments in prothrombin activation (186), and references therein), but specific functions for most individual sTyr are unknown.

The co-crystal X-ray structure for thrombin in complex with a non-sulfated FV peptide shows residues near sTyr-665 binding to exosite 1. Modeling suggests that the remainder of the FV segment wraps around the thrombin surface, potentially enabling sTyr-696 and sTyr-698 to interact with exosite 2 (134,187,188) (**Fig. S4B**).

Thus, although both exosites are implicated in FV recognition (114,115,133), details for how exosites interact with different parts of FV mostly are unknown.

References for supporting information

136. Rehfeld, J. F. (2017) Cholecystokinin-From Local Gut Hormone to Ubiquitous Messenger. Front Endocrinol (Lausanne) 8, 47
137. Vishnuvardhan, D., and Beinfeld, M. C. (2000) Role of tyrosine sulfation and serine phosphorylation in the processing of procholecystokinin to amidated cholecystokinin and its secretion in transfected AtT-20 cells. Biochemistry 39, 13825-13830
138. Miller, L. J., and Gao, F. (2008) Structural basis of cholecystokinin receptor binding and regulation. Pharmacol Ther 119, 83-95
139. Barnum, S. R. (2017) Complement: A primer for the coming therapeutic revolution. Pharmacol Ther 172, 63-72
140. 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
141. Matsuzaki, Y., Ogawa-Ohnishi, M., Mori, A., and Matsubayashi, Y. (2010) Secreted peptide signals required for maintenance of root stem cell niche in Arabidopsis. Science 329, 1065-1067
142. Nakayama, T., Shinohara, H., Tanaka, M., Baba, K., Ogawa-Ohnishi, M., and Matsubayashi, Y. (2017) A peptide hormone required for Casparian strip diffusion barrier formation in Arabidopsis roots. Science 355, 284-286
143. Amano, Y., Tsubouchi, H., Shinohara, H., Ogawa, M., and Matsubayashi, Y. (2007) Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in Arabidopsis. Proc Natl Acad Sci U S A 104, 18333-18338
144. Okamoto, S., Suzuki, T., Kawaguchi, M., Higashiyama, T., and Matsubayashi, Y. (2015) A comprehensive strategy for identifying long-distance mobile peptides in xylem sap. Plant J 84, 611-620
145. Patel, N., Mohd-Radzman, N. A., Corcilius, L., Crossett, B., Connolly, A., Cordwell, S. J., Ivanovici, A., Taylor, K., Williams, J., Binos, S., Mariani, M., Payne, R. J., and Djordjevic, M. A. (2018) Diverse peptide hormones affecting root growth identified in the Medicago truncatula secreted peptidome. Mol Cell Proteomics 17, 160-174
146. Bundgaard, J. R., Vuust, J., and Rehfeld, J. F. (1995) Tyrosine O-sulfation promotes proteolytic processing of progastrin. EMBO J 14, 3073-3079
147. 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
148. Chakrabarti, S., Hinczewski, M., and Thirumalai, D. (2017) Phenomenological and microscopic theories for catch bonds. J Struct Biol 197, 50-56
149. McEver, R. P., and Zhu, C. (2010) Rolling cell adhesion. Annu Rev Cell Dev Biol 26, 363-396
150. 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
151. 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

152. 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

153. Sanchez, J., Z, E. H., Lane, J. R., Liu, X., Bridgford, J. L., Payne, R. J., Canals, M., and Stone, M. J. (2019) Evaluation and extension of the two-site, two-step model for binding and activation of the chemokine receptor CCR1. J Biol Chem 294, 3464-3475

154. Preobrazhensky, A. A., Dragan, S., Kawano, T., Gavrilin, M. A., Gulina, I. V., Chakravarty, L., and Kolattukudy, P. E. (2000) Monocyte chemotactic protein-1 receptor CCR2B is a glycoprotein that has tyrosine sulfation in a conserved extracellular N-terminal region. J Immunol 165, 5295-5303

155. Zhu, J. Z., Millard, C. J., Ludeman, J. P., Simpson, L. S., Clayton, D. J., Payne, R. J., Widlanski, T. S., and Stone, M. J. (2011) Tyrosine sulfation influences the chemokine binding selectivity of peptides derived from chemokine receptor CCR3. Biochemistry 50, 1524-1534

156. Gutierrez, J., Kremer, L., Zaballos, A., Goya, I., Martinez, A. C., and Marquez, G. (2004) Analysis of post-translational CCR8 modifications and their influence on receptor activity. J Biol Chem 279, 14726-14733

157. Moussouras, N. A., Getschman, A. E., Lackner, E. R., Veldkamp, C. T., Dwinell, M. B., and Volkman, B. F. (2017) Differences in sulfotyrosine binding amongst CXCR1 and CXCR2 chemokine ligands. Int J Mol Sci 18

158. Colvin, R. A., Campanella, G. S., Manice, L. A., and Luster, A. D. (2006) CXCR3 requires tyrosine sulfation for ligand binding and a second extracellular loop arginine residue for ligand-induced chemotaxis. Mol Cell Biol 26, 5838-5849

159. 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

160. Fong, A. M., Alam, S. M., Imai, T., Haribabu, B., and Patel, D. D. (2002) CX3CR1 tyrosine sulfation enhances fractalkine-induced cell adhesion. J Biol Chem 277, 19418-19423

161. Choe, H., Moore, M. J., Owens, C. M., Wright, P. L., Vasilieva, N., Li, W., Singh, A. P., Shakri, R., Chitnis, C. E., and Farzan, M. (2005) Sulphated tyrosines mediate association of chemokines and Plasmodium vivax Duffy binding protein with the Duffy antigen/receptor for chemokines (DARC). Mol Microbiol 55, 1413-1422

162. Hewit, K. D., Fraser, A., Nibbs, R. J., and Graham, G. J. (2014) The N-terminal region of the atypical chemokine receptor ACKR2 is a key determinant of ligand binding. J Biol Chem 289, 12330-12342

163. Pejchal, R., Walker, L. M., Stanfield, R. L., Phogat, S. K., Koff, W. C., Poignard, P., Burton, D. R., and Wilson, I. A. (2010) Structure and function of broadly reactive antibody PG16 reveal an H3 subdomain that mediates potent neutralization of HIV-1. Proc Natl Acad Sci U S A 107, 11483-11488

164. McLellan, J. S., Pancera, M., Carrico, C., Gorman, J., Julien, J. P., Khayat, R., Louder, R., Pejchal, R., Sastry, M., Dai, K., O'Dell, S., Patel, N., Shahzad-ul-Hussan, S., Yang, Y., Zhang, B., Zhou, T., Zhu, J., Boyington, J. C., Chuang, G. Y., Diwanji, D., Georgiev, I., Kwon, Y. D., Lee, D., Louder, M. K., Moquin, S., Schmidt, S. D., Yang, Z. Y., Bonsignori, M., Crump, J. A., Kapiga, S. H., Sam, N. E., Haynes, B. F., Burton, D. R., Koff, W. C., Walker, L. M., Phogat, S., Wyatt, R., Orwenyo, J., Wang, L. X., Arthos, J., Bewley, C. A., Mascola, J. R., Nabel, G. J., Schief, W. R., Ward, A. B., Wilson, I. A., and Kwong, P. D. (2011) Structure of HIV-1 gp120 V1/V2 domain with broadly neutralizing antibody PG9. Nature 480, 336-343

165. Changela, A., Wu, X., Yang, Y., Zhang, B., Zhu, J., Nardone, G. A., O'Dell, S., Pancera, M., Gorny, M. K., Phogat, S., Robinson, J. E., Stamatatos, L., Zolla-Pazner, S., Mascola, J. R., and Kwong, P. D. (2011) Crystal structure of human antibody 2909 reveals conserved features of quaternary structure-specific antibodies that potently neutralize HIV-1. J Virol 85, 2524-2535

166. Chiang, J. J., Gardner, M. R., Quinlan, B. D., Dorfman, T., Choe, H., and Farzan, M. (2012) Enhanced recognition and neutralization of HIV-1 by antibody-derived CCR5-mimetic peptide variants. J Virol 86, 12417-12421

167. Fetzer, I., Gardner, M. R., Davis-Gardner, M. E., Prasad, N. R., Alfant, B., Weber, J. A., and Farzan, M. (2018) eCD4-Ig variants that more potently neutralize HIV-1. J Virol 92

168. Rydel, T. J., Tulinsky, A., Bode, W., and Huber, R. (1991) Refined structure of the hirudin-thrombin complex. J Mol Biol 221, 583-601

169. Corral-Rodriguez, M. A., Macedo-Ribeiro, S., Pereira, P. J., and Fuentes-Prior, P. (2010) Leech-derived thrombin inhibitors: from structures to mechanisms to clinical applications. J Med Chem 53, 3847-3861

170. Gandhi, P. S., Chen, Z., Appelbaum, E., Zapata, F., and Di Cera, E. (2011) Structural basis of thrombin-protease-activated receptor interactions. IUBMB Life 63, 375-382

171. Figueiredo, A. C., de Sanctis, D., Gutierrez-Gallego, R., Cereija, T. B., Macedo-Ribeiro, S., Fuentes-Prior, P., and Pereira, P. J. (2012) Unique thrombin inhibition mechanism by anophelin, an anticoagulant from the malaria vector. Proc Natl Acad Sci U S A 109, E3649-3658

172. Krishnaswamy, S. (2013) The transition of prothrombin to thrombin. J Thromb Haemost 11 Suppl 1, 265-276

173. Arni, R. K., Padmanabhan, K., Padmanabhan, K. P., Wu, T. P., and Tulinsky, A. (1993) Structures of the noncovalent complexes of human and bovine prothrombin fragment 2 with human PPACK-thrombin. Biochemistry 32, 4727-4737

174. Richardson, J. L., Kroger, B., Hoeffken, W., Sadler, J. E., Pereira, P., Huber, R., Bode, W., and Fuentes-Prior, P. (2000) Crystal structure of the human alpha-thrombin-haemadin complex: an exosite II-binding inhibitor. EMBO J 19, 5650-5660

175. 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

176. 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

177. 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

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

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

180. Camire, R. M., and Bos, M. H. (2009) The molecular basis of factor V and VIII procofactor activation. J Thromb Haemost 7, 1951-1961

181. Leyte, A., van Schijndel, H. B., Niehrs, C., Huttner, W. B., Verbeet, M. P., Mertens, K., and van Mourik, J. A. (1991) Sulfation of Tyr1680 of human blood coagulation factor VIII is essential for the interaction of factor VIII with von Willebrand factor. J Biol Chem 266, 740-746

182. Michnick, D. A., Pittman, D. D., Wise, R. J., and Kaufman, R. J. (1994) Identification of individual tyrosine sulfation sites within factor VIII required for optimal activity and efficient thrombin cleavage. J Biol Chem 269, 20095-20102

183. Dagil, L., Troelsen, K. S., Bolt, G., Thim, L., Wu, B., Zhao, X., Tuddenham, E. G. D., Nielsen, T. E., Tanner, D. A., Faber, J. H., Breinholt, J., Rasmussen, J. E., and Hansen, D. F. (2019) Interaction Between the a3 Region of Factor VIII and the TIL'E' Domains of the von Willebrand Factor. Biophys J 117, 479-489

184. Newell, J. L., and Fay, P. J. (2008) Acidic residues C-terminal to the A2 domain facilitate thrombin-catalyzed activation of factor VIII. Biochemistry 47, 8786-8795

185. Pittman, D. D., Tomkinson, K. N., Michnick, D., Selighsohn, U., and Kaufman, R. J. (1994) Posttranslational sulfation of factor V is required for efficient thrombin cleavage and activation and for full procoagulant activity. Biochemistry 33, 6952-6959

186. Hirbawi, J., Vaughn, J. L., Bukys, M. A., Vos, H. L., and Kalafatis, M. (2010) Contribution of amino acid region 659-663 of Factor Va heavy chain to the activity of factor Xa within prothrombinase. Biochemistry 49, 8520-8534

187. Corral-Rodriguez, M. A., Bock, P. E., Hernandez-Carvajal, E., Gutierrez-Gallego, R., and Fuentes-Prior, P. (2011) Structural basis of thrombin-mediated factor V activation: the Glu666-Glu672 sequence is critical for processing at the heavy chain-B domain junction. Blood 117, 7164-7173

188. Shim, J. Y., Lee, C. J., Wu, S., and Pedersen, L. G. (2015) A model for the unique role of factor Va A2 domain extension in the human ternary thrombin-generating complex. Biophys Chem 199, 46-50