



Structures of SALSA/DMBT1 SRCR domains reveal the conserved ligand-binding mechanism of the ancient SRCR fold

Martin P Reichhardt¹ , Vuokko Loimaranta², Susan M Lea^{1,3} , Steven Johnson¹ 

The scavenger receptor cysteine-rich (SRCR) family of proteins comprises more than 20 membrane-associated and secreted molecules. Characterised by the presence of one or more copies of the ~110 amino-acid SRCR domain, this class of proteins have widespread functions as antimicrobial molecules, scavenger receptors, and signalling receptors. Despite the high level of structural conservation of SRCR domains, no unifying mechanism for ligand interaction has been described. The SRCR protein SALSA, also known as DMBT1/gp340, is a key player in mucosal immunology. Based on detailed structural data of SALSA SRCR domains 1 and 8, we here reveal a novel universal ligand-binding mechanism for SALSA ligands. The binding interface incorporates a dual cation-binding site, which is highly conserved across the SRCR superfamily. Along with the well-described cation dependency on most SRCR domain–ligand interactions, our data suggest that the binding mechanism described for the SALSA SRCR domains is applicable to all SRCR domains. We thus propose to have identified in SALSA a conserved functional mechanism for the SRCR class of proteins.

DOI [10.26508/lsa.201900502](https://doi.org/10.26508/lsa.201900502) | Received 26 July 2019 | Revised 14 February 2020 | Accepted 14 February 2020 | Published online 25 February 2020

Introduction

The salivary scavenger and agglutinin (SALSA), also known as gp340, “deleted in malignant brain tumors 1” (DMBT1) and salivary agglutinin (SAG), is a multifunctional molecule found in high abundance on human mucosal surfaces (1, 2, 3, 4). SALSA has widespread functions in innate immunity, inflammation, epithelial homeostasis, and tumour suppression (5, 6, 7). SALSA binds and agglutinates a broad spectrum of pathogens including, but not limited to, human immunodeficiency virus type 1, *Helicobacter pylori*, *Salmonella enterica* serovar Typhimurium, and many types of streptococci (8, 9, 10, 11). In addition to its microbial scavenging function, SALSA has been suggested to interact with a wide array of endogenous immune defence molecules. These include secretory IgA, surfactant

proteins A (SP-A) and D (SP-D), lactoferrin, mucin-5B, and components of the complement system (1, 2, 12, 13, 14, 15, 16, 17, 18). SALSA thus engages innate immune defence molecules and has been suggested to cooperatively mediate microbial clearance and maintenance of the integrity of the mucosal barrier.

The 300- to 400-kD SALSA glycoprotein is encoded by the *DMBT1* gene. The canonical form of the gene encodes 13 highly conserved scavenger receptor cysteine-rich (SRCR) domains, followed by two C1r/C1s, urchin embryonic growth factor and bone morphogenetic protein-1 (CUB) domains that surround a 14th SRCR domain, and finally a zona pellucida domain at the C terminus (19, 20). The first 13 SRCRs are 109 aa domains found as “pearls on a string” separated by SRCR-interspersed domains (SIDs) (Fig 1A) (1, 21). The SIDs are 20- to 23-aa-long stretches of predicted disorder containing a number of glycosylation sites, which have been proposed to force them into an extended conformation of roughly 7 nm (7). In addition to this main form, alternative splicing and copy number variation mechanisms lead to expression of variants of SALSA containing variable numbers of SRCR domains in the N-terminal region.

The SRCR protein superfamily include a range of secreted and membrane-associated molecules, all containing one or more SRCR domains. For a number of these molecules, the SRCR domains have been directly implicated in ligand binding. These include CD6 signalling via CD166, CD163-mediated clearance of the haemoglobin-haptoglobin complex, Mac-2 binding protein’s (M2bp’s) interaction with matrix components, and the binding of microbial ligands by the scavenger receptors SR-A1, SP α , and MARCO (22, 23, 24, 25, 26, 27). Although the multiple SALSA SRCR domains likewise have been implicated in ligand binding, the molecular basis for its diverse interactions remains unknown.

To understand the multiple ligand-binding properties of the SALSA molecule, we undertook an X-ray crystallographic study to provide detailed information of the SALSA interaction surfaces. We here provide the atomic resolution structures of SALSA SRCR domains 1 and 8. We identify cation-binding sites and demonstrate their importance for ligand binding. By comparing our data to previously published structures of SRCR domains, we propose a

¹Sir William Dunn School of Pathology, University of Oxford, Oxford, UK ²Institute of Dentistry, University of Turku, Turku, Finland ³Central Oxford Structural Molecular Imaging Centre, University of Oxford, Oxford, UK

Correspondence: steven.johnson@path.ox.ac.uk; martinreichhardt@gmail.com

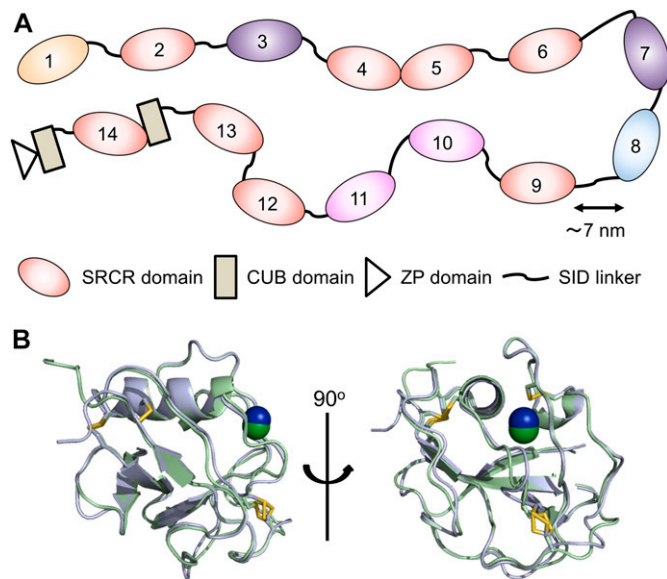


Figure 1. Crystal structure of SALSA domains SRCR1 and SRCR8. (A) Schematic representation of the domain organization of full-length SALSA. SRCR1 and SRCR8 are highlighted in green and blue, respectively. All SRCR domains share >88% sequence identity. 100% identity is shared by SRCR3 and 7 (yellow) and SRCR10 and 11 (purple). (B) Front and side views of an overlay of SRCR1 (green) and SRCR8 (blue), showing four conserved disulphide bridges (yellow). Both SRCR1 and SRCR8 were found to coordinate a metal ion, modelled as Mg^{2+} (dark green for SRCR1 and dark blue for SRCR8). The limited structural variation observed between SRCR1 and SRCR8 (92% sequence identity) imply that these are appropriate representations of all SALSA SRCR domains 1–13.

generalised binding mechanism for this ancient, evolutionarily conserved, fold.

Results

The scavenger receptor SALSA has a very wide range of described ligands, including microbial, host innate immune, and ECM molecules. To understand the very broad ligand-binding abilities of the SALSA molecule, we applied a crystallographic approach to determine the structure of the ligand-binding SRCR domains of SALSA. SRCR domains 1 and 8 (SRCR1 and SRCR8) were expressed in *Drosophila melanogaster* Schneider S2 cells with a C-terminal His-tag. The domains were purified by Ni-chelate and size-exclusion chromatography (SEC) and crystallised, and the structures were solved by molecular replacement. This yielded the structures of SRCR1 and SRCR8 at 1.77 and 1.29 Å, respectively (Fig 1B). (For crystallographic details, see Table 1).

The SALSA SRCR domains reveal a classic globular SRCR-fold, with four conserved disulphide bridges, as described for the SRCR type B domains. The fold contains one α -helix and one additional single helical turn. The N and C termini come together in a four-stranded β -sheet. SALSA SRCR1–13 are highly conserved, with 88–100% identity. Variation is only observed in 9 of the 109 aa residues, all of these observed in peripheral loops, without apparent structural significance. Combined, the data from SRCR1 and SRCR8 are thus valid representations of all SALSA SRCR domains. Both SRCR1 and SRCR8

are stabilized by a metal ion buried in the globular fold. The placement suggests the ion is bound during the folding of the domain and is modelled as Mg^{2+} , which is present in the original expression medium and in the crystallisation conditions of SRCR1.

So far, all described ligand-binding interactions of SALSA have been shown to be Ca^{2+} -dependent. We therefore proceeded to address the ligand-binding potential of the SRCR domains by adding Ca^{2+} , Mg^{2+} , and a cocktail of sugars to SRCR8 crystals before freezing. This yielded a second crystal form of SRCR8 with the original Mg^{2+} ion, site 1, and two additional cations bound, sites 2 and 3 (Fig 2). All three sites are class three cation-binding sites, with the coordination obtained from residues in distant parts of the sequence (28).

Assignment of the identity of the ions at the paired site was carried out by modelling Mg^{2+} or Ca^{2+} at each site, followed by refinement of the structure and analysis of the difference maps (Fig S1). These revealed that Mg^{2+} best satisfied the data at both sites, consistent with the 20-fold molar excess of Mg^{2+} over Ca^{2+} in the crystallisation solution. However, it is worth noting that either site could likely accommodate either cation depending on local concentration. Analysis of the bond lengths and coordination numbers suggest that site 2 is a canonical Mg^{2+} site, with octahedral geometry and average bond lengths of 2.1 Å, while site 3 displays a higher coordination number and longer bond lengths, more consistent with a Ca^{2+} -binding site (29). The Mg^{2+} at site 1 is coordinated by the backbone carbonyl groups of S1021 and V1060, as well as the side chains of D1023 and D1026, and two waters, and is buried in the domain fold (Fig 2C). The Mg^{2+} at site 2 is coordinated by D1019, D1020, and E1086 plus three waters (Fig 2D). The Mg^{2+} at site 3 is coordinated by the side chains of D1020, D1058, D1059, and N1081, with additional contributions from a water and an extra density (Fig 2E). Attempts to model this extra density as any of the sugars or alcohols present in the crystallisation solution failed to produce a satisfactory fit; therefore, it likely represents a superposition of a number of molecules.

In contrast to the Mg^{2+} at site 1, these cations at sites 2 and 3 are exposed on the surface of the domain, and the protein only contributes a fraction of the coordination sphere, with the remainder contributed by waters or small molecules from the crystallisation solution. According to the literature, the majority of described SALSA ligands are negatively charged. Thus, the surface-exposed cations likely provide a mechanism for ligand binding for the SALSA SRCR domains, whereby the anions of the ligand substitute for the waters or the density at site 3 observed in our structure. To test this hypothesis, site-directed mutagenesis was used, targeting the key residues coordinating sites 2 and 3. Included in the further analysis were single mutations D1019A and D1020A. While mutation of D1019 is expected to only disrupt binding of cations at site 2, mutation of the shared D1020, will likely affect binding of both cations.

As SALSA recognizes a very broad range of biological ligands, we set out to test the effect of SRCR domain point mutations on interactions with a wide array of biological ligands. These included binding to (1) hydroxyapatite, a phosphate-rich mineral essential for the binding of SALSA to the teeth surface, where it mediates antimicrobial effects (30); (2) heparin, a sulphated glycosaminoglycan as a mimic for the ECM/cell surface, for which binding of SALSA has been described to affect cellular differentiation and microbial colonisation (31); (3) Group A *Streptococcus* surface protein, Spy0843, a leucine-rich repeat protein demonstrated to bind to SALSA (32) (Fig 3).

Table 1. Data collection and refinement statistics (molecular replacement).

	SRCR1 (pdbid: 6sa4)	SRCR8 (pdbid: 6sa5)	SRCR8soak (pdbid: 6san)
Data collection			
Space group	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 1 2 ₁ 1
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)	36.77, 45.19, 69.37	32.82, 40.82, 62.99	27.24, 46.64, 93.63
α , β , γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00	90.00, 97.37, 90.00
Resolution (Å)	28.52–1.77 (1.80–1.77) ^a	40.82–1.29 (1.31–1.29)	46.64–1.36 (1.39–1.36)
<i>R</i> _{merge}	0.17 (1.36)	0.117 (1.12)	0.074 (0.801)
<i>I</i> / <i>σ</i>	8.1 (1.1)	8.6 (0.9)	14.9 (2.2)
Completeness (%)	99.8 (99.3)	100 (99.6)	98.3 (96.9)
Redundancy	6.3 (6.6)	11.4 (8.0)	6.6 (6.4)
Refinement			
Resolution (Å)	28.52–1.77 (1.95–1.77)	34.27–1.29 (1.35–1.29)	30.95–1.36 (1.39–1.36)
No. of reflections	11,730	21,958	49,166
<i>R</i> _{work} / <i>R</i> _{free}	0.186/0.229 (0.266/0.348)	0.155/0.188 (0.326/0.279)	0.186/0.226 (0.267/0.326)
No. of atoms			
Protein	824	829	3,192
Ligand/ion	26	18	34
Water	109	124	358
<i>B</i> -factors			
Protein	22.36	16.80	17.48
Ligand/ion	49.33	54.03	26.08
Water	31.15	33.66	35.55
R.m.s. deviations			
Bond lengths (Å)	0.006	0.009	0.007
Bond angles (°)	0.809	1.026	0.87
Ramachandran outliers	0	0	0
Rotamer outliers	0	0	0

Number of crystals was one for each structure.

^aValues in parentheses are for highest resolution shell. Data from SRCR1 and SRCR8 crystals were collected on Diamond beamline I04, while data for SRCR8soak were collected on beamline I03.

Different binding assays provide an understanding of a generalised binding mechanism of SALSA SRCR domains. While the WT SRCR domain bound to all three ligands, both of the cation-binding site mutations, D1019A and D1020A, abolished binding. This is consistent with the bound cations acting as a bridge for ligand interaction and thus provides a mechanistic explanation for the binding properties of the SALSA SRCR domains. In the literature, SALSA ligand binding has been described as specifically calcium dependent. To verify this, we conducted binding assays in an MgEGTA-containing buffer (Fig 3C). The exchange of magnesium for calcium abolished ligand interactions, thus supporting a calcium-specific mediation of binding. As mutation of site 2 alone (modelled as Mg²⁺ in our structure) abolished binding, our data suggest that Ca²⁺ may occupy site 2 under physiological conditions.

All known members of the SRCR superfamily share a very high degree of identity, both at the sequence and structural levels. An FFAS search (33) of the SRCR8 sequence showed highest similarities

to CD163 SRCR5 (score: -65.4, 46% identity), M2bp (score: -64.4, 54% identity), neurotrypsin (score: -61.5, 50% identity), MARCO (score: -59.4, 50% identity), CD5 SRCR1 (score: -48.8, 26% identity), and CD6 SRCR2 (score: -45.6, 60% identity). Using the Dali server (34), searches for the cation-binding SRCR8 soak structure identified two top hits as M2bp (pdbid: 1by2) and CD6 SRCR3 (pdbid: 5a2e). These were identified with respective Z-scores of 21.4 (r.m.s.d. of 1.1 Å with 106 of 112 residues aligned) and 20.4 (r.m.s.d. of 1.5 Å with 109 of 109 residues aligned). Despite the classical division of SRCR superfamily proteins into groups A and B, based on the conserved three versus four cysteine bridges, the SRCR fold is very highly conserved, and the SALSA SRCR domain structures correlate closely to both group A and group B SRCR superfamily domains (Fig 4A).

For members of the SRCR superfamily where the SRCR domain directly partakes in ligand binding, both microbial and endogenous protein ligands have been described. For MARCO, crystallographic structures identified a cation-binding site exactly corresponding to

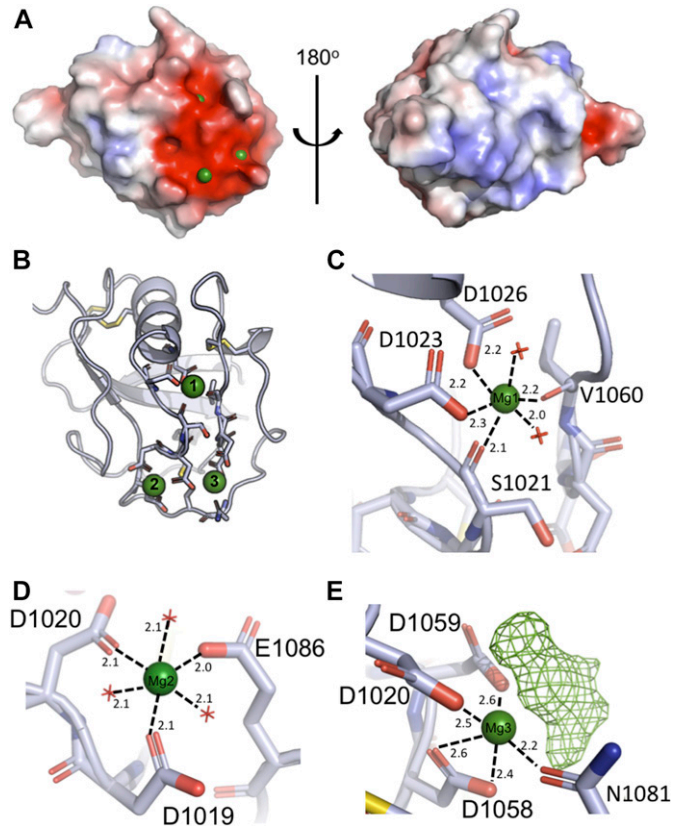


Figure 2. Crystal structure of SRCR8 with bound magnesium ions indicates mechanism of ligand binding. (A) Surface charge distribution of SRCR8 (calculated without the presence of cations) shows a positive cluster on one side with a strong negative cluster on the other. The negative cluster expands across ~300 Å² and mediates the binding of three cations (green). (B) Representation of the residues coordinating the three cations. The upper Mg²⁺, site 1, sits somewhat buried in the structure and may be essential for structural stability. The lower cations at sites 2 and 3 are more exposed. (C) Detailed view of the coordination of the upper Mg²⁺, site 1. The coordination number of six is achieved by two waters, two backbone carbonyls, and two side chain carboxylates. (D) Detailed view of the coordination of the cation at site 2, modelled as Mg²⁺ based on bond length and coordination number. Here, the coordination number of six is achieved by three waters and three side chain carboxylates. (E) Detailed view of the coordination of the cation at site 3, modelled as Mg²⁺. The coordination is achieved by three side chain carboxylates, one side chain amide, and an unmodelled density that is assumed to be a superposition of crystallisation condition compounds.

site 2 in the SALSA SRCR8 domain (26) and point mutations of this site abolished function. Common to the MARCO and SALSA SRCR domains is the cluster of negatively charged residues coordinating the functionally important cations. A similar cluster is also observed in the SRCR3 domain of CD6 and has been shown by mutagenesis to be directly involved in binding to the human surface receptor CD166. Indeed, a point mutation of D291A (corresponding to D1019 of SALSA-SRCR8) reduced the ligand-binding potential of CD6 to less than 10% (27). Furthermore, mutational studies of SRCR domains 2 and 3 from CD163 proved an involvement of this specific site in the binding of the haemoglobin-haptoglobin complex (36). All structural evidence from mutational studies of SRCR domains thus indicate a conserved surface-mediating ligand binding (Fig 4B).

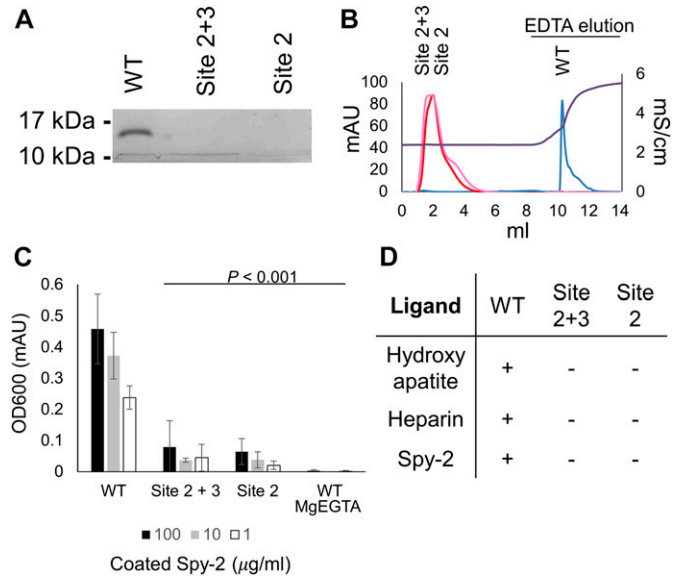


Figure 3. Mutating the cation-binding residues of scavenger receptor cysteine-rich (SRCR) domains abolishes function. Through multiple ligand-binding assays, we demonstrated the functional importance of cation binding by the SRCR domains. Mutations affecting site 2 (D1019A) and mutations affecting sites 2 and 3 (D1020A) both abolish function. (A) WT and mutant forms of SRCR8 were incubated with hydroxyapatite beads in a Ca²⁺-containing buffer. After extensive washing, bound protein was eluted with EDTA. Eluted fractions were run on a 4–20% SDS-PAGE gel and visualized by Coomassie staining. Only WT SRCR8 bound hydroxyapatite. (B) WT and mutant forms of SRCR8 were flown over a heparin (HiTrap HP, 1 ml) column in a Ca²⁺-containing buffer. Protein bound to the column was eluted with 0.5 M EDTA. Only WT SRCR8 bound the heparin column. Traces: SRCR8 (blue), D1020A (pink), D1019A (red), conductivity (brown). (C) In an ELISA-based setup, a concentration range of the Spy-2 domain of Spy0843 was coated (1–100 µg/ml). WT and mutant SRCR8 domains were added (100 µg/ml), and binding was detected with a monoclonal anti-SALSA antibody. Binding was only observed for WT SRCR8. (D) Overview of ligand-binding studies; + denotes binding, - denotes no binding.

Interestingly, various levels of calcium-dependency on ligand interactions have been described for all SRCR domains directly involved with binding. SR-A1, Spa, MARCO, CD5, and CD6 all rely on calcium for interactions with microbial ligands (24, 25, 26, 37, 38, 39). Furthermore, the binding of CD163 to the haemoglobin-haptoglobin complex is calcium-dependent, while CD6 also recognizes endogenous surface structures (other than CD166) in a calcium-dependent manner (40). This suggests that the cation-dependent binding mechanism identified for SALSA is a general conserved feature of all SRCR domains. Indeed, sequence alignment of SRCR domains from 10 different SRCR superfamily proteins, all with SRCR domains directly involved with ligand binding, reveals a very high level of conservation of the two cation-binding sites identified in the SALSA domains (Fig 4C). The ConSurf server is a tool to estimate (on a scale from 1 to 9) the level of evolutionary conservation of residues in a given fold (41). A search with the SRCR8 model shows that D1019, D1020, D1058, and E1086 all score 7 (highly conserved), while N1081 and D1059 score 6 and 4, respectively (thus less conserved). Whenever sequence identity is not conserved, substitutions are observed with other residues overrepresented in cation-binding sites (D, E, Q, and N) (28, 29). The cation-binding sites identified in the SALSA SRCR domains, thus, appear to be a highly conserved feature of the general SRCR fold.

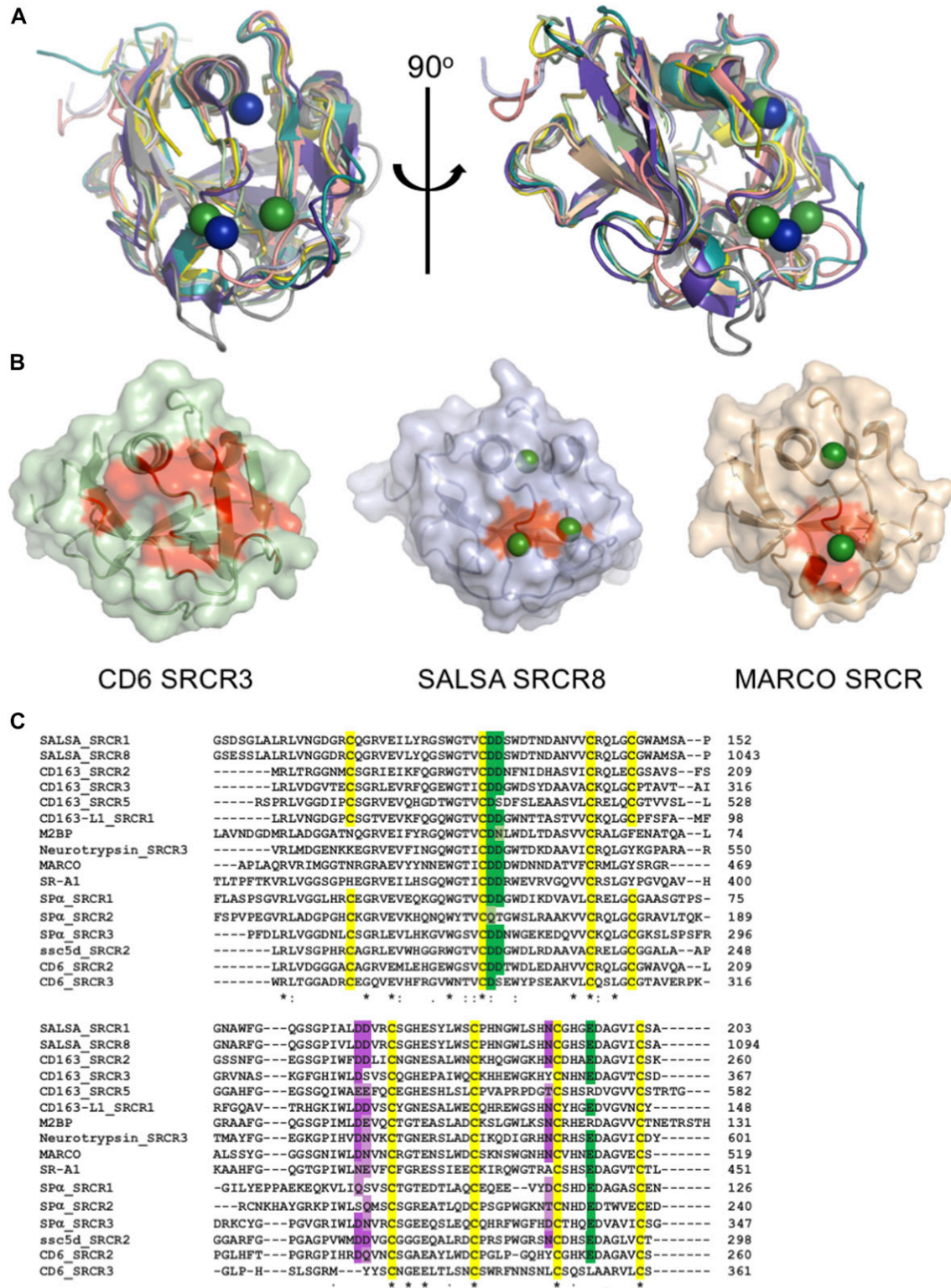


Figure 4. Conserved ligand-binding motif across scavenger receptor cysteine-rich (SRCR) domains.

SRCR domains from multiple proteins engage in cation-dependent ligand interactions. (A) Structural overlay of domains from seven SRCR superfamily proteins, all with ligand binding mediated through the SRCR domain. This reveals a highly conserved fold across both type A and type B SRCR domains. SRCR1 (pale green), SRCR8 (light blue), MARCO (pdbid: 2oy3, sand), CD163 (pdbid: 5jfb, purple), CD5 (pdbid: 2OTT, grey), CD6 (pdbid: 5a2e, pink), M2bp (pdbid: 1by2, yellow), and murine neurotrypsin (pdbid: 6h8m, teal) (35). SALSA magnesium: green, MARCO magnesium: blue. (B) Surface representation of CD6 SRCR3, SALSA SRCR8, and MARCO in same orientation. Point mutations with a verified impact on ligand binding are highlighted in red, indicating a conserved surface involved in ligand binding. Bound magnesium is highlighted in

Structural inspection of the overlay of SALSA SRCR8 with the corresponding region in the other known SRCR domain structures clearly shows the potential for cation binding at these sites. We thus propose that the cation-binding sites identified here are an essential feature of the ancient SRCR fold and are a conserved mechanism responsible for mediating ligand binding in the class of SRCR superfamily proteins.

Discussion

Although SALSA has previously been described to interact with a wide range of biological ligands, little has been known of the binding mechanisms. Furthermore, it has not been known if SALSA interacts with various ligands in a similar way or if distinct binding sites are used. Here, we demonstrate that mutations of a dual cation-binding site interrupt interactions with a representative selection of very different types of ligands. Specific disruption of site 2 was sufficient to abolish ligand binding. We modelled the cation at site 2 in our crystal as Mg^{2+} , based on an analysis of bond length, coordination number, and behaviour of crystallographic refinements with different cations modelled. However, experimental data demonstrated that binding to the ligands tested was only dependent on the presence of Ca^{2+} and not Mg^{2+} . This is in line with previous descriptions of most SALSA–ligand interactions (5, 6, 42). In the extracellular compartment, the molar concentration of Ca^{2+} is higher than Mg^{2+} (2.5 and 1 mM for calcium and magnesium, respectively), and it is therefore likely that both sites 2 and 3 will be occupied by Ca^{2+} in a physiological setting. The identification of this dual Ca^{2+} -binding site thus provides an explanation for the Ca^{2+} dependency of all SALSA–ligand interactions described in the literature, suggesting this mechanism of binding is applicable to all SALSA SRCR ligands. Multiple studies have proposed a role for the motif GRVEVxxxxW in ligand binding (43, 44, 45). The crystal data show that this peptide sequence is buried in the SRCR fold, and we found no validation for a role in ligand binding, although mutations within this sequence are likely to perturb the overall fold. This motif thus does not appear to have any physiological relevance as defining a ligand-binding site.

The conserved usage of a single ligand-binding area for multiple interactions suggests that each SALSA SRCR domain engages in one ligand interaction. A common feature of the ligands described here, as well as a number of other ligands such as DNA and LPS, is the presence of repetitive negatively charged motifs (31). We analysed ligand binding by individual SRCR domains in surface-plasmon resonance and isothermal calorimetry assays, but interactions were observed to be of very low affinity, making reliable measurements unfeasible. This is not surprising for a molecule such as SALSA, where the molecular makeup with the full extension of 13 repeated units, interspersed by predicted nonstructured flexible SIDs, provides a molecule that can generate high-avidity interactions with repetitive ligands, despite having only low-affinity interactions for an individual domain. Furthermore, it has been suggested that SALSA in body secretions may oligomerize into larger complexes (5, 46, 47, 48), probably via the C-terminal CUB and zona pellucida domains. The

repetitive nature and possible oligomerization allow SALSA to not only engage with a repetitive ligand on one surface (e.g., LPS or Spy0843 on microbes) but also engage in multiple ligand interactions simultaneously. This would be relevant for its interactions with other endogenous defence molecules, such as IgA, SPs, and complement components, where a cooperative effect on microbial clearance has been demonstrated (12, 16, 17, 49). In addition, this model of multiple ligand binding would be relevant for microbes described to use SALSA for colonisation of the teeth or the host epithelium (10, 50, 51) (Fig 5).

SALSA belongs to the SRCR superfamily, a family of proteins characterised by the presence of one or more copies of the ancient and evolutionarily highly conserved SRCR fold (52). Although a couple of SRCR domains, such as the ones found in complement factor I and hepsin, have not been described to bind ligands directly, most others have (53, 54). SRCR superfamily members, such as SALSA, SR-A1, Sp α , SSc5D, MARCO, CD6, and CD163, have broad scavenger-receptor functions, recognizing a broad range of microbial surface structures and mediate clearance (24, 25, 26, 37). Although this potentially is relevant for all SRCR superfamily proteins, some members of the family have distinct protein ligands, such as CD6, CD163, and M2bp (22, 23, 27, 55). With the exception of the CD6–CD166 interactions, most described SRCR–ligand interactions are calcium dependent, irrespective of the ligand (24, 25, 26, 37, 38, 39, 40). A cation-binding site is conserved across SRCR domains, and multiple studies support a role for this site in ligand binding. Even the specialised CD6–CD166 interaction uses the same surface for binding, despite “having lost” the calcium dependency (27).

Our studies have thus identified a dual cation-binding site as essential for SALSA–ligand interactions. Analysis of SRCR folds from various ligand-binding domains reveals a very high level of conservation of the residues at this dual site. The conservation of this site, along with the well-described cation dependency on most SRCR–ligand interactions, suggests that the binding mechanism described for the SALSA SRCR domains is applicable to all SRCR domains. We thus propose to have identified in SALSA a conserved functional mechanism for the SRCR class of proteins. This notion is further supported by the specific lack of conservation of these residues observed in the SRCR domains of complement factor I and hepsin, where no ligand binding has been shown. The SRCR domains in these two molecules may thus represent an evolutionary diversion from the common broad ligand-binding potential of the SRCR fold. The novel understanding of the SRCR domain generated here will allow for an interesting future targeting of other SRCR superfamily proteins, with the potential of modifying function.

Materials and Methods

Expression of recombinant proteins

Insect cell expression

Codon-optimized DNA (GeneArt; Thermo Fisher Scientific) was cloned into a modified pExpres2-2 vector (Expres2ion Biotechnologies) with a

green. (C) Clustal Omega (EMBL-EBI) sequence alignment of SRCR domains from 10 SRCR superfamily proteins. Conservation of the cation-binding sites are displayed in green (site 2) and purple (site 3). Dark colouring indicates 100% identity with the SALSA sites, and lighter colouring indicates conservation of residues commonly implicated in cation-binding (D, E, Q, or N). Cysteines are highlighted in yellow, and overall sequence identity is denoted by * (100%), : (strongly similar chemical properties), and . (weakly similar chemical properties).

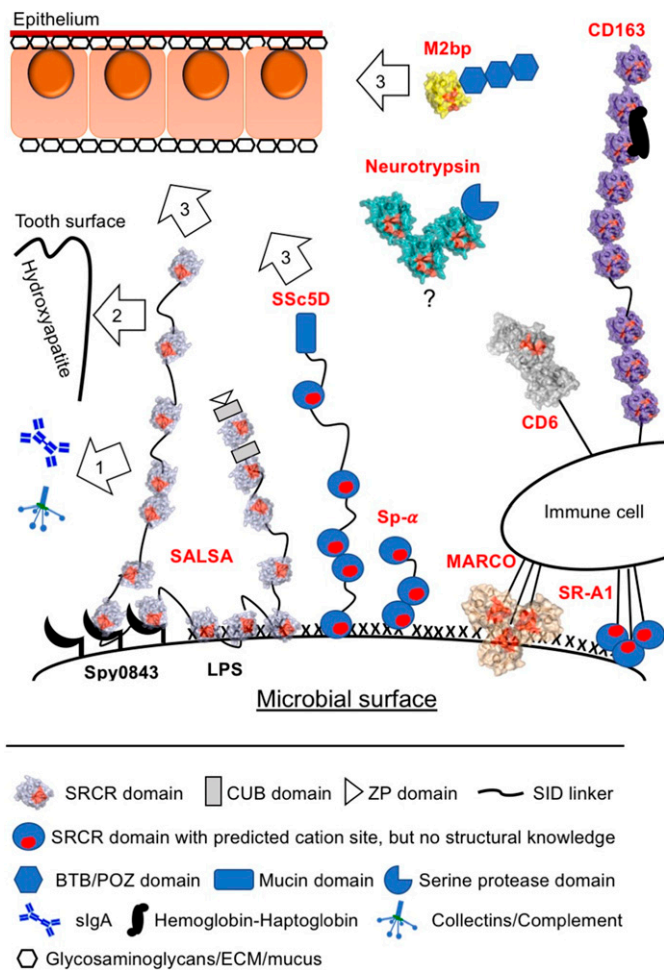


Figure 5. SALSAscavenger receptor cysteine-rich (SRCR) cation-binding motif reveals a conserved mechanism for broad-spectrum ligand interactions of SRCR superfamily molecules.

Based on mutational studies and structural information across SRCR proteins, we propose a generalised mechanism of ligand interaction mediated by the cation-binding surface motif of the evolutionarily ancient SRCR fold (left side). SALSAs has been described to bind a broad range of ligands, incorporating into a complex network of binding partners on the body surfaces and the colonizing microbiota. The multiple SRCR domains of full-length SALSAs bind repetitive targets (both protein and carbohydrate structures) on the surface of microbes. The secreted fluid-phase molecule may thus lead to microbial agglutination and clearance. However, the repetitive form of binding sites will allow for simultaneous binding to endogenous targets as well. This being, for example, 1) binding of IgA, collectins, and complement components to induce a cooperative antimicrobial effect; 2) binding of hydroxyapatite on the tooth surface; and 3) ECM proteins and glycosaminoglycans, as well as mucus components of the epithelial surface (such as heparin, galectin 3, and mucins) (right side). The cation-binding motif described in SALSAs is conserved in most other SRCR proteins. For CD6, CD163, and MARCO, mutational studies support a crucial role for this area in ligand interactions. CD163 binds the haemoglobin-haptoglobin complex and microbial surfaces. CD6 binds endogenous ligands but also engages in microbial binding. MARCO forms multimers and binds microbial surface structures. Other SRCR proteins with similar functions and conserved cation sites include SR-A1, Sp- α , SSc5D, and M2bp. The functional role of the neurotrypsin SRCR domains is not known. The remarkable repetitive formation of multiple SRCR domains in many SRCR superfamily proteins, with several domains containing a binding site with a broad specificity, would supposedly allow for interactions with multiple ligands simultaneously. The SRCR fold thus appears to be an important functional component of scavenging molecules engaging in complex network of interactions. The multiple SRCR domains shown for SALSAs, CD163, and

C-terminal His-6 tag. The purified plasmid was transformed into S2 cells grown in EX-CELL 420 (Sigma-Aldrich) with 25 μ l ExpreS2 Insect-TR 5X (ExpreS2ion Biotechnologies). Selection for stable cell lines (4 mg/ml geneticin [Thermo Fisher Scientific]) and expansion were carried out according to the manufacturer's instructions.

Escherichia coli expression

DNA strings (GeneArt; Thermo Fisher Scientific) were cloned into pETM-14 and transformed into M15pRep cells. Protein expression was carried out in LB media (with 30 μ g/ml kanamycin). Cells were induced with 1 mM IPTG. The cultures were centrifuged (3,220g, 15 min) and the cell pellets resuspended and lysed in PBS containing 1 mg/ml DNase and 1 mg/ml lysozyme.

Protein purification

SRPCR domains

Insect culture supernatant was collected by centrifugation (1,000g at 30 min), filtered and loaded onto a Roche cComplete Ni²⁺-chromatography column (1 ml, Cat. no. 06781543001; Sigma-Aldrich), washed in 20 CV buffer (50 mM Tris, pH: 9.0, 200 mM NaCl). Bound protein was eluted with 250 mM imidazole. Following this, SEC was carried out on a Superdex 75 16/60 HR column (GE Healthcare) equilibrated in 10 mM Tris, pH: 7.5, 200 mM NaCl.

Spy-2

Lysed cell pellets were homogenized and centrifuged at 20,000g for 30 min. The filtered supernatant was loaded onto a Ni²⁺-chromatography column (5 ml; QIAGEN) and washed in 20 CV buffer (50 mM Tris, pH: 8.5, 200 mM NaCl, 20 mM imidazole). Bound protein was eluted (in 50 mM Tris, pH: 8.5, 200 mM NaCl, 250 mM imidazole), concentrated, and subjected to SEC (Superdex 75 16/60 HR column; GE Healthcare).

Crystallisation, X-ray data collection, and structure determination

Purified SRCR1 and SRCR8 were concentrated to 20 mg/ml. SRCR1 was mixed with an equal volume of mother liquor containing 0.2 M MgCl₂ hexahydrate, 10% (wt/vol) PEG8000, 0.1 M Tris, pH: 7.0, and crystallised in 400 nl drops by the vapor diffusion method at 21°C. SRCR8 was mixed with an equal volume of mother liquor containing 0.1 M LiSO₄, 20% (wt/vol) PEG6000, 0.01 M HEPES, pH: 6.5, and crystallised in 800 nl drops. For SRCR8 + cation crystals were grown in 0.2 M MgCl₂ hexahydrate, 20% (vol/vol) isopropanol, 0.1 M HEPES, pH: 7.5, and crystallised in 400 nl drops. The crystallisation buffer was supplemented with 10 mM Mg²⁺ and 10 mM Ca²⁺, as well as 10 mM maltose, D-galactose, D-saccharose, D-mannose, D-glucose, and sucrose octasulphate (all Sigma-Aldrich), 24 h prior to freezing. All crystals were cryoprotected in mother liquor supplemented with 30% glycerol and flash frozen in liquid N₂. Data were collected at a temperature of 80 K

neurotrypsin are represented as copies of protein-specific SRCR domains with known structure. Conserved cation-coordinating residues are highlighted in red. SRCR8soak (light blue), MARCO (pdbid: 2oy3, sand), CD163 (pdbid: 5jfb, purple), CD6 (pdbid: 5a2e, grey), M2bp (pdbid: 1by2, yellow), and murine neurotrypsin (pdbid: 6h8m, teal).

on beamlines I04, at a wavelength of 1.0718 Å (for SRCR1 and SRCR8) and I03, at a wavelength of 0.9762 Å (for SRCR8cat) at the Diamond Light Source, as specified in Table 1. The structure of SRCR8 was solved by molecular replacement using MolRep within CCP4 (56) with the structure of CD6 SRCR domain 3 (PDB ID 5a2e (27)). The structures of SRCR1 and SRCR8 soaked in cations were solved by molecular replacement using the structure of SRCR8. Refinement and re-building were carried out in Phenix and Coot (57, 58). Assignment of metal ions was carried out by first refining the structure without anything in the metal binding sites, followed by addition of combinations of probable ligands, and re-refinement in phenix.refine using restraints generated by phenix.ready_set for each combination. The structures were characterised by the statistics shown in Table 1 with no Ramachandran outliers. Protein structure figures were prepared using Pymol version 2.0 (Schrödinger, LLC).

Hydroxyapatite binding assay

150 µl hydroxyapatite nanoparticle suspension (Cat. no. 702153; Sigma-Aldrich) was washed into buffer (10 mM Hepes, pH: 7.5, 150 mM NaCl, 1 mM Ca²⁺). Beads were incubated in 80 µl SRCR8, SRCR8 D34A, or SRCR8 D35A (all at 0.5 mg/ml in the same buffer) with shaking for 1 h at RT. Beads were spun and washed 6× in 1 ml buffer. Bound protein was eluted in 100 µl 0.5 M EDTA and visualized by SDS–PAGE (4–20%; Bio-Rad) and Coomassie staining (Instant Blue, Expedeon).

Heparin binding assay

SRCR8, SRCR8 D34A, or SRCR8 D35A in 10 mM Hepes, pH: 7.5, 10 mM NaCl, 1 mM Ca²⁺ were loaded onto a HiTrap Heparin HP column (1 ml; GE Healthcare), equilibrated in the same buffer. Bound protein was then eluted with 10 mM Hepes, pH: 7.5, 10 mM NaCl, 20 mM EDTA.

Spy-2 binding assay

On a MaxiSorp plate (Nunc), 100 µl purified Spy-2 was coated O/N at 4°C in a concentration ranging from 0.032 to 3.2 µM in coating buffer (100 mM NaHCO₃ buffer, pH: 9.5). The plate was blocked in 1% gelatine in PBS, and SRCR8, SRCR8 D34A, and SRCR8 D35A were added (all at 7.1 µM in 10 mM Hepes, pH 7.5, 150 mM NaCl, 1 mM Ca²⁺, 0.05% Tween20). Bound protein was detected with monoclonal anti-SALSA antibody diluted 1:10,000 (1G4; Novus Biologicals) and HRP-conjugated rabbit anti-mouse antibody 1:10,000 (W4028; Promega). The plate was developed with 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (Sigma-Aldrich) and analysed by spectrophotometry at 405 nm. To test calcium-specific dependency of the interaction, the WT assay above was repeated in a buffer containing 10 mM Hepes, pH 7.5, 150 mM NaCl, 1 mM Mg²⁺, 1 mM EGTA, and 0.05% Tween20.

Data availability

Structure factors and coordinates from this publication have been deposited to the PDB database <https://www.wwpdb.org> and assigned the identifiers: SRCR1 pdbid: 6sa4, SRCR8 pdbid: 6sa5, SRCR8soak with three cations pdbid: 6san.

Supplementary Information

Supplementary Information is available at <https://doi.org/10.26508/lsa.201900502>.

Acknowledgements

We acknowledge Diamond Light Source and the staff of beamlines I03 and I04 for access under proposal MX18069. The Central Oxford Structural Molecular and Imaging Centre is supported by the Wellcome Trust (201536). MP Reichhardt was financially supported by grants from the Wihuri Foundation and the Finnish Cultural Foundation. Staff and experimental costs in SM Lea laboratory were supported by a Wellcome Investigator Award (100298) and an Medical Research Council (UK) programme grant (M011984). V Loimaranta was supported by the Turku University Foundation.

Author Contributions

MP Reichhardt: conceptualization, data curation, formal analysis, funding acquisition, investigation, and writing—original draft, review, and editing.

V Loimaranta: resources.

SM Lea: conceptualization, data curation, formal analysis, funding acquisition, validation, investigation, methodology, project administration, and writing—review and editing.

S Johnson: conceptualization, data curation, formal analysis, supervision, validation, investigation, visualization, methodology, project administration, and writing—original draft, review, and editing.

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

References

- Holmskov U, Lawson P, Teisner B, Tornøe J, Willis AC, Morgan C, Koch C, Reid KB (1997) Isolation and characterization of a new member of the scavenger receptor superfamily, glycoprotein-340 (gp-340), as a lung surfactant protein-D binding molecule. *J Biol Chem* 272: 13743–13749. doi:10.1074/jbc.272.21.13743
- Thornton DJ, Davies JR, Kirkham S, Gautrey A, Khan N, Richardson PS, Sheehan JK (2001) Identification of a nonmucin glycoprotein (gp-340) from a purified respiratory mucin preparation: Evidence for an association involving the MUC5B mucin. *Glycobiology* 11: 969–977. doi:10.1093/glycob/11.11.969
- Schulz BL, Oxley D, Packer NH, Karlsson NG (2002) Identification of two highly sialylated human tear-fluid DMBT1 isoforms: The major high-molecular-mass glycoproteins in human tears. *Biochem J* 366: 511–520. doi:10.1042/bj20011876
- Reichhardt MP, Jarva H, de Been M, Rodriguez JM, Jimenez Quintana E, Loimaranta V, de Vos WM, Meri S (2014) The salivary scavenger and agglutinin in early life: Diverse roles in amniotic fluid and in the infant intestine. *J Immunol* 193: 5240–5248. doi:10.4049/jimmunol.1401631
- Madsen J, Mollenhauer J, Holmskov U (2010) Review: Gp-340/DMBT1 in mucosal innate immunity. *Innate Immun* 16: 160–167. doi:10.1177/1753425910368447

6. Reichhardt MP, Meri S (2016) SALSA: A regulator of the early steps of complement activation on mucosal surfaces. *Front Immunol* 7: 85. doi:[10.3389/fimmu.2016.00085](https://doi.org/10.3389/fimmu.2016.00085)
7. Reichhardt MP, Holmskov U, Meri S (2017) SALSA-A dance on a slippery floor with changing partners. *Mol Immunol* 89: 100–110. doi:[10.1016/j.molimm.2017.05.029](https://doi.org/10.1016/j.molimm.2017.05.029)
8. Prakobphol A, Xu F, Hoang VM, Larsson T, Bergstrom J, Johansson I, Frangmyr L, Holmskov U, Leffler H, Nilsson C, et al (2000) Salivary agglutinin, which binds *Streptococcus mutans* and *Helicobacter pylori*, is the lung scavenger receptor cysteine-rich protein gp-340. *J Biol Chem* 275: 39860–39866. doi:[10.1074/jbc.M006928200](https://doi.org/10.1074/jbc.M006928200)
9. Hartshorn KL, White MR, Mogues T, Ligtenberg T, Crouch E, Holmskov U (2003) Lung and salivary scavenger receptor glycoprotein-340 contribute to the host defense against influenza A viruses. *Am J Physiol Lung Cell Mol Physiol* 285: L1066–L1076. doi:[10.1152/ajplung.00057.2003](https://doi.org/10.1152/ajplung.00057.2003)
10. Loimaranta V, Jakubovics NS, Hytonen J, Finne J, Jenkinson HF, Stromberg N (2005) Fluid- or surface-phase human salivary scavenger protein gp340 exposes different bacterial recognition properties. *Infect Immun* 73: 2245–2252. doi:[10.1128/iai.73.4.2245-2252.2005](https://doi.org/10.1128/iai.73.4.2245-2252.2005)
11. Rosenstiel P, Sina C, End C, Renner M, Lyer S, Till A, Hellmig S, Nikolaus S, Folsch UR, Helmke B, et al (2007) Regulation of DMBT1 via NOD2 and TLR4 in intestinal epithelial cells modulates bacterial recognition and invasion. *J Immunol* 178: 8203–8211. doi:[10.4049/jimmunol.178.12.8203](https://doi.org/10.4049/jimmunol.178.12.8203)
12. Rundegren J, Arnold RR (1987) Differentiation and interaction of secretory immunoglobulin A and a calcium-dependent parotid agglutinin for several bacterial strains. *Infect Immun* 55: 288–292. doi:[10.1128/iai.55.2.288-292.1987](https://doi.org/10.1128/iai.55.2.288-292.1987)
13. Boackle RJ, Connor MH, Vesely J (1993) High molecular weight non-immunoglobulin salivary agglutinins (NIA) bind C1Q globular heads and have the potential to activate the first complement component. *Mol Immunol* 30: 309–319. doi:[10.1016/0161-5890\(93\)90059-k](https://doi.org/10.1016/0161-5890(93)90059-k)
14. Tino MJ, Wright JR (1999) Glycoprotein-340 binds surfactant protein-A (SP-A) and stimulates alveolar macrophage migration in an SP-A-independent manner. *Am J Respir Cell Mol Biol* 20: 759–768. doi:[10.1165/ajrcmb.20.4.3439](https://doi.org/10.1165/ajrcmb.20.4.3439)
15. Oho T, Bikker FJ, Nieuw Amerongen AV, Groenink J (2004) A peptide domain of bovine milk lactoferrin inhibits the interaction between streptococcal surface protein antigen and a salivary agglutinin peptide domain. *Infect Immun* 72: 6181–6184. doi:[10.1128/iai.72.10.6181-6184.2004](https://doi.org/10.1128/iai.72.10.6181-6184.2004)
16. Leito JT, Ligtenberg AJ, van Houdt M, van den Berg TK, Wouters D (2011) The bacteria binding glycoprotein salivary agglutinin (SAG/gp340) activates complement via the lectin pathway. *Mol Immunol* 49: 185–190. doi:[10.1016/j.molimm.2011.08.010](https://doi.org/10.1016/j.molimm.2011.08.010)
17. Reichhardt MP, Loimaranta V, Thiel S, Finne J, Meri S, Jarva H (2012) The salivary scavenger and agglutinin binds MBL and regulates the lectin pathway of complement in solution and on surfaces. *Front Immunol* 3: 205. doi:[10.3389/fimmu.2012.00205](https://doi.org/10.3389/fimmu.2012.00205)
18. Madsen J, Sorensen GL, Nielsen O, Tornoe I, Thim L, Fenger C, Mollenhauer J, Holmskov U (2013) A variant form of the human deleted in malignant brain tumor 1 (DMBT1) gene shows increased expression in inflammatory bowel diseases and interacts with dimeric trefoil factor 3 (TFF3). *PLoS One* 8: e64441. doi:[10.1371/journal.pone.0064441](https://doi.org/10.1371/journal.pone.0064441)
19. Holmskov U, Mollenhauer J, Madsen J, Vitved L, Gronlund J, Tornoe I, Kliem A, Reid KB, Poustka A, Skjodt K (1999) Cloning of gp-340, a putative opsonin receptor for lung surfactant protein D. *Proc Natl Acad Sci U S A* 96: 10794–10799. doi:[10.1073/pnas.96.19.10794](https://doi.org/10.1073/pnas.96.19.10794)
20. Mollenhauer J, Holmskov U, Wiemann S, Krebs I, Herberich S, Madsen J, Kioschis P, Coy JF, Poustka A (1999) The genomic structure of the DMBT1 gene: Evidence for a region with susceptibility to genomic instability. *Oncogene* 18: 6233–6240. doi:[10.1038/sj.onc.1203071](https://doi.org/10.1038/sj.onc.1203071)
21. Mollenhauer J, Wiemann S, Scheurlen W, Korn B, Hayashi Y, Wilgenbus KK, von Deimling A, Poustka A (1997) DMBT1, a new member of the SRCR superfamily, on chromosome 10q25.3-26.1 is deleted in malignant brain tumours. *Nat Genet* 17: 32–39. doi:[10.1038/ng0997-32](https://doi.org/10.1038/ng0997-32)
22. Inohara H, Akahani S, Kohts K, Raz A (1996) Interactions between galectin-3 and Mac-2-binding protein mediate cell-cell adhesion. *Cancer Res* 56: 4530–4534. <https://cancerres.aacrjournals.org/content/56/19/4530>
23. Kristiansen M, Graversen JH, Jacobsen C, Sonne O, Hoffman HJ, Law SK, Moestrup SK (2001) Identification of the haemoglobin scavenger receptor. *Nature* 409: 198–201. doi:[10.1038/35051594](https://doi.org/10.1038/35051594)
24. Santiago-Garcia J, Kodama T, Pitas RE (2003) The class A scavenger receptor binds to proteoglycans and mediates adhesion of macrophages to the extracellular matrix. *J Biol Chem* 278: 6942–6946. doi:[10.1074/jbc.M208358200](https://doi.org/10.1074/jbc.M208358200)
25. Sarrias MR, Rosello S, Sanchez-Barbero F, Sierra JM, Vila J, Yelamos J, Vives J, Casals C, Lozano F (2005) A role for human Sp alpha as a pattern recognition receptor. *J Biol Chem* 280: 35391–35398. doi:[10.1074/jbc.M505042200](https://doi.org/10.1074/jbc.M505042200)
26. Ojala JR, Pikkarainen T, Tuuttila A, Sandalova T, Tryggvason K (2007) Crystal structure of the cysteine-rich domain of scavenger receptor MARCO reveals the presence of a basic and an acidic cluster that both contribute to ligand recognition. *J Biol Chem* 282: 16654–16666. doi:[10.1074/jbc.M701750200](https://doi.org/10.1074/jbc.M701750200)
27. Chappell PE, Garner LI, Yan J, Metcalfe C, Hatherley D, Johnson S, Robinson CV, Lea SM, Brown MH (2015) Structures of CD6 and its ligand CD166 give insight into their interaction. *Structure* 23: 1426–1436. doi:[10.1016/j.str.2015.05.019](https://doi.org/10.1016/j.str.2015.05.019)
28. Pidcock E, Moore GR (2001) Structural characteristics of protein binding sites for calcium and lanthanide ions. *J Biol Inorg Chem* 6: 479–489. doi:[10.1007/s007750100214](https://doi.org/10.1007/s007750100214)
29. Dokmanic I, Sivic M, Tomic S (2008) Metals in proteins: Correlation between the metal-ion type, coordination number and the amino-acid residues involved in the coordination. *Acta Crystallogr D Biol Crystallogr* 64: 257–263. doi:[10.1107/S090744490706595X](https://doi.org/10.1107/S090744490706595X)
30. Haukioja A, Loimaranta V, Tenovuo J (2008) Probiotic bacteria affect the composition of salivary pellicle and streptococcal adhesion in vitro. *Oral Microbiol Immunol* 23: 336–343. doi:[10.1111/j.1399-302x.2008.00435.x](https://doi.org/10.1111/j.1399-302x.2008.00435.x)
31. End C, Bikker F, Renner M, Bergmann G, Lyer S, Blaich S, Hudler M, Helmke B, Gassler N, Autschbach F, et al (2009) DMBT1 functions as pattern-recognition molecule for poly-sulfated and poly-phosphorylated ligands. *Eur J Immunol* 39: 833–842. doi:[10.1002/eji.200838689](https://doi.org/10.1002/eji.200838689)
32. Loimaranta V, Hytonen J, Pulliainen AT, Sharma A, Tenovuo J, Stromberg N, Finne J (2009) Leucine-rich repeats of bacterial surface proteins serve as common pattern recognition motifs of human scavenger receptor gp340. *J Biol Chem* 284: 18614–18623. doi:[10.1074/jbc.M900581200](https://doi.org/10.1074/jbc.M900581200)
33. Jaroszewski L, Rychlewski L, Li Z, Li W, Godzik A (2005) FFAS03: A server for profile–profile sequence alignments. *Nucleic Acids Res* 33: W284–W288. doi:[10.1093/nar/gki418](https://doi.org/10.1093/nar/gki418)
34. Holm L, Laakso LM (2016) Dali server update. *Nucleic Acids Res* 44: W351–W355. doi:[10.1093/nar/gkw357](https://doi.org/10.1093/nar/gkw357)
35. Canciani A, Catucci G, Forneris F (2019) Structural characterization of the third scavenger receptor cysteine-rich domain of murine neurotrophin. *Protein Sci* 28: 746–755. doi:[10.1002/pro.3587](https://doi.org/10.1002/pro.3587)
36. Nielsen MJ, Andersen CB, Moestrup SK (2013) CD163 binding to haptoglobin-hemoglobin complexes involves a dual-point electrostatic receptor–ligand pairing. *J Biol Chem* 288: 18834–18841. doi:[10.1074/jbc.M113.471060](https://doi.org/10.1074/jbc.M113.471060)
37. Sarrias MR, Farnos M, Mota R, Sanchez-Barbero F, Ibanez A, Gimferrer I, Vera J, Fenutria R, Casals C, Yelamos J, et al (2007) CD6 binds to pathogen-associated molecular patterns and protects from LPS-induced septic shock. *Proc Natl Acad Sci U S A* 104: 11724–11729. doi:[10.1073/pnas.0702815104](https://doi.org/10.1073/pnas.0702815104)
38. Vera J, Fenutria R, Canadas O, Figueras M, Mota R, Sarrias MR, Williams DL, Casals C, Yelamos J, Lozano F (2009) The CD5 ectodomain interacts with

- conserved fungal cell wall components and protects from zymosan-induced septic shock-like syndrome. *Proc Natl Acad Sci U S A* 106: 1506–1511. doi:[10.1073/pnas.0805846106](https://doi.org/10.1073/pnas.0805846106)
39. Nonaka M, Ma BY, Imaeda H, Kawabe K, Kawasaki N, Hodohara K, Kawasaki N, Andoh A, Fujiyama Y, Kawasaki T (2011) Dendritic cell-specific intercellular adhesion molecule 3-grabbing non-integrin (DC-SIGN) recognizes a novel ligand, Mac-2-binding protein, characteristically expressed on human colorectal carcinomas. *J Biol Chem* 286: 22403–22413. doi:[10.1074/jbc.m110.215301](https://doi.org/10.1074/jbc.m110.215301)
 40. Patel DD, Wee SF, Whichard LP, Bowen MA, Pesando JM, Aruffo A, Haynes BF (1995) Identification and characterization of a 100-kD ligand for CD6 on human thymic epithelial cells. *J Exp Med* 181: 1563–1568. doi:[10.1084/jem.181.4.1563](https://doi.org/10.1084/jem.181.4.1563)
 41. Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, Ben-Tal N (2016) ConSurf 2016: An improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 44: W344–W350. doi:[10.1093/nar/gkw408](https://doi.org/10.1093/nar/gkw408)
 42. Ligtenberg AJ, Karlsson NG, Veerman EC (2010) Deleted in malignant brain tumors-1 protein (DMBT1): A pattern recognition receptor with multiple binding sites. *Int J Mol Sci* 11: 5212–5233. doi:[10.3390/ijms1112521](https://doi.org/10.3390/ijms1112521)
 43. Bikker FJ, Ligtenberg AJ, Nazmi K, Veerman EC, van't Hof W, Bolscher JG, Poustka A, Nieuw Amerongen AV, Mollenhauer J (2002) Identification of the bacteria-binding peptide domain on salivary agglutinin (gp-340/DMBT1), a member of the scavenger receptor cysteine-rich superfamily. *J Biol Chem* 277: 32109–32115. doi:[10.1074/jbc.m203788200](https://doi.org/10.1074/jbc.m203788200)
 44. Bikker FJ, Ligtenberg AJ, End C, Renner M, Blaich S, Lyer S, Wittig R, van't Hof W, Veerman EC, Nazmi K, et al (2004) Bacteria binding by DMBT1/SAG/gp-340 is confined to the VEVLXXXXW motif in its scavenger receptor cysteine-rich domains. *J Biol Chem* 279: 47699–47703. doi:[10.1074/jbc.m406095200](https://doi.org/10.1074/jbc.m406095200)
 45. Leito JT, Ligtenberg AJ, Nazmi K, de Blicke-Hogervorst JM, Veerman EC, Nieuw Amerongen AV (2008) A common binding motif for various bacteria of the bacteria-binding peptide SRCRP2 of DMBT1/gp-340/salivary agglutinin. *Biol Chem* 389: 1193–1200. doi:[10.1515/bc.2008.135](https://doi.org/10.1515/bc.2008.135)
 46. Young A, Rykke M, Smistad G, Rolla G (1997) On the role of human salivary micelle-like globules in bacterial agglutination. *Eur J Oral Sci* 105: 485–494. doi:[10.1111/j.1600-0722.1997.tb00235.x](https://doi.org/10.1111/j.1600-0722.1997.tb00235.x)
 47. Oho T, Yu H, Yamashita Y, Koga T (1998) Binding of salivary glycoprotein-secretory immunoglobulin A complex to the surface protein antigen of *Streptococcus mutans*. *Infect Immun* 66: 115–121. doi:[10.1128/iai.66.1.115-121.1998](https://doi.org/10.1128/iai.66.1.115-121.1998)
 48. Crouch EC (2000) Surfactant protein-D and pulmonary host defense. *Respir Res* 1: 93–108. doi:[10.1186/rr19](https://doi.org/10.1186/rr19)
 49. White MR, Crouch E, van Eijk M, Hartshorn M, Pemberton L, Tornoe I, Holmskov U, Hartshorn KL (2005) Cooperative anti-influenza activities of respiratory innate immune proteins and neuraminidase inhibitor. *Am J Physiol Lung Cell Mol Physiol* 288: L831–L840. doi:[10.1152/ajplung.00365.2004](https://doi.org/10.1152/ajplung.00365.2004)
 50. Jonasson A, Eriksson C, Jenkinson HF, Kallestal C, Johansson I, Stromberg N (2007) Innate immunity glycoprotein gp-340 variants may modulate human susceptibility to dental caries. *BMC Infect Dis* 7: 57. doi:[10.1186/1471-2334-7-57](https://doi.org/10.1186/1471-2334-7-57)
 51. Stoddard E, Cannon G, Ni H, Kariko K, Capodici J, Malamud D, Weissman D (2007) gp340 expressed on human genital epithelia binds HIV-1 envelope protein and facilitates viral transmission. *J Immunol* 179: 3126–3132. doi:[10.4049/jimmunol.179.5.3126](https://doi.org/10.4049/jimmunol.179.5.3126)
 52. Sarrias MR, Gronlund J, Padilla O, Madsen J, Holmskov U, Lozano F (2004) The scavenger receptor cysteine-rich (SRCR) domain: An ancient and highly conserved protein module of the innate immune system. *Crit Rev Immunol* 24: 1–37. doi:[10.1615/critrevimmunol.v24.i1.10](https://doi.org/10.1615/critrevimmunol.v24.i1.10)
 53. Roversi P, Johnson S, Caesar JJ, McLean F, Leath KJ, Tsiftoglou SA, Morgan BP, Harris CL, Sim RB, Lea SM (2011) Structural basis for complement factor I control and its disease-associated sequence polymorphisms. *Proc Natl Acad Sci U S A* 108: 12839–12844. doi:[10.1073/pnas.1102167108](https://doi.org/10.1073/pnas.1102167108)
 54. Koschubs T, Dengl S, Durr H, Kaluza K, Georges G, Hartl C, Jennewein S, Lanzendorfer M, Auer J, Stern A, et al (2012) Allosteric antibody inhibition of human hepsin protease. *Biochem J* 442: 483–494. doi:[10.1042/bj20111317](https://doi.org/10.1042/bj20111317)
 55. Burgueno-Bucio E, Mier-Aguilar CA, Soldevila G (2019) The multiple faces of CD5. *J Leukoc Biol* 105: 891–904. doi:[10.1002/JLB.MR0618-226R](https://doi.org/10.1002/JLB.MR0618-226R)
 56. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AG, McCoy A, et al (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67: 235–242. doi:[10.1107/s09074449100045749](https://doi.org/10.1107/s09074449100045749)
 57. Emsley P, Lohkamp B, Scott WG, Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66: 486–501. doi:[10.1107/s0907444910007493](https://doi.org/10.1107/s0907444910007493)
 58. Adams PD, Baker D, Brunger AT, Das R, DiMaio F, Read RJ, Richardson DC, Richardson JS, Terwilliger TC (2013) Advances, interactions, and future developments in the CNS, Phenix, and Rosetta structural biology software systems. *Annu Rev Biophys* 42: 265–287. doi:[10.1146/annurev-biophys-083012-130253](https://doi.org/10.1146/annurev-biophys-083012-130253)



License: This article is available under a Creative Commons License (Attribution 4.0 International, as described at <https://creativecommons.org/licenses/by/4.0/>).