The Structures of Secretory IgA in complex with *Streptococcus pyogenes* M4 and human CD89 provide insights on mucosal host-pathogen interactions

Qianqiao Liu[1] and Beth M. Stadtmueller[1, 2*]

[1]Department of Biochemistry, University of Illinois Urbana-Champaign, Urbana, Illinois 61801 USA

[2]Department of Biomedical and Translational Sciences, Carle Illinois College of Medicine, University of Illinois Urbana-Champaign, Urbana, Illinois 61801 USA

ABSTRACT

Immunoglobulin (Ig) A functions as monomeric IgA in the serum and Secretory (S) IgA in mucosal secretions. Host IgA Fc receptors (FcαRs), including human FcαR1/CD89, mediate IgA effector functions; however human pathogen *Streptococcus pyogenes* has evolved surface-protein virulence factors, including M4, that also engage the CD89 binding site on IgA. Despite human mucosa serving as a reservoir for pathogens, SIgA interactions with CD89 and M4 remain poorly understood. Here we report cryo-EM structures of M4-SIgA and CD89-SIgA complexes, which unexpectedly reveal different SIgA-binding stoichiometry for M4 and CD89. Structural data, supporting experiments, and modeling indicate that copies of SIgA bound to *S. pyogenes* M4 will adopt similar orientations on the bacterium surface and leave one host FcαR binding site open. Results suggest unappreciated functional consequences associated with SIgA binding to host and bacterial FcαRs relevant to understanding host-microbe co-evolution, IgA effector functions and to improving the outcomes of group A *Streptococcus* infection.

24 **INTRODUCTION**

25 Immunoglobulin (Ig) A is an antibody class expressed in mammals, where it functions in
26 monomeric (m) and secretory (S) forms; the mIgA is the predominant form in serum whereas
27 SIgA is the most abundant antibody present in mucosal secretions, where it protects the host
28 epithelium against microbes and toxins[1]. Both forms of IgA contain fragment antigen-binding (Fab)
29 and fragment crystallizable (Fc) regions. The Fab is responsible for binding to antigens, whereas
30 the Fc region interacts with other components of the immune system, including IgA Fc Receptors
31 (FcαRs) such as CD89 (a.k.a. FcαR1), that are expressed on the surface of some immune system
32 cells and can promote downstream effector functions such as phagocytosis and respiratory burst[2].

33 CD89-initiated effector functions are thought to occur when multiple copies of the receptor
34 bind to multiple IgA on IgA-antigen complexes, promoting clustering and cross-linking of the
35 receptors and associated cell signaling through FcR γ-chain complexes[3]. In turn, binding to free
36 IgA does not promote crosslinking has been proposed to transduce anti-inflammatory signals
37 instead[4]. CD89 effector functions have been studied primarily in the context of mIgA, consistent
38 with the observation that cells expressing CD89 are more prevalent in the serum than mucosa.
39 However, phagocyte populations also sample and/or patrol mucosal surfaces and SIgA has been
40 associated with CD89-expressing polymorphonuclear leukocytes (PMN) activity in the presence
41 of accessory molecule Mac-1 (CD11b/CD18)[5]. Whereas the binding of host FcαRs is likely to elicit
42 an effector immune response, IgA can also be bound by bacterial proteins, for example, M4 (a.k.a.
43 Arp4) and M22 (a.k.a Sir22) from *Streptococcus pyogenes,* the β-protein from group B
44 streptococci (GBS), and SSL7 from *Staphylococcus aureus*, interactions that may allow
45 pathogenic bacteria to evade host IgA responses[6–10]. Despite significance, the functional
46 consequences of IgA interactions with host FcαRs and bacterial proteins, sometimes referred to
47 as decoys, remains poorly understood.

48 The monomeric and secretory forms of IgA that mediate host interactions with antigens,
49 FcαRs, and bacterial proteins adopt at least two distinct molecular structures. The mIgA contains
50 two IgA heavy chains and two light chains, resulting in two Fabs and one Fcα. The Fcα contains
51 two copies of immunoglobulin constant domains, Cα2 and Cα3, and adopts a rigid structure linked
52 to two Fabs. Fcα contains two FcαR binding sites, one located at each Cα2 and Cα3 interface
53 (Fig. 1a). The Cα2/Cα3 interdomain has been recognized as a "hot spot" for host FcαRs and
54 bacterial IgA-binding proteins[6,11,12]. SIgA is polymeric and is typically composed of two mIgA that
55 assemble with one joining-chain (JC) to form dimeric (d) IgA, which is bound by the polymeric Ig-
56 receptor (pIgR) ectodomain, called secretory component (SC). The structures of dIgA and SIgA,
57 reported in 2020, revealed two Fcα connected through four C-terminal motifs called tailpieces
58 (Tps) and one JC, which contacts each heavy chain uniquely [13–15]. This results in two structurally
59 unique Fcαs that can be designated as Fc_{AB} (Fcα containing heavy chains A and B) and Fc_{CD}
60 (Fcα containing heavy chains C and D). Together, Fc_{AB} and Fc_{CD} adopt an asymmetric, bent and
61 tilted conformation with distinct concave and convex sides. In SIgA, the SC, which comprises five
62 Ig-like domains (D1-D5), is bound toward one side and protrudes from the center of the molecule
63 contributing further asymmetry. Overall, the SIgA conformation is predicted to constrain the
64 possible positions that Fabs can adopt and leave two FcαR binding sites accessible (Fig. 1b).
65 Notably, these two sites are located on the convex side, one on Fc_{AB} and one on $Fc_{CD,}$ whereas
66 sites on the concave side are occluded by the JC and the SC. Thus, SIgA is thought to have the
67 same number of accessible FcαR binding sites as mIgA but instead of two sites on one Fcα that
68 are horizontally oriented relative to each other, it has one site on two Fcαs that are vertically
69 oriented relative to each other (Fig. 1) [13]. The functional significance of these differences remains
70 a topic of investigation; however, it is notable that two receptor binding sites have been
71 evolutionary preserved in both mIgA and SIgA.
72

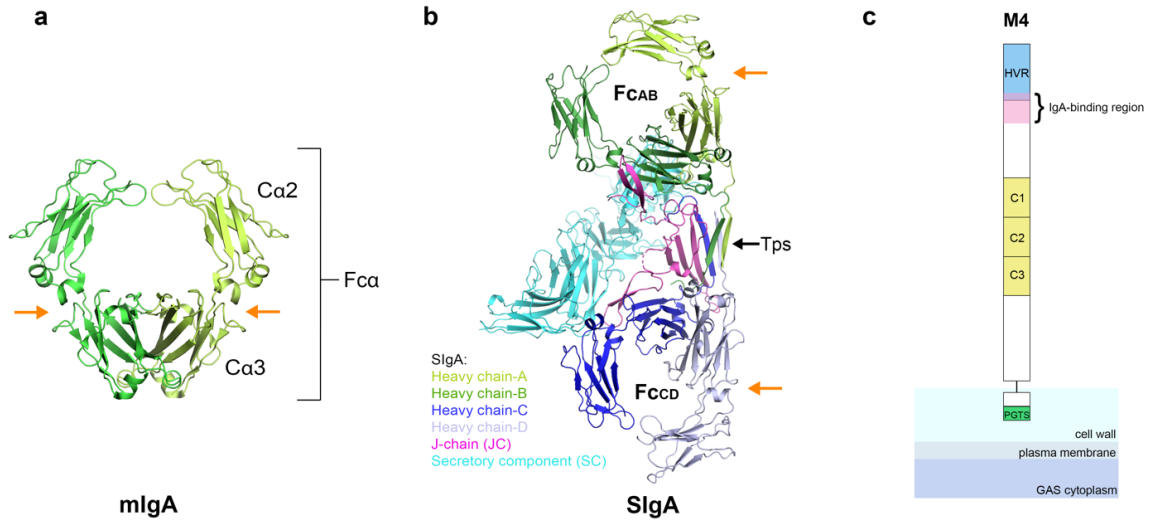


Figure 1. Structures of mIgA, SIgA, and schematic of M4. (a) The structure of mIgA lacking Fabs (Fcα; PDB code 1OW0 with CD89 removed). (b) The structure of human SIgA lacking Fabs (PDB code 6UE7). Structures are colored according to the key; one Cα2 and one Cα3 domain are labeled along with Fc_{AB} and Fc_{CD}. The SIgA tailpieces (Tps) are labeled and indicated with a black arrow and the FcαR binding sites are indicated by orange arrows. (c) Schematic of M4 depicted attached to the GAS cell wall through a C-terminal proline-glycine-threonine-serine (PGTS)-rich domain and a cell-wall-associated domain, which are followed by an ectodomain that includes three C-repeats, an IgA-binding region and an N-terminal hypervariable domain (HVR).

Both mIgA and SIgA can be bound by group A streptococcus (GAS; *Streptococcus pyogenes*)[16]. GAS has long been considered a significant human pathogen that continues to be one of the top ten causes of mortality worldwide from infectious diseases[17]. The clinical manifestations of this Gram-positive bacterium are wide-ranging and diverse, encompassing both mild and common local infections, such as tonsillitis (a.k.a. Strep throat) and impetigo, as well as life-threating systemic invasions like toxic shock and sepsis[18,19]. While non-invasive disease has ~ 1000-fold higher incidence than the invasive disease, approximately 1.78 million new invasive infections are reported each year and are associated over 160,000 deaths[17,20]. GAS asymptomatic carriers are also common, with a prevalence of 12% among healthy children[21]. Typically, invasive infections encounter mIgA in the serum, whereas non-invasive, region infections encounter SIgA originating from mucosal secretions (e.g., in the upper respiratory tract)

85 as well as mIgA that perfuses inflamed tissues. Exceptions to this include GAS skin infections
86 such as impetigo where likely, only mIgA would be encountered, and (asymptomatic) GAS
87 colonization of mucosal secretions where inflammation is limited and SIgA is expected to
88 dominate.

89 The most extensively studied GAS virulence factors are a family of serotype-specific
90 proteins called M proteins, which include IgA binding proteins M4 and M22. M proteins adopt a
91 coiled-coil structure that protrudes up to 60 nm from the bacterium cell and can bind a variety of
92 host extracellular matrix and serum proteins[22,23]. Variability in the M protein ectodomain is
93 correlated with the ability to bind specific host factors; for example, some M proteins bind IgG,
94 some IgA, some fibrinogen, and some other host factors[24,25]. These interactions are thought to
95 contribution to the bacterium's virulence primarily by modulating the host immune response (e.g.,
96 limiting phagocytosis).

97 M4 and M22 are found in GAS strains M4 and M22, respectively, both of which can bind
98 IgA and have widespread global distribution; M4 ranked in the top five most common serotypes
99 identified in a systemic review of global GAS distribution[8,9,26,27]. Mutational analysis of both M4
100 and M22 revealed a 29-residue, minimal IgA-binding region located within the ectodomain with
101 similar, but not identical, sequences[9] (Fig. 1c). M22 also binds IgG at a site overlapping with the
102 IgA binding site[8]. Despite significance, how M proteins interact with host ligands remains unclear.
103 Existing structural data includes crystal structures of an M1 fragment containing the HVR and a
104 repetitive sequence element within the ectodomain and the HVR of M2, M22, M28, and M49,
105 none of which detail the structure of the IgA-binding region or IgA-M protein interactions [28,29].
106 Published analysis indicated that M4 and M22 binding to IgA is dependent on residues located at
107 the Cα2/Cα3 interface, including residues bound by host CD89 and *S. aureus* SSL7[2,6,7]. The
108 overlapping interface of M4 and CD89 suggested that M4 may interfere with IgA effector functions
109 and studies suggested that M4 protein can block respiratory burst in neutrophils *in vitro*[11].

110 However, the role of M4 (and CD89) in disease pathogenesis remains poorly understood and
111 published studies have focused on mIgA despite mucosal routes of entry.

112 To investigate how M proteins interact with SIgA and how those interactions compare to
113 host FcαRs, we determined two Cryo-EM structures, one of SIgA in complex with M4 and one of
114 SIgA in complex with CD89. Together structures indicate that M4 and CD89 share only a subset
115 of IgA interface residues and bind through different angles of approach. This leads to differences
116 in binding stoichiometry, with M4 occupying only one FcαR binding site on SIgA, while leaving
117 the other site sterically accessible.

118

119 **RESULTS**

120 **Structure determination of the SIgA-M4 complex**

121 In order to better understand how bacterial virulence factors such as M4 engage SIgA, we
122 designed an M4 expression construct encoding the ectodomain (residues 1-314) and combined
123 resulting protein with purified, recombinant human SIgA1 to create an SIgA-M4 complex. Using
124 this complex, we determined the SIgA-M4 cryo-EM structure to an average resolution of 3.1 Å
125 (FSC=0.143) (Fig. 2a). Map quality and resolution were sufficient to refine the positions of main
126 chain and side chain atoms for the majority of residues in the complex with the exception of Fab
127 residues, which were disordered (Fig. S1). The final structure included a single M4 dimer bound
128 to one SIgA; M4 residues 45-73, which include the entire known IgA-binding domain, were built
129 whereas the positions of residues more distal from the binding site were poorly resolved or
130 disordered, likely due to the inherent flexibility of the M4 dimer's extended, coiled-coil structure.
131 SIgA is superimposable with a previously published human SIgA1 structure with a RMSD of
132 0.708[14].

133 The M4 adopts an α-helical, parallel coiled-coil homodimer as observed for other M
134 proteins[24]. The M4 fragment termini are stabilized by hydrophobic interactions including two ideal

135 heptads, characterized by having Tyr, Ile, or Leu, at both *a* and *d* positions[30]. Hydrophobic
136 residues at the *a* and *d* positions (Y48, L51, I69, L72) are packed against each other and stabilize
137 the coiled-coil (Fig. 2b and 2c). We also observed polar interactions stabilizing the center of the
138 M4 fragment, including E54 which forms a hydrogen bond with N55, and Y65 which forms a
139 hydrogen bond with Q66 (Fig. 2c). A subset of residues between P46-K68 in both helices contact
140 SIgA, with one (helix-a) forming the majority of contacts with the IgA heavy chain and the other
141 (helix-b) using equivalent residues to contact solvent and a smaller interface with IgA. Prior
142 studies, including mutational analysis, identified M4 residues 45-73 as the minimal IgA binding
143 region[9]; our observations indicate that residues 45, 70, 71, and 73 contribute indirectly by
144 stabilizing the M4 coiled-coil dimer.
145
146 **The structure of the SIgA-M4 interface**
147 Despite the reported accessibility of two FcαR binding sites on SIgA, our structure reveals M4
148 bound only to Fc_{AB}[13]. The M4 interface includes 11 residues on the Cα3 domain of both heavy
149 chain-A (HC-A) and heavy chain-B (HC-B) (Fig. 3), with a total buried surface area of 1000 $Å^2$.
150 Participating IgA residues are predominantly located along one β-strand in HC-A (F443, T444,
151 Q445, T447), forming hydrogen bonds with M4 helix-a residues R49 and E54. HC-A E389 and
152 R346 also reach out to interact with M4 helix-a residues Q47 and D57 respectively, and HC-A
153 D449 forms a salt bridge with R59 on helix-b of M4. Additionally, HC-A Tp residue K454 may
154 mediate longer-range electrostatic interactions (e.g. 4.5Å) with M4 helix-b residues D67 and E70.
155 The M4 interface with IgA HC-B includes only residues S356, E357 and A360, which form
156 hydrogen bonds with Q66 of helix-b and R61, K68 of helix-a respectively. All IgA interface
157 residues are conserved in human IgA subtypes (IgA1, IgA2m1, IgA2m2) and most are conserved
158 in mouse IgA, with the exception of R346 and L441(Fig. S2). Mouse IgA is not known to bind M4
159 thus R346 and L441 appear to be especially important for M4 binding and structurally equivalent

residues in mouse IgA, P341 and M436 respectively, do not appear to support the interaction [31]. Comparing M4 and M22 sequences, we find seven out of 11 IgA contact residues are identical and the other four residues are similar (Fig. 2b), consistent with prior analysis that identified similar IgA binding motifs in M4 and M22[9].

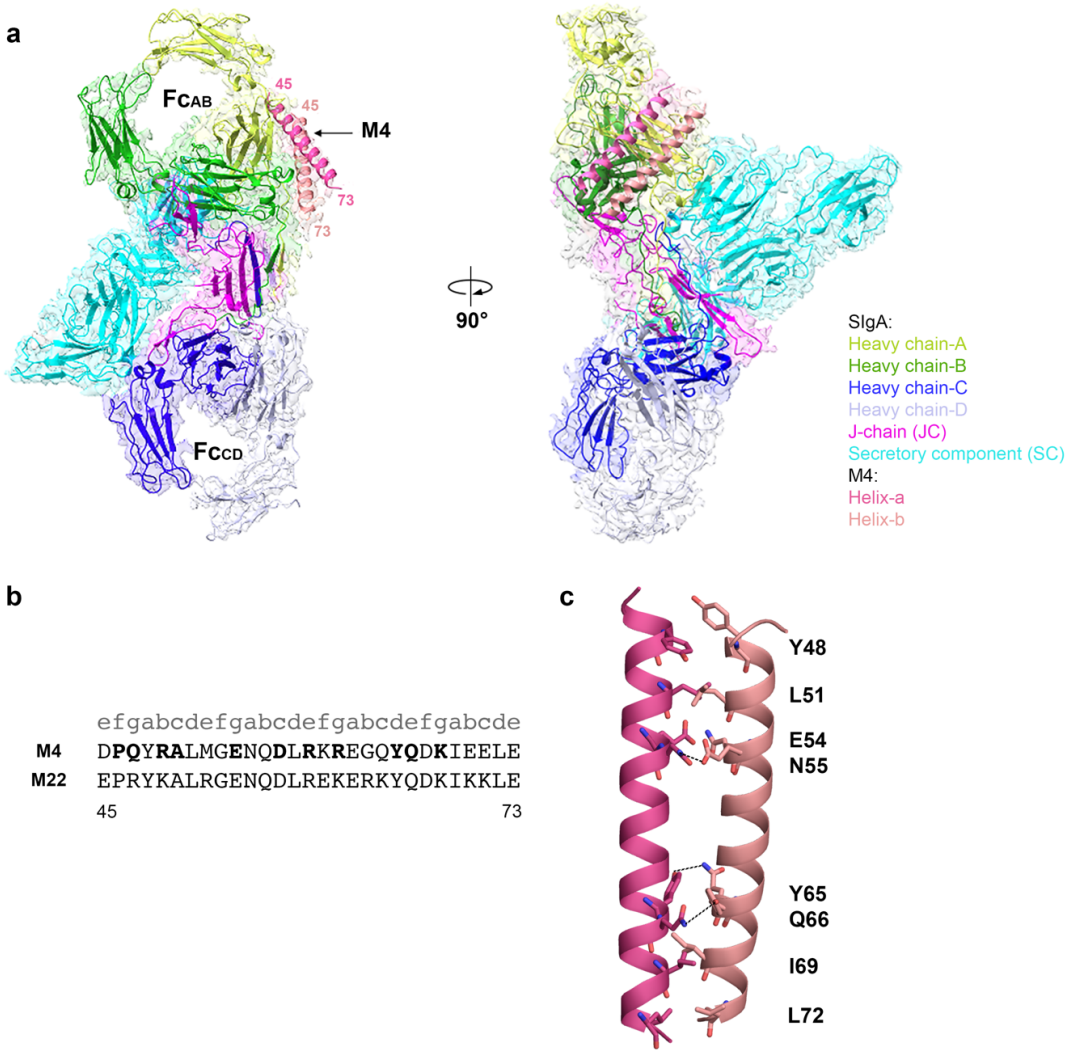
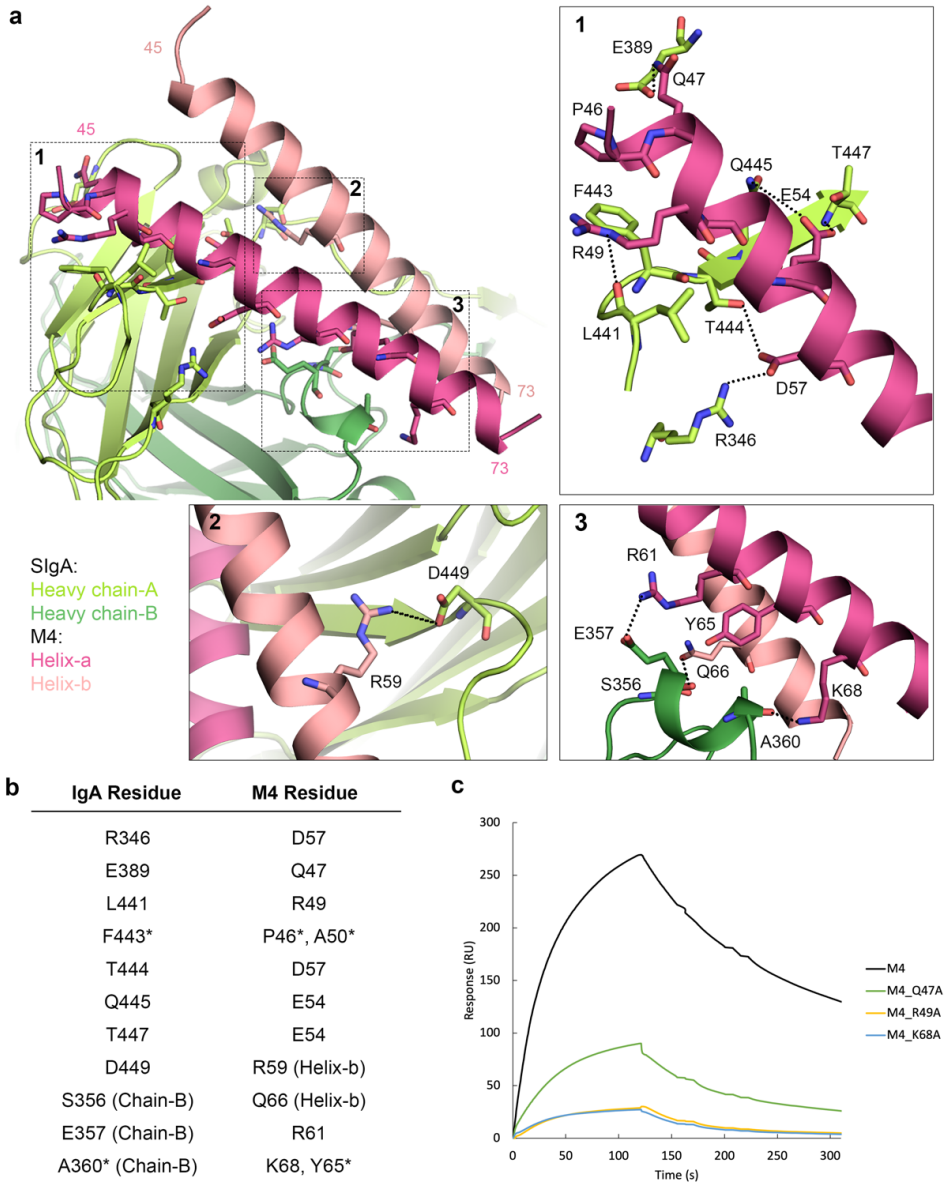


Figure 2. SIgA-M4 Structure. (a) The Cryo-EM structure of the M4 fragment (residues 45-73) in complex with human SIgA is shown as a cartoon with partially transparent map in two orientations. The model is colored according to the key, M4 is labeled, and Fabs are disordered. (b) Sequence alignment of the IgA binding regions of M4 and M22 with their heptad registrations shown above; numbering corresponds to both sequences and M4 residues interacting with IgA are depicted in bold. (c) Cartoon representation showing interchain interactions between M4 helix a and b. Polar interactions are shown as dotted lines.

165 **Mutagenesis confirms the key binding residues on M4**

166 Although side chain density was apparent in the structure, the N and C termini of the M4 coiled-
167 coil were not resolved and therefore to validate residue assignments and investigate the
168 contributions of individual M4 residues on SIgA binding, we conducted structure-based mutational
169 analysis and surface plasmon resonance (SPR) binding assays. Based on the pairwise
170 interactions between M4 and SIgA (Fig. 3b) and pairwise interactions between M4 helix-a and
171 helix-b, we made three individual M4 point mutations Q47A, R49A, K68A. A subset of residues,
172 including helix-b Q66, are positioned such that side chain atoms could interact with both IgA and
173 the other helix (e.g., helix-a); these were not mutated due to the possibility of disrupting the M4
174 dimer. All mutated M4 variants behaved equivalent to the wildtype M4 as assayed by SDS-PAGE
175 and size exclusion elution chromatography (not shown). M4 and mutant variant binding to SIgA
176 were determined by SPR. Sensorgrams revealed lower responses for all M4 mutants when
177 compared to concentration matched wild-type M4, as well as lower dissociation constants (Fig.
178 3c and Fig. S3). Compared to wild-type M4 (K_D=20 nM), mutant variants R49A (K_D=159 nM) and
179 K68A (K_D=156 nM) exhibited the most pronounce reduction in IgA-binding, consistent with the
180 observation that R49 interacts with one of the human-IgA specific binding site residues, L441.
181 K68 interacts with HC-B, indicating that contacts with both HC-A and HC-B chains are important.



IgA Residue	M4 Residue
R346	D57
E389	Q47
L441	R49
F443*	P46*, A50*
T444	D57
Q445	E54
T447	E54
D449	R59 (Helix-b)
S356 (Chain-B)	Q66 (Helix-b)
E357 (Chain-B)	R61
A360* (Chain-B)	K68, Y65*

Figure 3. The M4-SIgA interface. (a) Structure of the M4-SIgA interface. The boxed regions are enlarged (1-3) and depict the (1) M4 helix-a interaction with HC-A, (2) M4 helix-b interactions with HC-A, and (3) M4 interactions with HC-B. Black dotted lines indicate potential hydrogen bonds. (b) Polar and hydrophobic interactions between IgA and M4. Asterisks indicate residues contributing to hydrophobic interactions. (c) M4 binding analysis by surface plasmon resonance (SPR). The sensorgram shows the response of M4 (black) and individual M4 point mutants Q47A (green), R49A (yellow), and K68A (blue) binding to immobilized human SIgA1. All analyte concentrations shown are 0.125 uM. Sensorgrams, including a complete concentration series for each analyte, and binding models can be found in Figure S3.

183 **The structure of SIgA-CD89 complex**

184 The unexpected stoichiometry of the SIgA-M4 complex raised the question of how host FcαRs
185 bind SIgA. Modeling based on the structure of mIgA in complex with CD89 and SIgA structures
186 suggested that two CD89 can bind SIgA[14]; however, to verify this possibility, we determined the
187 Cryo-EM structure of SIgA with the CD89 ectodomain to a resolution of 3.2 Å (Fig. 4a). The
188 ectodomain of CD89 consists of two immunoglobulin-like domains (D1 and D2) arranged at an
189 angle of approximately 90 degrees to each other. Map quality and resolution was sufficient to
190 refine the positions of main chain and side chain atoms for the residues at the interface, while
191 density in the distal region of CD89, including the second Ig-like domain D2, was of lower
192 resolution. The refined structure revealed a 2:1 CD89:SIgA binding stoichiometry verifying that
193 two copies of CD89 can bind both mIgA and SIgA (Fig. 4a). The two copies of CD89 bound to
194 SIgA are separated by 108Å (distance between the two copies of D2 C-terminal residue T195)
195 and relative to each other are oriented differently than the two CD89 bound to mIgA for which C-
196 terminal residues are reportedly spaced by 124Å (Fig. 4a, and Fig. 4c)[2]. Alignment to the mIgA-
197 CD89 structure revealed superimposable binding sites (RMSD=1.165 at Fc_{AB}; RMSD=1.121 at
198 Fc_{CD}) on Cα2 and Cα3[2]. Consistent with the mIgA-CD89 structure, the CD89 D1 BC loop, the D
199 strand, the DE loop, and the FG loop (Ig-like domain nomenclature) contacted a total of 19 IgA
200 residues, including three Cα2 residues and 16 Cα3 residues, burying a total of 1656 $Å^2$ (Fig. 4b).
201

202 **CD89 and M4 share overlapping but distinct SIgA binding sites**

203 Comparison of the CD89 and M4 binding footprints on SIgA revealed five shared interface
204 residues (E389, L441, F443, T444, Q445) located on the Cα3 domain. These residues are among
205 those included in the Cα2/Cα3 "hot spot" for receptor and bacterial protein binding. CD89 binds
206 additional residues located toward and on Cα2, whereas M4 contacts only Cα3 and binds
207 residues closer to the Tps and JC, including R346, T447, D449 from HC-A and S356, E357, A360

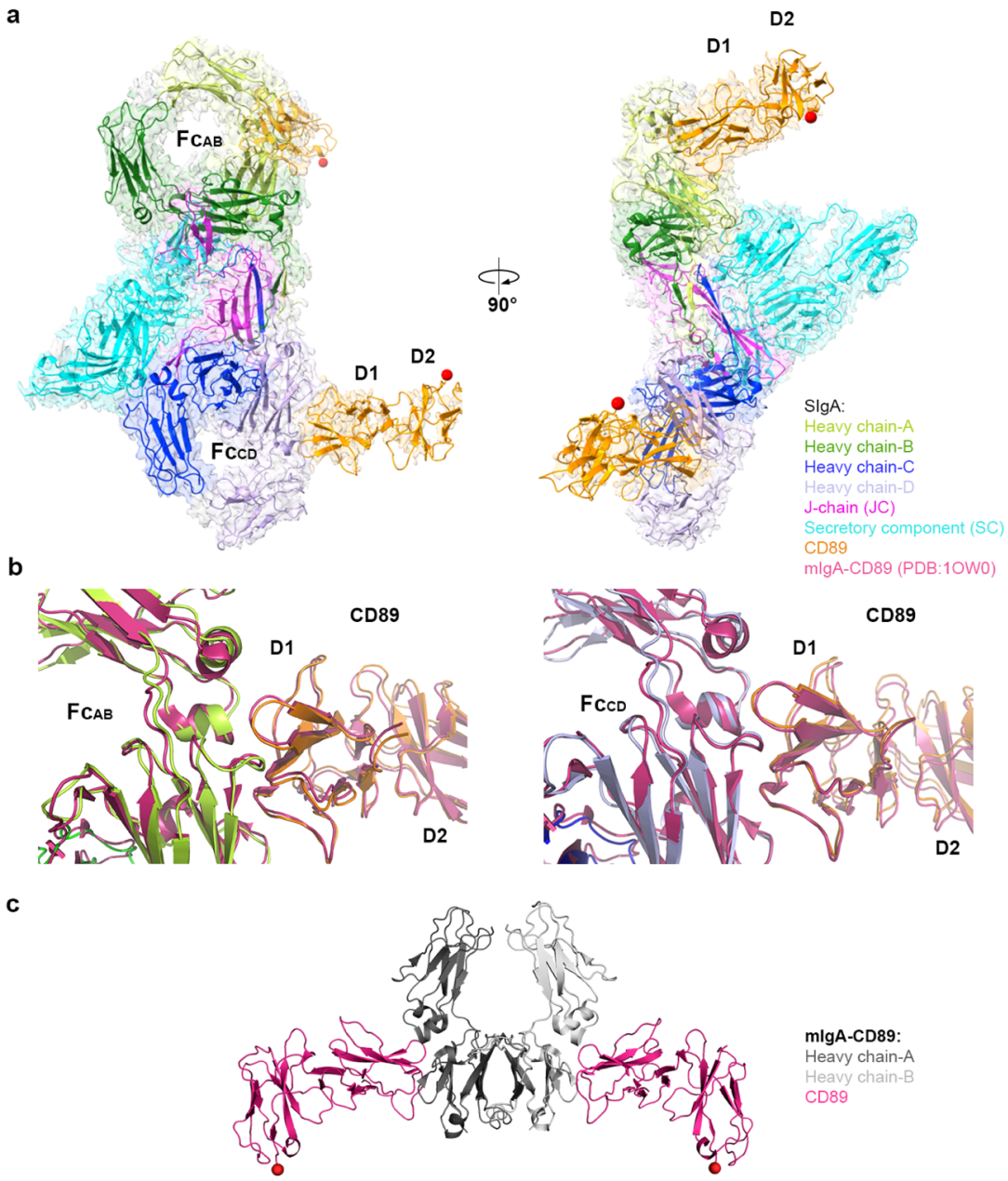


Figure 4. SIgA-CD89 structure. (a) The Cryo-EM structure of CD89 in complex with human SIgA is shown as a cartoon with partially transparent map in two orientations. The antigen binding fragments (Fabs) are disordered. The two Ig-like domains of CD89 are labeled D1, D2. CD89 D1 contacted a total of 19 IgA residues, including Cα2 residues L256, L257, L258 and Cα3 residues E348, R382, L384, S387, E389, M433, H436, E437, A438, L439, P440, L441, A442, F443, T444, Q445. (b) Alignment of the mIgA-CD89 crystal structure (PDB: 1OW0) to the Fc_{AB}-CD89 interface (left) and Fc_{CD}-CD89 interface (right) on SIgA-CD89. The BC, DE, FG loops are labeled on Fc_{AB}-CD89. (c) Crystal structure of CD89 in complex with human Fcα (PDB code: 1ow0). All models are colored according to the keys. The C-termini of CD89 D2 domains are indicated by red spheres, which on CD89-SIgA are 108Å apart.

209 from HC-B (Fig. 5a). CD89 and M4's distinct binding footprints and structures direct CD89 D2 C-
210 terminal residues away from SIgA and direct the M4 fragment C-terminal residues (and likely the
211 rest of the M4 coiled-coil), alongside Fc_{AB} in a line near parallel to $C\alpha3$ (Fig. 5b). Noted differences
212 contribute to the distinct stoichiometry of CD89 versus M4 binding to SIgA. In the case of M4, the
213 inclusion of S356, E357, A360 and R346 in the M4 binding interface, together with its extended
214 coiled-coiled topology, enforce the surprising 1M4:1SIgA stoichiometry because the steric
215 accessibility of these four residues and neighboring residues differs between Fc_{AB} and Fc_{CD}. In
216 Fc_{AB}, M4 can bind S356, E357, A360 and R346 with residues on both helices and extend the
217 coiled-coiled over neighboring residues and toward the Tps. In Fc_{CD}, S356, E357, A360 and R346
218 remain solvent accessible, but JC loop residues 24-32 bind the adjacent region, blocking the path
219 occupied by M4 (helix-b residues 69-72) and therefore occluding binding (Fig. 5b). The JC loop
220 that occludes M4 binding does not interfere with CD89 binding on Fc_{CD}, implying that sterically,
221 CD89 and M4 ectodomains could bind to the same copy of SIgA, with M4 binding Fc_{AB} and CD89
222 binding Fc_{CD} (Fig. 5c); however, whether this would occur in the context of bacterial and host cell
223 interactions remains to be determined.
224

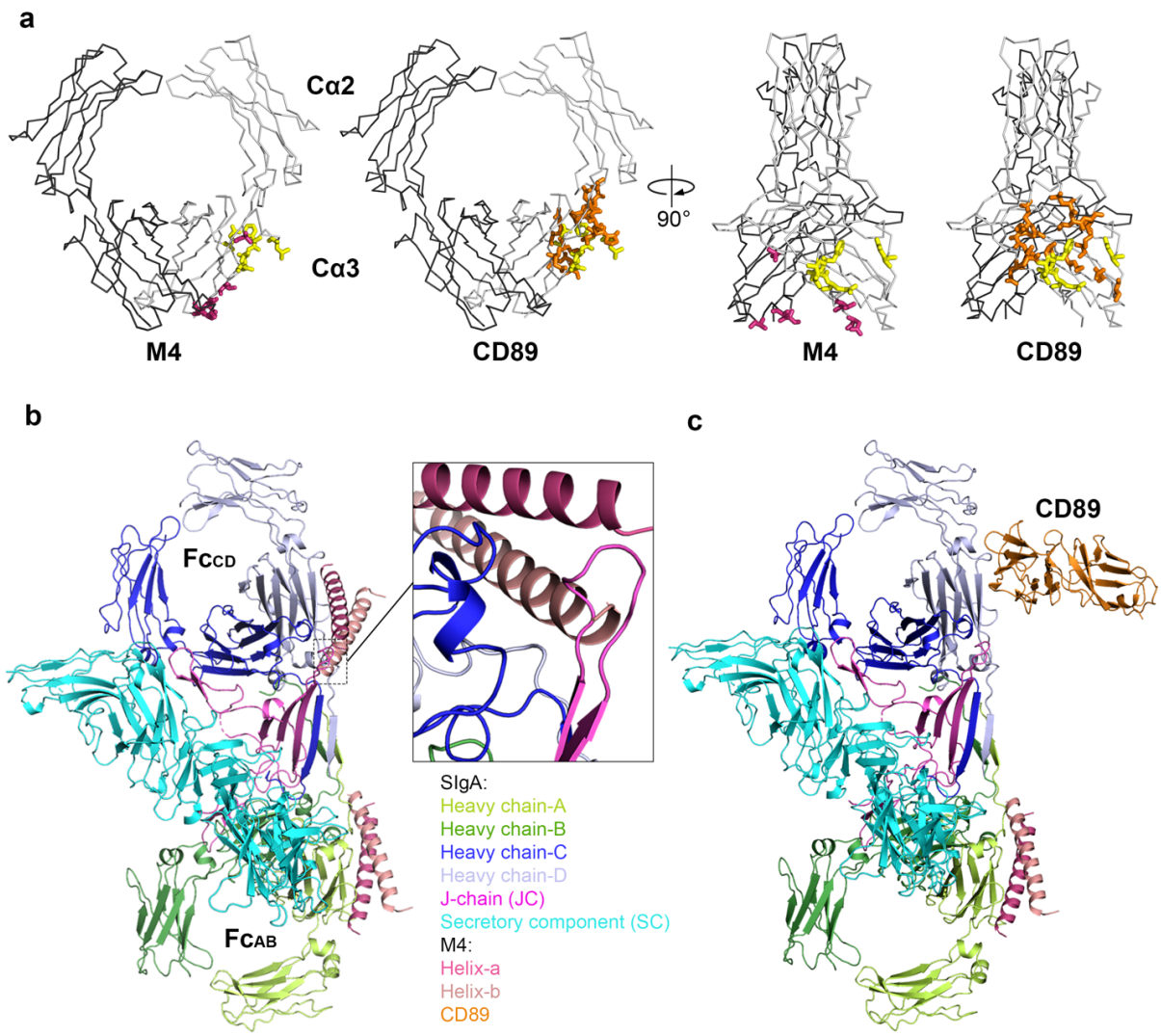


Figure 5. M4 and CD89 binding site comparison and modeling. (a) Comparison of M4 and CD89 binding site on Fcα shown in two orientations. Fcα is shown as ribbons and binding residues are shown as stick representation. The 11 residues bound by M4 are shown in pink and yellow and the 19 residues bound by CD89 are shown in orange and yellow. In all panels five shared binding site residues, E389, L441, F443, T444, Q445, are colored yellow. (b) M4 modeled at the second site (Fc_{CD}; blue) is partially overlapping with JC (magenta), indicating a steric clash. (c) Model of M4 (pink) and CD89 (orange) bound SIgA. CD89 bound to the Fc_{CD} site does not clash with SIgA residues.

225

226 **DISCUSSION**

227 IgA is associated with diverse effector functions mediated by monomeric and secretory
228 forms. Both forms have been shown to bind human CD89 and GAS; yet, interactions with SIgA
229 and potential differences between mIgA and SIgA engagement are poorly understood, limiting
230 knowledge of IgA effector functions and host-pathogen interplay[16,32]. Host FcαRs and bacterial
231 IgA-binding proteins have long been appreciated to bind mIgA through the conserved "hot spot",
232 however only recently were hot spot residues shown to be accessible on SIgA[6,11,12]. Our work
233 demonstrates that CD89 and M4 (and likely M22) can engage SIgA through just five common
234 residues and exhibit distinct binding footprints and stoichiometry, implying that while the hot spot
235 is conserved on mIgA and SIgA, its accessibility to different receptors and proteins is variable and
236 has potential to influence functional outcomes and provide unique selective advantages for both
237 the host and the pathogen.

238 It is notable that both mIgA and SIgA can bind two copies of CD89 *in vitro,* but that the
239 spacing and orientation of bound CD89 are different in CD89-mIgA and CD89-SIgA complexes.
240 Earlier reports suggested that the 124Å distance separating the CD89 D2 domains is too large to
241 initiate signaling (when two CD89 bind the same copy of mIgA). This may also be the case for
242 SIgA, in which we find the two copies of CD89 to be separated by 108Å (previously modeled to
243 99Å apart), although additional studies are needed to evaluate CD89 function in a host-cell
244 membrane [33]. If this is the case, then CD89 signaling will occur when copies of CD89 bind
245 adjacent IgA on IgA-antigen complexes. We envision that each monomeric/polymeric form of IgA
246 (and its secretory forms) may provide a unique set of CD89 binding sites, each with different
247 accessibility and each oriented differently relative to bound antigen and relative to copies of CD89
248 on the cell membrane[2]. This, in turn, may influence, CD89 clustering, potential for signaling, and
249 effector functions. Speculatively, this provides an explanation for how different forms of IgA in
250 different location (e.g. serum or mucosa) may elicit different outcomes upon encountering CD89,

251 or perhaps other FcαRs such as human Fcα/μR, and why additional factors (e.g. Mac-1) might
252 be involved in SIgA-CD89 dependent signaling[5,34].

253 Whereas mIgA and SIgA share the same number of CD89 binding sites, this does not
254 appear to be the case for M4 (and presumably, M22) since the JC, which occludes M4 binding to
255 SIgA Fc_{CD}, is not present in mIgA. Because differences in mIgA and SIgA engagement may
256 signify different functional outcomes from host-GAS interactions, we modeled the M4 structure
257 onto Fcα and tested soluble M4 binding to mIgA using SPR. Modeling verified the steric
258 accessibility of two M4 binding sites and SPR experiments revealed similar dissociation constants,
259 ~15nM and 20nM, for soluble M4 binding to mIgA and SIgA, respectively (Fig. S5). This suggests
260 that the two M4 binding sites on mIgA are molecularly equivalent to the single site on SIgA. While
261 we cannot completely rule out the possibility that the unliganded IgA Tp, which is absent from Fcα
262 structures, contributes to M4-mIgA binding in some way, its inclusion in our SPR experiments did
263 not markedly alter binding constants. These observations are consistent with publications
264 reporting similar affinities for M22 binding to mIgA and SIgA, and those reporting similar levels of
265 mIgA and SIgA binding to GAS cultures[16,35]. While M4 may share molecular interfaces with mIgA
266 and SIgA, the binding of two copies of M4 to one mIgA could enhance the avidity of IgA-GAS
267 interactions, effectively strengthening the interaction while also completely blocking CD89-binding
268 sites. This model is consistent with other reports indicating the M4 binds mIgA with higher affinity
269 than SIgA and with proposals that M4 functions to block CD89 effector functions (e.g.
270 phagocytosis)[11]. However, whether each mIgA bound to GAS can in fact be bound by two copies
271 of M4 *in vivo* remains to be tested; it is also unclear how many CD89- binding sites need to be
272 blocked to inhibit effector functions.

273 In the case of SIgA interactions with GAS, the sterically enforced 1:1 M4:SIgA
274 stoichiometry is intriguing because SIgA adopts an asymmetric structure, which published
275 modeling predicts will influence the positions of Fabs and accessibility of the SC and FcαR binding
276 sites[13]; therefore, M4 binding to Fc_{AB} positions SIgA differently than if it could bind to Fc_{CD}. To
277 visualize this possibility, we used Alphafold2-multimer to model the full-length M4, aligned it to
278 our M4-SIgA structure and created schematic representations of the complex on a bacterium
279 surface (Fig. 6)[36]. Assuming that M4 is oriented approximately perpendicular to the outer surface
280 of the membrane, it would extend ~400 Å, placing the bound SIgA ~ 250 Å from the surface where
281 one set of Fabs would be directed upward and away from the bacterium surface and the other set

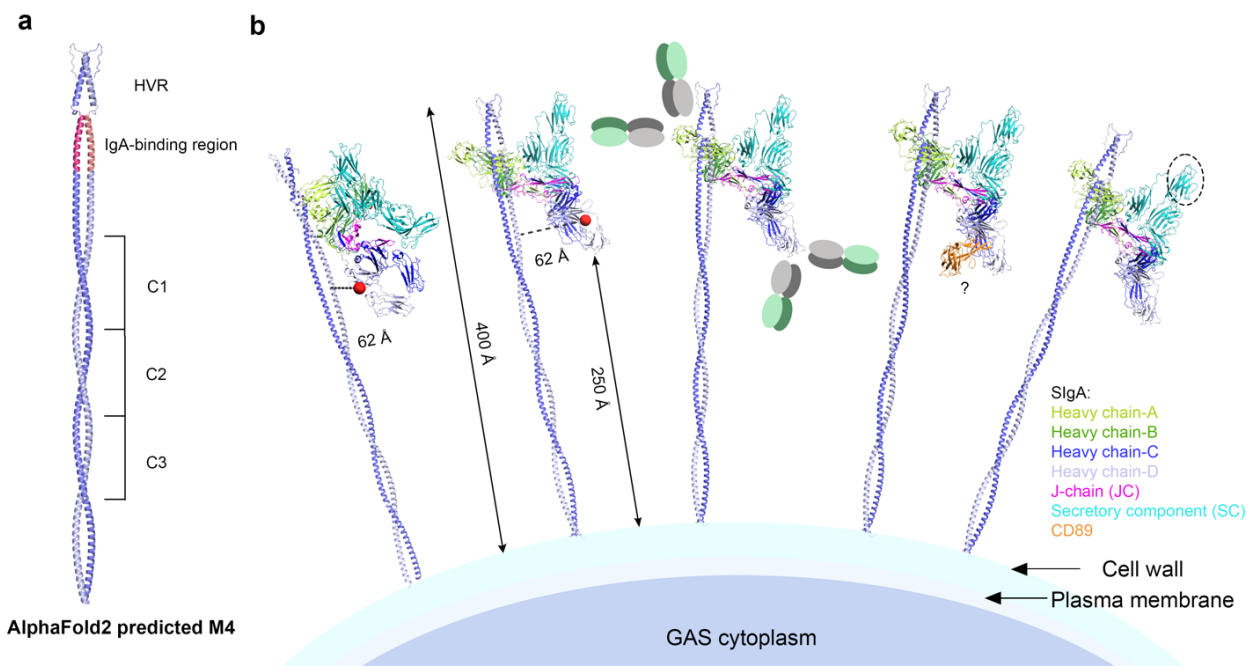


Figure 6. Schematic model of human SIgA binding to M4 on a GAS surface. (a) AlphaFold2-multimer predicted M4 dimer; residues 287-316 are predicted to be disordered and not shown. The IgA-binding region on the predicted model is colored in pink and has an RMSD of 0.640 when aligned to the SIgA-M4 structure. (b) Surface-associated M4, shown in two orientations, binds to one side of SIgA (Fc_{AB}), holding all SIgA at a similar orientation and distance from the bacterium surface while leaving the Fc_{CD} FcαR binding site (red sphere) open. Possible locations of disordered Fabs are shown as grey and green ovals. CD89 binding to M4 bound SIgA is sterically possible and modeled on one copy of the SIgA-M4 complex. SIgA SC domain D2 is indicated by a circle and is exposed near the outermost surface of bacterium.

282 directed toward the surface. In this model, the SC protrudes upward and away from the bacterium
283 surface similar to the M4 HVR, and the Fc_{CD} binding site is located closer to the bacterium surface
284 and approximately 62 Å from M4 (Fig. 6b). Other GAS surface proteins are expected to be shorter
285 than M4, implying that bound SIgA would occupy a position near the outermost surface of the
286 bacterium[37]. While information describing M4 flexibility, as well as the density of M4 and other
287 surface antigens, is needed to fully evaluate the geometric arrangement of SIgA-bound to the
288 bacterial surface, our data and analysis indicate that the 1:1 M4:SIgA stoichiometry will constrain
289 the location and orientation of bound SIgA, a possibility that should be considered when exploring
290 functional outcomes of bacterial-SIgA interactions (Fig. 6).

291 How might SIgA binding to GAS M4 (or M22) in a constrained orientation influence the
292 host or microbe? One possibility is that M4 binding to Fc_{AB} impacts the binding of SIgA Fabs to
293 other antigens on the surface of GAS, such as fibronectin-binding proteins (SFb), Streptococcus
294 pyogenes cell envelope proteinase (SpyCEP), streptococcal C5a peptidase (Scp)[38]. Bacteria that
295 can bind SIgA through M4 (or M22), may be able to limit antibody binding to other surface proteins,
296 although as suggested by others, the relative affinity of antibodies to other surface antigens and
297 their steric accessibility will likely impact the outcome[39]. Another possibility is that SIgA-bound M4
298 influences the localization of the bacterium (e.g. in host mucosal secretions). SIgA is covered in
299 host glycans, which have been implicated in binding to host mucus, lectins, and even microbial
300 proteins, something we speculate could be promoted by the position of the highly glycosylated
301 SC[40,41]. Our model predicts the D2 domain of SC, which is the most accessible among the five
302 domains, would be exposed near the M4 HVR toward the outermost surface of the bacterium
303 where similar to other SIgA (not bound to GAS), it might mediate interactions with other factors
304 (Fig. 6)[13].

305 The 1:1 M4:SIgA stoichiometry also leaves open the possibility that M4-bound SIgA could
306 bind CD89 (or another FcαR such as Fcα/μR) at the Fc_{CD} site and influence effector functions in
307 some way (Fig. 6b). The accessibility of these sites on mIgA and SIgA has not been verified on

308 a bacterial surface and it seems plausible that CD89 activity could be inhibited by occluding just
309 one of the two CD89-binding sites, especially since our model predicts that the Fc_{CD} site on SIgA-
310 M4 complexes would be pointed downwards toward the bacterial membrane, potentially occluding
311 the site. In this case, SIgA may act like mIgA, restricting CD89-dependent effector functions. Still,
312 we cannot rule out the possibility that a sterically accessible Fc_{CD} binding site and kinetics that
313 favor CD89 binding might allow one copy of CD89 to bind Fc_{CD} and subsequently, another copy
314 to compete with M4 for the Fc_{AB} site. Another possibility is that CD89 binding to Fc_{CD} could
315 promote other copies of CD89 binding to local SIgA (and/or mIgA) that is not bound by M4, but
316 rather is bound to another antigen with its Fabs.

317 The idea that IgA and its effectors have provided selective pressure driving GAS evolution
318 is reasonable given the existence of M4 and M22 virulence factors and prevalence of M4 and
319 M22 strains, which are abundant and geographically widespread[27,42]. Interestingly, both M4 and
320 M22 strains lack a Hyaluronic Acid (HA) capsule, found on other GAS strains, which is
321 antiphagocytic and has long been proposed as an essential virulence factor[43,44]. We speculate
322 that the Ig-binding proteins provide an alternative and compensatory mechanisms for M4 and
323 M22 strain virulence. Regardless, our work raises the possibility that SIgA has shaped GAS
324 evolution, perhaps providing different selective pressures than mIgA, and also highlights the
325 importance of further investigating GAS-SIgA interactions in the mucosa where M4/M22 might
326 block host $Fc\alpha R$ effector functions and/or block antibody binding to surface antigens and/or
327 provide another survival advantage for the bacteria (e.g. adhesion to mucosal factors).

328 The notion that host SIgA effector functions play a significant role in antimicrobial response
329 and may provide selective pressure extends beyond GAS, given that at least three distinct
330 pathogenic bacterial species shown to binding SIgA express specific IgA and/or SIgA binding
331 proteins, including *S. pyogenes, S. aureus and S. pneumoniae,* all of which utilize the
332 nasopharynx as their primary human reservoir where encounters with SIgA have potential to
333 influence virulence and host response[45,46]. To explore potential similarities and differences

334 between M4 and other IgA-binding proteins, we modeled a *S. aureus* SSL7-SIgA complex by
335 aligning the crystal structure of SSL7-Fcα (PDB:2QEJ) to each Fc on SIgA (Fig. S6)[7]. Resulting
336 models suggest that SSL7 can bind both the Fc_{AB} and Fc_{CD} "hot spots," resulting in 2:1 SSL7:SIgA
337 stoichiometry, as observed for CD89:SIgA. This stoichiometry appears possible because SSL7
338 adopts a flat "disk-like" topography that when bound to IgA or SIgA positions the C-terminus away
339 from Fcα, similar to CD89 and in contrast to M4[7]. This structure and trajectory likely allow SSL7
340 to bind SIgA Fc_{CD} without being occluded by the JC loop as observed for M4, signifying that
341 among Fcα-binding proteins, the sterically enforced 1:1 M4:SIgA stoichiometry appears to be
342 unique to M4 and hinting at functional differences relevant to understanding host interactions with
343 GAS and *S. aureus*. Unlike SSL7 and M4, which bind the IgA Fc, the *S. pneumoniae* SpsA (also
344 called CbpA) protein binds directly to pIgR or SC, meaning that it interacts only with SIgA and not
345 mIgA. This further supports the notion that microbial-SIgA interactions are functionally, and
346 broadly, relevant; together structural data provided here and in other recent reports provide a
347 framework to further elucidate the consequences of those interactions for both the host and the
348 microbe[2,6,13–15,47].
349

350 **METHODS**

351 **Protein expression and purification**

352 Genes encoding the human IgA1 heavy chain constant region and the lambda light chain constant
353 region were fused with STA121 VH and VL domain sequences [48]. The human tissue plasminogen
354 activator (tPA) signal sequence was inserted upstream of the IgA heavy chain sequence. These
355 sequences along with human JC and SC were cloned into mammalian expression vector
356 pD2610v1 (Atum). All constructs were transiently co-transfected into HEK Expi293F cells using
357 ExpiFectamine to produce SIgA complex (Thermo Fisher). Six days after transfection, SIgA was
358 purified from the cell supernatant using CaptureSelect™ Human IgA Affinity Resins. Superose 6

359 (Cytiva) size exclusion chromatography (SEC) was used to further purify SIgA. The gene
360 encoding the human CD89 ectodomain (uniprot ID: P24071) was fused to a C-terminal
361 hexahistine-tag and cloned into pD2610v1. The CD89 construct was transiently transfected into
362 HEK Expi293F cells using ExpiFectamine. CD89 was purified from the cell supernatant using Ni-
363 NTA affinity resin (Qiagen). Superose 6 SEC was used to further purify CD89. The gene encoding
364 the *Streptococcus pyogenes* M4 protein (uniprot ID P13050) was cloned into pET28_a (+) vector
365 (Novagen). The M4 endogenous N-terminal signal peptide and C-terminal cell wall sorting signal
366 were not included in the construct and a C-terminal hexahistine-tag was added. M4 mutants Q47A,
367 R49A, K68A were generated using the Q5 site-directed mutagenesis kit (NEB, E0554). The M4
368 mutant variant constructs were transformed and expressed in *E. coli* strain BL21 (DE3);
369 expression was induced using 1mM IPTG, added to the culture growing at 37°C when OD reaches
370 0.4-0.6. Cells were harvested 3-5 hours after IPTG induction by centrifugation, resuspended in
371 50 mM Tris-Cl, 150 mM NaCl, pH 7.5, 1mM phenylmethylsulfonyl fluoride and lysed by sonication.
372 The M4 was isolated from soluble fraction of cell lysate using Ni-NTA affinity resin. Superose 6
373 SEC was used to further purify M4. Purified SIgA and M4 or CD89 were mixed in a 1:2 molar ratio
374 and incubated in storage buffer (50 mM Tris-Cl, 150 mM NaCl, pH 7.5) at 4°C overnight and
375 complexes were purified using Superose 6 SEC.
376
377 **Cryo-EM grid preparation and data collection**
378 SIgA-M4: Quantifoil R2/2 300 mesh holey carbon grids were glow discharged for 60s; 3ul of SIgA-
379 M4 at 1.0 mg/ml was applied to each grid with Vitrobot Mark IV at SLAC with 1 blot of 1.5 s, 0
380 offset, 5 s wait, at 100% humidity and 4 C°. Movies were collected using EPU on a Titan Krios
381 (TEMALPHA at SLAC) operating at 300 kV and with a Gatan K4 direct electron detector. Both
382 untilted and 25-degree tilted movies were collected at 0.95 Å/Pixel, 40 frames per movie, total
383 exposure time of 4.95 s and total dose of 50 electrons/$Å^2$.

384 SIgA-CD89: Quantifoil R1.2/1.3 400 mesh holey carbon grids were glow discharged for
385 60s; 3ul of SIgA-CD89 at 1.0 mg/ml was applied to each grid with Vitrobot Mark IV with 1 blot of
386 2 s, 2 offset, 5 s wait, at 100% humidity and 4 C°. Movies were collected using SerialEM on a
387 Titan Krios at Purdue University operating at 300 kV and with a K3 direct electron detector. Both
388 untilted and 30-degree tilted movies were collected at 0.82 Å/Pixel, 60 frames per movie, total
389 exposure time of 2.41 s and total dose of 71.3 electrons/$Å^2$. Data collection statistics can be found
390 in Table SI.
391
392 **Cryo-EM data processing**
393 SIgA-M4: 2600 movies were collected with untilted stage and 3251 movies were collected with
394 stage tilted at 25 degrees. All data processing were done in CryoSPARC v.3.2.0[49]. Initial
395 processing of the two data sets were done separately. Movies were motion corrected, CTF
396 estimated and manually curated based on ice thickness, total motion and CTF fits, resulting in
397 2081 untilted and 2786 tilted movies. Particles were picked using blob picker in CryoSPARC with
398 minimum and maximum particle diameter of 100 Å and 200 Å. Particles were extracted with 2x
399 binning. 2D classification of these particles generated initial references for template-based particle
400 picking, which gave 970k particles from untilted movies and 1.3 million particles from tilted movies.
401 Those particles were combined and subjected to several rounds of 2D classification to remove
402 bad classes. A final set of 200K particles were used to generate the *ab-initio* model. Particles in
403 the final set were then re-extracted at full-pixel size and used for non-uniform refinements of the
404 model. The final refinement generated a map with an average resolution of 3.07 Å at FSC=0.143.
405 SIgA-CD89: 2955 movies were collected with untilted stage and 3696 movies were
406 collected with stage tilted at 30 degrees. All data processing were done in CryoSPARC v.4.1.2[49].
407 Initial processing of the two data sets were done separately. Movies were motion corrected, CTF
408 estimated and manually curated based on ice thickness, total motion and CTF fits, resulting in
409 2865 untilted and 3105 tilted movies. Particles were picked using blob picker in CryoSPARC with

410 minimum and maximum particle diameter of 150 Å and 250 Å. Particles were extracted with 2x
411 binning. 2D classification of these particles generated initial references for template-based particle
412 picking, which gave 507K particles from untilted movies and 732K particles from tilted movies.
413 Those particles were combined and went through several rounds of 2D classification to remove
414 bad classes. A final set of 249K particles were used to generate the *ab-initio* model. Particles in
415 the final set were then re-extracted at full-pixel size and used for non-uniform refinements of the
416 model. The final refinement generated a map with an average resolution of 3.18 Å at FSC=0.143.
417

418 **Structure building and refinement**

419 SIgA-M4: The structure of human SIgA (PDB: 6UE7) was docked into the map using UCSF
420 Chimera[50]. The M4 fragment was hand-built using the Coot Molecular Graphics Package[51]. All
421 chains of SIgA and M4 were refined as rigid bodies using Phenix real space refinement[52].
422 Subsequent iterations involved inspection of the map-model fit and manual adjustment of the
423 model followed by Phenix real space XYZ refinement. Iterations of Phenix refinements and
424 manual adjustments were repeated until model quality no longer improves.

425 SIgA-CD89: Human SIgA structure (PDB: 6UE7) and mIgA:CD89 structure (PDB:1OW0)
426 were docked into the map using UCSF Chimera[50]. All chains of SIgA and CD89 were refined as
427 rigid bodies using Phenix real space refienemnt[52]. Inspection of the map-model fit and manual
428 adjustment of the model were done after each Phenix real space XYZ refinement. Iterations of
429 Phenix refinements and manual adjustments were repeated until model quality no longer
430 improves. Model and refinement statistics can be found in Table SI.
431

432 **Surface plasmon resonance**

433 Surface plasmon resonance (SPR) binding experiments were done using a Bruker Sierra SPR-
434 32 Pro instrument. 0.2 uM human mIgA or SIgA in sodium acetate buffer (pH = 4.5) was coupled
435 to lanes B and D, respectively, on a high-capacity amine (HCA) chip using an amine coupling kit

436 (Bruker 1862634). Lane A and C were mock coupled (with only sodium acetate buffer PH 4.5)
437 and used as references. Responses were measured for a two-fold dilution series of each analyte
438 (M4 wild type and mutants Q47A, R49A, K68A) in HBS-EP+ buffer (0.01 M HEPES pH 7.4, 0.15
439 M NaCl, 3 mM EDTA, 0.005% v/v Surfactant P20) with a high concentration of 0.25 uM. All binding
440 experiments were done with 120 s contact time, 180 s dissociation time, and flow rate of 25 ul/min.
441 Surfaces were regenerated with 2.5M $MgCl_2$. All data were collected at room temperature. Data
442 were fitted to 1:1 binding model using Bruker Analyzer R4 and the K_D values were reported by
443 the Analyzer.
444
445 **Structure modelling, analysis, and sequence alignment**
446 Human SIgA (PDB: 6UE7) was aligned to the M4-bound SIgA in The PyMOL Molecular Graphics
447 System (Version 2.5.1 Schrödinger, LLC) and the RMSD between the bound and unbound state
448 of SIgA was calculated automatically during the alignment performed by PyMOL (A>align>align
449 to molecule). A list of all possible interactions between SIgA and M4 was generated using the
450 protein interaction calculator (PIC) webserver[53]. The interactions were visualized and manually
451 checked in PyMOL and contacts within 4 Å were selected and reported in Fig. 3b. The buried
452 surface area was calculated using the 'get_area' function in PyMOL.
453 AlphaFold2-multimer was used to predict the full length M4 dimer structure and ColabFold
454 1.5.2 was used to generate the prediction[36]. M4 residues 286-313 are predicted as disordered
455 and not shown in figures. Full length modeled M4 residues 45-73 were aligned to the
456 corresponding residues in the SIgA-M4 structure using PyMOL, with an RMSD of 0.640. The
457 aligned structure of full length M4 with SIgA were used in Fig. 6. The distance of M4 to GAS
458 surface was estimated by measuring the distance between M4 residues 15 and 285. The distance
459 from the FcαR site on Fc_{CD} to M4 was measure from HC-D residue L441 to M4 helix-b residue
460 H97.

461 The sequence alignment between human and mouse IgA heavy chain (human: Uniprot
462 P10876, P10877, A0A0G2JMB2; mouse: Uniprot P01878) was carried out using ClustalOmega
463 and the alignment figure was made using ESPript3[54,55]. The alignment of mIgA:CD89 (PDB:
464 1OW0) to SIgA:CD89 and RMSD calculation were done in the same way as described above.
465

466 **Figures**

467 Structural figures were made using UCSF Chimera and PyMOL; SPR data were plotted using
468 Bruker Sierra-32 Analyzer and Microsoft Excel[50]. All figures were assembled using Adobe
469 Photoshop.
470

471 **DATA AVAILABILITY**

472 The Cryo-EM density maps have been deposited in the EM databank (www.ebi.ac.uk/emdb) with
473 the accession codes 40568, and 40567 for M4-SIgA and CD89-SIgA structures, respectively, and
474 the refined coordinates have been deposited in the Protein Data Bank (www.rcsb.org) with
475 accession codes 8SKV (M4-SIgA) and 8SKU (CD89-SIgA).
476

477 **AUTHOR CONTRIBUTIONS**

478 The study was conceived by B.M.S and Q.L.; experiments were conducted by Q.L.; B.M.S and
479 Q.L. analyzed data and wrote the manuscript.
480

481 **ACKNOWLEDGEMENTS**

482 This work was supported by NIH R01 AI165570 (PI Beth Stadtmueller) and University of Illinois
483 start-up funding. Cryo-Electron microscopy data were collected at the Purdue Cryo-EM facility
484 (http://cryoem.bio.purdue.edu) with assistance from Thomas Klose and the Stanford-SLAC Cryo-
485 EM Center (S^2C^2) supported by the National Institutes of Health Common Fund Transformative
486 High Resolution Cryo-Electron Microscopy program (U24 GM129541) with assistance from Htet

487 A. Khant. We thank members of the Stadtmueller lab for insightful discussion and Sarah Leonard
488 for supporting SPR experiments.

489 **MATERIALS AND CORRESPONDENCE.**

490 Correspondence and requests for materials should be addressed to Beth Stadtmueller,
491 bethms@illinois.edu.

492

493 **COMPETING INTERESTS.**

494 Authors declare no competing interests.

495

496 **REFERENCES**

497 1. Mantis, N. J., Rol, N. & Corthésy, B. Secretory IgA's complex roles in immunity and
498 mucosal homeostasis in the gut. *Mucosal Immunol* **4**, 603–611 (2011).
499 2. Herr, A. B., Ballister, E. R. & Bjorkman, P. J. Insights into IgA-mediated immune
500 responses from the crystal structures of human FcαRI and its complex with IgA1-Fc.
501 *Nature 2003 423:6940* **423**, 614–620 (2003).
502 3. Shen, L. *et al.* Presentation of ovalbumin internalized via the immunoglobulin-A Fc
503 receptor is enhanced through Fc receptor γ-chain signaling. *Blood* **97**, 205–213 (2001).
504 4. Mkaddem, S. Ben *et al.* Lyn and Fyn function as molecular switches that control
505 immunoreceptors to direct homeostasis or inflammation. *Nat Commun* **8**, (2017).
506 5. Spriel, A. B. van, Leusen, J. H. W., Vilé, H. & Winkel, J. G. J. van de. Mac-1
507 (CD11b/CD18) as Accessory Molecule for FcαR (CD89) Binding of IgA. *The Journal of*
508 *Immunology* **169**, 3831–3836 (2002).
509 6. Wines, B. D., Willoughby, N., Fraser, J. D. & Mark Hogarth, P. A competitive mechanism
510 for staphylococcal toxin SSL7 inhibiting the leukocyte IgA receptor, FcαRI, is revealed by
511 SSL7 binding at the Cα2/Cα3 interface of IgA. *Journal of Biological Chemistry* **281**,
512 1389–1393 (2006).
513 7. Ramsland, P. A. *et al.* Structural basis for evasion of IgA immunity by Staphylococcus
514 aureus revealed in the complex of SSL7 with Fc of human IgA1. *PNAS* **104**, 15051–15056
515 (2007).
516 8. Stenberg, L., Otoole, P. W., Mesteckm, J. & Lindahl, G. Molecular Characterization of
517 Protein Sir, a Streptococcal Cell Surface Protein That Binds Both Immunoglobulin A and
518 Immunoglobulin G*. *Journal of Biological Chemistry* **269**, 13458–13464 (1994).
519 9. Johnsson, E., Andersson, G., Lindahl, G. & Hedén, L. O. Identification of the IgA-binding
520 region in streptococcal protein Arp. *J. Immunol* **153**, 3557–3564 (1994).
521 10. Russell-Jones, G. J., Gotschlich, E. C. & Blake, M. S. A surface receptor specific for
522 human iga on group B streptococci possessing the Ibc protein antigen. *Journal of*
523 *Experimental Medicine* **160**, 1467–1475 (1984).

11. Pleass, R. J., Areschoug, T., Lindahl, G. & Woof, J. M. Streptococcal IgA-binding Proteins Bind in the Cα2-Cα3 Interdomain Region and Inhibit Binding of IgA to Human CD89. *Journal of Biological Chemistry* **276**, 8197–8204 (2001).
12. Pleass, R. J., Dunlop, J. I., Anderson, C. M. & Woof, J. M. Identification of Residues in the CH2/CH3 Domain Interface of IgA Essential for Interaction with the Human Fcα Receptor (FcαR) CD89. *Journal of Biological Chemistry* **274**, 23508–23514 (1999).
13. Kumar Bharathkar, S. *et al.* The Structures of Secretory and Dimeric Immunoglobulin A. *Elife* 61801 (2020).
14. Kumar, N., Arthur, C. P., Ciferri, C. & Matsumoto, M. L. Structure of the secretory immunoglobulin A core. *Science (1979)* **367**, 1008–1014 (2020).
15. Wang, Y. *et al.* Structural insights into secretory immunoglobulin A and its interaction with a pneumococcal adhesin. *Cell Res* **30**, 602 (2020).
16. Lindahl, G. & Akerstrom, B. Receptor for IgA in group A streptococci: cloning of the gene and characterization of the protein expressed in Escherichia coii. *Mol Microbiol* **3**, 239–247 (1989).
17. Ralph, A. P. & Carapetis, J. R. Group a streptococcal diseases and their global burden. *Curr Top Microbiol Immunol* **368**, 1–27 (2013).
18. Mitchell, T. J. The pathogenesis of streptococcal infections: from Tooth decay to meningitis. *Nature Reviews Microbiology 2003 1:3* **1**, 219–230 (2003).
19. Happonen, L. *et al.* A quantitative Streptococcus pyogenes–human protein–protein interaction map reveals localization of opsonizing antibodies. *Nature Communications 2019 10:1* **10**, 1–15 (2019).
20. Carapetis, J. R., Steer, A. C., Mulholland, E. K. & Weber, M. The global burden of group A streptococcal diseases. *Lancet Infect Dis* **5**, 685–694 (2005).
21. Shaikh, N., Leonard, E. & Martin, J. M. Prevalence of streptococcal pharyngitis and streptococcal carriage in children: a meta-analysis. *Pediatrics* **126**, (2010).
22. McNamara, C. *et al.* Coiled-coil irregularities and instabilities in group A Streptococcus M1 are required for virulence. *Science (1979)* **319**, 1405–1408 (2008).
23. Phillips, G. N., Flicker, P. F., Cohen, C., Manjula, B. N. & Fischetti, V. A. Streptococcal M protein: alpha-helical coiled-coil structure and arrangement on the cell surface. *Proc Natl Acad Sci U S A* **78**, 4689–4693 (1981).
24. Frost, H. R., Sanderson-Smith, M., Walker, M., Botteaux, A. & Smeesters, P. R. Group A streptococcal M-like proteins: From pathogenesis to vaccine potential. *FEMS Microbiol Rev* **42**, 193–204 (2018).
25. Smeesters, P. R., McMillan, D. J. & Sriprakash, K. S. The streptococcal M protein: a highly versatile molecule. *Trends Microbiol* **18**, 275–282 (2010).
26. Steer, A. C., Law, I., Matatolu, L., Beall, B. W. & Carapetis, J. R. Global emm type distribution of group A streptococci: systematic review and implications for vaccine development. *Lancet Infect Dis* **9**, 611–616 (2009).
27. Rafei, R., AL Iaali, R., Osman, M., Dabboussi, F. & Hamze, M. A global snapshot on the prevalent macrolide-resistant emm types of Group A Streptococcus worldwide, their phenotypes and their resistance marker genotypes during the last two decades: A systematic review. *Infect Genet Evol* **99**, (2022).
28. Buffalo, C. Z. *et al.* Conserved patterns hidden within group A Streptococcus M protein hypervariability recognize human C4b-binding protein. *Nat Microbiol* **1**, (2016).

569 29. MacHeboeuf, P. *et al.* Streptococcal M1 protein constructs a pathological host fibrinogen
570 network. *Nature 2011 472:7341* **472**, 64–68 (2011).
571 30. Kolesinski, P., Wang, K. C., Hirose, Y., Nizet, V. & Ghosh, P. An M protein coiled coil
572 unfurls and exposes its hydrophobic core to capture LL-37. *Elife* **11**, (2022).
573 31. Akerstrom, B., Lindqvist, A. & Lindahl, G. *BINDING PROPERTIES OF PROTEIN ARP,*
574 *A BACTERIAL IgA-RECEPTOR*. *Mokcular Immunology* vol. 28 (1991).
575 32. Oortwijn, B. D. *et al.* Monomeric and polymeric IgA show a similar association with the
576 myeloid FcαRI/CD89. *Mol Immunol* **44**, 966–973 (2007).
577 33. Wang, Y. *et al.* Structural insights into secretory immunoglobulin A and its interaction
578 with a pneumococcal adhesin. *Cell Research 2020 30:7* **30**, 602–609 (2020).
579 34. Breedveld, A. & Van Egmond, M. IgA and FcαRI: Pathological roles and therapeutic
580 opportunities. *Front Immunol* **10**, 553 (2019).
581 35. Sandin, C. *et al.* Isolation and Detection of Human IgA Using a Streptococcal IgA-
582 Binding Peptide. *The Journal of Immunology* **169**, 1357–1364 (2002).
583 36. Bryant, P., Pozzati, G. & Elofsson, A. Improved prediction of protein-protein interactions
584 using AlphaFold2. *Nature Communications 2022 13:1* **13**, 1–11 (2022).
585 37. Fischetti, V. A. M Protein and Other Surface Proteins on Streptococci. in *Streptococcus*
586 *pyogenes : Basic Biology to Clinical Manifestations* (2016).
587 38. Castro, S. A. & Dorfmueller, H. C. A brief review on Group A Streptococcus
588 pathogenesis and vaccine development. *R Soc Open Sci* **8**, (2021).
589 39. Mills, J. O. & Ghosh, P. Nonimmune antibody interactions of Group A Streptococcus M
590 and M-like proteins. *PLoS Pathog* **17**, e1009248 (2021).
591 40. Royle, L. *et al.* Secretory IgA N- and O-Glycans Provide a Link between the Innate and
592 Adaptive Immune Systems. *Journal of Biological Chemistry* **278**, 20140–20153 (2003).
593 41. Kaetzel, C. S. The polymeric immunoglobulin receptor: bridging innate and adaptive
594 immune responses at mucosal surfaces. *Immunol Rev* **206**, 83–99 (2005).
595 42. Brouwer, S. *et al.* Pathogenesis, epidemiology and control of Group A Streptococcus
596 infection. *Nat Rev Microbiol* 1 (2023) doi:10.1038/S41579-023-00865-7.
597 43. Husmann, L. K., Yung, D. L., Hollingshead, S. K. & Scott, J. R. Role of putative
598 virulence factors of Streptococcus pyogenes in mouse models of long-term throat
599 colonization and pneumonia. *Infect Immun* **65**, 1422 (1997).
600 44. Wessels, M. R., Moses, A. E., Goldberg, J. B. & Dicesare, T. J. Hyaluronic acid capsule is
601 a virulence factor for mucoid group A streptococci. *Proc Natl Acad Sci U S A* **88**, 8317
602 (1991).
603 45. Kluytmans, J., Van Belkum, A. & Verbrugh, H. Nasal Carriage of Staphylococcus aureus:
604 Epidemiology, Underlying Mechanisms, and Associated Risks. **10**, 505–520 (1997).
605 46. Brooks, L. R. K. & Mias, G. I. Streptococcus pneumoniae's virulence and host immunity:
606 Aging, diagnostics, and prevention. *Front Immunol* **9**, 376210 (2018).
607 47. Langley, R. *et al.* The staphylococcal superantigen-like protein 7 binds IgA and
608 complement C5 and inhibits IgA-Fc alpha RI binding and serum killing of bacteria. *J*
609 *Immunol* **174**, 2926–2933 (2005).
610 48. Moor, K. *et al.* High-avidity IgA protects the intestine by enchaining growing bacteria.
611 *Nature 2016 544:7651* **544**, 498–502 (2017).
612 49. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for
613 rapid unsupervised cryo-EM structure determination. *Nature Methods 2017 14:3* **14**, 290–
614 296 (2017).

615 50. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and
616 analysis. *J Comput Chem* **25**, 1605–1612 (2004).
617 51. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Biological Crystallography Features
618 and development of Coot. doi:10.1107/S0907444910007493.
619 52. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and
620 electrons: recent developments in Phenix. *urn:issn:2059-7983* **75**, 861–877 (2019).
621 53. Tina, K. G., Bhadra, R. & Srinivasan, N. PIC: Protein Interactions Calculator. *Nucleic*
622 *Acids Res* **35**, 473–476 (2007).
623 54. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new
624 ENDscript server. *Nucleic Acids Res* **42**, W320–W324 (2014).
625 55. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence
626 alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
627

628

SUPPLEMENTARY INFORMATION

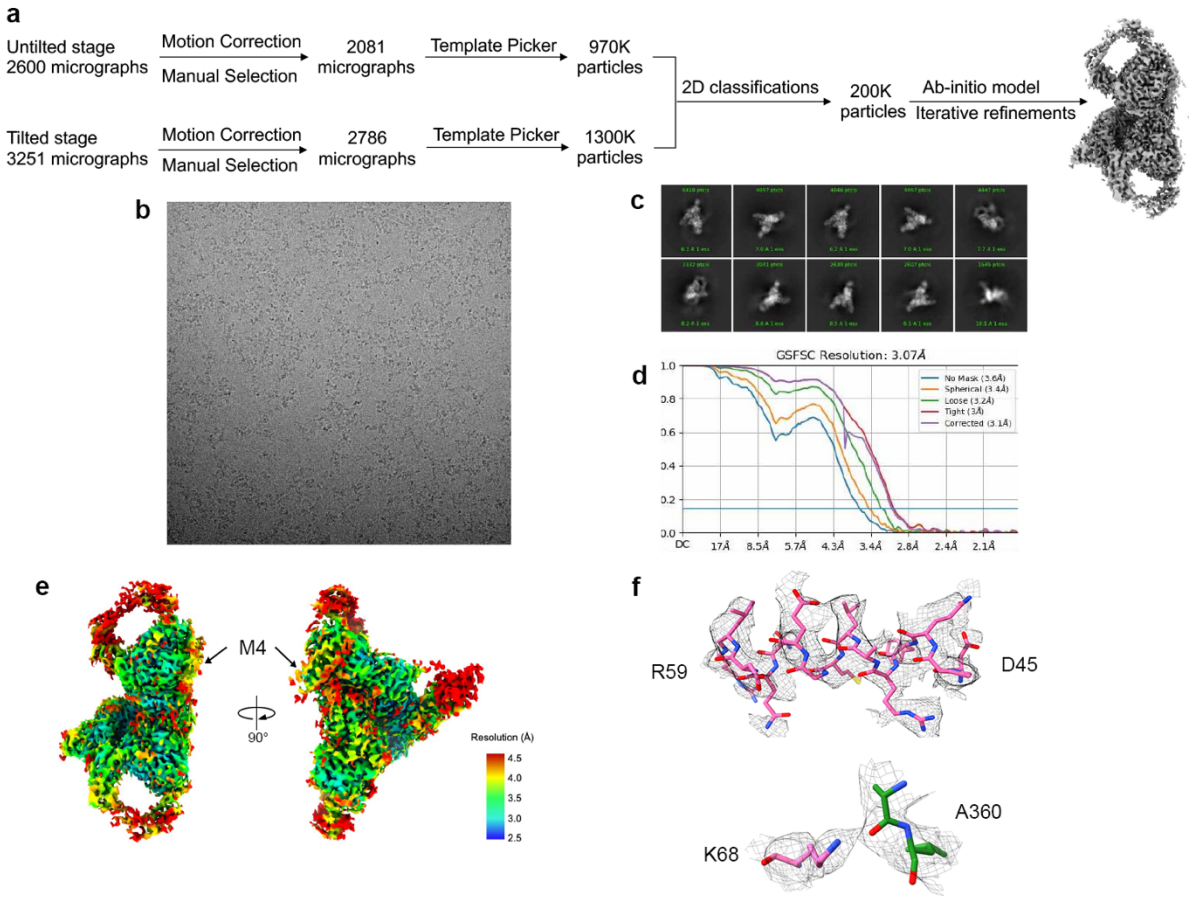


Figure S1. CryoEM data collection and CryoSPARC processing pipeline. (a) Schematic summary of data processing pipeline of SIgA-M4 complex in CryoSPARC; (b) representative micrograph; (c) representative 2D class averages; (d) FSC curve for the final reconstruction with reported resolution at FSC=0.143 shown by the blue horizontal line. (e) Local resolution map calculated by cryoSPARC and colored in Chimera. (f) Density from M4 helix-a residues 45-49 and interactions between M4 K68 and SIgA A360 contoured to 3σ and carved at 2 Å cutoff.

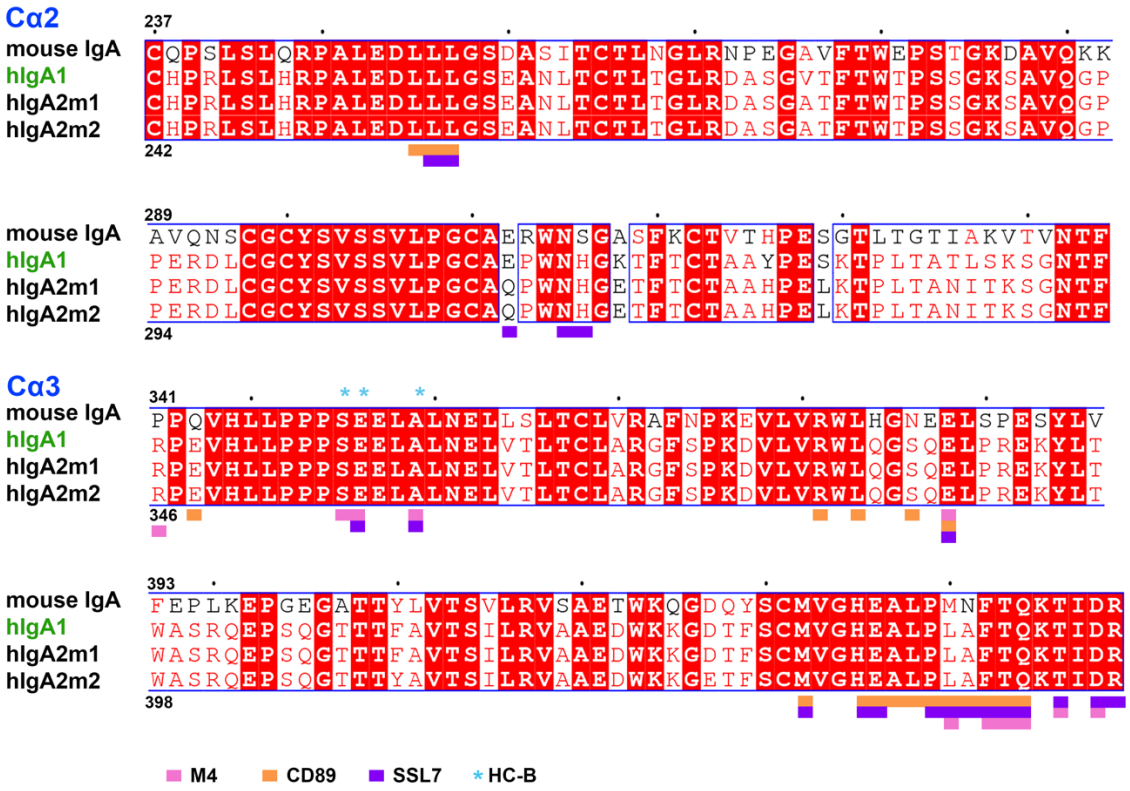


Figure S2. Sequence alignment of human and mouse IgA heavy chains. Sequence alignment including human IgA subtypes (hIgA1, hIgA2m1 and hIgA2m2) and mouse IgA. Identical residues are colored in red. Each square denotes an interfacing residue described in the main text. Squares are colored according to the key at the bottom of the figure, blue asterisks indicate that the residue is on heavy chain-B.

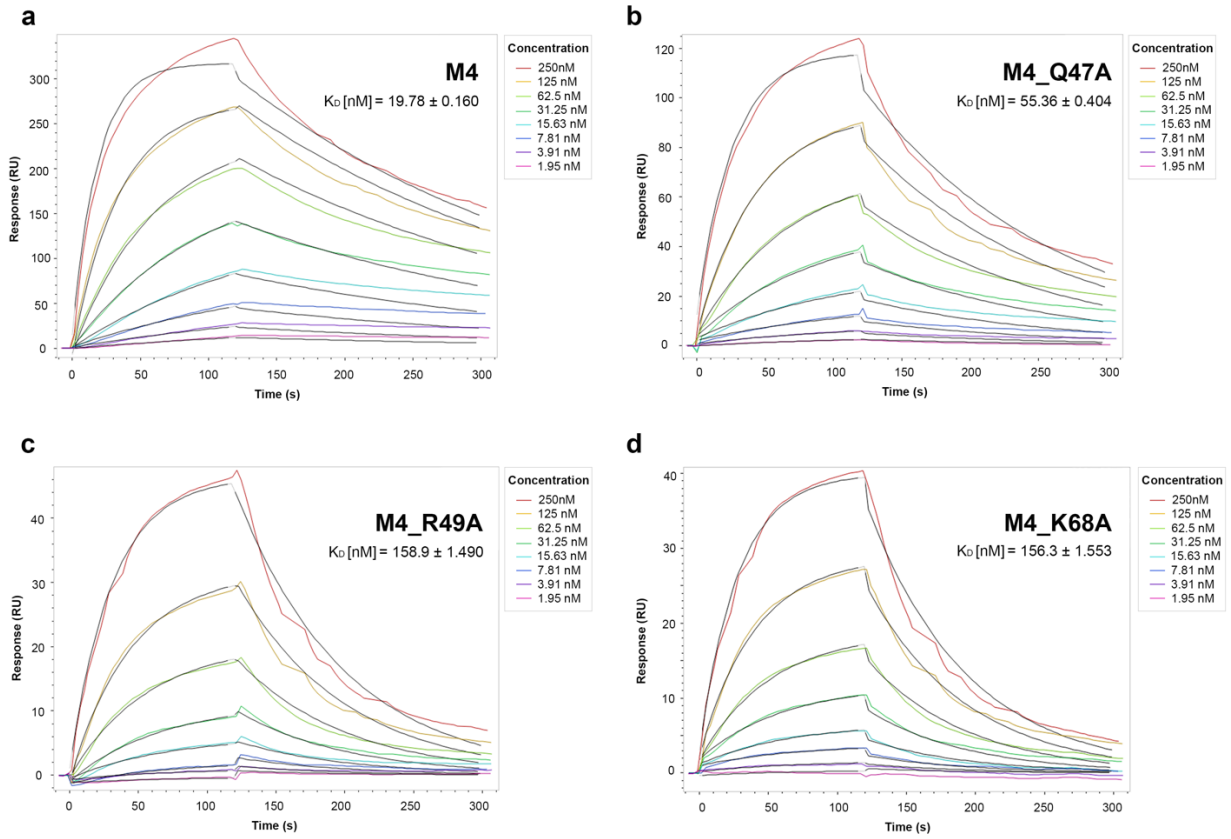


Figure S3. M4 binding analysis by surface plasmon resonance (SPR). Sensorgrams showing the response of M4 (a) and individual M4 point mutants, Q47A (b), R49A (c), and K68A (d) binding to immobilized hSIgA1. All data were collected on a surface immobilized with 0.2uM hSIgA1 and all analytes were tested with a two-fold dilution series starting from 0.25uM. The experimental curves were fit with 1:1 binding models, shown in black.

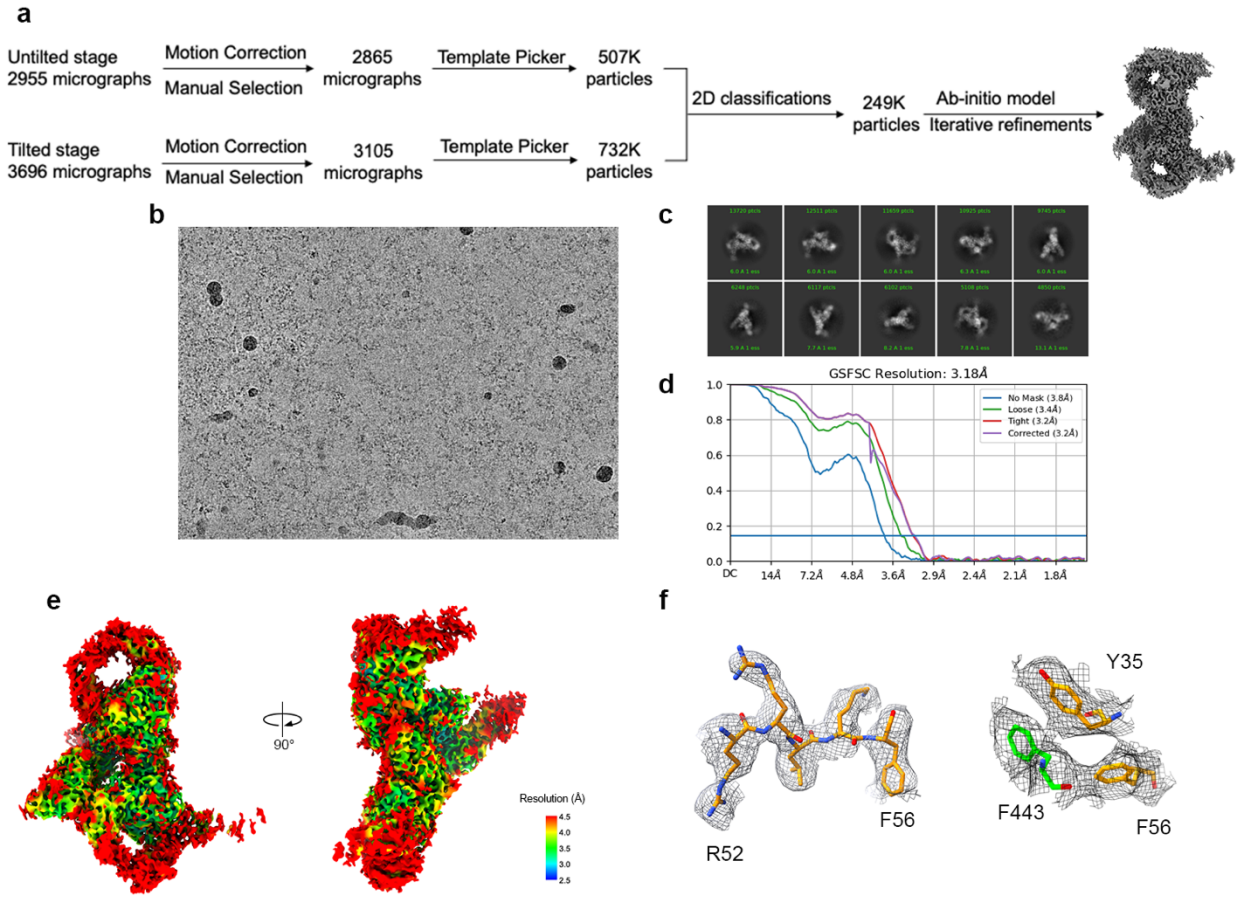


Figure S4. SIgA-CD89 CryoEM data collection and CryoSPARC processing pipeline. (a) Schematic summary of data processing pipeline of SIgA-CD89 complex in CryoSPARC; (b) representative micrograph; (c) representative 2D class averages; (d) FSC curve for the final reconstruction with reported resolution at FSC=0.143 shown by the blue horizontal line. (e) Local resolution map calculated by cryoSPARC and colored in Chimera; (f) density from CD89 residues 52-56 and part of the hydrophobic core of the SIgA-CD89 interface contoured to 3σ and carved at 2 Å cutoff.

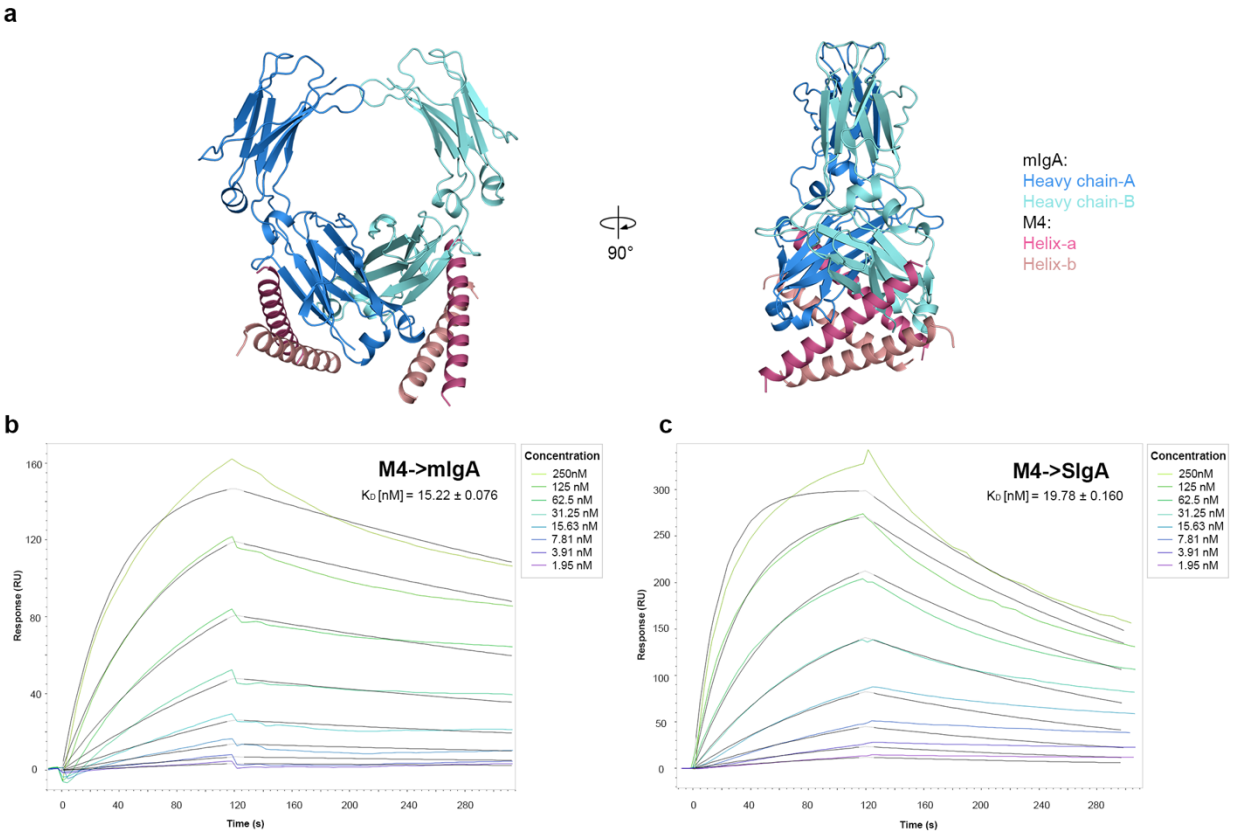


Figure S5. M4 binding to mIgA and SIgA (a) Model of M4 (pink) bound mIgA (blue) shown in two orientations. The model was made by aligning two copies of HC-A (bound to M4) from the M4-SIgA structure to the two IgA heavy chains from the Fcα-CD89 crystal structure (PDB code 1OW0 with CD89 removed). M4-HC-A contacts observed in the model are equivalent to those observed in the SIgA-M4 structure; however, Tp residues, including HC-A K454, which are ~ 4.5Å from M4 helix b residues D67 and E70 and could contribute to electrostatic interactions, are not ordered in 1ow0. (b) SPR Sensorgrams showing the response of M4 binding to immobilized hmIgA1. (c) SPR Sensorgrams showing the response of M4 binding to immobilized hSIgA1. Data were collected on a surface immobilized with 0.2uM hSIgA1 or 0.2uM hmIgA1 and all analytes were tested with a two-fold dilution series starting from 0.25uM. The experimental curves were fit using a 1:1 binding model, shown in black.

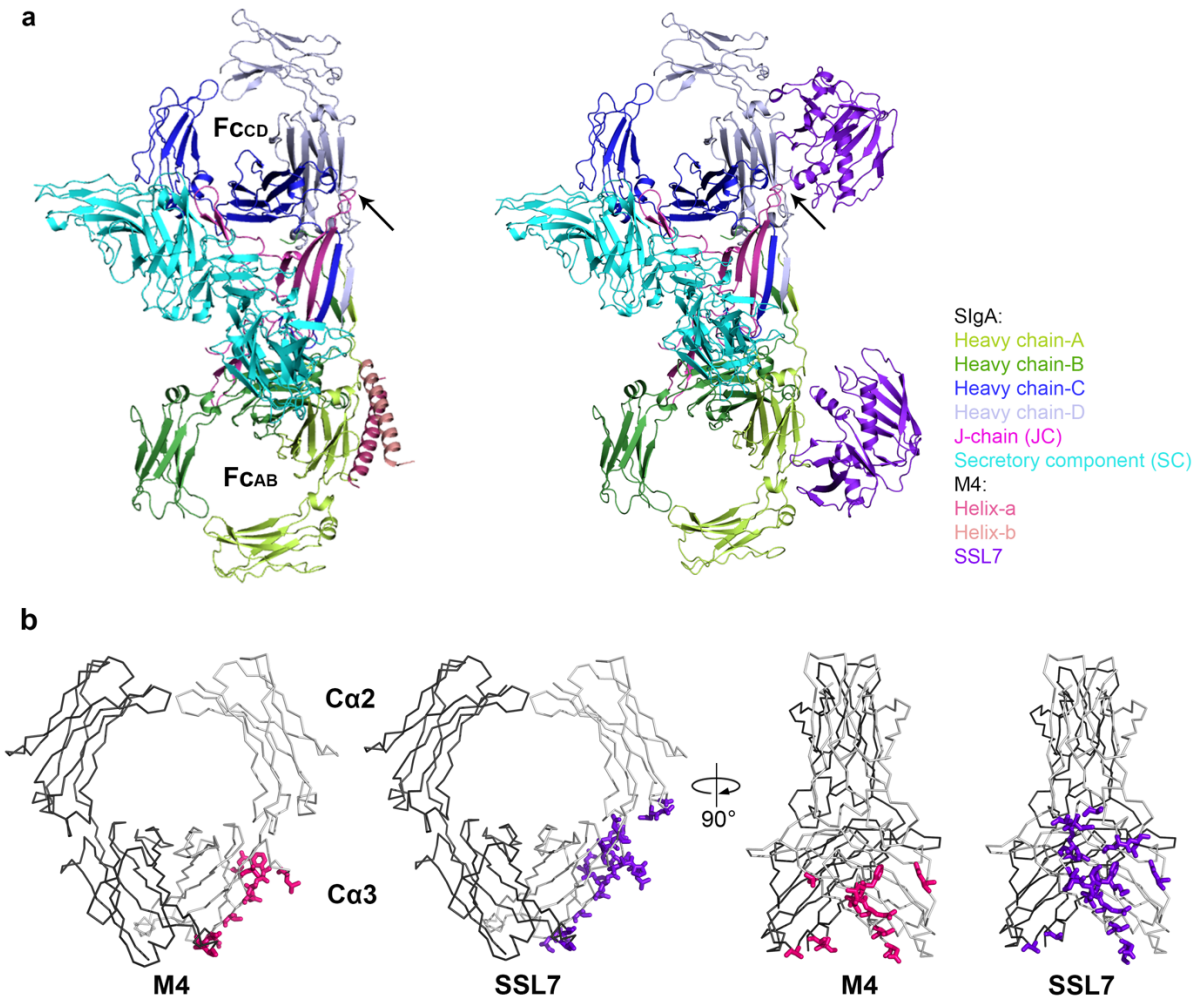


Figure S6. Comparison of M4 and SSL7 binding to SIgA. (a) The structure of M4-SIgA and model of SSL7 bound to SIgA. The SSL7-SIgA model was built by aligning the cocrystal structure of an SSL7-Fcα complex (PDB: 2QEJ) to both copies of SIgA Fcα. M4 only binds at Fc_{AB}, as shown in Fig. 5 and the second site (Fc_{CD}) is occluded by JC (magenta) whereas SSL7 (purple) bound to the second site (Fc_{CD}) is not overlapping with any part of SIgA, indicating that binding is feasible. The JC loop that occluded M4 binding is indicated by an arrow. (b) Comparison of M4 and SSL7 binding sites on Fcα. Monomeric Fcαs (PDB code: 1OW0 with CD89 removed) are shown as ribbons and binding residues are shown as stick representations. The 11 residues bound by M4 are shown in pink and the 20 residues bound by SSL7 are shown in purple. Nine Cα3 residues are bound by both M4 and SSL7 including: E357 (HC-B), A360 (HC-B), E389, L441, F443, T444, Q445, T447, D449. The other 11 residues bound by SSL7 interface are: L257, L258, E313, N316, H317, M433, H436, E437, P440, A442, R450.

688 **Table S1. Cryo-EM data collection, refinement, and validation statistics**
689

	#1 SIgA:M4 (EMDB-40568) (PDB 8SKV)	#2 SIgA:CD89 (EMDB-40567) (PDB 8SKU)
Data collection and processing		
Magnification	130K	165K
Voltage (kV)	300	300
Electron exposure (e–/Å^2)	50	71
Defocus range (μm)	-0.8 to -2.0	-0.8 to -2.0
Pixel size (Å)	0.95	0.82
Symmetry imposed	C1	C1
Initial particle images (no.)	2270K	1240K
Final particle images (no.)	200K	249K
Map resolution (Å)	3.07	3.18
FSC threshold	0.143	0.143
Map resolution range (Å)	2-5	2-5
Refinement		
Initial model used (PDB code)	6ue7	6ue7, 1ow0
Model resolution (Å)	3.3	3.7
FSC threshold	0.5	0.5
Map sharpening *B* factor (Å^2)	185	170
Model composition		
Non-hydrogen atoms	12220	14476
Protein residues	1632	1951
Ligands	NAG:15	NAG:12
B factors (Å^2)		
Protein	40.39	52.26
Ligand	67.15	74.18
R.m.s. deviations		
Bond lengths (Å)	0.002	0.005
Bond angles (°)	0.544	0.642
Validation		
MolProbity score	1.67	1.83
Clashscore	5.26	7.13
Poor rotamers (%)	0.32	0.2
Ramachandran plot		
Favored (%)	94.32	93.15
Allowed (%)	5.77	6.85
Disallowed (%)	0	0

690