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 accessible. 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.

INTRODUCTION

Immunoglobulin (Ig) A is an antibody class expressed in mammals, where it functions in monomeric (m) and secretory (S) forms; the mIgA is the predominant form in serum whereas SIgA is the most abundant antibody present in mucosal secretions, where it protects the host epithelium against microbes and toxins[1]. Both forms of IgA contain fragment antigen-binding (Fab) and fragment crystallizable (Fc) regions. The Fab is responsible for binding to antigens, whereas the Fc region interacts with other components of the immune system, including IgA Fc Receptors (FcαRs) such as CD89 (a.k.a. FcαR1), that are expressed on the surface of some immune system cells and can promote downstream effector functions such as phagocytosis and respiratory burst[2]. Whereas the binding of host FcαRs is likely to elicit a favorable immune response, IgA can also be bound by bacterial proteins, for example, M4 (a.k.a. Arp4) and M22 (a.k.a Sir22) from *Streptococcus pyogenes* and SSL7 from *Staphylococcus aureus*, interactions that allow pathogenic bacteria to evade host IgA responses[3–6]. Despite significance, the functional consequences of IgA interactions with host FcαRs and bacterial proteins, sometimes referred to as decoys, remains poorly understood.

The monomeric and secretory forms of IgA that mediate host interactions with antigens, FcαRs and bacterial proteins adopt at least two distinct molecular structures. The mIgA contains two IgA heavy chains and two light chains, resulting in two Fabs and one Fcα. The Fcα contains two copies of immunoglobulin constant domains, Cα2 and Cα3, and adopts a rigid structure linked to two Fabs. Fcα contains two FcαR binding sites, one located at each Cα2 and Cα3 interface (Fig. 1a). In contrast, SIgA is polymeric and is typically composed of two mIgA that assemble with one joining-chain (JC) to form dimeric (d) IgA, which is bound by the polymeric Ig-receptor ectodomain, called secretory component (SC). The structures of dIgA and SIgA, reported in 2020, revealed two Fcα connected through four C-terminal motifs called tailpieces (Tps) and one JC, which contacts each heavy chain uniquely [7–9]. This results in two structurally unique Fcαs that

can be designated as Fc_{AB} (Fcα containing heavy chains A and B) and Fc_{CD} (Fcα containing heavy chains C and D). Together, Fc_{AB} and Fc_{CD} adopt an asymmetric bent and tilted conformation with distinct concave and convex sides. In SIgA, the SC is bound toward one side and protrudes from the center of the molecule contributing further asymmetry. Overall, the SIgA conformation is predicted to constrain the possible positions that Fabs can adopt and leave two FcαR binding sites accessible (Fig. 1b). Notably, these two sites are located on the convex side, one on Fc_{AB} and one on Fc_{CD}, whereas sites on the concave side are occluded by the JC and the SC. Thus, SIgA is presumed to have the same number of accessible FcαR binding sites as mIgA but instead of two sites on one Fcα that are horizontally oriented relative to each other, it has one site on two Fcαs that are vertically oriented relative to each other (Fig. 1)[7]. The functional significance of these differences remains a topic of investigation; however, it is notable that two receptor binding sites have been evolutionary preserved in both mIgA and SIgA.

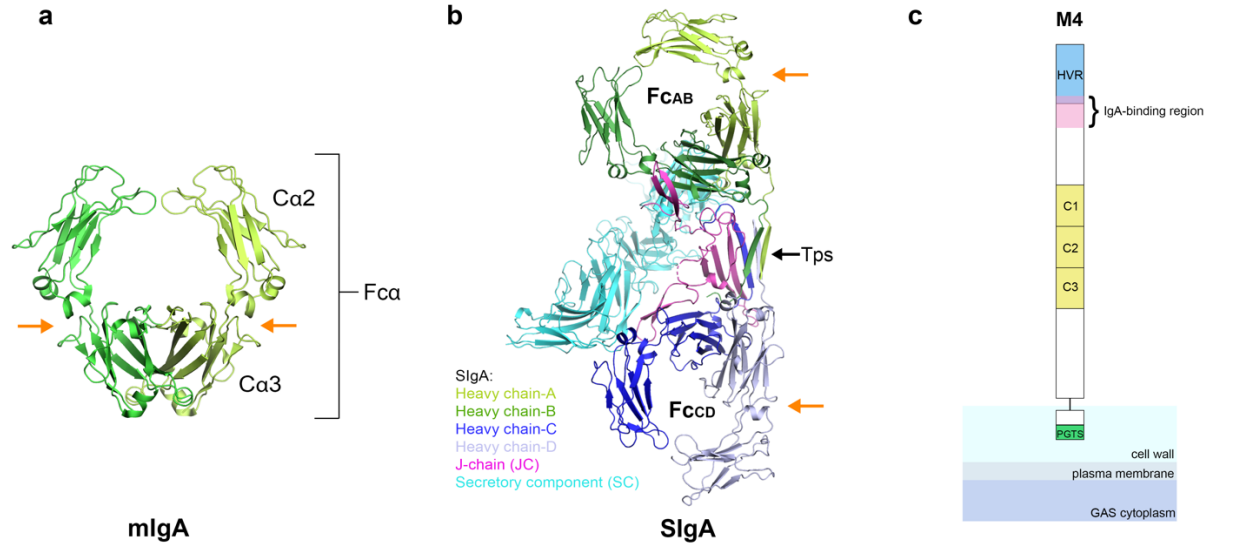


Figure 1. Structures of mIgA, SIgA, and schematic of M4. (a) The structure of mIgA lacking Fabs (Fcα; PDB code 1OW0 with CD89 removed). (b) The structure of human SIgA lacking Fabs (PDB code 6UE7). Structures are colored according to the key; one Cα2 and one Cα3 domain are labeled along with Fc_{AB} 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).

61 Both mIgA and SIgA can be bound by group A streptococcus (GAS; *Streptococcus*
62 *pyogenes*). GAS has long been considered a significant human pathogen that continues to be
63 one of the top ten causes of mortality worldwide from infectious diseases[10]. The clinical
64 manifestations of this Gram-positive bacterium are wide-ranging and diverse, encompassing both
65 mild and common local infections, such as tonsillitis, impetigo, as well as life-threating systemic
66 invasions like toxic shock and sepsis[11,12]. Although tonsillitis caused by GAS is much more
67 common than invasive GAS infections, with a 1000-fold higher incidence, it still contributes
68 significantly to the global burden of disease, resulting in over 160,000 deaths each year[13]. The
69 prevalence of severe GAS disease is at least 18.1 million cases, with 1.78 million new cases each
70 year[10]. Globally, GAS poses a significant threat to public health, causing both morbidity and
71 mortality and thus efforts to understand and limit its virulence are of high interest.

72 The most extensively studied GAS virulence factors are a family of serotype-specific
73 proteins called M proteins, which include IgA binding proteins M4 and M22. M proteins adopt a
74 coiled-coil structure that protrudes up to 60 nm from the bacterium cell and can bind a variety of
75 host extracellular matrix and serum proteins[14,15]. M proteins are attached to the bacterial cell wall
76 through a C-terminal proline-glycine-threonine-serine (PGTS)-rich domain and a cell-wall-
77 associated domain, which are followed by an ectodomain that includes several repetitive amino
78 acid sequence elements and a N-terminal hypervariable domain (HVR) (Fig. 1c). Variability in the
79 M protein ectodomain is correlated with the ability to bind specific host factors; for example, some
80 M proteins bind IgG, some IgA, some fibrinogen, and some other host factors[16,17]. These
81 interactions are thought to contribution to the bacterium's virulence primarily by modulating the
82 host immune response (e.g. limiting phagocytosis). Although M proteins have been shown to be
83 immunogenic, developing a vaccine that targets M proteins has proven difficult due to the
84 extensive diversity in the protein's most immunogenic region located at the N-terminus and
85 therefore developing other approaches to block the functions of M proteins is of interest [18].

86 M4 and M22 are the two M proteins found in GAS strains M4 and M22, respectively, both
87 of which can bind IgA[5,19–21]. Mutational analysis of both M4 and M22 revealed a 29-residue,
88 minimal IgA-binding region located within the ectodomain and including five C-terminal residues
89 of the HVR with similar, but not identical, sequences[6] (Fig. 1c). M22 also binds IgG at a site
90 overlapping with the IgA binding site[5]. Despite significance, how M proteins interact with host
91 ligands remains unclear. Existing structural data includes crystal structures of an M1 fragment
92 containing the HVR and a repetitive sequence element within the ectodomain and the HVR of M2,
93 M22, M28, and M49, none of which detail the structure of the IgA-binding region or IgA-M protein
94 interactions [22,23]. Published mutational analysis indicated that M4 and M22 binding to IgA is
95 dependent on residues located at the Cα2/Cα3 interface, including residues bound by host CD89
96 and *S. aureus* SSL7[2–4]. The overlapping interface of M4 and CD89 suggested that M4 may
97 interfere with IgA effector functions and studies suggested that M4 protein can block respiratory
98 burst in neutrophils *in vitro*[24]. However, the role of M4 (and CD89) in disease pathogenesis in the
99 host remains poorly understood and published studies have focused on mIgA despite mucosal
100 routes of entry. Indeed, pathogen interactions with mIgA are most likely to occur in the context of
101 invasive disease whereas interactions with SIgA are most likely to occur during mucosal
102 colonization.
103 To investigate how M proteins interact with SIgA and how those interactions compare to
104 host FcαRs, we determined two Cryo-EM structures, one of SIgA in complex with M4 and one of
105 SIgA in complex with CD89. Together structures indicate that in contrast to previously proposed
106 models, M4 and CD89 share only a subset of IgA interface residues and bind through different
107 angles of approach. This leads to differences in binding stoichiometry, with M4 occupying only
108 one FcαR binding site on SIgA, while leaving the other site sterically accessible to CD89 and
109 likely other FcαRs.
110

111 **RESULTS**

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

113 In order to better understand how bacterial virulence factors such as M4 engage SIgA, we
114 designed an M4 expression construct encoding the ectodomain (residues 1-314) and combined
115 resulting protein with purified, recombinant human SIgA1 to create an SIgA-M4 complex. Using
116 this complex, we determined the SIgA-M4 cryo-EM structure to an average resolution of 3.1 Å
117 (FSC=0.143) (Fig. 2a). Map quality and resolution were sufficient to refine the positions of main
118 chain and side chain atoms for the majority of residues in the complex with the exception of Fab
119 residues, which were disordered (Fig. S1). The final structure included a single M4 dimer bound
120 to one SIgA; M4 residues 46-74, which include the entire known IgA-binding domain, were built
121 whereas the positions of residues more distal from the binding site were poorly resolved or
122 disordered, likely due to the inherent flexibility of the M4 dimer's extended, coiled-coil structure.
123 SIgA is superimposable with a previously published human SIgA1 structure with a RMSD of
124 0.708[8].

125 The M4 adopts an α-helical, parallel coiled-coil homodimer as observed for other M
126 proteins [16]. The M4 fragment termini are stabilized by hydrophobic interactions including two ideal
127 heptads, characterized by having Tyr, Ile, or Leu, at both *a* and *d* positions[25]. Hydrophobic
128 residues at the *a* and *d* positions (Y49, L52, I70, L73) are packed against each other and stabilize
129 the coiled-coil (Fig. 2b and 2c). We also observed polar interactions stabilizing the center of the
130 M4 fragment, including E55 which forms a hydrogen bond with N56, and Y66 which forms a
131 hydrogen bond with Q67 (Fig. 2c). A subset of residues between P47-K69 in both helices contact
132 SIgA, with one (helix-a) forming the majority of contacts with the IgA heavy chain and the other
133 (helix-b) using equivalent residues to contact solvent and a smaller interface with IgA. Prior
134 studies, including mutational analysis, identified M4 residues 46-74 as the minimal IgA binding
135 region[6]; our observations indicate that residues 46, 71, 72, and 74 contribute indirectly by
136 stabilizing the M4 coiled-coil dimer.

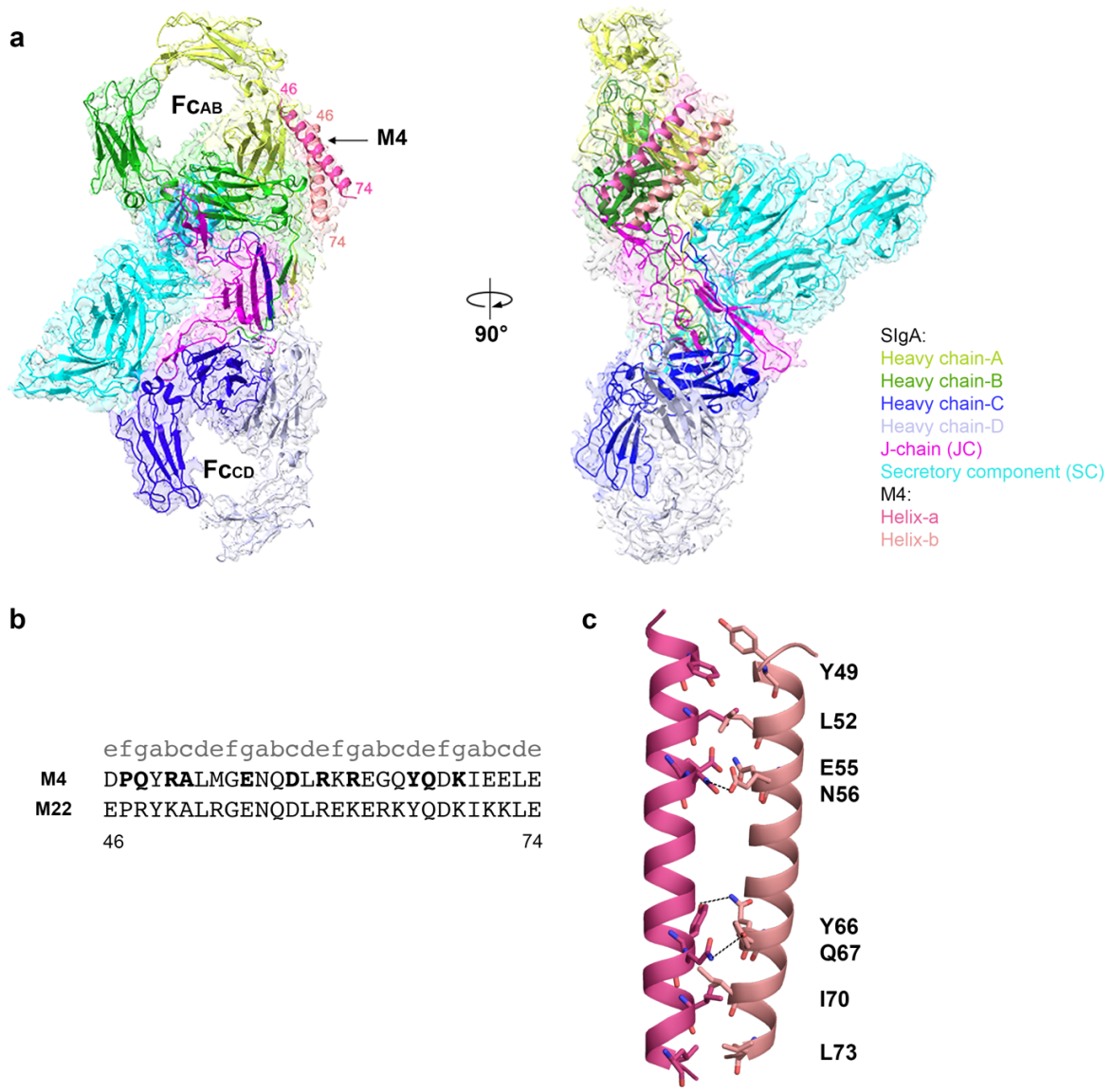


Figure 2. SIgA-Apr4 Structure. (a) The Cryo-EM structure of the M4 fragment (residues 46-74) in complex with human SIgA is shown as a cartoon with partially transparent map in two orientations. The model is colored according to the key, M4 is labeled, and Fabs are disordered. (b) Sequence alignment of 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) Interchain interactions between helix a and b of M4. Polar interactions are shown as dotted lines.

137

138 **The structure of the SIgA-M4 interface**

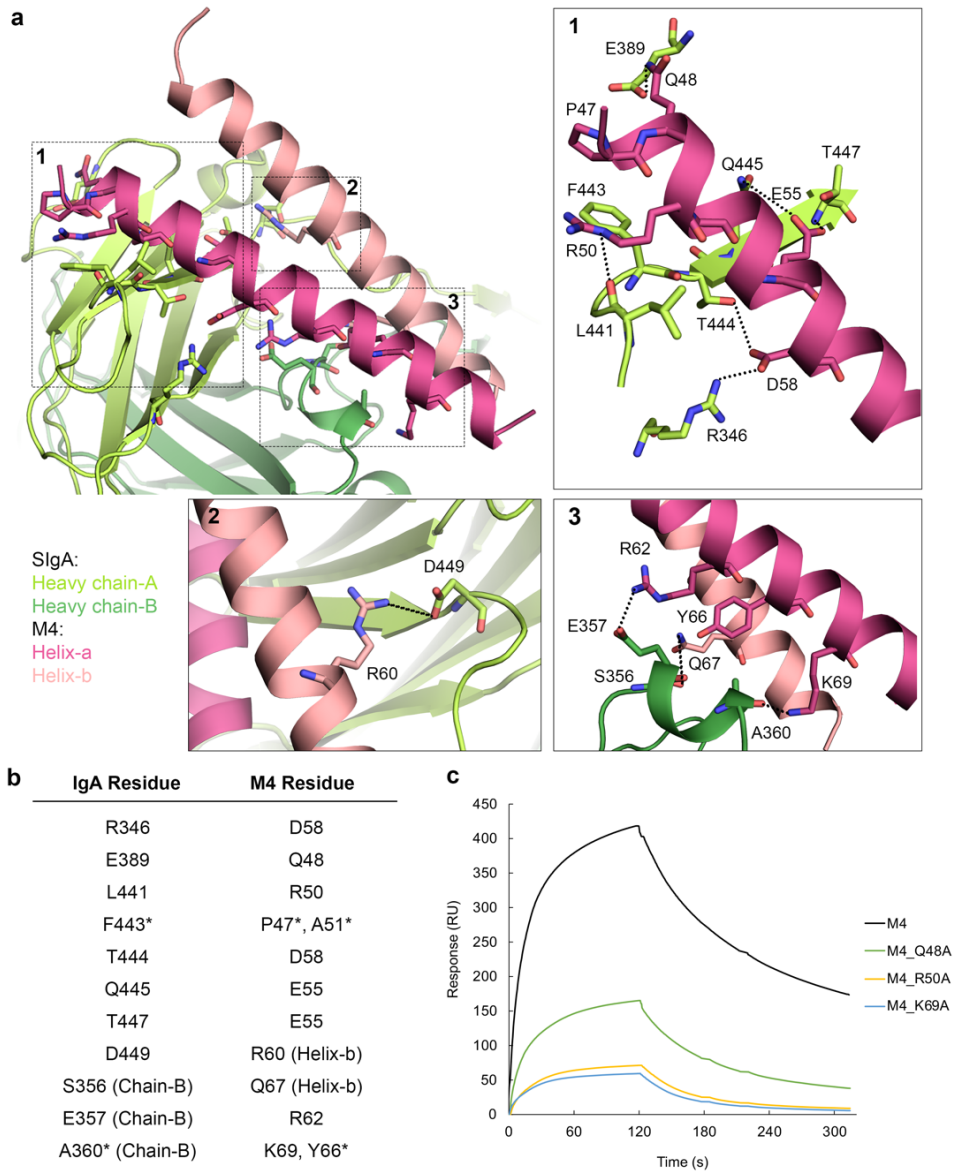
139 Despite the reported accessibility of two FcαR binding sites on SIgA, our structure reveals M4

140 bound only to Fc_{AB}[7]. The M4 interface includes 11 residues on the Cα3 domain of both heavy

141 chain-A (HC-A) and heavy chain-B (HC-B) (Fig. 3), with a total buried surface area of 1000 Å^2.
142 Participating IgA residues are predominantly located along one b-strand in HC-A (F443, T444,
143 Q445, T447), forming hydrogen bonds with M4 helix-a residues R50 and E55. HC-A E389 and
144 R346 also reach out to interact with M4 helix-a residues Q48 and D58 respectively, and HC-A
145 D449 forms a salt bridge with R60 on helix-b of M4. The Apr4 interface with IgA HC-B includes
146 only residues S356, E357 and A360, which form hydrogen bonds with Q67 of helix-b and R62,
147 K69 of helix-a respectively. All IgA interface residues are conserved in human IgA subtypes (IgA1,
148 IgA2m1, IgA2m2) and most are conserved in mouse IgA, with the exception of R346 and L441(Fig.
149 S2). Mouse IgA is not known to bind M4 thus R346 and L441 appear to be especially important
150 for M4 binding and structurally equivalent residues in mouse IgA, P341 and M436 respectively,
151 do not appear to support the interaction [26].
152
153 **Mutagenesis confirms the key binding residues on M4**
154 Although side chain density was apparent in the structure, the N and C termini of the M4 coiled-
155 coil were not resolved and therefore to validate residue assignments and investigate the
156 contributions of individual M4 residues on SIgA binding, we conducted structure-based mutational
157 analysis and surface plasmon resonance (SPR) binding assays. Based on the pairwise
158 interactions between M4 and SIgA (Fig. 3b) and pairwise interactions between M4 helix-a and
159 helix-b, we made three individual M4 point mutations Q48A, R50A, K69A. The subset of residues,
160 in which one copy was interacting with IgA and the other copy was stabilizing the coiled-coil were
161 not mutated due to the possibility of disrupting the M4 dimer. All mutated M4 variants behaved
162 equivalent to the wildtype M4 as assayed by SDS-PAGE and size exclusion elution
163 chromatography (not shown). M4 and mutant variant binding to SIgA were determined by SPR.
164 Sensorgrams revealed lower responses for all M4 mutants when compared to concentration
165 matched wild-type M4, indicating markedly reduced binding affinity (Fig. 3c and Fig. S3). Mutant
166 variants R50A and K69A exhibited the most pronounce reduction in IgA-binding, consistent with

167 the observation that R50 interacts with one of the human-IgA specific binding site residues, L441.

168 K69 interacts with HC-B, indicating that contacts with both HC-A and HC-B chains are important.



IgA Residue	M4 Residue
R346	D58
E389	Q48
L441	R50
F443*	P47*, A51*
T444	D58
Q445	E55
T447	E55
D449	R60 (Helix-b)
S356 (Chain-B)	Q67 (Helix-b)
E357 (Chain-B)	R62
A360* (Chain-B)	K69, Y66*

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 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 Q48A (green), R50A (yellow), and K69A (blue) binding to immobilized human SIgA1. All analyte concentrations shown are 0.5 uM. Sensorgrams including a complete concentration series for each analyte can be found in Figure S3.

169

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

171 The unexpected stoichiometry of the SIgA-M4 complex raised the question of how host FcαRs
172 bind SIgA. Modeling based on the structure of mIgA in complex with CD89 and SIgA structures
173 suggested that two CD89 can bind SIgA [8]; however, to verify this possibility, we determined the
174 Cryo-EM structure of SIgA with CD89 to a resolution of 3.2 Å (Fig. 4a). The ectodomain of CD89
175 consists of two immunoglobulin-like domains (D1 and D2) arranged at an angle of approximately
176 90 degrees to each other. Map quality and resolution was sufficient to refine the positions of main

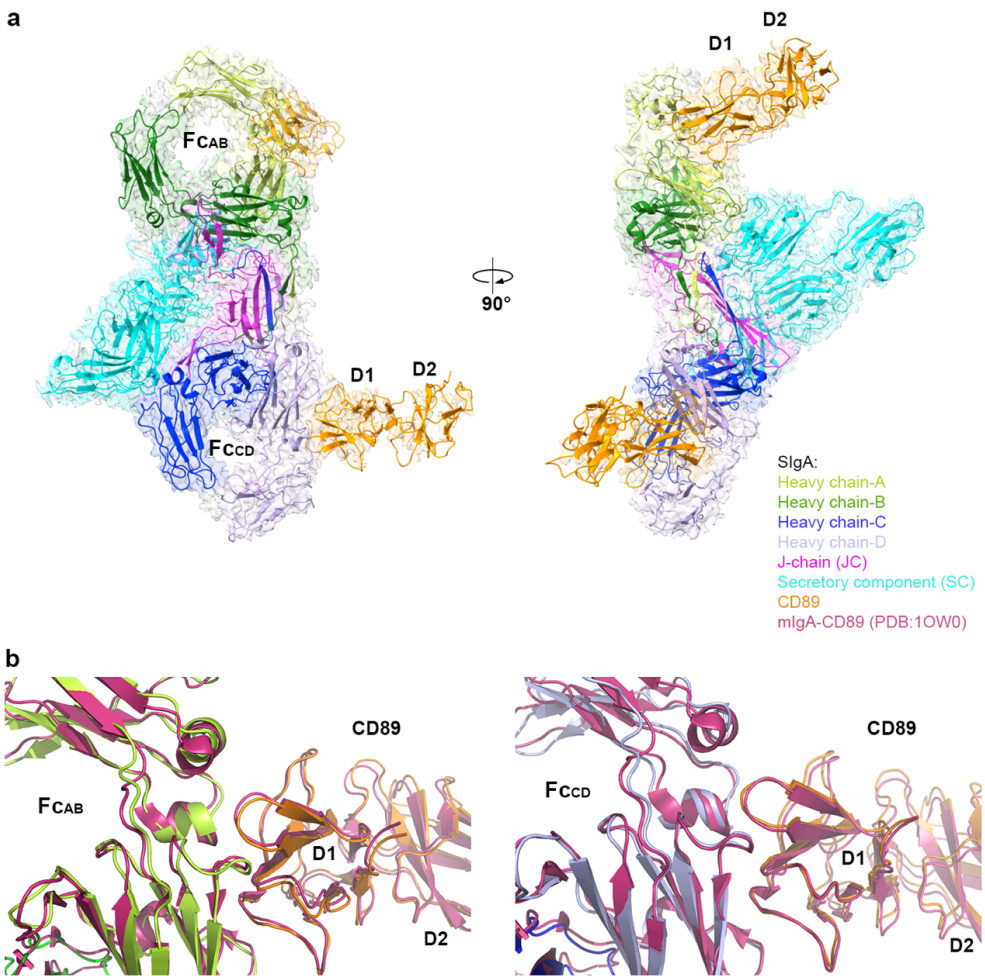


Figure 4. SIgA-CD89 structure. (a) The Cryo-EM structure of CD89 in complex with human SIgA is shown as a cartoon with partially transparent map in two orientations. The antigen binding fragments (Fabs) are disordered. The two Ig-like domains of CD89 are labeled D1, D2. CD89 D1 contacted a total of 19 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. All models are colored according to the key.

177 chain and side chain atoms for the residues at the interface, while density in the distal region of
178 CD89, including the second Ig-like domain D2, was of lower resolution (Fig. S5). The refined
179 structure revealed a 2:1 CD89:SIgA binding stoichiometry verifying that two copies of CD89 can
180 bind both mIgA and SIgA (Fig. 4a). Alignment to the mIgA-CD89 structure revealed
181 superimposable binding sites (RMSD=1.165 at Fc_{AB}; RMSD=1.121 at Fc_{CD}) on Cα2 and Cα3 [2].
182 Consistent with the mIgA-CD89 structure, the CD89 D1 BC loop, the D strand, the DE loop, and
183 the FG loop (Ig-like domain nomenclature) contacted a total of 19 IgA residues, including three
184 Cα2 residues and 16 Cα3 residues, burying a total of 1656 $Å^2$ (Fig. 4b).
185
186 **CD89 and M4 share overlapping but distinct SIgA binding sites**
187 Comparison of the CD89 and M4 binding footprints on SIgA revealed five shared interface
188 residues (E389, L441, F443, T444, Q445) located on the Cα3 domain. CD89 binds additional
189 residues located toward and on Cα2, whereas M4 contacts only Cα3 and binds residues closer
190 to the Tps and JC, including R346 from HC-A and S356, E357, A360 from HC-B (Fig. 5a). CD89
191 and M4's distinct binding footprints and structures direct CD89 D2 C-terminal residues away from
192 SIgA and direct the M4 fragment C-terminal residues (and likely the rest of the M4 coiled-coil),
193 alongside Fc_{AB} in a line near parallel to Cα3 (Fig. 5b). Noted differences contribute to the distinct
194 stoichiometry of CD89 versus M4 binding to SIgA. In the case of M4, the inclusion of S356, E357,
195 A360 and R346 in the M4 binding interface, together with its extended coiled-coiled topology,
196 enforce the surprising 1M4:1SIgA stoichiometry because the steric accessibility of these four
197 residues and neighboring residues differs between Fc_{AB} and Fc_{CD}. In Fc_{AB}, M4 can bind S356,
198 E357, A360 and R346 with residues on both helices and extend the coiled-coiled over neighboring
199 residues and toward the Tps. In Fc_{CD}, S356, E357, A360 and R346 remain solvent accessible,
200 but JC loop residues 24-32 bind the adjacent region, blocking the path occupied by M4 (helix-b
201 residues 70-73) and therefore occluding binding (Fig. 5b). The JC loop that occludes M4 binding

202 does not interfere with CD89 binding on Fc_{CD}, implying that sterically, CD89 and M4 ectodomains
203 could bind to the same copy of SIgA, with M4 binding Fc_{AB} and CD89 binding Fc_{CD} (Fig. 5c);
204 however, whether this would occur in the context of bacterial and host cell interactions remains
205 to be determined.

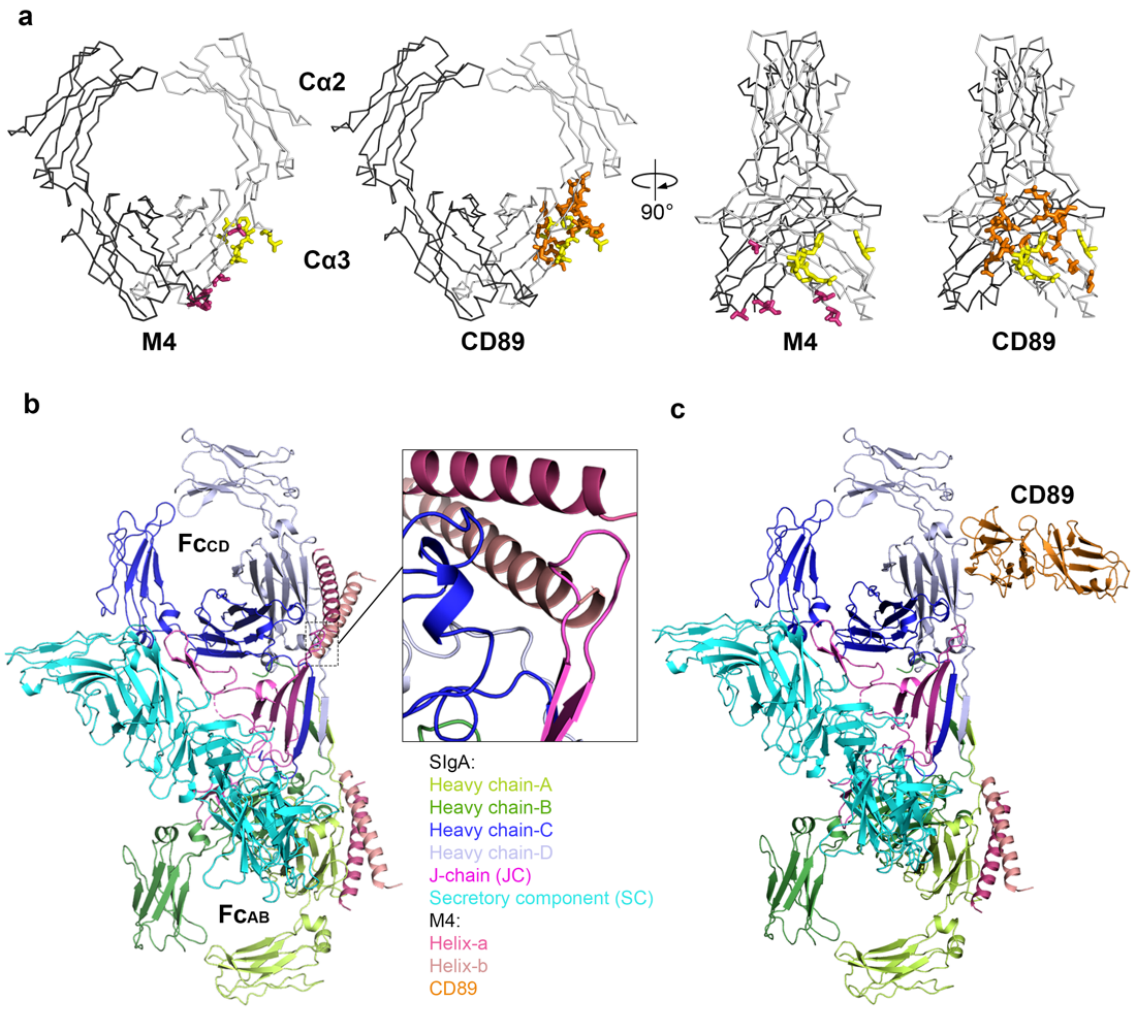


Figure 5. M4 and CD89 binding site comparison and modeling. (a) Comparison of M4 and CD89 binding site on Fcα shown in two orientations. Fcα is shown as ribbons and binding residues are shown as stick representation. The 11 residues bound by M4 are shown in pink and yellow and the 19 residues bound by 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.

206

207

208 **DISCUSSION**

209 Data presented here provide new structural and mechanistic insights on how host FcαRs
210 and bacterial surface proteins engage SIgA, informing models for IgA effector functions, pathogen
211 virulence and host-microbe coevolution. Host and bacterial protein binding to IgA has long been
212 appreciated to occur through a conserved "hot spot," localized to the Cα2/Cα3 domain interface
213 found on mIgA, and more recently SIgA. The region supports binding of human FcαRs, including
214 CD89, as well as GAS M22, the β-protein from group B streptococci (GBS) and SSL7 from *S.*
215 *aureus*, presumably through overlapping binding sites [2–6,21,24,27]. Our work validates that CD89
216 can bind SIgA using the same site as reported for mIgA and confirms that both forms of IgA are
217 able to bind two copies of CD89 *in vitro*[2]. Our work also demonstrates that CD89 and M4 (and
218 likely M22) engage SIgA through the "hot spot" but share only five IgA contact residues and exhibit
219 distinct stoichiometry. Thus, while each receptor occludes a similar region on IgA, the modes of
220 binding are different and likely associated with unique functional consequences. In particular, the
221 five shared residues are all located on Cα3 in what is topologically described as the C-strand
222 (residue E389) and FG loop (residues L441, F443, T444 and Q445), indicating that the Cα3 FG
223 loop is critical for the interaction and likely represents the center of the hotspot. As noted, the M4
224 binding footprint extends even further toward the Tps to engage Cα3 residues R346, S356, E357,
225 and A360, which in the context of binding to SIgA appear to enforce the unexpected 1:1 M4:SIgA
226 binding stoichiometry. Notably, R346 from HC-A has not been reported to bind human FcαRs or
227 bacterial IgA binding proteins making it unique to the M4 binding footprint. The inclusion of this
228 site may in part explain why mouse IgA reportedly does not bind M4 or M22 because residues
229 R346 (and neighboring contact residue L441) are not conserved in mouse IgA (Fig. S2) [26,28].

230 Data reported here are applicable to multiple GAS strains. Epidemiology surveillance
231 reports revealed that serotype M4 and M22 strains are abundant and geographically widespread,
232 especially M4, being one of the top five most common serotypes in a systemic review of global

233 distribution of GAS types[29]. As noted, the M22 strain expresses M22, another M protein specific
234 for IgA. We find that M4 residues making direct contact with IgA are highly conserved in M22.
235 Seven out of 11 direct contact residues are identical and the other four residues are similar (Fig.
236 2b), consistent with prior mutational analysis that identified similar IgA binding motifs in M4 and
237 M22. The IgA binding fragment of M22 has been produced as a peptide called Sap, which
238 corresponds to residues 35-83 of M22 with the addition of a C-terminal cysteine residue[28]. Sap is
239 likely superimposable with our structure but extending 10 residues in N and C- terminal directions.
240 M22 also binds human IgG, and the binding of IgA and IgG are mutually exclusive, indicating
241 shared or contiguous binding regions[5].

242 Our work also sheds new light on *S. aureus* SSL7 interactions with IgA, which we explored
243 by comparing the IgA binding footprints of M4 and SSL7, and modeling SSL7 binding to SIgA.
244 The published SSL7- human Fcα structure revealed SSL7 binding 20 residues on both Cα2 and
245 Cα3, nine of which are shared with M4 (Fig S5). Notably, SSL7 binds HC-B Cα3 residues E357
246 and A360 (adjacent and overlapping with M4-bound S356, E357 and A360) positioning part of its
247 binding site closer to the Tps, similar to M4 and in contrast to CD89. The structure of the SSL7-
248 human Fcα complex revealed two copies of SSL7; however, the inclusion of E357 and A360 in
249 the binding site raises the question of whether two copies of SSL7 can bind SIgA, or if the nearby
250 JC loop residues would occlude binding to the Fc_{CD} as observed for M4. To explore this, we
251 modeled SSL7-SIgA by aligning the crystal structure of SSL7-Fcα (PDB:2QEJ) to each Fc on
252 SIgA (Fig. S5a)[4]. Resulting models suggest that SSL7 can bind both Fc_{AB} and Fc_{CD} resulting in
253 2:1 SSL7:SIgA stoichiometry, as observed for CD89:SIgA. This stoichiometry appears possible
254 because SSL7 adopts a flat "disk-like" topography, which at binds IgA using two loops from its
255 Oligonucleotide/Oligosaccharide-Binding (OB)-fold, that positions the C-terminus away from Fcα,
256 similar to CD89 and in contrast to M4[4]. This structure and trajectory likely allows SSL7 to bind
257 SIgA Fc_{CD} E357 and A360 without being occluded by the JC loop as observed for M4 (Fig. S5).

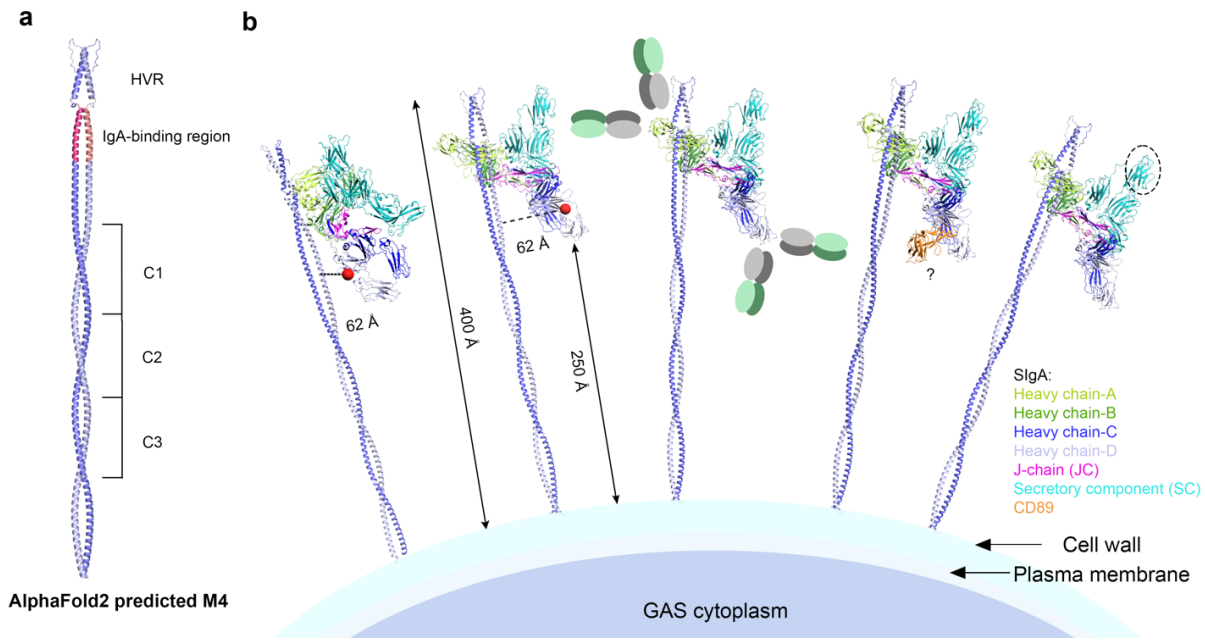


Figure 6. Schematic model of human SIgA binding to M4 on a GAS surface. (a) AlphaFold2-multimer predicted M4 dimer; residues 287-316 are predicted to be disordered and not shown. The IgA-binding region on the predicted model is colored in pink and has an RMSD of 0.640 when aligned to the SIgA-M4 structure. (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.

258

259 Our data and analysis provide unexpected insights on potential functional outcomes of
260 M4-IgA interactions, most notably that among characterized FcαRs and Fcα-binding proteins the
261 sterically enforced 1:1 M4:SIgA stoichiometry appears to be unique to M4 (and presumably, M22).
262 This finding is intriguing because SIgA adopts an asymmetric structure, which published modeling
263 predicts will influence the positions of Fabs and accessibility of the SC and FcαR binding sites[7];
264 therefore, M4 binding to Fc_{AB} positions SIgA differently than if it could bind to Fc_{CD}. To visualize
265 this possibility, we used Alphafold2-multimer to model the full-length M4, aligned it to our M4-SIgA
266 structure and created schematic representations of the complex on a bacterium surface (Fig. 6)[30].
267 Assuming that M4 is oriented approximately perpendicular to the outer membrane surface, M4

268 would extend ~400 Å, placing the bound SIgA ~ 250 Å from the surface where one set of Fabs
269 would be directed upward and away from the bacterium surface and the other set directed toward
270 the surface. In this model, the SC protrudes upward and away from the bacterium surface similar
271 to the M4 HVR, and the Fc_{CD} binding site is located closer to the bacterium surface and
272 approximately 62 Å from M4 (Fig. 6b). Other GAS surface proteins are expected to be shorter
273 than M4, implying that bound SIgA would occupy a position near the outermost surface of the
274 bacterium[31]. While information describing M4 flexibility, as well as the density of M4 and other
275 surface antigens, is needed to fully evaluate the geometric arrangement of SIgA-bound to the
276 bacterial surface, our data and analysis indicate that the 1:1 M4:SIgA stoichiometry will constrain
277 the location and orientation of bound SIgA, a possibility that should be considered when exploring
278 functional outcomes of bacterial-SIgA interactions (Fig. 6).

279 How might SIgA binding to GAS M4 (or M22) in a constrained orientation influence the
280 host or microbe? One possibility is that M4 binding to Fc_{AB} impacts the binding of SIgA Fabs to
281 other streptococcal surface antigens, such as streptococcal pyrogenic exotoxin (Spe), fibronectin-
282 binding proteins (FbI), Streptococcus pyogenes cell envelope proteinase (SpyCEP),
283 streptococcal C5a peptidase (ScpA) and streptolysin O (SLO)[32]. Bacteria that can bind SIgA
284 through M4 (or M22), may be able to limit antibody binding to other surface proteins, although as
285 suggested by others, the relative affinity of antibodies to other surface antigens and their steric
286 accessibility will likely impact the outcome[33]. Another possibility is that SIgA-bound M4 influences
287 the localization of the bacterium (e.g. in host mucosal secretions). SIgA is covered in host glycans,
288 which have been implicated in binding to host lectins, something we speculate could be promoted
289 by the position of SC, which our model predicts would be exposed near the M4 HVR toward the
290 outermost surface of the bacterium (Fig. 6).

291 Another way that M4/M22 binding to Fc_{AB} may impact virulence is by modulating
292 interactions between host FcαR and SIgA bound to GAS. For example, M4 has been shown to
293 reduce IgA-triggered respiratory burst in neutrophils[24]; although the precise mechanism of

294 inhibition remains unclear, CD89 activity is thought to occur when multiple copies of the receptor
295 bind to the antibody Fcα and thus, it has been proposed that M4 occludes CD89 binding thereby
296 blocking downstream effector functions. Our data indicate that mIgA has two accessible M4
297 binding sites and two accessible CD89 binding sites whereas SIgA has one accessible M4 binding
298 site and two accessible, CD89 binding sites. Although the accessibility of these sites on mIgA
299 and SIgA has not been verified on a bacterial surface and it seems plausible that CD89 activity
300 could be inhibited by occluding just one of the two CD89-binding sites, our data raise the intriguing
301 possibility that CD89 (and human FcαRs such as Fcα/μR) might bind Fc_{CD} on SIgA-M4 complexes
302 and promote effector functions in some way (Fig 6b). For example, kinetics that favor CD89
303 binding might allow one copy of CD89 to bind Fc_{CD} and subsequently, another copy to compete
304 with M4 for the Fc_{AB} site. Another possibility is that CD89 binding to Fc_{CD} could promote other
305 copies of CD89 binding to local SIgA that is not bound by M4, but rather is bound to another
306 antigen with its Fabs.

307 The question of how M4 (and M22) binding to SIgA impacts both the host and bacteria
308 must also consider the relative location of interactions with mIgA and SIgA and host immune
309 effector cells. As noted, GAS will be most likely to encounter mIgA in the serum during invasive
310 disease and SIgA in mucosal secretions during non-invasive disease. In the first example, the
311 binding of two copies of M4 to one mIgA could block both CD89 binding sites, limiting the functions
312 of CD89-expressing effector cells (e.g., neutrophils). CD89 effector functions have been studied
313 primarily in the context of mIgA and cells expressing CD89 are more prevalent in the serum than
314 mucosa, raising the possibility that CD89 binding to SIgA or an SIgA-M4 complex on a bacterium
315 surface may be a rare event. However, SIgA has been associated with CD89-expressing
316 polymorphonuclear leukocytes (PMN) activity with the presence of an accessory molecule Mac-
317 1 (CD11b/CD18)[34], and SIgA structures presented here reveal two CD89-accessible FcαR
318 binding sites, suggesting that throughout evolution, SIgA interactions with host FcαRs have

319 proved functionally advantageous. Similarly, GAS primarily colonize the human nasopharynx
320 mucosa, with comparatively few infections becoming invasive and M4 being one of the dominant
321 serotypes for non-invasive GAS. Given this and the prevalence of SIgA in the human nasopharynx
322 mucosa, it is tempting to speculate that host-GAS co-evolution has been driven by interactions
323 between virulence factors (e.g. M4 and M22) and SIgA. Interestingly, both M4 and M22 strains
324 lack a Hyaluronic Acid (HA) capsule, found on other GAS strains, which is antiphagocytic and has
325 long been proposed as an essential virulence factor[35,36]. We speculate that the Ig-binding proteins
326 provide an alternative and compensatory mechanisms for M4 and M22 strain virulence.
327 Regardless of whether the function of M4/M22 is to block host FcαR effector functions and/or
328 block antibody binding to surface antigens and/or provide another survival advantage for the
329 bacteria, our data highlight the potential selective pressure of SIgA and its effectors and the
330 importance of understanding their mechanisms of action in the mucosa. This extends beyond
331 GAS, given that at least two distinct pathogenic bacterial species expressing IgA-binding proteins,
332 *S. pyogenes* and *S. aureus,* utilize the human nasopharynx as their primary reservoir where
333 encounters with SIgA have potential to influence virulence and host response, and also serve as
334 a point for therapeutic intervention.
335

336 **METHODS**

337 **Protein expression and purification**

338 Genes encoding the human IgA1 heavy chain constant region and the lambda light chain constant
339 region were fused with STA121 VH and VL domain sequences [37]. The human tissue plasminogen
340 activator (tPA) signal sequence was inserted upstream of the IgA heavy chain sequence. These
341 sequences along with human JC and SC were cloned into mammalian expression vector
342 pD2610v1 (Atum). All constructs were transiently co-transfected into HEK Expi293F cells using
343 ExpiFectamine to produce SIgA complex (Thermo Fisher). Six days after transfection, SIgA was
344 purified from the cell supernatant using CaptureSelect™ Human IgA Affinity Resins. Superose 6

345 (Cytiva) size exclusion chromatography (SEC) was used to further purify SIgA. The gene
346 encoding the human CD89 ectodomain (uniprot ID: P24071) was fused to a C-terminal
347 hexahistine-tag and cloned into pD2610v1. The CD89 construct was transiently transfected into
348 HEK Expi293F cells using ExpiFectamine. CD89 was purified from the cell supernatant using Ni-
349 NTA affinity resin (Qiagen). Superose 6 SEC was used to further purify CD89. The gene encoding
350 the *Streptococcus pyogenes* M4 protein (uniprot ID P13050) was cloned into pET28_a (+) vector
351 (Novagen). The M4 endogenous signal peptide and propeptide were not included in the construct
352 and a C-terminal hexahistine-tag was added. M4 mutants Q48A, R50A, K69A were generated
353 using the Q5 site-directed mutagenesis kit (NEB, E0554). The M4 mutant variant constructs were
354 transformed and expressed in *E. coli* strain BL21 (DE3); expression was induced using 1mM
355 IPTG, added to the culture growing at 37°C when OD reaches 0.4-0.6. Cells were harvested 3-5
356 hours after IPTG induction by centrifugation, resuspended in 50 mM Tris-Cl, 150 mM NaCl, pH
357 7.5, 1mM phenylmethylsulfonyl fluoride and lysed by sonication. The M4 was isolated from soluble
358 fraction of cell lysate using Ni-NTA affinity resin. Superose 6 SEC was used to further purify M4.
359 Purified SIgA and M4 or CD89 were mixed in a 1:2 molar ratio and incubated in storage buffer
360 (50 mM Tris-Cl, 150 mM NaCl, pH 7.5) at 4°C overnight and complexes were purified using
361 Superose 6 SEC.
362
363 **Cryo-EM grid preparation and data collection**
364 SIgA-M4: Quantifoil R2/2 300 mesh holey carbon grids were glow discharged for 60s; 3ul of SIgA-
365 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
366 offset, 5 s wait, at 100% humidity and 4 C°. Movies were collected using EPU on a Titan Krios
367 (TEMALPHA at SLAC) operating at 300 kV and with a Gatan K4 direct electron detector. Both
368 untilted and 25-degree tilted movies were collected at 0.95 Å/Pixel, 40 frames per movie, total
369 exposure time of 4.95 s and total dose of 50 electrons/$Å^2$.

370 SIgA-CD89: Quantifoil R1.2/1.3 400 mesh holey carbon grids were glow discharged for
371 60s; 3ul of SIgA-CD89 at 1.0 mg/ml was applied to each grid with Vitrobot Mark IV with 1 blot of
372 2 s, 2 offset, 5 s wait, at 100% humidity and 4 C°. Movies were collected using SerialEM on a
373 Titan Krios at Purdue University operating at 300 kV and with a K3 direct electron detector. Both
374 untilted and 30-degree tilted movies were collected at 0.82 Å/Pixel, 60 frames per movie, total
375 exposure time of 2.41 s and total dose of 71.3 electrons/$Å^2$. Data collection statistics can be found
376 in Table SI.
377
378 **Cryo-EM data processing**
379 SIgA-M4: 2600 movies were collected with untilted stage and 3251 movies were collected with
380 stage tilted at 25 degrees. All data processing were done in CryoSPARC v.3.2.0[38]. Initial
381 processing of the two data sets were done separately. Movies were motion corrected, CTF
382 estimated and manually curated based on ice thickness, total motion and CTF fits, resulting in
383 2081 untilted and 2786 tilted movies. Particles were picked using blob picker in CryoSPARC with
384 minimum and maximum particle diameter of 100 Å and 200 Å. Particles were extracted with 2x
385 binning. 2D classification of these particles generated initial references for template-based particle
386 picking, which gave 970k particles from untilted movies and 1.3 million particles from tilted movies.
387 Those particles were combined and subjected to several rounds of 2D classification to remove
388 bad classes. A final set of 200K particles were used to generate the *ab-initio* model. Particles in
389 the final set were then re-extracted at full-pixel size and used for non-uniform refinements of the
390 model. The final refinement generated a map with an average resolution of 3.07 Å at FSC=0.143.
391 SIgA-CD89: 2955 movies were collected with untilted stage and 3696 movies were
392 collected with stage tilted at 30 degrees. All data processing were done in CryoSPARC v.4.1.2[38].
393 Initial processing of the two data sets were done separately. Movies were motion corrected, CTF
394 estimated and manually curated based on ice thickness, total motion and CTF fits, resulting in
395 2865 untilted and 3105 tilted movies. Particles were picked using blob picker in CryoSPARC with

396 minimum and maximum particle diameter of 150 Å and 250 Å. Particles were extracted with 2x
397 binning. 2D classification of these particles generated initial references for template-based particle
398 picking, which gave 507K particles from untilted movies and 732K particles from tilted movies.
399 Those particles were combined and went through several rounds of 2D classification to remove
400 bad classes. A final set of 249K particles were used to generate the *ab-initio* model. Particles in
401 the final set were then re-extracted at full-pixel size and used for non-uniform refinements of the
402 model. The final refinement generated a map with an average resolution of 3.18 Å at FSC=0.143.
403

404 **Structure building and refinement**

405 SIgA-M4: The structure of human SIgA (PDB: 6UE7) was docked into the map using UCSF
406 Chimera[39]. The M4 fragment was hand-built using the Coot Molecular Graphics Package[40]. All
407 chains of SIgA and M4 were refined as rigid bodies using Phenix real space refinement[41].
408 Subsequent iterations involved inspection of the map-model fit and manual adjustment of the
409 model followed by Phenix real space XYZ refinement. Iterations of Phenix refinements and
410 manual adjustments were repeated until model quality no longer improves.

411 SIgA-CD89: Human SIgA structure (PDB: 6UE7) and mIgA:CD89 structure (PDB:1OW0)
412 were docked into the map using UCSF Chimera[39]. All chains of SIgA and CD89 were refined as
413 rigid bodies using Phenix real space refienemnt[41]. Inspection of the map-model fit and manual
414 adjustment of the model were done after each Phenix real space XYZ refinement. Iterations of
415 Phenix refinements and manual adjustments were repeated until model quality no longer
416 improves. Model and refinement statistics can be found in Table SI.
417

418 **Surface plasmon resonance**

419 Surface plasmon resonance (SPR) binding experiments were done using a Bruker Sierra SPR-
420 32 Pro instrument. 0.2 uM human dIgA or SIgA in sodium acetate buffer (pH = 4.5) was coupled
421 to lanes B and D, respectively, on a high-capacity amine (HCA) chip using an amine coupling kit

422 (Bruker 1862634). Lane A and C were mock coupled (with only sodium acetate buffer PH 4.5)
423 and used as references. Responses were measured for a two-fold dilution series of each analyte
424 (M4 wild type and mutants Q48A, R50A, K69A) in HBS-EP+ buffer (0.01 M HEPES pH 7.4, 0.15
425 M NaCl, 3 mM EDTA, 0.005% v/v Surfactant P20) with a high concentration of 2 uM. All binding
426 experiments were done with 120 s contact time, 180 s dissociation time, and flow rate of 25 ul/min.
427 Surfaces were regenerated with 2.5M $MgCl_2$. All data were collected at room temperature.
428

429 **Structure modelling, analysis, and sequence alignment**

430 Human SIgA (PDB: 6UE7) was aligned to the M4-bound SIgA in The PyMOL Molecular Graphics
431 System (Version 2.5.1 Schrödinger, LLC) and the RMSD between the bound and unbound state
432 of SIgA was calculated automatically during the alignment performed by PyMOL (A>align>align
433 to molecule). A list of all possible interactions between SIgA and M4 was generated using the
434 protein interaction calculator (PIC) webserver[42]. The interactions were visualized and manually
435 checked in PyMOL and contacts within 4 Å were selected and reported in Fig. 3b. The buried
436 surface area was calculated using the 'get_area' function in PyMOL.

437 AlphaFold2-multimer was used to predict the full length M4 dimer structure and ColabFold
438 1.5.2 was used to generate the prediction[30]. M4 residues 287-314 are predicted as disordered
439 and not shown in figures. Full length modeled M4 residues 46-74 were aligned to the
440 corresponding residues in the SIgA-M4 structure using PyMOL, with an RMSD of 0.640. The
441 aligned structure of full length M4 with SIgA were used in Fig. 6. The distance of M4 to GAS
442 surface was estimated by measuring the distance between M4 residues 16 and 286. The distance
443 from the FcαR site on Fc_{CD} to M4 was measure from HC-D residue L441 to M4 helix-b residue
444 H98.

445 The sequence alignment between human and mouse IgA heavy chain (human: Uniprot
446 P10876, P10877, A0A0G2JMB2; mouse: Uniprot P01878) was carried out using ClustalOmega

447 and the alignment figure was made using ESPript3[43,44]. The alignment of mIgA:CD89 (PDB:
448 1OW0) to SIgA:CD89 and RMSD calculation were done in the same way as described above.
449
450 **Figures**
451 Structural figures were made using UCSF Chimera and PyMOL; SPR data were plotted using
452 Bruker Sierra-32 Analyzer and Microsoft Excel[39]. All figures were assembled using Adobe
453 Photoshop.
454
455 **DATA AVAILABILITY**
456 The Cryo-EM density maps have been deposited in the EM databank (www.ebi.ac.uk/emdb) with
457 the accession codes 40568, and 40567 for M4-SIgA and CD89-SIgA structures, respectively, and
458 the refined coordinates have been deposited in the Protein Data Bank (www.rcsb.org) with
459 accession codes 8SKV (M4-SIgA) and 8SKU (CD89-SIgA).
460
461 **AUTHOR CONTRIBUTIONS**
462 The study was conceived by B.M.S and Q.L.; experiments were conducted by Q.L.; B.M.S and
463 Q.L. analyzed data and wrote the manuscript.
464
465 **ACKNOWLEDGEMENTS**
466 This work was supported by NIH R01 AI165570 (PI Beth Stadtmueller) and University of Illinois
467 start-up funding. Cryo-Electron microscopy data were collected at the Purdue Cryo-EM facility
468 (http://cryoem.bio.purdue.edu) with assistance from Thomas Klose and the Stanford-SLAC Cryo-
469 EM Center (S^2C^2) supported by the National Institutes of Health Common Fund Transformative
470 High Resolution Cryo-Electron Microscopy program (U24 GM129541) with assistance from Htet
471 A. Khant. We thank members of the Stadtmueller lab for insightful discussion and Sarah Leonard
472 for supporting SPR experiments.

473 **MATERIALS AND CORRESPONDENCE.**

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

476

477 **COMPETING INTERESTS.**

478 Authors declare no competing interests.

479

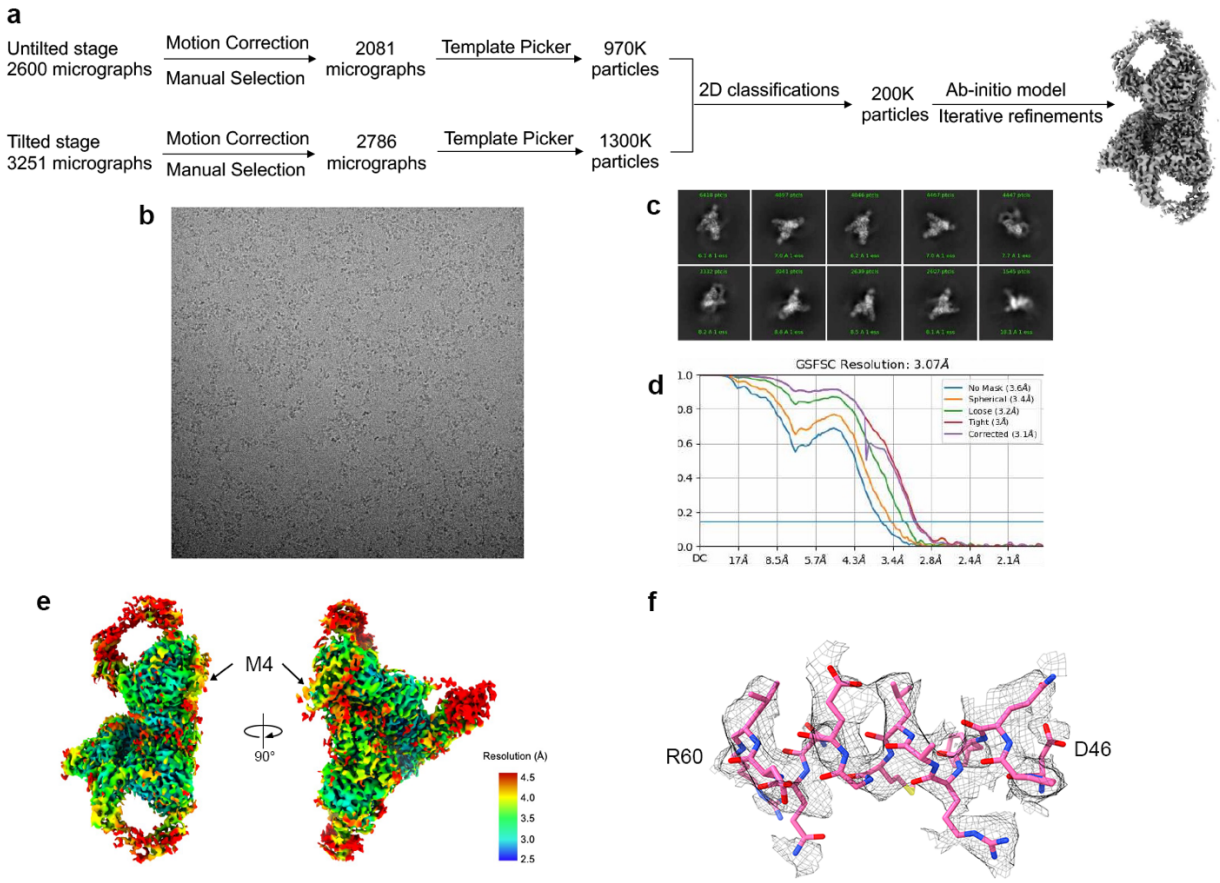
480 **REFERENCES**

481 1. Mantis, N. J., Rol, N. & Corthésy, B. Secretory IgA's complex roles in immunity and
482 mucosal homeostasis in the gut. *Mucosal Immunol* **4**, 603–611 (2011).
483 2. Herr, A. B., Ballister, E. R. & Bjorkman, P. J. Insights into IgA-mediated immune
484 responses from the crystal structures of human FcαRI and its complex with IgA1-Fc.
485 *Nature 2003 423:6940* **423**, 614–620 (2003).
486 3. Wines, B. D., Willoughby, N., Fraser, J. D. & Mark Hogarth, P. A competitive mechanism
487 for staphylococcal toxin SSL7 inhibiting the leukocyte IgA receptor, FcαRI, is revealed by
488 SSL7 binding at the Cα2/Cα3 interface of IgA. *Journal of Biological Chemistry* **281**, 1389–
489 1393 (2006).
490 4. Ramsland, P. A. *et al.* Structural basis for evasion of IgA immunity by Staphylococcus
491 aureus revealed in the complex of SSL7 with Fc of human IgA1. *PNAS* **104**, 15051–
492 15056 (2007).
493 5. Stenberg, L., Otoole, P. W., Mesteckm, J. & Lindahl, G. Molecular Characterization of
494 Protein Sir, a Streptococcal Cell Surface Protein That Binds Both Immunoglobulin A and
495 Immunoglobulin G*. *Journal of Biological Chemistry* **269**, 13458–13464 (1994).
496 6. Johnsson, E., Andersson, G., Lindahl, G. & Hedén, L. O. Identification of the IgA-binding
497 region in streptococcal protein Arp. *J. Immunol* **153**, 3557–3564 (1994).
498 7. Kumar Bharathkar, S. *et al.* The Structures of Secretory and Dimeric Immunoglobulin A.
499 *Elife* 61801 (2020).
500 8. Kumar, N., Arthur, C. P., Ciferri, C. & Matsumoto, M. L. Structure of the secretory
501 immunoglobulin A core. *Science* **367**, 1008–1014 (2020).
502 9. Wang, Y. *et al.* Structural insights into secretory immunoglobulin A and its interaction with
503 a pneumococcal adhesin. *Cell Res* **30**, 602 (2020).
504 10. Ralph, A. P. & Carapetis, J. R. Group a streptococcal diseases and their global burden.
505 *Curr Top Microbiol Immunol* **368**, 1–27 (2013).
506 11. Mitchell, T. J. The pathogenesis of streptococcal infections: from Tooth decay to
507 meningitis. *Nature Reviews Microbiology 2003 1:3* **1**, 219–230 (2003).
508 12. Happonen, L. *et al.* A quantitative Streptococcus pyogenes–human protein–protein
509 interaction map reveals localization of opsonizing antibodies. *Nature Communications*
510 *2019 10:1* **10**, 1–15 (2019).
511 13. Carapetis, J. R., Steer, A. C., Mulholland, E. K. & Weber, M. The global burden of group
512 A streptococcal diseases. *Lancet Infect Dis* **5**, 685–694 (2005).
513 14. McNamara, C. *et al.* Coiled-coil irregularities and instabilities in group A Streptococcus
514 M1 are required for virulence. *Science* **319**, 1405–1408 (2008).

515 15. Phillips, G. N., Flicker, P. F., Cohen, C., Manjula, B. N. & Fischetti, V. A. Streptococcal M
516 protein: alpha-helical coiled-coil structure and arrangement on the cell surface. *Proc Natl*
517 *Acad Sci U S A* **78**, 4689–4693 (1981).
518 16. Frost, H. R., Sanderson-Smith, M., Walker, M., Botteaux, A. & Smeesters, P. R. Group A
519 streptococcal M-like proteins: From pathogenesis to vaccine potential. *FEMS Microbiol*
520 *Rev* **42**, 193–204 (2018).
521 17. Smeesters, P. R., McMillan, D. J. & Sriprakash, K. S. The streptococcal M protein: a
522 highly versatile molecule. *Trends Microbiol* **18**, 275–282 (2010).
523 18. Steer, A. C. *et al.* Status of research and development of vaccines for Streptococcus
524 pyogenes. *Vaccine* **34**, 2953–2958 (2016).
525 19. R. Maliszewski, C., J. March, C., A. Schoenborn, M., Gimpel, steve & Shen, L.
526 Expression Cloning of a Human Fc Receptor for IgA. *J. Exp. Med.* **172**, 1665–1672
527 (1990).
528 20. Lindahl, G. & Akerstrom, B. Receptor for IgA in group A streptococci: cloning of the gene
529 and characterization of the protein expressed in Escherichia coii. *Mol Microbiol* **3**, 239–
530 247 (1989).
531 21. Frithz, E., Heden, L.-O. & Lindahl, G. Extensive sequence homology between IgA
532 receptor and M proteins in Streptococcus pyogenes. *Mol Microbiol* **3**, 1111–1119 (1989).
533 22. Buffalo, C. Z. *et al.* Conserved patterns hidden within group A Streptococcus M protein
534 hypervariability recognize human C4b-binding protein. *Nat Microbiol* **1**, (2016).
535 23. MacHeboeuf, P. *et al.* Streptococcal M1 protein constructs a pathological host fibrinogen
536 network. *Nature 2011 472:7341* **472**, 64–68 (2011).
537 24. Pleass, R. J., Areschoug, T., Lindahl, G. & Woof, J. M. Streptococcal IgA-binding
538 Proteins Bind in the Cα2-Cα3 Interdomain Region and Inhibit Binding of IgA to Human
539 CD89. *Journal of Biological Chemistry* **276**, 8197–8204 (2001).
540 25. Kolesinski, P., Wang, K. C., Hirose, Y., Nizet, V. & Ghosh, P. An M protein coiled coil
541 unfurls and exposes its hydrophobic core to capture LL-37. *Elife* **11**, (2022).
542 26. Akerstrom, B., Lindqvist, A. & Lindahl, G. *BINDING PROPERTIES OF PROTEIN ARP, A*
543 *BACTERIAL IgA-RECEPTOR*. *Mokcular Immunology* vol. 28 (1991).
544 27. Pleass, R. J., Dunlop, J. I., Anderson, C. M. & Woof, J. M. Identification of Residues in
545 the CH2/CH3 Domain Interface of IgA Essential for Interaction with the Human Fcα
546 Receptor (FcαR) CD89. *Journal of Biological Chemistry* **274**, 23508–23514 (1999).
547 28. Sandin, C. *et al.* Isolation and Detection of Human IgA Using a Streptococcal IgA-Binding
548 Peptide. *The Journal of Immunology* **169**, 1357–1364 (2002).
549 29. Steer, A. C., Law, I., Matatolu, L., Beall, B. W. & Carapetis, J. R. Global emm type
550 distribution of group A streptococci: systematic review and implications for vaccine
551 development. *Lancet Infect Dis* **9**, 611–616 (2009).
552 30. Bryant, P., Pozzati, G. & Elofsson, A. Improved prediction of protein-protein interactions
553 using AlphaFold2. *Nature Communications 2022 13:1* **13**, 1–11 (2022).
554 31. Fischetti, V. A. M Protein and Other Surface Proteins on Streptococci. in *Streptococcus*
555 *pyogenes : Basic Biology to Clinical Manifestations* (2016).
556 32. Castro, S. A. & Dorfmueller, H. C. A brief review on Group A Streptococcus pathogenesis
557 and vaccine development. *R Soc Open Sci* **8**, (2021).
558 33. Mills, J. O. & Ghosh, P. Nonimmune antibody interactions of Group A Streptococcus M
559 and M-like proteins. *PLoS Pathog* **17**, e1009248 (2021).
560 34. Spriel, A. B. van, Leusen, J. H. W., Vilé, H. & Winkel, J. G. J. van de. Mac-1
561 (CD11b/CD18) as Accessory Molecule for FcαR (CD89) Binding of IgA. *The Journal of*
562 *Immunology* **169**, 3831–3836 (2002).
563 35. Husmann, L. K., Yung, D. L., Hollingshead, S. K. & Scott, J. R. Role of putative virulence
564 factors of Streptococcus pyogenes in mouse models of long-term throat colonization and
565 pneumonia. *Infect Immun* **65**, 1422 (1997).

566 36. Wessels, M. R., Moses, A. E., Goldberg, J. B. & Dicesare, T. J. Hyaluronic acid capsule
567 is a virulence factor for mucoid group A streptococci. *Proc Natl Acad Sci U S A* **88**, 8317
568 (1991).
569 37. Moor, K. *et al.* High-avidity IgA protects the intestine by enchaining growing bacteria.
570 *Nature 2016 544:7651* **544**, 498–502 (2017).
571 38. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for
572 rapid unsupervised cryo-EM structure determination. *Nature Methods 2017 14:3* **14**, 290–
573 296 (2017).
574 39. Pettersen, E. F. *et al.* UCSF Chimera--a visualization system for exploratory research and
575 analysis. *J Comput Chem* **25**, 1605–1612 (2004).
576 40. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Biological Crystallography Features
577 and development of Coot. doi:10.1107/S0907444910007493.
578 41. Liebschner, D. *et al.* Macromolecular structure determination using X-rays, neutrons and
579 electrons: recent developments in Phenix. *urn:issn:2059-7983* **75**, 861–877 (2019).
580 42. Tina, K. G., Bhadra, R. & Srinivasan, N. PIC: Protein Interactions Calculator. *Nucleic*
581 *Acids Res* **35**, 473–476 (2007).
582 43. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new
583 ENDscript server. *Nucleic Acids Res* **42**, W320–W324 (2014).
584 44. Sievers, F. *et al.* Fast, scalable generation of high-quality protein multiple sequence
585 alignments using Clustal Omega. *Mol Syst Biol* **7**, 539 (2011).
586

587 **SUPPLEMENTARY INFORMATION**



588

589 **Figure S1. CryoEM data collection and CryoSPARC processing pipeline.** (a) Schematic
590 summary of data processing pipeline of SIgA-M4 complex in CryoSPARC; (b) representative
591 micrograph; (c) representative 2D class averages; (d) FSC curve for the final reconstruction with
592 reported resolution at FSC=0.143 shown by the blue horizontal line. (e) Local resolution map
593 calculated by cryoSPARC and colored in Chimera. (f) Density from M4 helix-a residues 46-60
594 contoured to 3σ and carved at 2 Å cutoff.

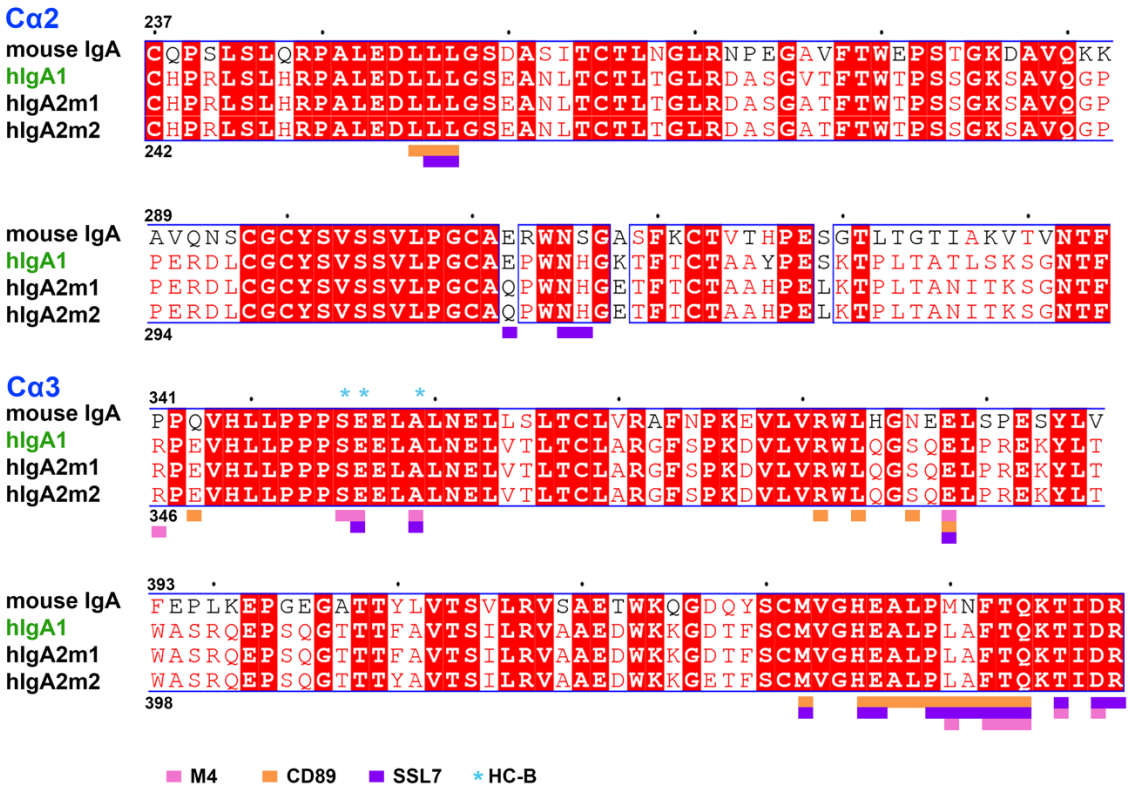


Figure S2. Sequence alignment of human and mouse IgA heavy chains. Sequence alignment including human IgA subtypes (hIgA1, hIgA2m1 and hIgA2m2) and mouse IgA. Identical residues are colored in red. Each square denotes an interfacing residue described in the main text. Squares are colored according to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

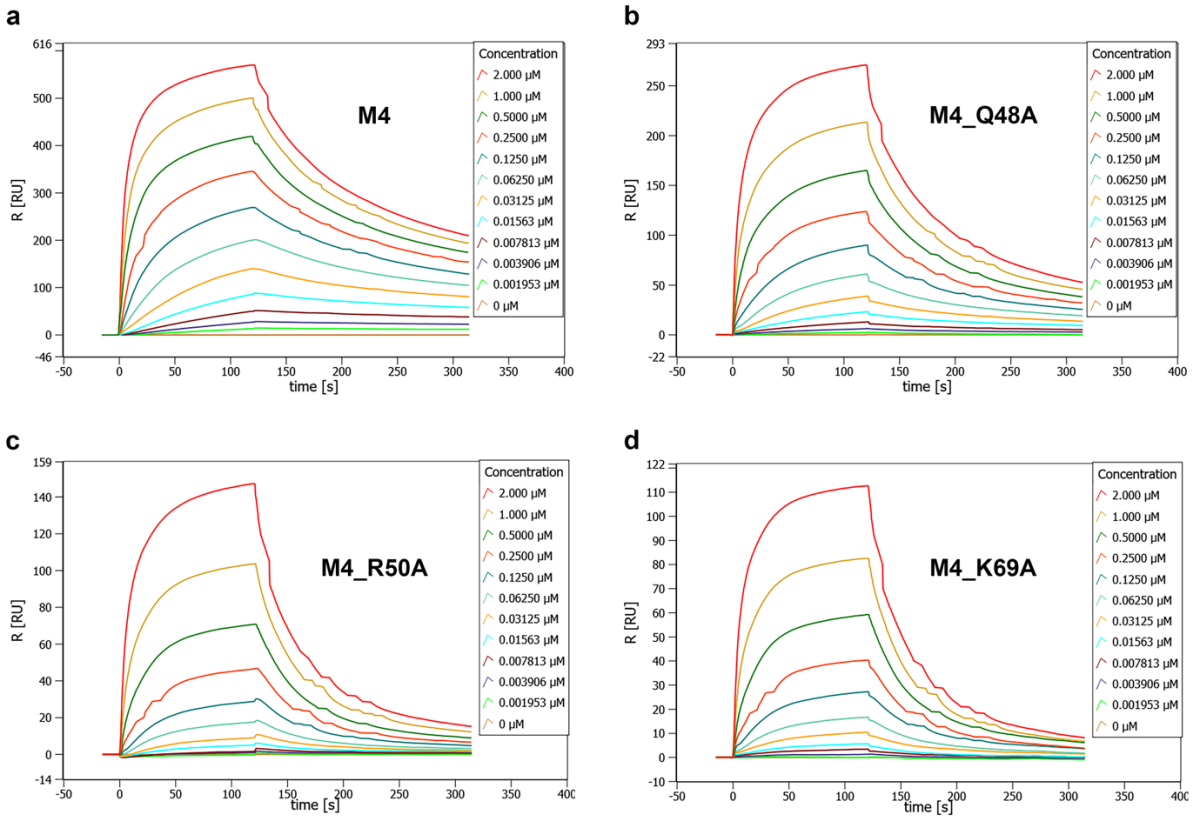


Figure S3. M4 binding analysis by surface plasmon resonance (SPR). Sensorgrams showing the response of M4 (a) and individual M4 point mutants, Q48A (b), R50A (c), and K69A (d) binding to immobilized hSIgA1. All data were collected on a surface immobilized with 0.2uM hSIgA1 and all analytes were tested with a two-fold dilution series starting from 2uM.

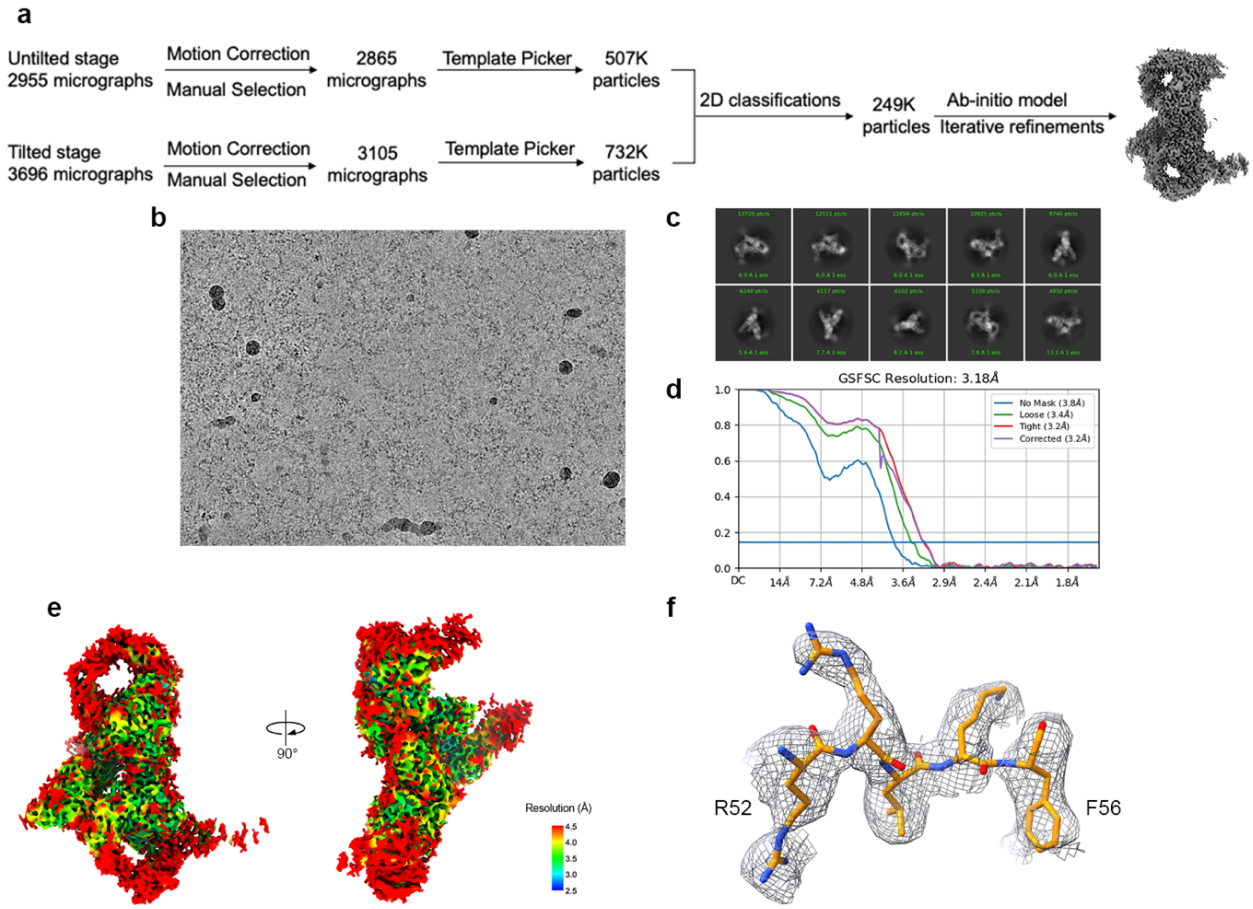


Figure S4. SIgA-CD89 CryoEM data collection and CryoSPARC processing pipeline. (a) Schematic summary of data processing pipeline of SIgA-CD89 complex in CryoSPARC; (b) representative micrograph; (c) representative 2D class averages; (d) FSC curve for the final reconstruction with reported resolution 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 contoured to 3σ and carved at 2 Å cutoff.

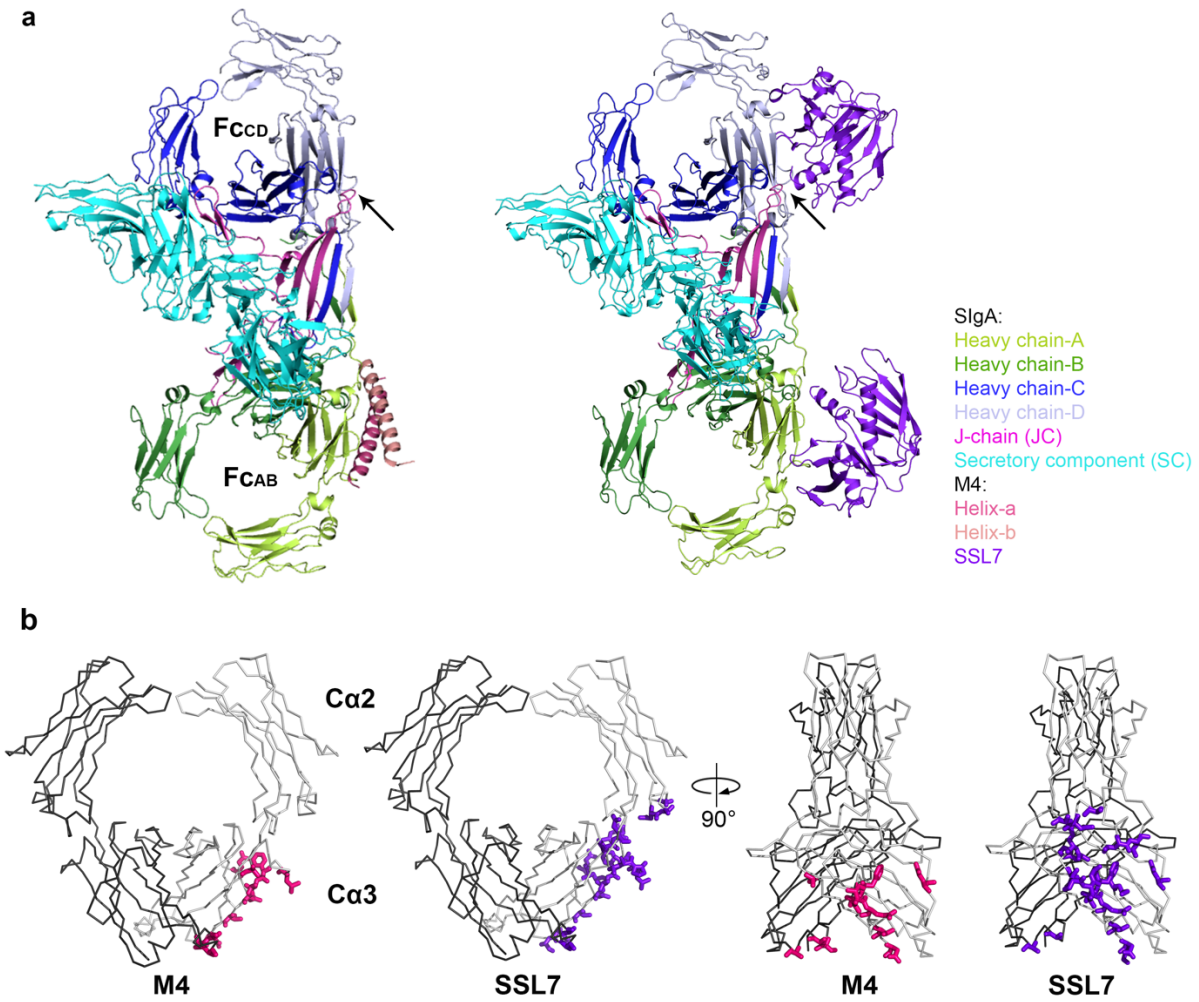


Figure S5. Comparison of M4 and SSL7 binding to SIgA. (a) The structure of M4-SIgA and model of SSL7 bound to SIgA. The SSL7-SIgA model was built by aligning the cocrystal structure of an SSL7-Fcα complex (PDB: 2QEJ) to both copies of SIgA Fcα. M4 only binds at Fc_{AB}, as shown in Fig. 5 and the second 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.

626 **Table S1. Cryo-EM data collection, refinement, and validation statistics**
627

	#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)	3.0	3.2
FSC threshold	0.143	0.143
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	0	0
Ligand	51.53	41.21
R.m.s. deviations		
Bond lengths (Å)	0	0.005
Bond angles (°)	0	0.642
Validation		
MolProbity score	1.67	1.93
Clashscore	5.26	9.20
Poor rotamers (%)	0.32	0.2
Ramachandran plot		
Favored (%)	94.32	93.15
Allowed (%)	5.77	6.85
Disallowed (%)	0	0

628