The structures of secretory and dimeric immunoglobulin A

Sonya Kumar Bharathkar†, Benjamin W Parker†, Andrey G Malyutin²,³, Nandan Haloi⁴,⁵, Kathryn E Huey-Tubman², Emad Tajkhorshid¹,⁴,⁵, Beth M Stadtmueller*¹

¹Department of Biochemistry, University of Illinois Urbana-Champaign, Urbana, United States; ²Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, United States; ³Beckman Institute, California Institute of Technology, Pasadena, United States; ⁴Center for Biophysics and Quantitative Biology, University of Illinois at Urbana-Champaign, Urbana, United States; ⁵NIH Center for Macromolecular Modeling and Bioinformatics, Beckman Institute for Advanced Science and Technology, Urbana, United States

Abstract Secretory (S) Immunoglobulin (Ig) A is the predominant mucosal antibody, which binds pathogens and commensal microbes. SlgA is a polymeric antibody, typically containing two copies of IgA that assemble with one joining-chain (JC) to form dimeric (d) IgA that is bound by the polymeric Ig-receptor ectodomain, called secretory component (SC). Here, we report the cryo-electron microscopy structures of murine SlgA and dlgA. Structures reveal two IgAs conjoined through four heavy-chain tailpieces and the JC that together form a β-sandwich-like fold. The two IgAs are bent and tilted with respect to each other, forming distinct concave and convex surfaces. In SlgA, SC is bound to one face, asymmetrically contacting both IgAs and JC. The bent and tilted arrangement of complex components limits the possible positions of both sets of antigen-binding fragments (Fabs) and preserves steric accessibility to receptor-binding sites, likely influencing antigen binding and effector functions.

Introduction

The vertebrate mucosa is a vast extracellular environment that mediates host interactions with a broad range of antigens including toxins, pathogens, and commensal organisms. The diversity of these antigens, some of which are beneficial to the host and some of which are harmful, has driven complex evolutionary interplay between mucosal immune system molecules and mucosal antigens, resulting in antibodies with novel structural and functional mechanisms compared to circulatory counterparts (Flajnik, 2010; Kaetzel, 2014).

Secretory (S) Immunoglobulin (Ig) A is the predominant, mammalian mucosal antibody; it is a polymeric antibody assembled in plasma cells that link two IgA monomers and a single copy of the joining chain (JC) to form dimeric (d) IgA (and to a lesser extent, higher order polymers) (Woof and Russell, 2011). Following secretion in the lamina propria, dIgA is bound by the polymeric Ig-receptor (plgR), a transcytotic Fc receptor (FcR) expressed on the basolateral surface of epithelial cells. The plgR uses five Ig-like domains (D1-D5) and a cytoplasmic tail to bind and transcytose JC-containing antibodies to the apical surface of epithelial cells; there, the plgR ectodomain, called secretory component (SC), is proteolytically cleaved, releasing the SC-dlgA complex into the mucosa where it is called SlgA (Figure 1A; Stadtmueller et al., 2016a). In the mucosa, SlgA functions are thought to be dominated by physical mechanisms including coating, cross-linking, agglutination, and enchained growth of mucosal antigens; outcomes are diverse and, typically, not associated with inflammation (Woof and Russell, 2011; Pabst and Slack, 2020). For example, SlgA-dependent enchained growth
of Salmonella promotes its clearance from the gut, whereas SlgA-interactions with B. fragilis can promote its colonization of the gut (Moor et al., 2017; Donaldson et al., 2018). Furthermore, colostrum SlgA can provide passive immunity to newborns and has been shown to have a life-long influence on microbiome composition (Rogier et al., 2014).

SlgA functions are supported by a unique, yet poorly understood molecular structure. Similar to other antibody classes, the IgA monomer is made up of two copies of the heavy chain and two copies of the light chain, each with variable and constant domains. Together, these chains form two antigen-binding fragments (Fabs), each containing variable domains with complementary determining regions (CDRs) that facilitate binding to a large repertoire of antigens. In contrast to monomeric IgA, SlgA is an antibody dimer and has four Fabs that are presumed to enhance binding avidity and antigen cross-linking potential when compared to antibodies with two Fabs; yet, a mechanistic understanding of how SlgA engages antigens remains a topic of investigation. IgA Fabs connect to the antibody Fc, which is made up of two copies of the IgA constant heavy chain domains, C\text{H}2-C\text{H}3 (Figure 1B). The IgA C\text{H}3 contains a C-terminal extension called the tailpiece (Tp), which along with the JC, is required for IgA dimerization and subsequent plgR binding. The Fc regions of dlgA provide binding sites for SC as well other host and microbial FcRs; however, the conformations of the dlgA Fc and plgR domains have remained elusive along with the structures of the Tps and JC (Stadtmueller et al., 2016a).

To investigate SlgA structure and how it supports diverse mucosal functions, we targeted structures of SlgA from the mouse, which has been long been used as a model organism for immunological research (Masopust et al., 2017). Unlike humans, which have two IgA heavy chain isotypes, IgA1 and IgA2 (and two IgA2 allotypes IgA2m1 and IgA2m2) mice express a single IgA heavy chain that shares 79–80% sequence similarity and 66% sequence identity with human variants among the C\text{H}2-C\text{H}3 domains; the sequences of mouse SC and JC also share a high degree of sequence similarity.
with human counterparts (Figure 1—figure supplement 1). Accordingly, we co-expressed a mouse monoclonal IgA with mouse JC and SC in transiently transfected mammalian cells, which resulted in monodisperse SlgA (see Materials and methods). Using cryoelectron microscopy (CryoEM), we determined a ~3.3 Å resolution structure of SlgA, which revealed a pseudosymmetric arrangement of the two IgA Fcs, bound asymmetrically to JC and SC. Comparisons with a ~3.3 Å resolution cryoEM structure of dIgA, revealed a dominant role for JC in maintaining the conformation of the dimer interface and geometric relationship between the two IgA Fcs. Finally, we modeled plausible positions that could be adopted by SlgA Fabs. Together, results suggest that the pseudosymmetric arrangement of SlgA core components will constrain the positions of the Fabs and influence binding of host and microbial factors.

**Results**

**The structure of SlgA**

We determined the CryoEM structure of mouse SlgA to a final average resolution of 3.3 Å (Figure 2). Local map resolution was variable, revealing side chain density for many residues, particularly at the interfaces between complex components and domains; however, some regions at the periphery of the complex were not well resolved, likely due to inherent flexibility. Local resolution was lowest for Fabbs, which are mostly disordered and were not built. Additionally, sidechains were not well-resolved for the majority of residues in SC D2 and some residues in the C\textsubscript{i}2 domains (Figure 2—figure supplement 1, Figure 2—figure supplement 2, Figure 2—figure supplement 3). The refined structure revealed a pseudosymmetric assembly of two IgA monomers conjoined at the center by the JC and bound by SC. The C\textsubscript{i}2-C\textsubscript{i}3 domains in both Fcs aligned with the published monomeric IgA Fc structure lacking the Tp (Herr et al., 2003) (not shown); however, despite shared sequences, the four IgA heavy chains formed structurally unique contacts with the JC and SC. To distinguish these differences, we designated a unique ID for each heavy chain: A, B, C, or D, which we also use to describe corresponding C\textsubscript{H} domains and Fcs (Figure 2). The FC\textsubscript{AB} and FC\textsubscript{CD} are bent and tilted with respect to each other, resulting in distinct concave and convex surfaces on the complex. To describe the conformation, we defined the angle between the centroid axes of FC\textsubscript{AB} and FC\textsubscript{CD} (97 degrees) as ‘bend’, and the angle between the two non-intersecting centroid planes of FC\textsubscript{AB} and FC\textsubscript{CD} (30 degrees) as ‘tilt’. SC asymmetrically contacts the same face of both FC\textsubscript{AB} and FC\textsubscript{CD} along what we define as the ‘front face’ of the molecule. Potential N-linked glycosylation sites (PNGS) are distributed throughout the complex; the seven, PNGS located on SC, are clustered on the front face of SlgA (Figure 2). The map revealed well-ordered carbohydrates at a subset of these sites, which were modeled in the structure (Figure 2—figure supplement 4).

**The structure of the Fc-JC interface**

The SlgA structure reveals numerous molecular interactions that effectively lock the Fc regions from two IgA monomers with a single JC to form a centrally located interface. The interface is dominated by the four heavy chain Tps (TP\textsubscript{A}, TP\textsubscript{B}, TP\textsubscript{C}, TP\textsubscript{D}) and the first half of the JC sequence, or, ‘core’ (JC\textsubscript{core}), which together fold into a single \(\beta\)-sandwich-like domain at the center of the complex. The interface is further stabilized by the second half of the JC sequence, which folds into two beta-hairpin ‘wings’ (JC\textsubscript{W1} and JC\textsubscript{W2}), each extending outward from the center of the complex and binding one C\textsubscript{i}3 from each IgA monomer (Figure 3).

Notably, whereas the four Tps share identical sequence, they form unique interactions with each other and the JC. TP\textsubscript{C} and TP\textsubscript{D} and the first three JC \(\beta\)-strands form one side of the \(\beta\)-sandwich, oriented toward the front face of SlgA, while TP\textsubscript{A}, TP\textsubscript{B} and the subsequent \(\beta\)-strand from the JC form the other side, oriented toward the back face of SlgA (Figure 3A,B). Despite their location on opposing sides of the complex, TP\textsubscript{A} and TP\textsubscript{C} adopt similar conformations that are stabilized though \(\beta\)-sheet interactions with TP\textsubscript{B} and TP\textsubscript{D}, respectively, as well as interactions between C-terminal residues and two unique pockets formed in part by the JC\textsubscript{core} (Figure 3C). The cornerstones of these interactions are TP\textsubscript{A} and TP\textsubscript{C} penultimate cysteine residues (Cys 466), which disulfide bond to JC\textsubscript{core} Cys 68 and JC\textsubscript{core} Cys 14, respectively, and thereby lock both IgA monomers to the JC. Additional stability is provided by TP\textsubscript{A} and TP\textsubscript{C} ultimate tyrosine residues (Try 467), which in addition to binding JC\textsubscript{core} residues form contacts with C\textsubscript{i}3C, and SC D1, respectively (Figure 3C). TP\textsubscript{B} and TP\textsubscript{D} form \(\beta\)-
sheet interactions with TpA and TpC, respectively, but in contrast to TpA and TpC, form limited interactions between their C-terminal residues and other complex components. For example, although TpD Cys466 is visible in the cryoEM map, we fail to find evidence that TpB or TpD Cys466 are involved in disulfide bonds and the seven C-terminal residues of TpB are disordered. Notably, however, TpD Tyr467 is positioned to contact two arginine residues in the JCcore as well as conserved residues in SC D1, signifying another type of conserved interaction for the Tp ultimate tyrosine residues that complements those observed for TpA and TpC Tyr467 (Figure 3C).
Figure 3. Fc dimer interface and JC structure. (A) Schematic depicting the topology that arises from JC and its interactions with four Tps and colored as in Figure 2. JC regions (JC\textsubscript{W1}, JC\textsubscript{W2}, JC\textsubscript{core}, and N- and C termini) are labeled along with Tp\textsubscript{A-D} and their location relative to the front face or back face of SIgA; disulfide bonds are indicated in yellow and the residue boundaries of each \beta-strand are labeled. (B) SIgA molecular surface representation (SC removed) colored as in Figure 2 and indicating the location of JC\textsubscript{W1}, JC\textsubscript{W2} and Tps relative to each HC\textsubscript{A-D} in two SIgA orientations. (C) Cartoon Figure 3 continued on next page
The two JC wings, JC\textsubscript{W1} and JC\textsubscript{W2}, represent a second set of unique interactions that help to lock the two IgA Fc monomers together. JC\textsubscript{W1} and JC\textsubscript{W2} are unique in sequence yet both form antiparallel β-sheets, enclosed on the N- and C-terminal ends by intra-JC disulfide bonds. JC\textsubscript{W1} is composed primarily of hydrophobic and acidic residues and uses hydrophobic interactions to bind the front face of C\textsubscript{H3}, whereas acidic residues in JC\textsubscript{W1} mediate interactions with SC (Figure 3D). JC\textsubscript{W2} is five residues longer than JC\textsubscript{W1} and shares a larger, more chemically diverse interface with the back face of C\textsubscript{H3A}, involving several electrostatic interactions (Figure 3D). Notably, although residues forming the interface between each JC wing and its respective C\textsubscript{H3} domain differ, both wings bind a similar location on C\textsubscript{H3A} and C\textsubscript{H3B}, illustrating a mechanism by which a single JC can bind two identical antibodies uniquely.

**Secretory component**

SC, the plgR ectodomain, has five domains, D1-D5, each having an Ig-like fold with loops structurally similar to antibody CDRs. The crystal structure of unliganded human SC demonstrated that these domains adopt a compact arrangement, in which a subset of residues in three D1 CDR-like loops (hereafter CDR) interface with residues in D4 and D5 to form a closed conformation (Figure 4A; Stadtmueller et al., 2016a). In contrast to unliganded SC, the SlgA structure reveals SC domains bound to dIgA in an elongated, open conformation contacting both antibodies and the JC asymmetrically on the front face of the complex (Figure 4A). D1 contacts Fc\textsubscript{AB}, Fc\textsubscript{CD}, and the JC, D2 projects away from other SC domains and the front face of the molecule, and D3 is clamped to D1, positioning D4 and D5 to contact Fc\textsubscript{CD} and JC\textsubscript{W1} (Figure 4A).

Interactions between SC D1 and dIgA components are of special interest because residues in D1 CDR loops are reportedly necessary for binding to JC-containing antibodies, yet are also involved in stabilizing SC in its unliganded conformation (Brandtzaeg, 2013; Stadtmueller et al., 2016a). Indeed, the SlgA structure reveals residues in or adjacent to CDR1, CDR2, and CDR3 bridging interactions with JC, C\textsubscript{H3A}, C\textsubscript{H3B}, TP\textsubscript{C}, and TP\textsubscript{D}, rather than binding to SC D4 or D5 as observed in the closed conformation (Figure 4A,B). The D1 CDR1 appears to play a key role binding JC, primarily through a conserved Try-Pro-Asp motif located in the JC C-terminal loop following JC\textsubscript{W2}. This interaction is dominated by D1 CDR1 Arg31 and His32, which hydrogen bond to JC Asp137 thereby linking D1 to the C-terminus of JC. An additional contact, between D1 Asn30 and JC 106Arg, appears to link D1 to the loop connecting JC\textsubscript{W1} and JC\textsubscript{W2} (Figure 4B). In turn, residues flanking CDR1, and residues in and adjacent to CDR2, mediate contacts with two, non-overlapping sites on C\textsubscript{H3A} and C\textsubscript{H3B} (Figure 4B). In contrast to CDR2, which binds the Fc\textsubscript{AB}, D1 CDR3 residues bind the Fc\textsubscript{CD}, contacting the C-terminal regions of TP\textsubscript{C} and TP\textsubscript{D}. Notably, CDR3 Leu101 contacts the ultimate residue, Tyr467, in both TP\textsubscript{C} and TP\textsubscript{D}, highlighting the dual role that the conserved ultimate tyrosine residues in the Tps play in binding to JC and SC residues (Figures 3C and 4B).

Compared to D1, D2-D5 form limited interactions with dIgA components; yet, we find each engaged in a unique position that stabilizes the SlgA structure. D2 does not contact dIgA directly; instead, the D1-D2 linker is extended, positioning D2 at the outer edge of the front face where it shares a minimal, $\sim$250 Å$^2$, interface with D3. The map surrounding D2 is poorly ordered suggesting that the position of D2 is flexible to the extent allowed by the linkers between D1-D2 and D2-D3. Together with D2, D3-D4-D5 adopt a near-linear arrangement, or arm, that is bent approximately 48 degrees relative to D1, positioning D4-D5 to contact the Fc\textsubscript{CD} front face and the JC\textsubscript{W1} (Figure 4A,B). This conformation is stabilized by a 324 Å$^2$ interface between D1 and D3, involving conserved hydrophobic residues from both domains (Figure 4C). The interface appears to function as a
keystone in a D1-D3-D4-D5 bridge connecting Fc\textsubscript{AB} and Fc\textsubscript{CD} and contrasts with SC’s unliganded structure, in which D1 and D3 do not share an interface (Figure 4A). Although D3 does not contact either of the Fcs or the JC, its position is further stabilized by a 646 Å\textsuperscript{2} interface with D4 (distinct from the D3-D4 interface in the closed conformation) and several D3-D4 linker residues that are positioned to contact JC\textsubscript{W1} (Figures 3D and 4D). D4 and D5 share a similar interface in both unliganded SC and SIgA; however, the domains are repositioned relative to other domains such that residues in D4 and D5 that contact D1 in SC’s unliganded conformation, are solvent exposed. Together, D4-D5 engage the front face of Fc\textsubscript{CD} through a patch of residues that surround C\textsubscript{H}2Cys306 and the tip of JC\textsubscript{W1}. The interface is dominated by residues in D5 CDR1, including Cys470, which disulfide bonds to C\textsubscript{H}2Cys306, covalently linking SC to the antibody, as well as D5 CDR1 Lys471 and Tyr473, which form hydrogen bonding and hydrophobic interactions with C\textsubscript{H}2Cys306.
Additional stability at this interface is provided by a handful basic residues located in D5 and linkers connecting D3-D4 and D4-D5, which form electrostatic interactions with acidic residues in JCW1 and effectively sandwich JCW1 in-between Cβ2 and SC (Figures 3D and 4D).

**Structure of dlgA**

The SlgA structure revealed a bent and tilted relationship between two IgA monomers, which was stabilized in part by a bridge of interactions between SC domains. To investigate SC’s contribution to the conformation of SlgA, we determined a 3.3 Å average resolution cryoEM structure of dlgA, which contained the same components as SlgA, except for SC (Figure 5, Figure 5—figure supplement 1, Figure 5—figure supplement 2). The dlgA structure revealed two IgA monomers arranged similar to those in SlgA; Fabs were disordered. The centroid axes of FcAB and FcCD were

![Figure 5. dlgA structure.](image-url)

**Figure 5.** dlgA structure. (A) Cartoon representation (with semi-transparent molecular surface) of the unliganded dlgA structure shown in two orientations and colored as in **Figure 2**. The bend and tilt between the two Fcs is indicated with a line and the angle measured in the structure; the concave and convex sides are labeled, along with the front face and the back face and Cβ domains. PNGS are shown as orange spheres. (B) Molecular surface representation of SlgA (gray; SC removed) aligned to dlgA (green) on the JC from each complex structure. (C) CryoEM maps and structures for SlgA (right) and dlgA (left) detailing the region surrounding TpD. (D) CryoEM maps and structures for SlgA (right) and dlgA (left) detailing the region surrounding the JC C-terminus.

The online version of this article includes the following figure supplement(s) for figure 5:

**Figure supplement 1.** dlgA cryoEM data collection and processing strategy.

**Figure supplement 2.** dlgA cross-validation FSC curves.
bent 98 degrees relative to each other (compared to 97 degrees in SIgA), and the centroid planes were tilted at an angle of 19 degrees (compared to 30 degrees in SIgA) (Figure 5A). These results suggest that the bend in SIgA is conferred by JC binding to heavy chains and/or interactions between the Tps from both antibodies. The 11-degree difference in tilt between the two structures suggests that SC binding can influence the conformational relationship between the two IgAs. Aligned individually, the structures of the heavy chains and JC were largely superimposable with equivalent chains in SIgA; however, differences were apparent in regions that are bound by SC in SIgA (Figure 5B). In particular, TpD, which forms part of the SC D1-JC interface in SIgA, adopts a different conformation, and the three C-terminal residues are disordered (Figure 5C). The three C-terminal residues of JC (Tyr-Pro-Asp), which bind SC D1 CDR1 in SIgA, are also disordered in dIgA (Figure 5D).

**SIgA structure impact on antigen binding**

The SIgA and dIgA structures revealed bent and tilted relationships between two IgA monomers, which are supported by stabilizing interactions that are likely to limit flexibility between the two IgA Fcs. Seeking to better visualize how this geometry could influence the flexible positions of SlgA Fabs, we modeled potential Fab locations onto SlgA structures using a computational conformational-search approach, which approximated the position of each Fab by mapping a vector (beginning at N-terminus of C\textsubscript{H}2 and ending at the center of mass of Fab CDRs) onto a Fibonacci spherical lattice (FSL) (Marques et al., 2013; Figure 6—figure supplement 1). We reasoned that this approach would broadly survey possible Fab positions without the constraints of the diverse Fc-Fab linkers and contacts between heavy chains and light chains found in SIgA among mammals (Woof and Kerr, 2006). This strategy identified 8000 possible conformations for each Fab, from which those that clashed with Fc, JC, or SC were eliminated. Inspecting all SlgA models revealed a distribution, in which Fab CDRs dominated positions on the concave side of SIgA. Notably, this distribution revealed a large exposed surface area on the convex side of SIgA (C\textsubscript{H}2B-C\textsubscript{H}3B and C\textsubscript{H}2D-C\textsubscript{H}3D) that was never occupied by the Fabs (Figure 6A). To quantify distributions for each of the four Fab’s CDRs, we measured the angle of each vector relative to the center of mass of the JC and Fcs along the FC\textsubscript{AB} plane (Figure 6B). The distribution revealed a dominance of angles less than 90 degrees, consistent with Fabs occupying more positions on the concave side of SIgA, but also a unique distribution for each Fab, collectively illustrating how the complex pseudosymmetry (bend and tilt relative to the FC\textsubscript{AB}) might influence the positions of CDRs and associated interactions with antigen.

**Discussion**

Our work describes the structures of mouse SIgA and its precursor, dIgA, which lacks the SC. Together, these structures provide a glimpse of one of the most abundant and long-studied mammalian antibodies. Among other discoveries, decades of prior work determined the structures of unliganded SC and monomeric IgA Fc lacking the Tps, provided biochemical insights on SlgA components, and also revealed many key elements of plgR-dependent transcytosis, mucosal antibody evolution and function. (Herr et al., 2003; Bonner et al., 2009; Woof and Russell, 2011; Brandtzæg, 2013; Stadtmueller et al., 2016a). Yet without the structures of the JC, dIgA and SlgA, interpreting biological data, understanding functional mechanisms and designing IgA-based therapeutics has remained a challenge. Our structures and related, human IgA structures published in the same timeframe as this work (Kumar et al., 2020; Wang et al., 2020) provide an opportunity to address many outstanding questions surrounding SlgA and its functions as a mucosal antibody; we discuss a subset of them below and provide a model for formation, transport, and function of SlgA in mammals.

**The role of JC in dimeric IgA assembly**

The JC plays an essential and early-stage role in SlgA production by linking two or more IgA monomers and conferring the ability of the resulting antibody polymer (e.g. dIgA) to bind the plgR ectodomain (Woof and Russell, 2011). Yet, its evolutionary origin, its structure (even its fold), and how it facilitates antibody polymerization remained elusive through decades of research. Our dIgA and SlgA structures resolve some of these uncertainties by revealing the JC fold, how it is integrated.
with the Tps from four IgA heavy chains, and how it is stabilized by contacts between the JC wings
and IgA C3 domains. The structure does not appear to be closely related to proteins or complexes
with other functions because our searches of the protein data bank have failed to identify structures
with high similarity to the JC or the JC-Tp assembly; although, we cannot rule out the possibility that
the unliganded JC might adopt a different conformation with similarity to determined structures.
For now, this leaves the JC relationship to other proteins, and structural clues on the JC evolutionary
origin, uncertain.

The molecular contacts observed in the dIgA and SlgA structures reveal, in part, how the JC facil-
itates antibody polymerization. Central to this mechanism appears to be the JC’s ability to bind
each of the four identical heavy chain sequences uniquely, resulting in pseudosymmetric, dimeric
antibody. This mechanism may represent an efficient way for the JC to link two antibodies; yet, it
also appears likely to influence the function of the final product, SlgA, because our data indicate
that JC plays a dominant role in inducing and/or maintaining the bent relationship between the two
FcS, which our modeling suggests can influence Fab positions. Despite these observations, many
details of JC-dependent antibody polymerization remain outstanding; in particular, the order of
events during which the JC binds two IgA monomers to assemble dIgA is unclear. The process
appears likely to involve conformation changes in both the JC and the IgA monomers because in the
absence of JC, monomeric IgA (the form typically found in serum) is functionally active (Woof and
Russell, 2011), suggesting that the interactions we observe between Tps from two IgA monomers
do not form, or are not stable, without the JC. Similarly, in the absence of contacts with IgA, the
conformation of JC we observe in SlgA and dIgA is not likely to be stable, suggesting that the JC’s
unliganded structure is unknown. These observations point to a polymerization process that involves
conformational changes in the JC and two IgA heavy chains, which promote β-sheet like interactions
between Tps and/or Tps and the JC as well as disulfide bond formation between chains. Recent
reports also suggest that molecular chaperons are involved in dIgA assembly (Suzuki et al., 2019;
Xiong et al., 2019).

**plgR binding to dIgA**

Following assembly, dIgA binding to the plgR ectodomain, or SC, is required for SlgA delivery to
the mucosa. When unliganded, SC domains adopt a compact conformation, in which conserved resi-
dues in D1 form an interface with D4 and/or D5 (Stadtmueller et al., 2016a). The SlgA structure
reveals a marked rearrangement of SC domains that positions residues in the D1 CDRs in direct con-
tact with C3α3, C3β3, TpC, TpD and JC and positions D5 CDRs in contact with C4β3C. This confor-
mational change is consistent with published double electron-electron resonance (DEER) distance
measurements showing a 70 Å separation between nitroxide spin labels attached to SC D1-resi-
due 67 and D5 residue 491 in human SlgA (Stadtmueller et al., 2016a); the distance between Cα atoms
of equivalent residues in the mouse SlgA structure, D1 Ile67 and D5 His493, is 61 Å. The SC crystal
structure and DEER experiments led to a model, in which accessible motifs in D1 contact dIgA in a
recognition binding event that triggers a conformational change allowing previously buried D1
motifs and accessible D5 motifs to bind dIgA (Stadtmueller et al., 2016a). Our structure is consis-
tent with this model, revealing that conserved D1 residues Arg31, His32, Arg34, Thr48 and Tyr 55,
which bind D4 or D5 in unliganded SC, mediate interactions with C4β3 and JC. It is yet unclear which
residues mediate the recognition binding event; however, one possibility is that SC D1-CDR3 resi-
dues, which are exposed in unliganded SC, form the initial contacts with the tailpieces (TpC and TpD)
and C4β3α3, which are exposed in dIgA (Figure 5C). Our structures suggest that during the conforma-
tional change, D2 moves away from D1 while D3 forms hydrophobic interactions with D1 in a pro-
cess that positions the D4-D5 arm near the FC and D5 in contact with C3α2C.

The plgR (and SC) has long been known to bind only JC-containing molecules (Woof and Russell,
2011), a requirement that is validated by the SC-JC interfaces in the SlgA structure. However, it also
appears likely that the JC indirectly supports SC binding by inducing the bend that positions C4β3α3
(which binds D1) and C4β3α3 (which binds D5) optimally to be bound by the SC bridge comprising D1-
D3-D4 and D5. D2 is absent from this bridge, and does not appear to contact the heavy chains or
JC, suggesting that despite its reported contribution to dIgA-binding kinetics (Stadtmueller et al.,
2016a), its role is indirect. The distal location of D2 in the SlgA structure is also consistent with a
model in which SC from birds, reptiles and amphibians, which are lacking the D2 domain
(Stadtmueller et al., 2016b), would bind analogous to D1-D3-D4-D5 in our SlgA structure.
The functional significance of SC

Although the pIgR plays a critical role in delivering SIgA to mucosal secretions, functionally, why its ectodomain (SC) remains attached to SIgA is less clear. While SC may play a role in stabilizing the conformation of SIgA, the dIgA structure suggests that SC binding induces relatively small changes in the bent and tilted relationship between the two IgAs. SC has been reported to protect SIgA from degradation and to play a role in binding to host and bacterial factors (Kaetzel, 2005). Indeed, our structure reveals SC forming extensive interfaces with the heavy chains and the JC where it may protect areas especially vulnerable to proteolysis; however, roughly 16% of the SIgA core (not including Fabs) is occluded by SC, leaving the majority of the molecule exposed. Being located on one face of the molecule, SC exhibits significant accessible surface area (in excess of 25,000 Å²) leaving it well-positioned to interact with host or microbial factors. D2 is particularly accessible, being located distal from SIgA’s center and exhibiting evidence of flexibility. It also includes four of the seven PNGS which could facilitate carbohydrate-mediated binding events, although the number of PNGS on each SC domain is variable among species. It is unclear why mammalian SC evolved to include D2; our observations point toward a role in mediating interactions with host and/or microbial factors (Figure 7).

Comparisons with human SIgA

SIgA is found in all mammals; however, some species including humans, express multiple isoforms, and allotypes of the IgA heavy chain, which are associated with unique expression patterns and functions in vivo (Woof and Russell, 2011). Furthermore, differences in IgA heavy chain C\textsubscript{H}2 and C\textsubscript{H}3, JC and SC sequences have also been associated with variable glycosylation patterns, variable interactions with FcRs and differences in potential to form higher order polymers (e.g. tetrameric SIgA) (Woof and Russell, 2011), suggesting that on some level, SIgA structures are heterogeneous. Understanding SIgA structural heterogeneity is important for understanding function, especially between mouse and human IgAs given the long history of using mouse model systems for immunological research (Masopust et al., 2017). Our structures include the C\textsubscript{H}2 and C\textsubscript{H}3 domains from the single IgA heavy chain found in the mouse genome, which along with JC and SC sequences share 65–77% identity and 79–86% similarity with human counterpart sequences (Figure 1—figure supplement 1A). Analysis of our structures and comparison of human and mouse SIgA component sequences revealed a handful of differences at interfaces between complex components, including contacts between the JC wings and SC or C\textsubscript{H}3 (Figure 1—figure supplement 1B). However, the majority of interfacing residues appear to be conserved (Figure 1—figure supplement 1B). Consistent with these observations, comparison of our mouse SIgA with structures of dimeric forms of the human SIgA1 core (no Fabs), reported in the same timeframe as this work (Kumar et al., 2020; Wang et al., 2020) revealed largely superimposable structures with some variability in contacts between complex components (not shown). The modeled positions of SC D2 are variable among structures, likely due to inherent flexibility in the domain’s position, however, the structures have similar angles of bend (99 degrees in human SIgA and 97 degrees in mouse SIgA) and tilt (31 degrees in human SIgA and to 30 degrees tilt in mouse SIgA). Although additional studies will be needed to determine if and how sequence variability in SIgA components might uniquely influence SIgA assembly and functions, the similarities we observe suggest that the core Fc region of SIgA variants found in mammalian mucosa adopt largely conserved structures characterized by limited conformational flexibility, which our modeling suggests will constrain the positions that Fabs can occupy.

SIgA functions

Our data provide a broad range of structural findings, addressing long-outstanding questions regarding the structure of the IgA Tps, the JC and SC, yet how the structure supports SIgA’s divergent roles in pathogen clearance and microbial homeostasis remains an open topic of investigation. In the mucosa, SIgA encounters a broad range of binding partners and antigens. Besides pIgR, a number of FcR are known to interact with IgA, including pathogen receptors such as SSL7, Arp4 or Sir22, and host receptors such as FcεR1, although FcεR1 is absent from the mouse genome. IgA FcRs typically bind a canonical site located near C\textsubscript{H}2-C\textsubscript{H}3 elbow (Herr et al., 2003; Ramsland et al., 2007; Kazeeva and Shevelev, 2009). In the SIgA and dIgA structures, two of the four accessible
sites are occluded by JC wing interactions with C_{3A} and C_{3C}. This leaves two sites, on C_{3B} and C_{3D}, accessible for FcR binding. Notably, these sites are located on the convex edge of the complex, which our computational modeling predicts, is never occluded by Fabs (Figures 6 and 7). This arrangement would provide an unobstructed approach for FcRs to bind SlgA, which in the context of host cell receptors would promote favorable outcomes. On the other hand, it could leave SlgA especially vulnerable to pathogen FcRs, the binding of which might lead to Fab cleavage or other outcomes that would be detrimental to the host. We also anticipate that the bent and tilted relationship between the two IgAs, and predicted directionality of the two sets of Fabs, would influence how SlgA engages antigen. In the context of endogenous mammalian antibodies, unique properties of different C_{4}Z - C_{4}I linkers and stabilizing interactions between the heavy chains and light chains are likely to further dictate Fab positions; however, perhaps both sets of Fabs being directed toward the concave side would stabilize binding to flexible antigens, to carbohydrates, or perhaps influence the strength of antigen crosslinking (Figure 7).

Figure 6. Modeling of Fabs and CDRs on SlgA. (A) SlgA structure shown with four Fabs, each modeled in a single position as well as all possible positions and shown in multiple orientations. Complex components are colored according to the key (center). (B) The number of positions sampled by Fab CDRs were quantified by measuring the angle (Φ) between the Fab vector and the y-axis, a vector parallel to the FcAB plane and passing through the center of mass of Fc and JC. The frequency (normalized count) of each angle is shown as a histogram. The online version of this article includes the following figure supplement(s) for figure 6: Figure supplement 1. Modeling strategy for the computational search of SlgA Fab positions.
### Key resources table

| Reagent type (species) or resource | Designation                        | Source or reference                  | Identifiers                  | Additional information                                                                 |
|-----------------------------------|------------------------------------|-------------------------------------|-----------------------------|---------------------------------------------------------------------------------------|
| Other                             | IGHA_MOUSE                         | UniProtKB                           | UniProtKB: P01878           | Amino acid sequence of *Mus musculus* IgA heavy chain constant regions                  |
| Other                             | LAC1_MOUSE                         | UniProtKB                           | UniProtKB: P01843           | Amino acid sequence of *Mus musculus* lambda-1 light chain constant region               |
| Other                             | IGJ_MOUSE                          | UniProtKB                           | UniProtKB: P01592           | Amino acid sequence of *Mus musculus* joining chain                                     |
| Other                             | PIGR_MOUSE                         | UniProtKB                           | UniProtKB: O70570           | Amino acid sequence of *Mus musculus* polymeric Ig Receptor                             |
| Other                             | STA121                             | Moor et al., 2017                   |                             | Variable region of human IgA2 antibody STA121                                         |
| Cell line (Homo sapiens)          | HEK Expi293F                       | Thermo Fisher                       | Cat#: A14635 RRID: CVCL_D615| Construct encoding STA121 heavy chain (HC) variable region fused to *Mus musculus* IgA HC constant domains |
| Recombinant DNA reagent           | pD2610-v1                          | ATUM                                 | Cat#: pD2610-v1-03          | Mammalian expression plasmid                                                          |
| Transfected construct (Homo sapiens) | STA121 IgA Heavy chain (HC) in pD2610v1 | Materials and methods of this paper |                             | Construct encoding STA121 heavy chain (HC) variable region fused to *Mus musculus* IgA HC constant domains |
| Transfected construct (Homo sapiens) | STA121 IgA light chain (LC) lambda in pD2610v1 | Materials and methods of this paper |                             | Construct encoding STA121 light chain (LC) variable region fused to *Mus musculus* lambda LC constant domain |
| Transfected construct (Homo sapiens) | Mouse Secretory component (SC) in pD2610v1 | Materials and methods of this paper |                             | Construct encoding residues 1–567 of *Mus musculus* pIgR (a.k.a. secretory component, SC) |
| Transfected construct (Homo sapiens) | Mouse joining chain (JC) in pD2610v1 | Materials and methods of this paper |                             | Construct encoding *Mus musculus* joining chain (JC)                                   |
| Peptide, recombinant protein      | STA121 Secretory IgA               | Materials and methods of this paper |                             | Protein complex produced from transfected constructs and containing: STA121 HC, STA121 LC, JC, SC |
| Peptide, recombinant protein      | STA121 dimeric IgA                  | Materials and methods of this paper |                             | Protein complex produced from transfected constructs and containing: STA121 HC, STA121 LC, JC, SC |
| Other                             | CaptureSelect LC-lambda (Mouse) Affinity Matrix | Thermo Fisher RRID: SCR_008452 | Cat#: 194323005           | Affinity matrix for protein purification                                               |
| Other                             | Superose 6 Increase 10/300 GL      | GE Healthcare Life Sciences RRID: SCR_000004 | Cat#: 29091596           | Size exclusion column for protein purification                                         |
| Software, algorithm               | Rosetta CryoEM refinement package  | Wang et al., 2016 RRID: SCR_015701 |                             |                                                                                        |
| Software, algorithm               | Phenix                             | Afonine et al., 2018a Afonine et al., 2018b RRID: SCR_014224 Phenix.refine RRID: SCR_016736 |                             |                                                                                        |
| Software, algorithm               | Pymol Molecular Graphics System    | Schrodinger LLC RRID: SCR_014879 RRID: SCR_000305 |                             |                                                                                        |

Continued on next page
### Construct design and protein expression

Genes encoding the *Mus musculus* IgA heavy chain constant region (Uniprot P01878) and the lambda light chain constant region (Uniprot A0A0G2JE99) were fused with the STA121 V\textsubscript{H} and V\textsubscript{L} domain sequences (Moor et al., 2017) to create complete heavy chain and light chain sequences. The TPA signal sequence (residues MDAMKRGLCCVLLLCGAVFVSPAGA) was encoded at the start of the heavy chain sequence and the mouse IgKappa signal sequence (residues METD TLLLWVLLLWVPGSTG) was encoded at the start of the light chain sequence. These sequences, along with *Mus musculus* JC (Uniprot P01592; native signal peptide) and *Mus musculus* pIgR ectodomain (SC) residues 1–567 (Uniprot O70570; native signal peptide), were codon optimized, synthesized (Integrated DNA Technologies, Inc) and each cloned into mammalian expression vector pD2610v1 (Atum). Resulting expression constructs were transiently co-transfected into HEK Expi293-F cells with ExpiFectamine, according to company protocol (Thermo Fisher). Co-transfection to produce SIgA and dIgA utilized equal amounts of each DNA expression construct; SlgA was produced by co-transfecting all four constructs, whereas dIgA was produced by co-transfecting all constructs except for the construct encoding the pIgR ectodomain (SC). Five days following transfection cellular supernatants were harvested and SlgA complexes were purified using CaptureSelect LC-lambda (Mouse) Affinity Matrix (Thermo Fisher) and Superose 6 (GE Healthcare Life Sciences) size exclusion chromatography (SEC). SEC fractions corresponding to the expected size of SlgA (containing two IgA) and dIgA were maintained in buffer containing 20 mM Tris-HCl pH 7.4 and 150 mM NaCl and utilized for cryoEM. Requests for reagents generated in this paper should be directed to the corresponding author, Beth Stadtmueller. Permission to use any reagent containing the STA121 sequence must be obtained from Institute for Research in Biomedicine (Bellinzona, Switzerland).

### CryoEM grid preparation and data collection

Quantifoil R2/2 300 mesh grids were glow discharged using a Pelco easiGlow system for 1 min at 20 mA current. A 3 μL drop of the SlgA or dIgA sample (1.5 mg/ml and 1.0 mg/ml respectively) was applied to the grids and blotted (Whatman #1 paper) for 2 to 8 s using Vitrobot Mark IV (Thermo Fisher) with a blot force of 5, 0 s wait and drain time, at 4°C and 100% RH. Grids were plunged into liquid nitrogen cooled ethane. Movies were collected using SerialEM (Schorb et al., 2019) on a Titan Krios (Thermo Fisher) operating at 300 kV, equipped with BioQuantum Energy Filter (20 eV slit width, Gatan) and a K3 direct electron detector (Gatan). SlgA movies were collected at 130,000 magnification in super resolution mode with calibrated pixel size of 0.326 Å/pixel, 40 frames per movie, 0.03 s per frame, and total dose of ~60 electrons/Å\(^2\). The dIgA movies were collected at 105,000 magnification in super resolution mode with calibrated pixel size of 0.418 Å/pixel, 40 frames per movie, 0.05 s per frame, and total dose of ~60 electrons/Å\(^2\).

### CryoEM data processing

For the SlgA, 1512 movies were collected. The dataset was processed independently in cryoSparc v. 2.X (Punjani et al., 2017) and Relion 3.1-beta. (Scheres, 2012; Zivanov et al., 2018; Figure 2—figure supplement 1, Figure 2—figure supplement 2). Maps resulting from the cryoSparc and Relion three processing pipelines were combined (phenix.combine_focused_maps). The resulting...
Figure 7. Model for the formation, transport, and function of SlgA. Schematic summary depicting the unliganded SC structure (pdb code 5D4K) as plgR bound to basolateral surface of an epithelial cell in its closed conformation and recognizing bent dlgA from the lamina propria. The plgR binding to dlgA triggers a conformation change that repositions its domains to facilitate numerous stabilizing contacts with dlgA. The dlgA:plgR complex transcytoses to the apical membrane where the plgR is proteolytically cleaved, releasing SlgA into the mucosa. In the mucosa, SlgA Fabs (shown in all possible modeled positions) are directed toward the concave side of the antibody ‘looking’ for potential antigens while its Fc-receptor-binding regions are exposed on the convex side and accessible to potential host or microbial receptors. SC domains are also partially exposed; the D2 domain is almost completely accessible, protruding out of the SlgA where it may bind host and bacterial factors. Upon encountering antigen, SlgA Fabs bind, promoting antigen coating, agglutination, or enchained growth.
‘combined map’ was used for initial model building (described below); however, rebuilding and refinement utilized the cryoSparc processing pipeline, which is designated as the primary data in the EMDB.

CryoSparc Live beta was used for the initial processing of the data - movies were motion corrected, binned by 2, and CTF estimation was performed. Particles were initially picked by blob picker with minimum and maximum particle diameter of 100 Å and 120 Å, respectively. Streaming 2D classification was used to generate initial 2D references, which were then used as templates for particle picking. After completion of the collection session, images were screened by limiting total motion to 100 px or less and CTF fit estimation to 10 Å or better, resulting in 1167 selected images or 956,981 picked particles. Particles were exported into cryoSparc for further processing. They were 2D classified into 100 classes and the best classes were selected (623,071 particles). A smaller subset of only large views was also selected (123,533) and four ab-initio structures were generated. The best structure (79,904 particles) was refined against the 123,533 particle set to 3.6 Å resolution. The volume was low pass filtered to 60 and both the 3.6 Å map and the filtered volume were used in heterogeneous refinement with the all good particles from 2D classification. The two classes generated a 4.4 Å structure with 368,541 particles and a noise class at 9 Å with 254,530 particles. Good class was refined to 3.4 Å using non-uniform refinement job type. The particles were then used in heterogeneous refinement job type using the output of non-uniform refinement as two input volumes. Once again, one of the classes reached 4.2 Å resolution with 229,252 particles and the other class served as a trap for bad particles (7.22 Å with 139,289 particles). The good particle set was refined using homogeneous refinement job type to 3.5 Å and 3.2 Å with non-uniform refinement.

At the same time, all the movies were motion corrected, binned by 2, and dose-weighted in Relion v.3. CTFs were estimated with CTFFind4 (v. 4.1.13) (Rohou and Grigorieff, 2015). Images were screened for ice thickness, contamination, and CTF fits, resulting in 1266 remaining images. Particles were auto-picked with LoG picker in Relion 3 (minimum and maximum diameters of 60 and 180 Å respectively). Approximately 1.4 million particles were extracted, binned four times, and subjected to several rounds of 2D classification. At each step, good classes were selected, and remaining particles were classified again with a smaller mask diameter. Selected particles (310,626) were merged into a single file and re-extracted with 2X binning and further underwent 2D classification. Remaining 184,737 particles were used in a 3D refinement (using an initial model generated in a parallel cryoSparc processing pipeline). After the refinement, particles were re-extracted at full pixel size and refined. After this initial refinement, the particles underwent Bayesian polishing and per-particle-CTF estimation. Particles were once again refined and the map was post-processed and B-factor sharpened, resulting in the final average resolution of 3.7 Å at FSC = 0.143.

For the dIgA, 2951 movies were initially collected using beam-image shift across nine holes in a 3 × 3 pattern. All the processing steps were done in cryoSparc v2. Movies were imported, motion corrected, and CTFs were estimated. Exposures were then curated to remove bad micrographs, resulting in 2258 remaining images. Particles were picked based on five templates generated from a 4.5 Å structure generated from an earlier, smaller dataset. Particles (1.81 million) were extracted and binned 4X for an initial round of 2D classification. All classes showing secondary structure were selected (688,183). Remaining particles were 2D classified with a smaller circular mask, to try and capture projections along the length of the particle. This round of classification was repeated twice, resulting in additional 168,530 and 59,876 particles. The initial particle subset was used to generate six ab-initio structures and then processed with heterogeneous refinement with three classes. The best class was combined with the remaining two subsets and another round of heterogeneous refinement was performed. 462,707 particles belonging to the good class were re-extracted at full pixel size and another round of heterogeneous refinement with three classes was performed. Resulting set of 288,823 particles were refined (Legacy). The particle stack was then split into nine groups based on the pattern of collection. Per-particle CTF refinement, beam tilt, trefoil, and spherical aberration corrections were performed for each group. A final round of homogeneous refinement was performed, resulting in the final overall resolution of 3.3 Å at FSC = 0.143.

**Structure building, refinement, and validation**

Starting models for SIgA structure determination were made using SWISS-MODEL (Waterhouse et al., 2018). Briefly, homology models of mouse IgA Fc and SC were generated using the amino acid sequences corresponding to individual components, heavy chain and SC, and
reference pdb files 1OW0 chains A and B (human monomeric IgA Fc) and 5D4K (unliganded human SC), respectively (Herr et al., 2003; Stadtmueller et al., 2016a). Tps and JC sequences did not exhibit homology to any known structure and were not modeled. Homology models were docked into to real-space electron density using UCSF Chimera (Pettersen et al., 2004; Yang et al., 2012); each Fcs was docked as a single unit, while SC domains D1, D2, D3, D4, D5 were docked individually. Domain positions were refined as rigid bodies using Phenix (Afonine et al., 2018b). Inspection of the map and preliminary SlgA model fit, revealed numerous loops that fit density poorly as well as unaccounted for density at the center of the molecule (Tp and JC). Poorly fitting loops were manually re-built into density using the Coot Molecular Graphics Package (Emsley and Cowtan, 2004), along with preliminary placement of a subset of JC and Tp residues. The resulting unrefined SlgA model contained four HC, one SC, and one JC (14 folded domains).

To facilitate building the complete JC and Tp sequences, we implemented a strategy to build and refine the structure using a combination of the Rosetta CryoEM refinement package and Rosetta comparative modeling (RosettaCM) (Wang et al., 2016). Briefly, the unrefined starting model and combined map, including data from Relion and Cryosparc processing pipelines, were used to build 2913 Rosetta-modified structures using Rosetta 2019.35.60890, running on a 64 CPU server. All 2913 structures were scored based on geometry (determined by MolProbity score [Williams et al., 2018a]) and fit, which was approximated with a Rosetta-determined Fourier-shell correlation (FSC) between the structure’s calculated map and the experimental density map. We scored each structure by dividing the FSC by the MolProbity score; the four highest-scoring structures were compared to the unrefined SlgA structure and the experimental map. Subsequent iterations of this process were used to rebuild portions of the C\(_2\) and C\(_3\) domains, during which the Phenix phenix.map_input_model_cc was used to determine model fit to density. The map surrounding SC D2 was poorly ordered and contained extra density likely representing an ensemble of D2 positions and/or partially ordered carbohydrates attached to the four PNGS on D2. We positioned the domain with the best average fit to the density using distance constraints of the linkers connecting it to D1 and D3. The resulting structure was rebuilt by hand and refined against the map from the cryoSparc processing pipeline using Coot and Phenix (Emsley and Cowtan, 2004; Afonine et al., 2018b). This process included the addition of carbohydrates at PNGS where well-ordered density for at least one base was visible, as well as refinement of atomic displacement parameters (ADPs).

The dIgA structure was determined using a similar approach as used for the SlgA structure. Briefly, homology models of the mouse IgA Fc were docked into the dIgA map along with the JC from the SlgA structure. Domain positions were refined as rigid bodies using Phenix (Afonine et al., 2018b) and poorly fitting regions were re-built by hand. The resulting structure, including carbohydrates, was refined using the Rosetta CryoEM refinement package and Phenix (Wang et al., 2016; Afonine et al., 2018a).

The final SlgA and dIgA structures and their fit to cryoEM maps were evaluated by hand and validated using Phenix EM Validation, Molprobity, and EMRinger (Barad et al., 2015; Afonine et al., 2018a; Williams et al., 2018a); results are summarized in Supplementary file 1. Additionally, each final models’ fit to unsharpened data was evaluated by calculating map to model FSC curves against all data (FSCmap-model) and against two half maps, to access overfitting. Briefly, the final refined models were displaced by 0.5 Å after setting the B-factor to 99 using Phenix pdbtools. Each displaced model was refined against half of the data (half map 1) and the resulting refined model and half map 1 were used to calculate a map to model FSC termed ‘FSCwork’; the refined model and the half of the data not used in refinement calculation (half map 2) were used to calculate a map to model FSC termed ‘FSCfree’.

SlgA and dIgA cryoEM maps and structure coordinate files have been deposited in the EM database with accession codes EMD-22309 (dlgA) and EMD-22310 (SlgA) and the protein databank with accession codes 7JG1(dlgA) and 7JG2 (SlgA).

**Structure analysis and figures**

The sequence alignment between human and mouse counterparts of JC (human: Uniprot P01591, mouse: Uniprot P01592), plgR ectomain (human: Uniprot P01833, mouse: Uniprot O70570) and heavy chain- C\(_{\mu}2\)-C\(_{\mu}3\) regions (human IgA1: Uniprot P01876, human IgA2: Uniprot P01877, human IgA two allotypes [Lombana et al., 2019], mouse IgA: Uniprot Q99LA6) were carried out using
Contacts between individual Tps and the JC and all other S IgA components were evaluated by inspecting all interfacing residues within approximately 7 Å. The probability of conserved contacts feature (planes were then measured relative to one another from the Tools > analysis > centroid/axes/planes feature (Pettersen et al., 2004; Yang et al., 2012). The RMSD differences between S IgA and dIgA structures were calculated using PyMol. To visualize the difference between S IgA and dIgA, structures were aligned on JC Cα atoms, which have a RMSD of 0.891. Figures were made using the PyMol Molecular Graphics System (Schrodinger LLC), UCSF Chimera (Pettersen et al., 2004; Yang et al., 2012), UCSF ChimeraX (Goddard et al., 2018) and Inkscape (https://inkscape.org/).

### Acknowledgements

We thank Pamela Bjorkman and members of the Bjorkman lab at Caltech for supporting preliminary work related to this study as well as Jost Vielmetter and the Caltech Protein Expression Center (housed and funded in part by the Caltech Beckman Institute) for guidance on IgA protein expression strategies. We also thank members of the Stadtmueller Laboratory, as well as Emma Slack (ETH Zurich), for insightful conversations regarding the SIgA and dIgA structures. Cryo Electron microscopy was done in the Beckman Institute cryo-EM resource center at Caltech. The STA121 antibody sequence was provided by Luca Piccoli and the Institute for Research in Biomedicine (Bellinzona, Switzerland). This study was funded by University of Illinois Urbana-Champaign start-up funding, National Institutes Health (United States) grant P41-GM10460, and National Institutes Health (United States) grant R01 AI041239. Molecular graphics and analyses performed with UCSF Chimera, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from NIH P41-GM103311. Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-
GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

Additional information

Funding

| Funder                                      | Grant reference number | Author                                      |
|---------------------------------------------|------------------------|---------------------------------------------|
| University of Illinois at Urbana-Champaign  | Start-up funding       | Sonya Kumar Bharathkar, Benjamin W Parker, Beth Stadtmueller |
| National Institute of General Medical Sciences | P41-GM10460           | Nandan Haloi, Emad Tajkhorshid             |
| National Institutes of Allergy and Infectious Diseases | R01 AI041239         | Kathryn E Huey-Tubman, Beth Stadtmueller    |

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Sonya Kumar Bharathkar, Benjamin W Parker, Conceptualization, Resources, Formal analysis, Validation, Investigation, Visualization, Methodology, Writing - original draft, Writing - review and editing; Andrey G Malyutin, Validation, Investigation, Visualization, Writing - original draft; Nandan Haloi, Formal analysis, Visualization, Methodology, Writing - original draft; Kathryn E Huey-Tubman, Resources, Methodology, Writing - review and editing; Emad Tajkhorshid, Supervision, Funding acquisition, Methodology; Beth M Stadtmueller, Conceptualization, Resources, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Visualization, Writing - original draft, Project administration, Writing - review and editing

Author ORCIDs

Sonya Kumar Bharathkar [https://orcid.org/0000-0002-3637-4990](https://orcid.org/0000-0002-3637-4990)
Benjamin W Parker [https://orcid.org/0000-0002-1464-1189](https://orcid.org/0000-0002-1464-1189)
Emad Tajkhorshid [http://orcid.org/0000-0001-8434-1010](http://orcid.org/0000-0001-8434-1010)
Beth M Stadtmueller [https://orcid.org/0000-0003-0637-3206](https://orcid.org/0000-0003-0637-3206)

Decision letter and Author response

Decision letter [https://doi.org/10.7554/eLife.56098.sa1](https://doi.org/10.7554/eLife.56098.sa1)
Author response [https://doi.org/10.7554/eLife.56098.sa2](https://doi.org/10.7554/eLife.56098.sa2)

Additional files

Supplementary files

- Supplementary file 1. CryoEM data collection and refinement statistics associated with SlgA and dlgA structures.
- Transparent reporting form

Data availability

SlgA and dlgA cryoEM maps and structure coordinate files have been deposited in the EM databank with accession codes EMD-22309 (dlgA) and EMD-22310 (SlgA) and the protein databank with accession codes 7JG1(dlgA) and 7JG2 (SlgA).

The following datasets were generated:

| Author(s)         | Year | Dataset title                  | Dataset URL               | Database and Identifier |
|-------------------|------|--------------------------------|---------------------------|-------------------------|
| Bharathkar SK, Par- | 2020 | Secretory Immunoglobin A (SlgA) | [https://www.rcsb.org/](https://www.rcsb.org/) | RCSB Protein Data |

Kumar Bharathkar, Parker, et al. eLife 2020;9:e56098. DOI: [https://doi.org/10.7554/eLife.56098](https://doi.org/10.7554/eLife.56098)
Kumar Bharathkar, Parker, et al. eLife 2020;9:e56098. DOI: https://doi.org/10.7554/eLife.56098

References

Afonine PV, Klaholz BP, Moriarty NW, Poon BK, Sobolev OV, Terwilliger TC, Adams PD, Urzhumtsev A. 2018a. New tools for the analysis and validation of cryo-EM maps and atomic models. Acta Crystallographica Section D, Structural Biology 74:814–840. DOI: https://doi.org/10.1107/S2059798318009324

Afonine PV, Poon BK, Read RJ, Sobolev OV, Terwilliger TC, Urzhumtsev A, Adams PD. 2018b. Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallographica. Section D, Structural Biology 74:531–544. DOI: https://doi.org/10.1107/S2059798318006551, PMID: 29872004

Altshul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. Journal of Molecular Biology 215:403–410. DOI: https://doi.org/10.1016/0022-2836(95)00018-5

Barad BA, Echols N, Wang RY, Cheng Y, DiMaio F, Adams PD, Fraser JS. 2015. EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. Nature Methods 12:943–946. DOI: https://doi.org/10.1038/nmeth.3541, PMID: 26280328

Bloom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S. 2004. Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics 4:1633–1649. DOI: https://doi.org/10.1002/pmic.200300771, PMID: 15174133

Bonner A, Almogren A, Furtado PB, Kerr MA, Perkins SJ. 2009. The nonplanar secretory IgA2 and near planar secretory IgA1 solution structures rationalize their different mucosal immune responses. Journal of Biological Chemistry 284:5077–5087. DOI: https://doi.org/10.1074/jbc.M807529200

Brandtzaeg P. 2013. Secretory IgA: designed for Anti-Microbial defense. Frontiers in Immunology 4:222. DOI: https://doi.org/10.3389/fimmu.2013.00222, PMID: 23964273

Donaldson GP, Ladinsky MS, Yu KB, Sanders JG, Yoo BB, Chou WC, Conner ME, Earl AM, Knight R, Bjorkman PJ, Mazmanian SK. 2018. Gut Microbiota utilize immunoglobulin A for mucosal colonization. Science 360:795–800. DOI: https://doi.org/10.1126/science.aao9026, PMID: 29724905

Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. Acta Crystallographica. Section D, Biological Crystallography 60:2126–2132. DOI: https://doi.org/10.1107/S0907444904019158, PMID: 15572765

Flajnik MF. 2010. All GOD’s creatures got dedicated mucosal immunity. Nature Immunology 11:777–779. DOI: https://doi.org/10.1038/nmi0910-777, PMID: 20720582

Goddard TD, Huang CC, Meng EC, Petterson EF, Couch GS, Morris JH, Ferrin TE. 2018. UCSF ChimeraX: meeting modern challenges in visualization and analysis. Protein Science 27:14–25. DOI: https://doi.org/10.1002/pro.3235

Gouet P, Courcelle E, Stuart D, Metoz F. 1999. ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics 15:305–308. DOI: https://doi.org/10.1093/bioinformatics/15.3.305

Herr AB, Ballister ER, Bjorkman PJ. 2003. Insights into IgA-mediated immune responses from the crystal structures of human FcαRI and its complex with IgA1-Fc. Nature 423:614–620. DOI: https://doi.org/10.1038/nature01685, PMID: 12768205

Humphrey W, Dalke A, Schulten K. 1996. VMD: visual molecular dynamics. Journal of Molecular Graphics 14:33–38. DOI: https://doi.org/10.1016/0727-0926(96)00018-5

Kaetzel CS. 2005. The polymeric immunoglobulin receptor: bridging innate and adaptive immune responses at mucosal surfaces. Immunological Reviews 206:83–99. DOI: https://doi.org/10.1111/j.1058-4838.2005.00278.x, PMID: 16048543

Kaetzel CS. 2014. Coevolution of mucosal immunoglobulins and the polymeric immunoglobulin receptor: evidence that the commensal Microbiota provided the driving force. ISRN Immunology 2014:1–20. DOI: https://doi.org/10.1155/2014/541537

Kazevea TN, Shevelev AV. 2009. IgA-specific proteins of pathogenic Bacteria. Biochemistry 74:12–21. DOI: https://doi.org/10.1134/S00062979090100027, PMID: 19232043

Krispin E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. Journal of Molecular Biology 372:774–797. DOI: https://doi.org/10.1016/j.jmb.2007.05.022, PMID: 17681537

Kumar N, Arthur SP, Ciferri C, Matsumoto ML. 2020. Structure of the secretory immunoglobulin A core. Science 367:1008–1014. DOI: https://doi.org/10.1126/science.aaaz5807

Lombana TN, Rajan S, Zom JA, Mandikian D, Chen EC, Estevez A, Yip V, Bravo DD, Phung W, Farahi F, Viajar S, Lee S, Gill A, Sandoval W, Wang J, Ciferri C, Boswell CA, Matsumoto ML, Spiess C. 2019. Production,
characterization, and in vivo half-life extension of polymeric IgA molecules in mice. mAbs 11:1122–1138. DOI: https://doi.org/10.1080/19420862.2019.1622940, PMID: 31122132

Marques R, Bouville C, Ribardiêre M, Santos LP, Bouatouch K. 2013. Spherical fibonacci point sets for illumination integrals. Computer Graphics Forum 32:134–143. DOI: https://doi.org/10.1111/cgf.12190

Masopust D, Sivula CP, Jameson SC. 2017. Of mice, dirty mice, and men: using mice to understand human immunology. The Journal of Immunology 199:383–388. DOI: https://doi.org/10.4049/jimmunol.1700453, PMID: 28896328

Moore K, Diard M, Sellin ME, Felmy B, Wotzka SY, Toska A, Bakkeren E, Arnoldini M, Bansept F, Co AD, Vollér T, Minola A, Fernandez-Rodriguez B, Agatic G, Barbieri S, Piccoli L, Casiragi C, Corti D, Lanzavecchia A, Regoes RR, et al. 2017. High-avidity IgA protects the intestine by enchaining growing bacteria. Nature 544:498–502. DOI: https://doi.org/10.1038/nature22058, PMID: 28405025

Pabst O, Slack E. 2020. IgA and the intestinal microbiota: the importance of being specific. Mucosal Immunology 13:12–21. DOI: https://doi.org/10.1038/s41385-019-0227-4, PMID: 31201222

Randall PA, Willoughby N, Trist HM, Farrugia W, Hogarth PM, Fraser JD, Wines BD. 2007. Structural basis for evasion of IgA immnity by Staphylococcus aureus revealed in the complex of SSL7 with fc of human IgA1. PNAS 104:15051–15056. DOI: https://doi.org/10.1073/pnas.0706028104, PMID: 17848512

Regier EW, Frantz AL, Bruno ME, Wedlund L, Cohen DA, Stromberg AJ, Kaetzel CS. 2014. Secretory antibodies in breast milk promote long-term intestinal homeostasis by regulating the gut microbiota and host gene expression. PNAS 111:3074–3079. DOI: https://doi.org/10.1073/pnas.1315792111, PMID: 24569806

Rohou A, Grigorieff N. 2015. CTFFIND4: fast and accurate defocus estimation from electron micrographs. Journal of Structural Biology 192:216–221. DOI: https://doi.org/10.1016/j.jsb.2015.08.008, PMID: 26278980

Scheres SH. 2012. RELION: implementation of a bayesian approach to cryo-EM structure determination. Journal of Structural Biology 180:519–530. DOI: https://doi.org/10.1016/j.jsb.2012.09.006, PMID: 23000701

Schorb M, Haberbosch I, Hagen WJH, Schwab Y, Mastronarde DN. 2019. Software tools for automated transmission electron microscopy. Nature Methods 16:471–476. DOI: https://doi.org/10.1038/s41592-019-0396-9, PMID: 31086343

Sievers F, Wilm A, Dineen G, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Soding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using clustal omega. Molecular Systems Biology 7:539. DOI: https://doi.org/10.1038/msb.2011.75, PMID: 21988835

Stadtmueller BM, Huey-Tubman KE, López CJ, Yang Z, Hubbell WL, Bjorkman P.J. 2016a. The structure and dynamics of secretory component and its interactions with polymeric immunoglobulins. eLife 5:e10640. DOI: https://doi.org/10.7554/eLife.10640, PMID: 26943617

Stadtmueller BM, Yang Z, Huey-Tubman KE, Roberts-Matarić H, Hubbell WL, Bjorkman P.J. 2016b. Biophysical and biochemical characterization of avian secretory component provides structural insights into the evolution of the polymeric Ig receptor. The Journal of Immunology 197:1408–1414. DOI: https://doi.org/10.4049/jimmunol.1600463

Suzuki K, Vogelzang A, Fagarasan S. 2019. MZB1 folding and unfolding the role of IgA. PNAS 116:13163–13165. DOI: https://doi.org/10.1073/pnas.1908021116, PMID: 31201222

Tina KG, Bhadra R, Srinivasan N. 2007. PIC: protein interactions calculator. Nucleic Acids Research 35:W473–W476. DOI: https://doi.org/10.1093/nar/gkm423

Wang RY, Song Y, Barad BA, Cheng Y, Fraser JS, DiMaio F. 2016. Automated structure refinement of macromolecular assemblies from cryo-EM maps using rosetta. eLife 5:e17219. DOI: https://doi.org/10.7554/eLife.17219, PMID: 27669148

Wang Y, Wang G, Li Y, Zhu Q, Shen H, Gao N, Xiao J. 2020. Structural insights into secretory immunoglobulin A and its interaction with a pneumococcal adhesin. Cell Research 30:602–609. DOI: https://doi.org/10.1038/s41422-020-0336-3, PMID: 32398862

Waterhouse A, Berton M, Bienert S, Studer G, Tauriello G, Gumienny R, Heer FT, de Beer TAP, Rempfer C, Bordoli L, Lepore R, Schwede T. 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Research 46:W296–W303. DOI: https://doi.org/10.1093/nar/gky427, PMID: 29788355

Williams CJ, Headd JJ, Moriarty NW, Prisant MG, Vidal VE, Deis LN, Verma V, Keedy DA, Chen VB, Jain S, Lewis SM, Arendall WB, Snoeyink J, Adams PD, Lovell SC, Richardson JS, Richardson DC. 2018. MolProbity: more and better reference data for improved all-atom structure validation. Protein Science 27:293–315. DOI: https://doi.org/10.1002/pro.3330, PMID: 29067766

WooF JM, Kerr MA. 2006. The function of immunoglobulin A in immunity. The Journal of Pathology 208:270–282. DOI: https://doi.org/10.1002/path.1877, PMID: 16362985

WooF JM, Russell MW. 2011. Structure and function relationships in IgA. Mucosal Immunology 4:590–597. DOI: https://doi.org/10.1038/mi.2011.39, PMID: 21937984

Xiong E, Li Y, Min Q, Cui C, Liu H, Hong R, Lai N, Wang Y, Sun J, Matsumoto R, Takahashi D, Hase K, Shinkura R, Tsubata T, Wang JY. 2019. MZB1 promotes the secretion of J-chain-containing dimeric IgA and is critical for
the suppression of gut inflammation. *PNAS* **116**:13480–13489. DOI: https://doi.org/10.1073/pnas.1904204116, PMID: 31127044

Yang Z, Lasker K, Schneidman-Duhovny D, Webb B, Huang CC, Pettersen EF, Goddard TD, Meng EC, Sali A, Ferrin TE. 2012. UCSF chimera, MODELLER, and IMP: an integrated modeling system. *Journal of Structural Biology* **179**:269–278. DOI: https://doi.org/10.1016/j.jsb.2011.09.006, PMID: 21963794

Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, Scheres SH. 2018. New tools for automated high-resolution cryo-EM structure determination in RELION-3. eLife **7**:e42166. DOI: https://doi.org/10.7554/eLife.42166, PMID: 30412051