The Structures of Secretory and Dimeric Immunoglobulin A

Sonya Kumar Bharathkar[1†], Benjamin W. Parker[1†], Andrey Malyutin[2], Nandan Haloi[3,4], Emad Tajkhorshid[1,3,4], and Beth M. Stadtmueller[1*]

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

[2]Division of Biology and Biological Engineering, California Institute of Technology, Pasadena, CA 91125 USA

[3]Center for Biophysics and Quantitative Biology, University of Illinois at Urbana–Champaign, Urbana, 61801

[4]NIH Center for Macromolecular Modeling and Bioinformatics, Beckman Institute for Advanced Science and Technology, Urbana, 61801

[†] These authors contributed equally to this work

*Correspondence to: Beth M. Stadtmueller; bethms@illinois.edu

The authors declare that no competing interests exist.

Abstract:

Secretory (S) Immunoglobulin (I) A is the predominant mucosal antibody, which binds pathogens and commensal microbes. SIgA 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 SIgA and dIgA. 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 SIgA, 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 mucosa mediates vertebrate 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 molecules and mucosal antigens, resulting in antibodies with novel architecture and functional mechanisms compared to circulatory counterparts. The most prominent mammalian mucosal antibody is Secretory (S) Immunoglobin (Ig) A, with the average human secreting grams each day. SIgA is a polymeric antibody, typically composed of two IgA monomers (although higher order polymers have been reported) linked through one joining-chain (JC) and bound by one secretory component (SC) (Donaldson et al., 2018; Flajnik, 2010; Kaetzel, 2014; Pabst and Slack, 2020).

SIgA assembly begins in plasma cells, which link two IgA monomers, each having two heavy chains (HC) and two light chains (LC), and one JC to form dimeric (d) IgA. The polymeric Ig receptor (pIgR) binds dIgA on the basolateral surface of epithelial cells and transports it to the apical surface. Subsequently, the pIgR ectodomain, called secretory component (SC), is proteolytically cleaved, releasing the SC-dIgA complex into the mucosa where it is called SIgA (Figure 1A). Usually, SIgA contains four antigen binding fragments (Fabs) and two Fc regions. Each Fab is a dimer of variable LC and constant LC domains (V_L1-C_L1) bound to variable HC and constant HC domains (V_H1-C_H1); each Fc is a dimer of two constant HC domains (C_H), C_H2-C_H3 (Figure 1B). IgA C_H3 domains have a unique C-terminal extension called a tailpiece (Tp), which is also found on IgM C_H4, and is required for antibody oligomerization. In dIgA, the Tp is known to form inter-chain disulfide bonds with the JC, which is required for pIgR binding (Woof and Russell, 2011). The pIgR (and SC) contains five Ig-like domains (D1-D5), each having loops structurally similar to antibody complementarity determined regions (CDRs); D1 is necessary and sufficient for binding to dIgA, yet D5 is also known to bind through covalent and non-covalent interactions (Figure 1)(Stadtmueller et al., 2016a).

In contrast to circulatory antibodies, which typically induce inflammation, SIgA functions are thought to be dominated by physical mechanisms such as coating, cross-linking, agglutination and enchained growth of mucosal antigens (Moor et al., 2017). Orally administered SIgA is also protective and can influence the diversity in gut microbiota; for example, SIgA in colostrum has been shown to provide passive immunity to newborns and to have a life-long influence on microbiome composition (Rogier et al., 2014). Antigen binding is facilitated by four Fabs, which are presumed to increase binding avidity and antigen cross-linking, when compared to monomeric antibodies such as IgG. Additionally, the two Fcs bind host and microbial Fc receptors (FcR) that regulate effector and unknown functions, although the number of accessible receptor binding sites on each SIgA is uncertain. SC is also known to mediate direct interactions with host and bacterial proteins and carbohydrates, while also protecting the antibody from proteases (Woof and Russell, 2011). Despite significance, the structures of the JC, dIgA and SIgA have remained elusive.

To investigate how SIgA structure supports diverse mucosal functions, we targeted structures of mouse SIgA. Mice express a single IgA HC, with C_H domains sharing 78% sequence similarity with human IgA2m1 (Figure S1) and have long been a predominant model system for studying immune system function. Accordingly, we developed a method to co-express a mouse monoclonal IgA with mouse JC and SC in transiently transfected mammalian cells (see Materials and Methods). This resulted in monodisperse SIgA containing dIgA and little

evidence for higher-order polymers. Using cryoelectron microscopy (CryoEM) we determined a ~3.7Å resolution structure of SIgA, which revealed a pseudosymmetric arrangement of two IgA Fcs, bound asymmetrically to JC and SC. Comparison 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 SIgA Fabs. Together, results suggest that the pseudosymmetric arrangement of SIgA core components will constrain the positions of the Fabs and influence binding of host and microbial factors.

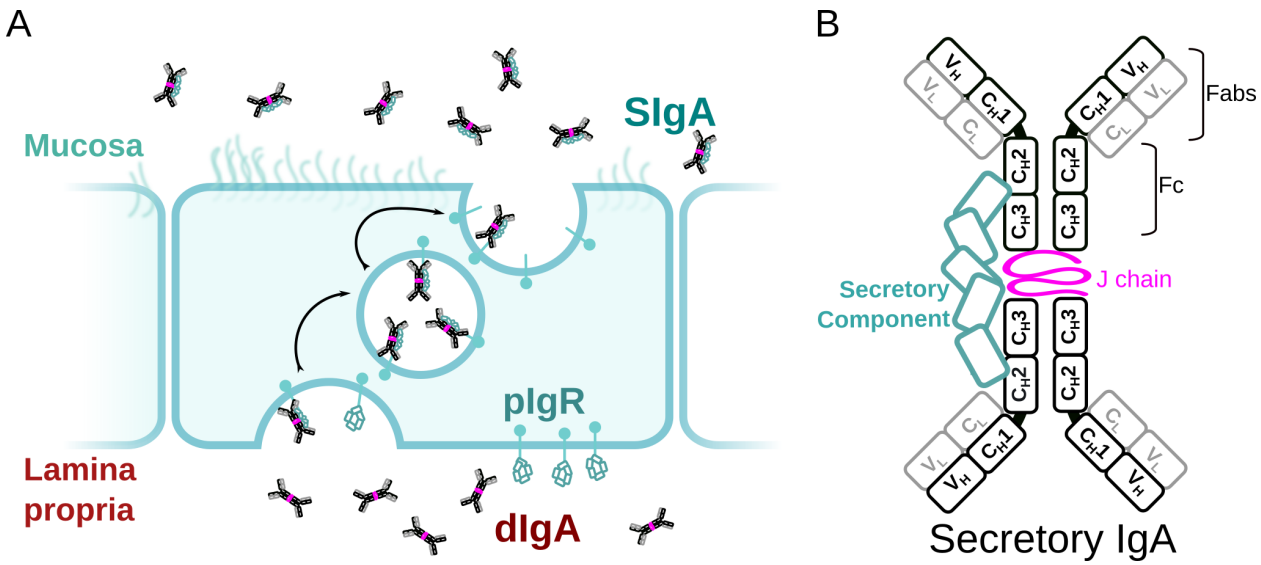

Figure 1. SIgA delivery to the mucosa. (A) Schematic depicting unliganded pIgR binding to dIgA from the lamina propria on the basolateral surface of an epithelial cell followed by transcytosis to the apical membrane and SIgA release into the mucosa. (B) Schematic showing protein components of SIgA, including two IgA monomers, joining chain (JC) and secretory component (SC). IgA Heavy chain constant (C_H), heavy chain variable (V_H), light chain variable (V_L) and light chain constant (C_L) domains are indicated along with antigen binding fragments (Fabs).

RESULTS

The Structure of SIgA

We determined the CryoEM structure of mouse SIgA to a final average resolution of 3.7Å; however, local map resolution was variable and was well-resolved at interfaces between complex components, revealing side chain density for many residues. Resolution was lowest for C_H2 domains and SC D2, located on outer edges of the complex (Figure S2). Although portions of the C_H1 domains were visible in the map, Fabs were mostly disordered and were not built. The refined structure revealed a pseudosymmetric assembly of two IgA monomers conjoined at the center by the JC and bound by SC. The C_H2-C_H3 domains in both Fcs aligned with published monomeric IgA Fc structure lacking the Tp (Herr et al., 2003); however despite shared sequences, the four IgA HCs formed structurally unique contacts with the JC and SC. To distinguish these differences, we designated a unique ID for each HC: A, B, C, or D, which we also use to describe corresponding C_H domains and Fcs. The

relationship between the Fc_{AB} and Fc_{CD} is non-linear, appearing bent and twisted with distinct concave and convex surfaces (Figure 2). To describe the conformation, we defined the angle between the centroid axes of Fc_{AB} and Fc_{CD} (97 degrees) as "bend", and the angle between the two non-intersecting centroid planes of Fc_{AB} and Fc_{CD} (30 degrees) as "tilt". SC asymmetrically contacts the same face of both Fc_{AB} and Fc_{CD} along what we define as the "front face" of the molecule. Potential N-linked glycosylation sites (PNGS) are distributed throughout the complex; however, the seven, PNGS located on SC, are clustered on the front face of SIgA (Figure 2).

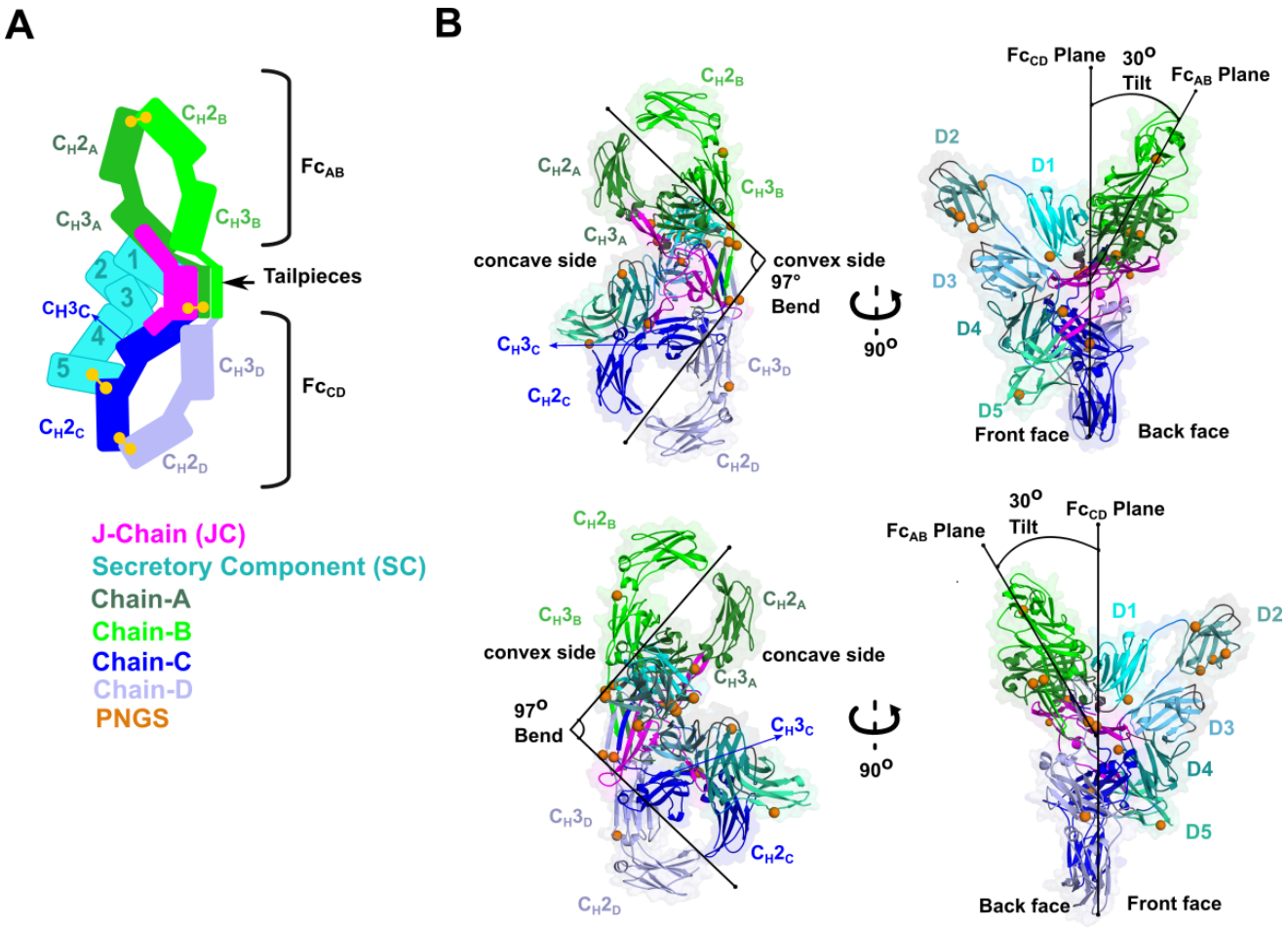

Figure 2. SIgA structure. (A) Schematic representing ordered domains in the SIgA structure. The relative position of each SIgA component is approximated based on the structure. Chain IDs and corresponding C_H domains and Fcs are labeled along with SC domains (1-5). Each SIgA component is depicted in a unique color; J chain (JC), magenta; secretory component (SC), teal; heavy chain (HC) HC_A, dark green, HC_B light green, HC_C dark blue, and HC_D light blue, respectively. (B) Cartoon representation (with semi-transparent molecular surface) of the SIgA structure shown in four orientations and colored as in (A). C_H and SC domains (D1-D5) are labeled and PNGS are shown as orange spheres. 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.

The structure of the dimer interface

The SIgA structure reveals an interface between the two IgA monomers mediated by the JC and the four HC Tps (Tp_A, Tp_B, Tp_C, Tp_D), which are integrated into a β-sandwich-like fold that separates the base of Fc_{AB} and Fc_{CD} by ~ 45Å. The interface is stabilized by three intra-JC disulfides and two inter-chain disulfides between the JC and the penultimate cysteines of the Tp_A and Tp_C, as well as contacts between two JC beta hairpins "wings" and two C_H3 domains. The central β-sandwich-like fold is formed by the first 68 residues, or "core", of the JC (JC_{core}) and Tp_A, Tp_B, Tp_C, and Tp_D, which each adopt a β-strand that extends away from the base of the corresponding Fc. The first three JC β-strands and Tp_C and Tp_D form one side of the β -sandwich, oriented toward the front face of SIgA whereas the subsequent two JC sheets and Tp_A and Tp_B form the other side of the β -sandwich, oriented toward the back face of SIgA. Tp_A and Tp_C disulfide bond to the JC and hydrogen bond to Tp_B and Tp_D (Figure 3A, B). Tp_A and Tp_C are pseudosymmetric in conformation, using identical residues and similar geometry to interface with unique JC core residues. The C-terminal four residues of Tp_A are buried within the JC. Tp_A Cys466 forms a disulfide bond with JC Cys68 and ultimate Tyr467 contacts JC residues Val7 and Lys70 and C_H3_C Lys441 and His345. Like Tp_A, the Tp_C penultimate Cys466 forms a disulfide with JC, in this case JC Cys14, and Tp_C Tyr467 contacts JC Arg106 as well as SC Leu101 (Figure 3C). In contrast, Tp_B and Tp_D do not form contacts with the JC but instead contact Tp_A and Tp_C using their β-sheet-forming regions (Figure 3A, C). We fail to find evidence that Tp_B or Tp_D disulfide bond with any other structure components. While the three C-terminal residues of Tp_D are observed interfacing with SC D1, the seven C-terminal residues of Tp_B are disordered in the structure (Figure 3C).

The second half of the JC sequence folds into "wings" (JC_W), each forming contacts with C_H3 that are distinct from the Tp-JC_{Core} assembly (Figure 3D). While both wings form antiparallel β-sheets and bind the same site on C_H3 domains, they are otherwise unique in terms of sequence and interfacing residues, conferring pseudosymmetry on complex. Wing 1 (JC_{W1}) is enclosed on the N- and C-terminal ends by the Cys71-Cys91 disulfide bond and is bound to the core through the Cys12-Cys101 disulfide; JC_{W1} interacts with C_H3_C (Fc_{CD}) on the front face of the molecule primary through six residues, which are mostly hydrophobic. Also unique to JC_{W1}, acidic residues Asp79 and Glu80 at the tip of the wing, which appear to form electrostatic interactions with SC-D5 Arg443. Wing 2 (JC_{W2}) is constrained within the intra-JC Cys109-Cys134 disulfide bond and forms extensive contacts with the C_H3_A (Fc_{AB}) on the back face of SIgA. JC_{W2} is five residues longer than JC_{W1} and shares a larger, more chemically diverse interface with C_H3_A. The interface is less hydrophobic than JC_{W1}-C_H3, with interaction strength likely dominated by electrostatic interactions between JC Arg117, JC His119, and several negatively charged patches on C_H3_A. JC_{W2} Lys123 also caps a helix formed from C_H3_A residues 248-253. JC_{W2} is followed by a 3-residue loop, containing conserved residues Tyr-Pro-Asp, which interacts with the SC as described below (Figure 3D).

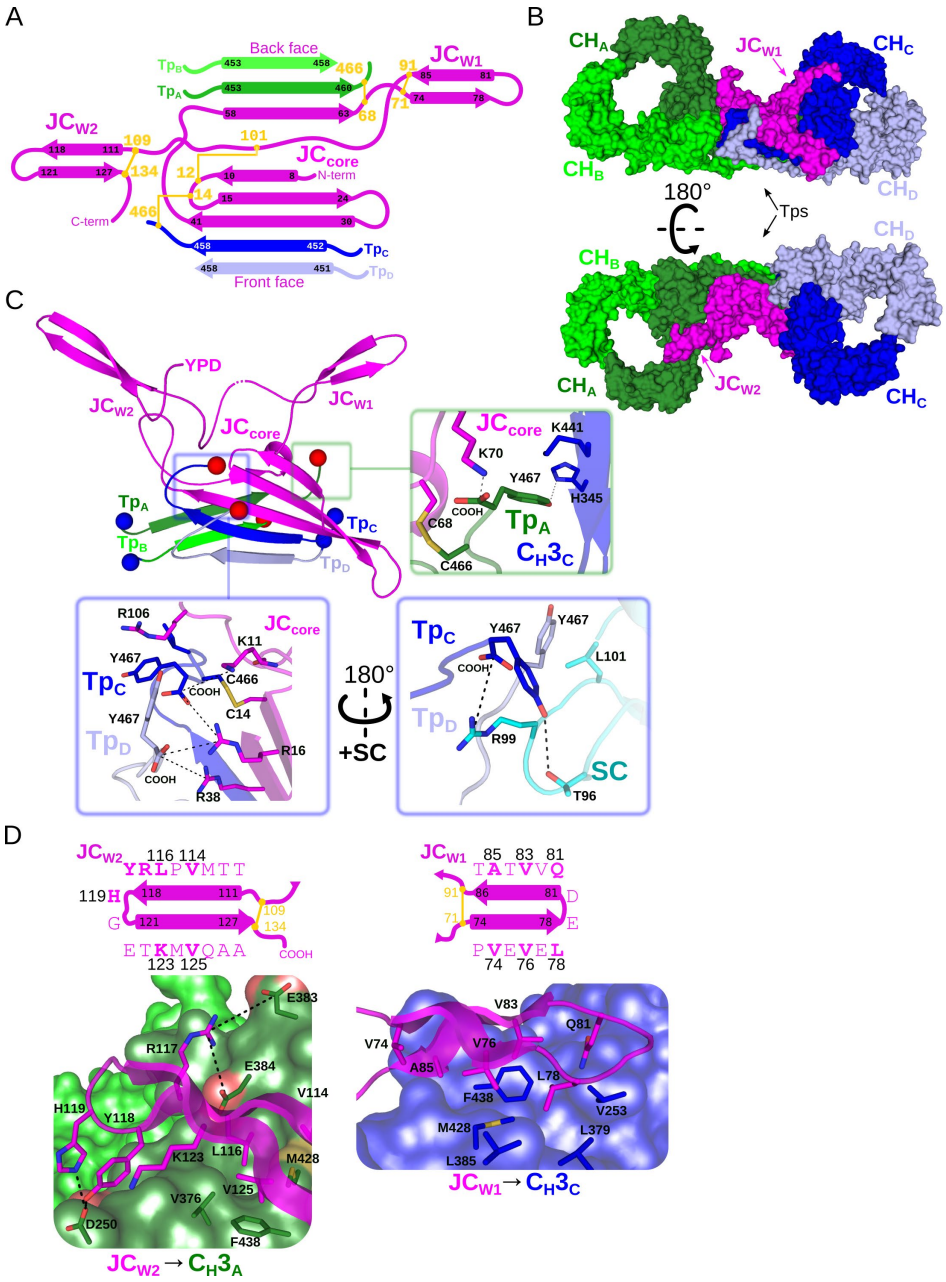

Figure 3. Fc dimer interface and JC structure. (A) Schematic depicting the topology that arises from JC and its interactions with four HC Tps and colored as in Figure 2. JC regions (JC_{W1}, JC_{W2}, JC_{core}, and N- and C termini) are labeled along with Tp_{A-D} and their location relative to the front face or back face of SIgA; disulfide bonds are indicated in yellow and the residue boundaries of each β-strand are labeled. (B) SIgA molecular surface representation (SC removed) colored as in Figure 2 and indicating the location of JC_{W1}, JC_{W2} and Tps relative to each HC_{A-D} in two SIgA orientations. (C) Cartoon representation of JC and Tps complex (same region as shown in panel A) with the N- and C termini of each Tp shown as blue and red spheres, respectively. Regions surrounding the three C-terminal residues of Tp are boxed, Tp_A (green box), and Tp_C and Tp_D (blue boxes), and enlarged. Enlargements show Tp carboxy termini (COOH) and side chain sticks involved in interactions with adjacent Tp, JC, C_H3 and SC. The three C-terminal residues of Tp_B are disordered and not shown. (D) Topology diagram and sequence (top) and structure (bottom) detailing interactions between JC_{W2} and C_H3_A (left) and JC_{W1} C_H3_C (right) and colored as in A-C. JC residues interacting with C_H3 are indicated in bold in the topology diagram (top) and shown as sticks on a cartoon representation (bottom); C_H3 domains are shown as molecular surface representations with interfacing residues shown as sticks and negatively charged atoms shown as a red surface.

Secretory Component

When unliganded, SC domains adopt a compact arrangement, in which a subset of residues in D1 CDR-like loops interface with residues in D4 and D5 to form a closed conformation. In SIgA, these domains are repositioned relative to each other, forming an elongated conformation that contacts Fc_{AB}, Fc_{CD} and JC asymmetrically on the front face of SIgA (Figure 4A).

D1 is reportedly necessary for pIgR and SC binding to JC-containing antibodies (Brandtzaeg, 2013b) and consistent with this requirement, the SIgA structure reveals SC-D1 contacting the JC and all four HC. D1 buries 1344.8 A^2 of surface area on SIgA and bridges interactions between SC, C_H3_B, C_H3_A, Tp_C, Tp_D and JC using residues in or adjacent to CDR loops (Figure 4). The D1-FC_{AB} interface is stabilized by numerous hydrogen bonding and hydrophobic interactions; conserved interfaces involve D1 Arg34, which binds C_H3_B Leu356 and C_H3_B Glu358, D1 Thr48, which binds C_H3_B Leu356 and C_H3_B Asn357, and D1 Tyr55, which binds C_H3_B Leu359 (Figure 4B). Additionally, D1 CDR2 Gly54 mediates the primary contact with C_H3_A, forming a main chain interaction with C_H3_A Ala403. The D1 FC_{CD} interface is stabilized largely by residues in D1 CDR1, CDR3 and Tp_C and Tp_D. Specifically, D1 Asn98 and D1 99Arg (CDR3) are positioned to bind Tp_D Gly464, and Tp_C Tyr467, respectively, while also contacting neighboring residues; conserved D1 Leu101 contacts ultimate Tyr467 in both Tp_C and Tp_D (Figures 3C,4B). These interactions signify a second type of conserved interaction for ultimate Tyrosine residues in the Tps, which in Tp_C and Tp_D bind SC and in Tp_A and Tp_B bind JC. The SC D1-JC interface is dominated by residues in D1 CDR1 and the three C-terminal residues of JC, which follow JC_{W2}. Conserved stabilizing contacts include D1 Arg31 and D1 His32, which bind the JC C-terminal residue, Asp137 (Figure 4B). The D1-JC interaction is further stabilized by other hydrogen bonds involving D1 Thr33 and D1 Ser28, which bind JC Asp132 and JC Asp137 respectively. An additional conserved contact, between D1 Asn30 and JC 106Arg, appears to link D1 to the loop connecting JC_{W1} and JC_{W2}.Notably, a subset of D1 residues mediating the C_H3_B-D1-JC interface, Arg31, His32, Arg34, Thr48 and Tyr 55 residues bind D4 or D5 when SC is unliganded (Stadtmueller et al., 2016a).

In contrast to SC D1, D2 does not contact dIgA or JC; the D1- D2 linker is extended, positioning D2 at the outer edge of the front face where it shares a minimal, ~250Å^2, interface with D3. The map surrounding D2 is poorly ordered suggesting that its position is flexible to the extent allowed by the D1-D2 and D2-D3 linkers. Together with D2, D3-D4-D5 adopt a near-linear arrangement that is bent approximately 48 degrees relative to D1 and the D1-2 linker, positioning D4-D5 to contact the FC_{CD} front face and the JC_{W1} (Figure 4A, B). This dIgA-bound conformation is stabilized by a 324 Å^2 hydrophobic interface between D1 and D3, involving conserved D1 residues Pro3, Phe5, Pro7, Tyr23 and Pro25 and D3 residues Leu226, Tyr228 and Phe326 (Figure 4B). The interface appears to function as a keystone in a D1-D3-D4-D5 bridge connecting Fc_{AB} and Fc_{CD} and contrasts with SC's unliganded structure, in which D1 and D3 do not share an interface and D3 Leu226 and D3 Tyr228 stabilize interactions with D4. Although D3 does not directly contact the HC or JC, D3-D4 linker residues Arg337 and D4 Lys341 are positioned to contact JC_{W1} residues Glu75 and Glu77 respectively, which along with a 646Å^2 interface shared with D4, appear to stabilize its position at the center of the SIgA front face above the JC core (Figure 4). The D3-D4 contacts in SIgA are distinct from those in the closed SC (not shown).

Consistent with an established role in dIgA binding (Stadtmueller et al., 2016a), our structure reveals SC D5 contacting the front face of Fc_{CD} (C_H2_C) and JC_{W1} (Figure 4B). The position of D5 is stabilized by a 692Å^2 interface with D4, which is overlapping but distinct from the D4-D5 interface in unliganded SC; the D4-D5 linker also appears to contact contacts JC_{W1} residues Glu79 and Asp80. D5 contacts with C_H2_C are mediated by CDR1 residues and include Cys470, which disulfide bonds to C_H2_C Cys306 and hydrogen bonds with neighboring residues C_H2_C Ala307 and C_H2_C Ser255. Consistent with published data demonstrating that D5 CDR1 mediates non-covalent binding to Fc (Stadtmueller et al., 2016a), D5 Lys471 contacts C_H2_C Asp256 (and JC Asp80) and D5 Tyr473 contacts C_H2_C Cys306 and C_H2_C Glu308.

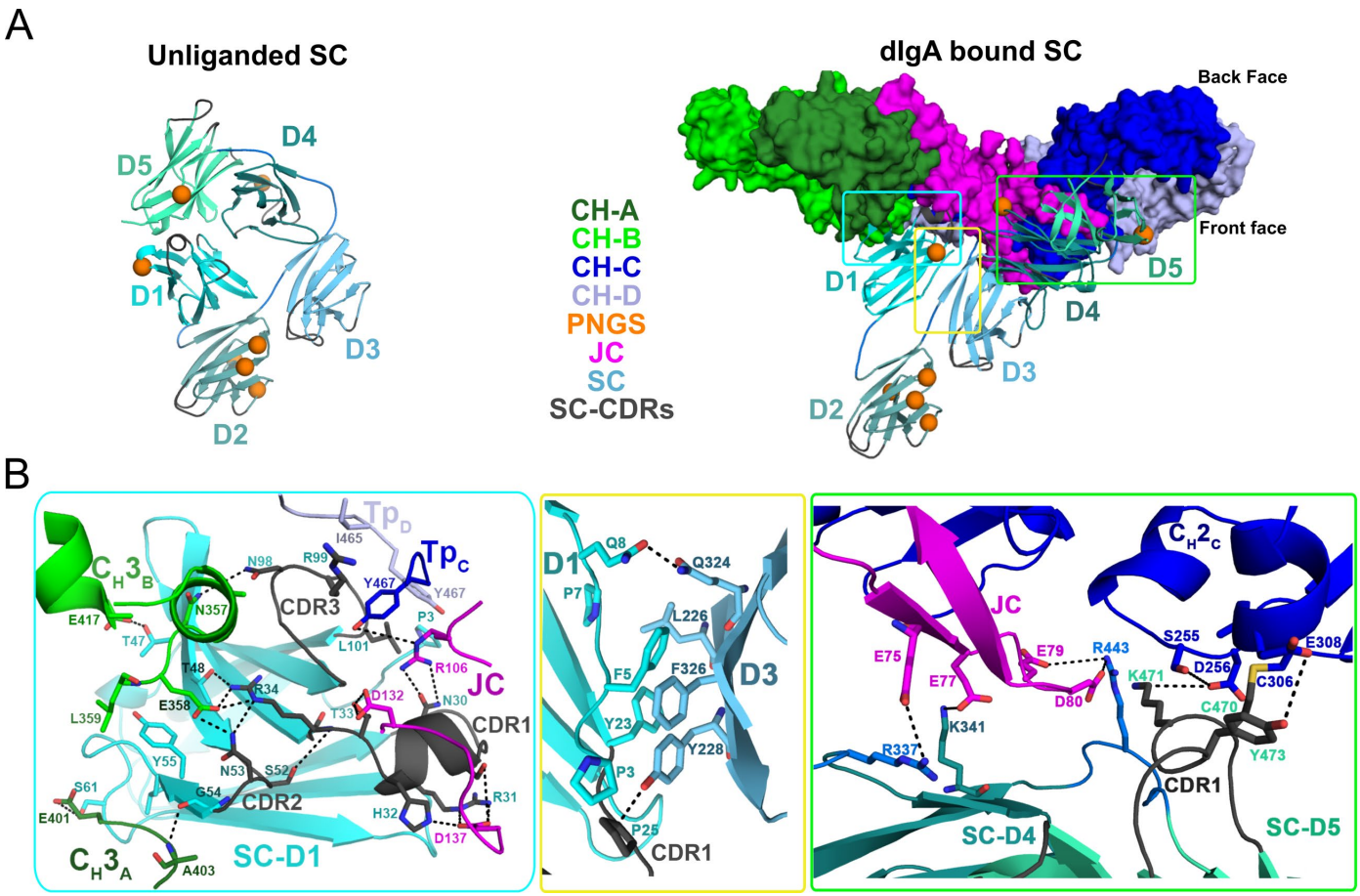

Figure 4. SC Structure. (A) Cartoon representations of the unliganded human SC crystal structure (PDB code 5D4K; left) and the dIgA-bound mouse SC (right). The dIgA is shown as a molecular surface representation; SC domains D1-D5 are show in color gradient from cyan (D1) to pale green (D5), with CDR loops colored grey; PNGS are shown as orange spheres. Regions where SC interfaces with dIgA and other SC domains are boxed. (B) Boxed, enlarged views of stabilizing interfaces between SC Fc, JC and other SC domains; colored as in panel A. (Left panel) SC D1 interface with C_H3_A, C_H3_B, Tp_C, Tp_D. (Center panel) SC D1-D3 interface. (Right panel) SC D4-D5 interface with JC and C_H2_C. In all panels conserved, interfacing residues are shown as sticks and hydrogen bonds and salt bridges are show as black dashes.

Structure of dIgA

The SIgA 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 SIgA we determined a 3.3Å (average) resolution CryoEM structure of dIgA, which contained the same components as SIgA, except for SC (Figure S3). The dIgA structure revealed two IgA monomers in an arrangement similar to those in SIgA; Fabs were disordered. The centroid axes of Fc_{AB} and Fc_{CD} were 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 and/or interactions between the Tps. 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 HC and JC were largely superimposable with equivalent chains in SIgA; however, differences were apparent in regions bound by SC in SIgA (Figure 5B). In particular, Tp_D, which forms part of the SC D1-JC interface in SIgA, adopts a different conformation, and the three C-terminal residues are disordered. The three C-terminal residues of JC (Tyr-Pro-Asp), which bind SC D1 CDR1, are also disordered (Figure 5C).

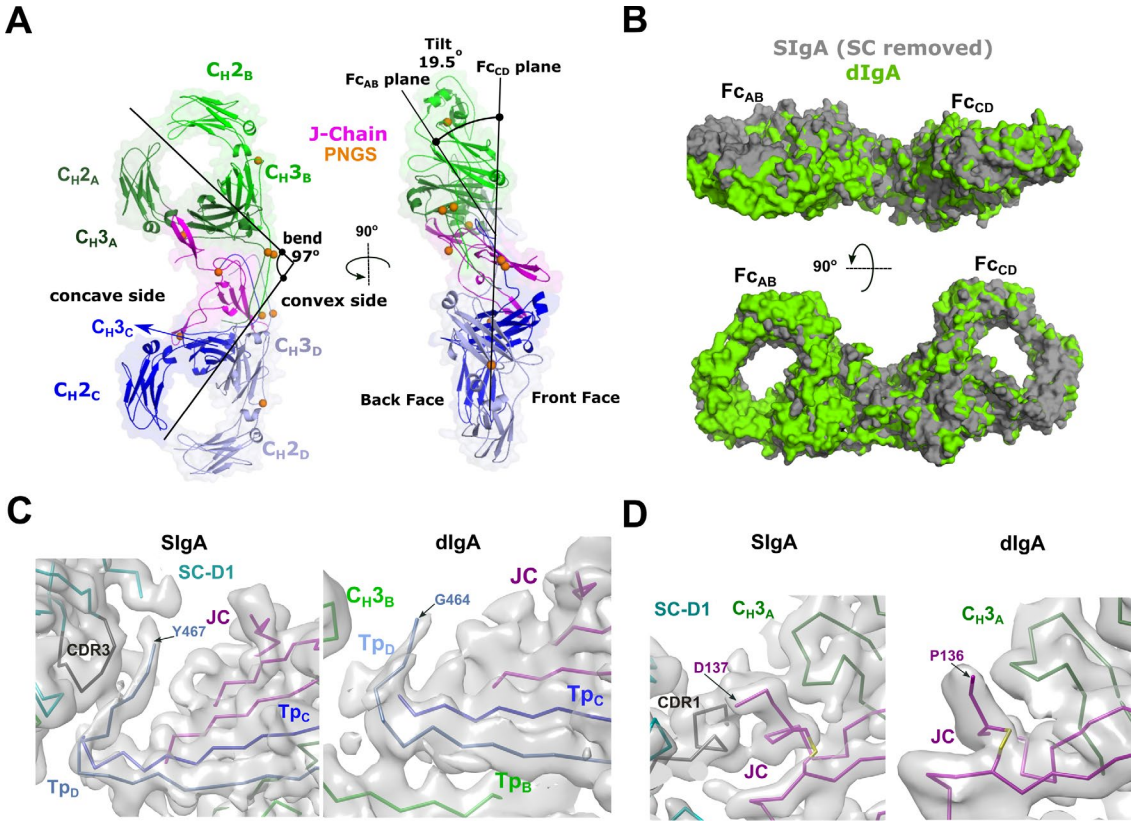

Figure 5. dIgA Structure. (B) Cartoon representation (with semi-transparent molecular surface) of the unliganded dIgA structure shown in two orientations and colored as in Figure 2. The bend and tilt between the two Fcs is indicated with a line and the angle measured in the structure; the concave and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

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 SIgA Fabs, we modeled potential Fab location onto SIgA structures using a computational conformational-search approach, which approximated the position of each Fab by mapping a vector (beginning at N-terminus of C_H2 and ending at the center of mass of Fab CDRs) onto a Fibonacci spherical lattice (FSL)(Marques, 2013) (Figure S4). We reasoned that this approach would broadly survey possible Fab positions without the constraints of the diverse Fc-Fab linkers and HC-LC contacts 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 SIgA 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_H2_B-C_H3_B and C_H2_D-C_H3_D) 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 the center of mass of the JC and Fcs along the Fc_{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_{AB}) might influence the positions of CDRs and associated interactions with antigen.

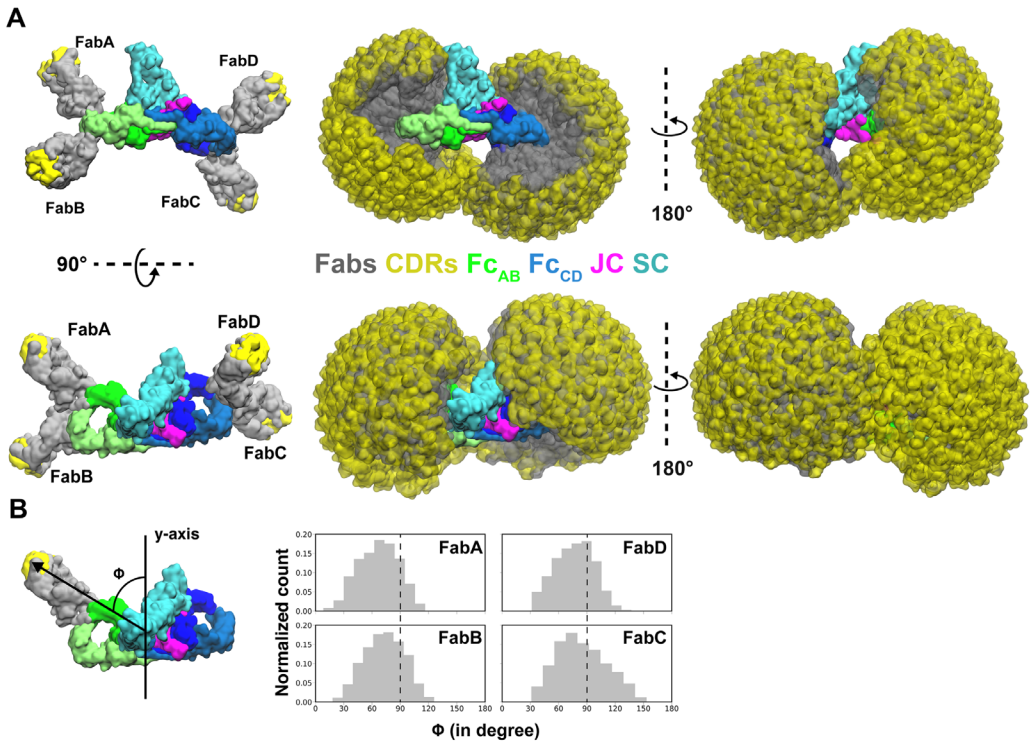

Figure 6. Modeling of Fabs and CDRs on SIgA. (A) SIgA structure shown with four Fabs modeled in a single position, as well as all possible positions, and shown in multiple orientations. Complex components are colored according to the key (center). (B) The number of positions sampled by Fab CDRs were quantified by measuring the angle (ϕ) between the Fab vector and the y-axis, a vector parallel to the Fc_{AB} plane and passing through the center of mass of Fc and JC. The frequency (normalized count) of each angle is shown as a histogram.

DISCUSSION

Our work describes the structure of mouse SIgA and its precursor, dIgA, which lacks the SC. Together, these structures provide a first glimpse at one of the most abundant mammalian antibodies. Based on a high degree of sequence conservation among IgA HC, JC and SC across mammalian species (Figure S1), we anticipate that our structures will allow interpretation of biological data resulting from various immunological model systems (e.g. mouse and rabbit) while also providing a model for human SIgA, a potential therapeutic target. Indeed, our SIgA structure appears similar to the structure of a human SIgA1 core, published in the same timeframe as this work (Kumar et al., 2020), although further analysis will be needed to determine similarities and differences between the structures. SIgA is a long-studied molecule. Among other discoveries, prior work determined the structure of a monomeric IgA Fc lacking the Tps, the structure of unliganded SC, and also provided insight on the overall shape and identified roles for many conserved residues involved in complex formation, transcytosis and function in the mucosa (Bonner et al., 2009; Brandtzaeg, 2013a; Herr et al., 2003; Stadtmueller et al., 2016a; Woof and Russell, 2011). Our structures provide a framework for interpreting many of these results, some of which we discuss here and summarize in Figure 7.

pIgR binding to dIgA. When unliganded, SC domains adopt a compact conformation, in which conserved residues in D1 form an interface with D4 and/or D5 (Stadtmueller et al., 2016a). The SIgA structure reveals a marked rearrangement of SC domains that positions residues in the D1 CDRs in direct contact with C_H3_A, C_H3_B, Tp_C, Tp_D and JC and positions D5 CDRs in contact with C_H3_C. This conformational change is consistent with published double electron-electron resonance (DEER) distance measurements showing a 70Å separation between nitroxide spin labels attached to SC D1- residue 67 and D5 residue 491 in human SIgA (Stadtmueller et al., 2016a); the distance between Cα atoms of equivalent residues, D1 Ile67 and D5 His493, in the murine SIgA structure is 61Å. 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 consistent 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 C_H3_B and JC. It is yet unclear which residues mediate the recognition binding event; however, one possibility is that SC D1-CDR3 residues, which are exposed in unliganded SC, form the initial contacts with the tailpieces (Tp_C and Tp_D) and C_H3_B, which are exposed in dIgA (Figure 5C). Our structures suggest that during the conformational change, D2 moves away from D1 while D3 forms hydrophobic interactions with D1 in a process that positions the D4-D5 arm near the FC_{CD} and D5 in contact with C_H2_C (Figure 7).

The pIgR (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 SIgA structure. However, it also appears likely that the JC indirectly supports SC binding by inducing the bend that positions C_H3_B (which binds D1) and C_H2_C (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 HCs 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 SIgA structure is also

consistent with a model in which SC from birds, reptiles and amphibians, which are all lacking the D2 domain (Stadtmueller et al., 2016b), would bind analogous to D1-D3-D4-D5 in our SIgA structure.

The functional significance of SC

Whereas 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 IgA. 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 HCs and the JC where it may protect areas especially vulnerable to proteolysis; however, roughly 16 percent 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 $Å^2$) 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 dominant role in mediating interactions with host and/or microbial factors (Figure 7).

The role of JC

The JC structure has long remained elusive, and understandably so given that we find its fold integrated with four additional chains (HC_{A-D}); without HC contacts, this particular conformation is not likely to be stable. The JC's unique structure provides an example of how a single polypeptide can use unique sequence to assemble four identical chains into a pseudosymmetric structure. As noted, the JC appears likely to play a dominant role inducing the bent relationship between the two Fcs, which we speculate may influence SIgA function and facilitate pIgR binding. However, it is important to note that JC also facilitates polymerization of IgM into pentamers, and to a lesser extent IgA into higher order polymers. In the context of higher order polymers, such as pentametric IgM, a similar conformation may facilitate polymerization in some manner, for example by stabilizing interactions between domains in adjacent Fcs (Muller et al., 2013).

SIgA functions

Our data provides 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 FcαR1, although FcαR1 is absent from the mouse genome. IgA FcRs typically bind a canonical site located near C_H2-C_H3 elbow (Herr et al., 2003; Kazeeva and Shevelev, 2009; Ramsland et al., 2007). In the SIgA and dIgA structures, two of the four accessible sites are occluded by JC wing interactions with C_H3_A and C_H3_C. This leaves two sites, on C_H3_B and

C_H3_D, 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, 7). This arrangement would provide an unobstructed approach for FcRs to bind SIgA, which in the context of host cell receptors would promote favorable outcomes. On the other hand, it could leave SIgA 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 SIgA interactions engages antigen. In the context of endogenous mammalian antibodies, unique properties of different C_H2 -C_H1 linkers and stabilizing interactions between the HCs and LCs are likely to further constrain 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).

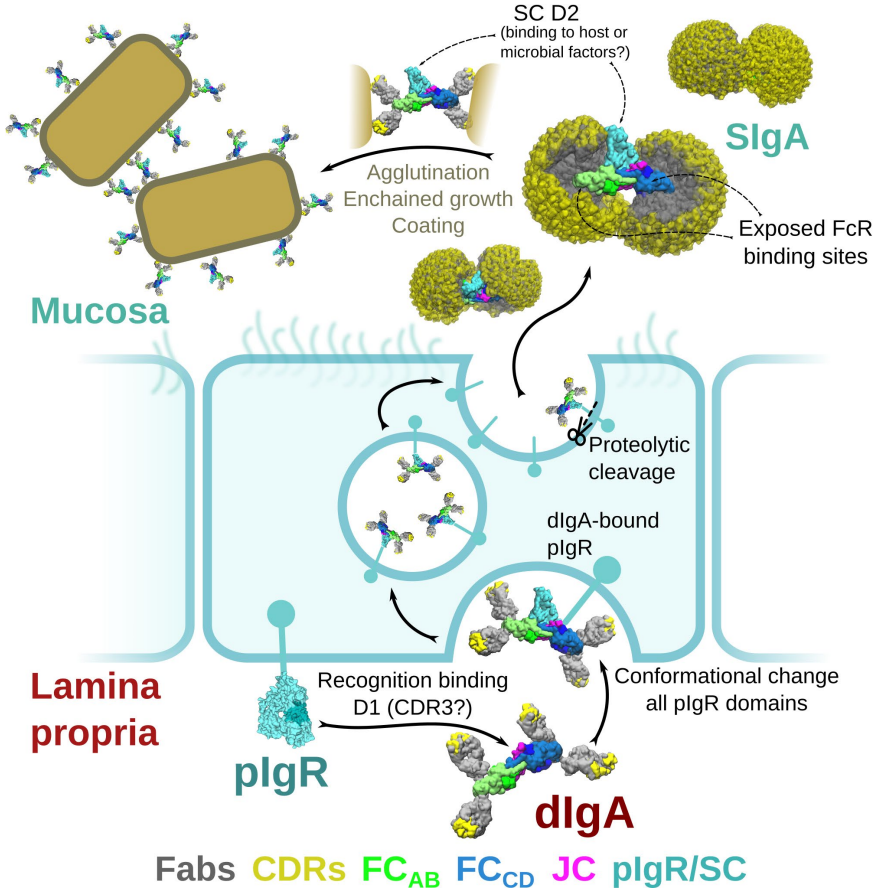

Figure 7. Model for the formation, transport and function of SIgA. (A) Schematic summary depicting the unliganded SC structure (pdb code 5D4K) as pIgR bound to basolateral surface of an epithelial cell in its closed conformation and recognizing bent dIgA from the lamina propria. The pIgR binding to dIgA triggers a conformation change that repositions its domains to facilitate numerous stabilizing contacts with dIgA. The dIgA-pIgR complex transcytoses to the apical membrane where the pIgR is proteolytically cleaved, releasing SIgA into the mucosa. In the mucosa, SIgA 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 SIgA where it may bind host and bacterial factors. Upon encountering antigen, SIgA Fabs bind, promoting antigen coating, agglutination or enchained growth.

MATERIALS AND METHODS

Construct design and Protein Expression

Genes encoding the *mus musculus* IgA HC constant region (Uniprot P01878) and the lambda LC constant region (Uniprot A0A0G2JE99) were fused with proprietary HC and LC variable region sequences to create complete HC and LC sequences. The TPA signal sequence (residues MDAMKRGLCCVLLLCGAVFVSPAGA) was encoded at the start of the HC sequence and the mouse IgKappa signal sequence (residues METDTLLLWVLLLWVPGSTG) was encoded at the start of the LC sequence. These sequences, along with *mus musculus* JC (Uniprot P01592; native signal peptide) and *mus musculus* pIgR ectodomain residues 1-567 (Uniprot P01592; native signal peptide), were codon optimized, synthesized (Integrated DNA Technologies, Inc.) and cloned into mammalian expression vector pD2610v1 (Atum). Resulting expression constructs were transiently co-transfected into HEK Expi-293-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; SIgA 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 SIgA complexes were purified using CaptureSelect™ LC-lambda (Mouse) Affinity Matrix (Thermo Fisher) and Superose 6 (GE Healthcare Life Sciences) size exclusion chromotography (SEC). SEC fractions corresponding to the expected size of SIgA (containing two IgA) and dIgA were utilized for CryoEM.

CryoEM Grid Preparation and Data Collection

Quantifoil R2/2 300 mesh grids were glow discharged using a Pelco easiGlow system for 1 minute at 20 mA current. A 3µL drop of the SIgA or dIgA sample (concentration 1.5 mg/ml) 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, 0s 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) operation at 300 kV, equipped with BioQuantum Energy Filter (20 eV slit width, Gatan) and a K3 direct electron detector (Gatan). SIgA 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. 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, 0.05s per frame, and total dose of ~60 electrons/Å^2.

CryoEM Data Processing

For the SIgA, 1512 movies were initially collected. Images 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 (Scheres, 2012; Zivanov et al., 2018) (minimum and maximum diameters of 60 and 180Å respectively). Approximately 1.4 million particles were extracted, binned 4 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 (310626) were merged into a single file and re-extracted with 2X binning and further underwent 2D classification. Remaining 184737 particles were used in a 3D refinement (using an initial model generated in a parallel cryoSparc processing pipeline (Punjani et al., 2017)). 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 9 holes in a 3 x 3 pattern. All of the processing steps were done in cryoSparc v.2. 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 5 templates generated from a 4.5Å 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 6 ab-initio classes and then processed with heterogeneous refinement with 3 classes. The best class was combined with the remaining 2 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 3 classes was performed. Resulting set of 288,823 particles were refined (Legacy). The particle stack was then split into 9 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, HC and SC, and reference pdbs 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 SIgA 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 the JC and Tp residues. The resulting unrefined SIgA model contained four HC, one SC, and one JC (14 folded domains).

To counter the challenges of hand building five chains, adopting a previously undetermined fold, into a moderate resolution map, we implemented a strategy to rebuilt and refine the structure using a combination of

the Rosetta CryoEM refinement package and Rosetta comparative modeling (RosettaCM) (Wang et al., 2016). Briefly, 2913 Rosetta-modified structures were produced using the Rosetta CryoEM refinement package in 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., 2018)) and density 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 SIgA structure and the experimental map and were further rebuilt manually using Coot. Subsequent iterations of this process were used to rebuild portions of the CH2 and CH3 domains, during which the Phenix phenix.map_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 dIgA structure was determined using a similar approach as used for the SIgA structure. Breifly, homology models of the mouse IgA Fc were docked into the dIgA map along with the JC from the SIgA 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 was refined using the Rosetta CryoEM refinement package (Wang et al., 2016). The final SIgA and dIgA structures and their fit to cryoEM maps were evaluated by hand and validated using Phenix EM Validation, Molprobity and EMRinger (Afonine et al., 2018a; Barad et al., 2015; Williams et al., 2018); results are summarized in Table S1. CryoEM maps and structure coordinate files will be deposited in the protein databank and the EM databank.

Structure Analysis

The sequence alignment between human and mouse counterparts of J-Chain (human: Uniprot P01591,mouse: Uniprot P01592), pIgR ectomain (human: Uniprot P01833, mouse: Uniprot O70570) and Heavychain- C_H2-C_H3 regions (human: Uniprot P01877, mouse: Uniprot Q99LA6) were carried using ClustalOmega (Sievers et al., 2011) and figures were made with EsPript 3 (Gouet et al., 1999). Percent Identity and similarity values were determined using NCBI-BLAST (Altschul et al., 1990).

Contacts between individual Tps and the JC and all other SIgA components were evaluated by inspecting all interfacing residues within approximately 7Å. The probability of conserved contacts was evaluated by hand inspection of the cryoEM map. A list of all possible interactions between SC and the components of dIgA were made by using the Protein Interaction Calculator (PIC) webserver (Tina et al., 2007). The list of interactions were visualized in Pymol Molecular Graphics System (Schrodinger LLC) , to validate the distances and additional unspecified interactions within 4A. PDBePISA server (https://www.ebi.ac.uk/pdbe/pisa/) (Krissinel and Henrick, 2007) was used for the calculation of interface surface area, between various elements of the structure. Pymol Molecular Graphics System (Schrodinger LLC) was used with and without SC to calculate the percentage of accessible surface area of dIgA. The potential N-linked glycosylation sites were determined using the NetNGlyc 1.0 Server (Blom et al., 2004). The centroid axes and centroid planes for Fc_{AB} and Fc_{CD} (residues 237-445) were individually determined using UCSF Chimera. The angles between the axes and 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 differences between SIgA and dIgA structures were calculated using PyMol. To visualize the difference between Fc_{AB} in SIgA and dIgA (SC removed) 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) and UCSF Chimera (Pettersen et al., 2004; Yang et al., 2012).

Computational search for potential Fab positions:

The pdbs files used in Fab modeling were generated using the template mode of SWISS-MODEL (Waterhouse et al., 2018) with reference pdb 4EOW chain A and chain B and C_H1-V_H1 and C_L1-V_L1 sequences, respectively. In order to evenly sample all the potential positions, each Fab was rotated such that the center of mass (C.O.M) of the CDR was arranged on a Fibonacci spherical lattice (FSL) of one thousand points (Marques, 2013). During the rotation, the N-terminal residue in the C_H2 domain (pivot) was set to be in the center of the FSL. Each of 1,000 generated orientations was further rotated along the axis spanning from the pivot to the CDR, with an interval of $\gamma=45^0$. In total, 8,000 structures were generated for each Fab. Any position in which two or more Fab amino acids had a clash was removed from the data set. Clashes were defined if more than 8 atoms of Fab (except the linker) were within 1 Å of any atom of the Fc, JC and SC domains of SIgA. A minimum of 8 atoms was chosen to define a clash because one amino acid has on average 8 heavy atoms. The linker region of the Fab was not considered for the clash calculation because we assumed that the linker is flexible. The number of positions sampled by Fab CDRs were quantified by measuring ϕ, defined to be the angle between the Fab vector and the y-axis, a vector parallel to the Fc_{AB} plane and passing through the center of mass of Fc and JC. Figure were made using VMD (Humphrey et al., 1996).

ACKNOWLEDGEMENTS

We thank members of the Stadtmueller Laboratory (UIUC) and Emma Slack (ETH Zurich) for insightful conversations and thank Kathryn Huey-Tubman (Caltech) for assistance developing IgA protein expression strategies. This study was funded by University of Illinois Urbana-Champaign start-up funding to BMS. 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.

AUTHOR CONTRIBUTIONS

BMS, SKB, BWP conceived the study; AM collected and processed cryoEM data; BMS, SKB, BWP solved and analyzed structures; NH and ET modeled Fab positions; BMS wrote the paper with contributions from all authors.

REFERENCES

Afonine, P.V., Klaholz, B.P., Moriarty, N.W., Poon, B.K., Sobolev, O.V., Terwilliger, T.C., Adams, P.D., and Urzhumtsev, A. (2018a). New tools for the analysis and validation of cryo-EM maps and atomic models. Acta Crystallogr D Struct Biol *74*, 814-840.

Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018b). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr D Struct Biol *74*, 531-544.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. J Mol Biol *215*, 403-410.

Barad, B.A., Echols, N., Wang, R.Y., Cheng, Y., DiMaio, F., Adams, P.D., and Fraser, J.S. (2015). EMRinger: side chain-directed model and map validation for 3D cryo-electron microscopy. Nature methods *12*, 943-946.

Blom, N., Sicheritz-Ponten, T., Gupta, R., Gammeltoft, S., and Brunak, S. (2004). Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence. Proteomics *4*, 1633-1649.

Bonner, A., Almogren, A., Furtado, P.B., Kerr, M.A., and Perkins, S.J. (2009). The nonplanar secretory IgA2 and near planar secretory IgA1 solution structures rationalize their different mucosal immune responses. J Biol Chem *284*, 5077-5087.

Brandtzaeg, P. (2013a). Secretory IgA: Designed for Anti-Microbial Defense. Frontiers in immunology *4*, 222.

Brandtzaeg, P. (2013b). Secretory IgA: Designed for Anti-Microbial Defense. Front Immunol *4*, 222.

Donaldson, G.P., Ladinsky, M.S., Yu, K.B., Sanders, J.G., Yoo, B.B., Chou, W.C., Conner, M.E., Earl, A.M., Knight, R., Bjorkman, P.J.*, et al.* (2018). Gut microbiota utilize immunoglobulin A for mucosal colonization. Science *360*, 795-800.

Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr *60*, 2126-2132.

Flajnik, M.F. (2010). All GOD's creatures got dedicated mucosal immunity. Nat Immunol *11*, 777-779.

Gouet, P., Courcelle, E., Stuart, D.I., and Metoz, F. (1999). ESPript: analysis of multiple sequence alignments in PostScript. Bioinformatics *15*, 305-308.

Herr, A.B., Ballister, E.R., and Bjorkman, P.J. (2003). Insights into IgA-mediated immune responses from the crystal structures of human FcalphaRI and its complex with IgA1-Fc. Nature *423*, 614-620.

Humphrey, W., Dalke, A., and Schulten, K. (1996). VMD: visual molecular dynamics. J Mol Graph *14*, 33-38, 27-38.

Kaetzel, C.S. (2005). The polymeric immunoglobulin receptor: bridging innate and adaptive immune responses at mucosal surfaces. Immunol Rev *206*, 83-99.

Kaetzel, C.S. (2014). Coevolution of Mucosal Immunoglobulins and the Polymeric Immunoglobulin Receptor: Evidence That the Commensal Microbiota Provided the Driving Force. ISRN Immunology *2014*, 20.

Kazeeva, T.N., and Shevelev, A.B. (2009). IgA-specific proteins of pathogenic bacteria. Biochemistry (Mosc) *74*, 12-21.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J Mol Biol *372*, 774-797.

Kumar, N., Arthur, C.P., Ciferri, C., and Matsumoto, M.L. (2020). Structure of the secretory immunoglobulin A core. Science.

Marques, R., C. Bouville, M. Ribardière, L. P. Santos, and K. Bouatouch (2013). Spherical Fibonacci Point Sets for Illumination Integrals. Computer Graphics Forum *32.8*, 134-143.

Moor, K., Diard, M., Sellin, M.E., Felmy, B., Wotzka, S.Y., Toska, A., Bakkeren, E., Arnoldini, M., Bansept, F., Co, A.D.*, et al.* (2017). High-avidity IgA protects the intestine by enchaining growing bacteria. Nature *544*, 498-502.

Muller, R., Grawert, M.A., Kern, T., Madl, T., Peschek, J., Sattler, M., Groll, M., and Buchner, J. (2013). High-resolution structures of the IgM Fc domains reveal principles of its hexamer formation. Proc Natl Acad Sci U S A *110*, 10183-10188.

Pabst, O., and Slack, E. (2020). IgA and the intestinal microbiota: the importance of being specific. Mucosal Immunol *13*, 12-21.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem *25*, 1605-1612.

Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature methods *14*, 290-296.

Ramsland, P.A., Willoughby, N., Trist, H.M., Farrugia, W., Hogarth, P.M., Fraser, J.D., and Wines, B.D. (2007). Structural basis for evasion of IgA immunity by Staphylococcus aureus revealed in the complex of SSL7 with Fc of human IgA1. Proc Natl Acad Sci U S A *104*, 15051-15056.

Rogier, E.W., Frantz, A.L., Bruno, M.E., Wedlund, L., Cohen, D.A., Stromberg, A.J., and Kaetzel, C.S. (2014). Secretory antibodies in breast milk promote long-term intestinal homeostasis by regulating the gut microbiota and host gene expression. Proc Natl Acad Sci U S A *111*, 3074-3079.

Rohou, A., and Grigorieff, N. (2015). CTFFIND4: Fast and accurate defocus estimation from electron micrographs. J Struct Biol *192*, 216-221.

Scheres, S.H. (2012). RELION: implementation of a Bayesian approach to cryo-EM structure determination. J Struct Biol *180*, 519-530.

Schorb, M., Haberbosch, I., Hagen, W.J.H., Schwab, Y., and Mastronarde, D.N. (2019). Software tools for automated transmission electron microscopy. Nature methods *16*, 471-477.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J.*, et al.* (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology *7*, 539.

Stadtmueller, B.M., Huey-Tubman, K.E., Lopez, C.J., Yang, Z., Hubbell, W.L., and Bjorkman, P.J. (2016a). The structure and dynamics of secretory component and its interactions with polymeric immunoglobulins. Elife *5*.

Stadtmueller, B.M., Yang, Z., Huey-Tubman, K.E., Roberts-Mataric, H., Hubbell, W.L., and Bjorkman, P.J. (2016b). Biophysical and Biochemical Characterization of Avian Secretory Component Provides Structural Insights into the Evolution of the Polymeric Ig Receptor. J Immunol *197*, 1408-1414.

Tina, K.G., Bhadra, R., and Srinivasan, N. (2007). PIC: Protein Interactions Calculator. Nucleic acids research *35*, W473-476.

Wang, R.Y., Song, Y., Barad, B.A., Cheng, Y., Fraser, J.S., and DiMaio, F. (2016). Automated structure refinement of macromolecular assemblies from cryo-EM maps using Rosetta. Elife *5*.

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L.*, et al.* (2018). SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic acids research *46*, W296-W303.

Williams, C.J., Headd, J.J., Moriarty, N.W., Prisant, M.G., Videau, L.L., Deis, L.N., Verma, V., Keedy, D.A., Hintze, B.J., Chen, V.B.*, et al.* (2018). MolProbity: More and better reference data for improved all-atom structure validation. Protein Sci *27*, 293-315.

Woof, J.M., and Kerr, M.A. (2006). The function of immunoglobulin A in immunity. J Pathol *208*, 270-282.

Woof, J.M., and Russell, M.W. (2011). Structure and function relationships in IgA. Mucosal Immunol *4*, 590-597.

Yang, Z., Lasker, K., Schneidman-Duhovny, D., Webb, B., Huang, C.C., Pettersen, E.F., Goddard, T.D., Meng, E.C., Sali, A., and Ferrin, T.E. (2012). UCSF Chimera, MODELLER, and IMP: an integrated modeling system. J Struct Biol *179*, 269-278.

Zivanov, J., Nakane, T., Forsberg, B.O., Kimanius, D., Hagen, W.J., Lindahl, E., and Scheres, S.H. (2018). New tools for automated high-resolution cryo-EM structure determination in RELION-3. Elife *7*.

SUPPLEMENTAL FIGURES AND TABLES

Figure S1. Sequence alignment. Sequences corresponding to the mouse SIgA structure components aligned to homologous human sequences. Sequences for protein domains are separated by line and labeled.

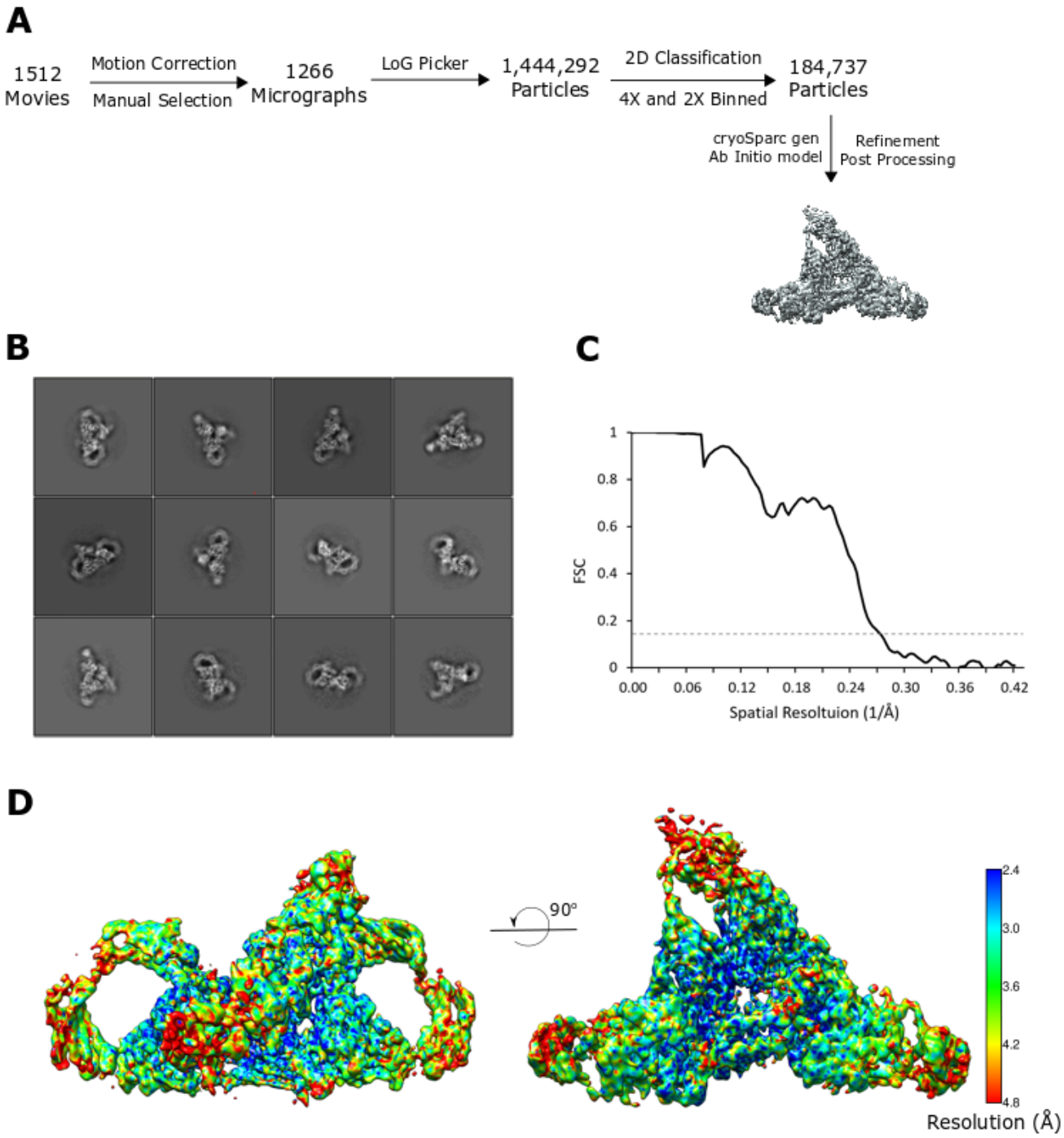

Figure S2. SIgA cryoEM data collection and processing. (A) Schematic summary of data reduction process; (B) representative 2D class averages; (C) FSC resolution estimate; (D) CryoEM map depicting local resolution between 2.4-4.8Å.

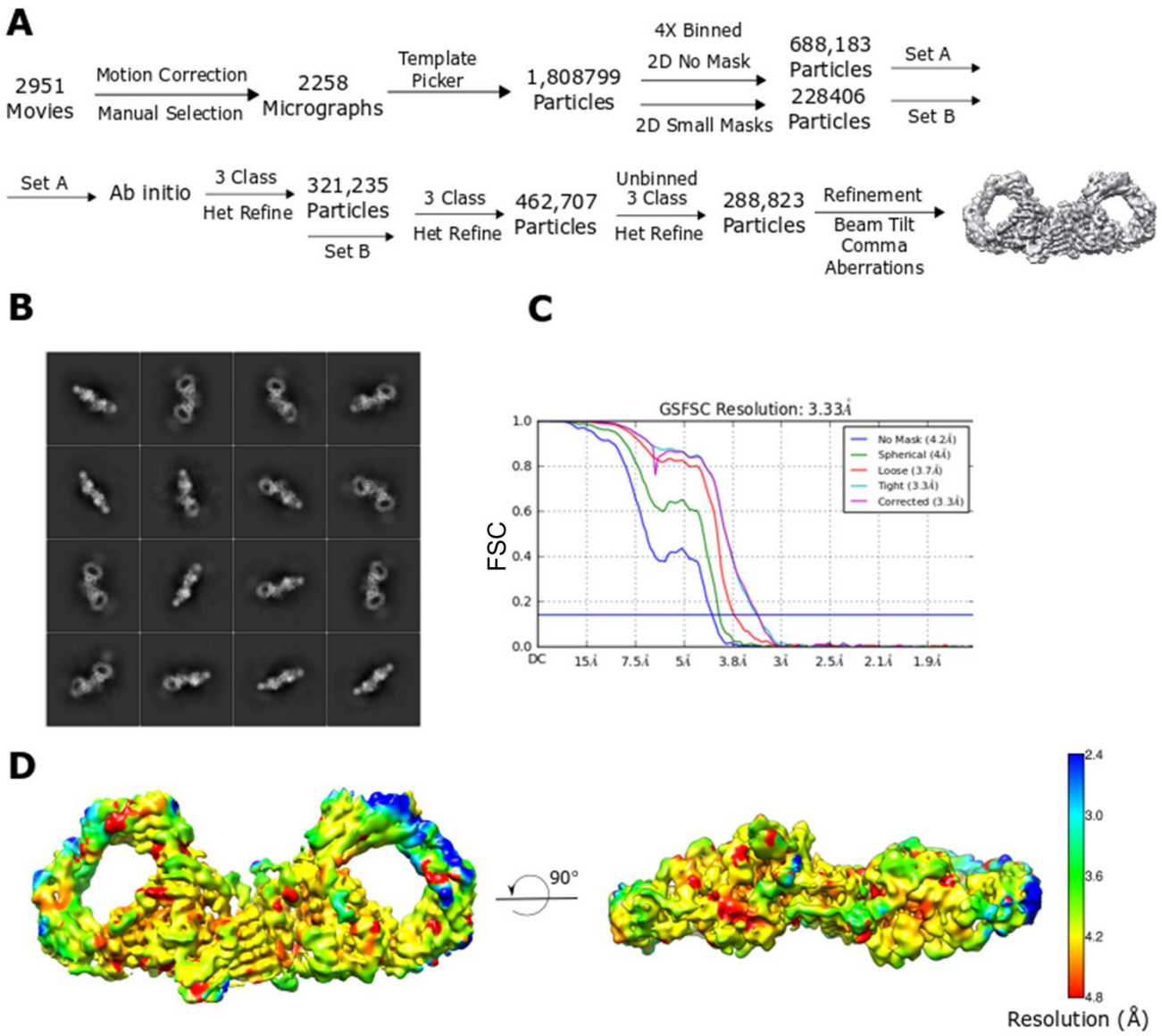

Figure S3. dIgA cryoEM data collection and processing strategy. (A) Schematic summary of data reduction process; (B) representative 2D class averages; (C) FSC resolution estimate; (D) CryoEM map depicting local resolution between 2.4-4.8Å.

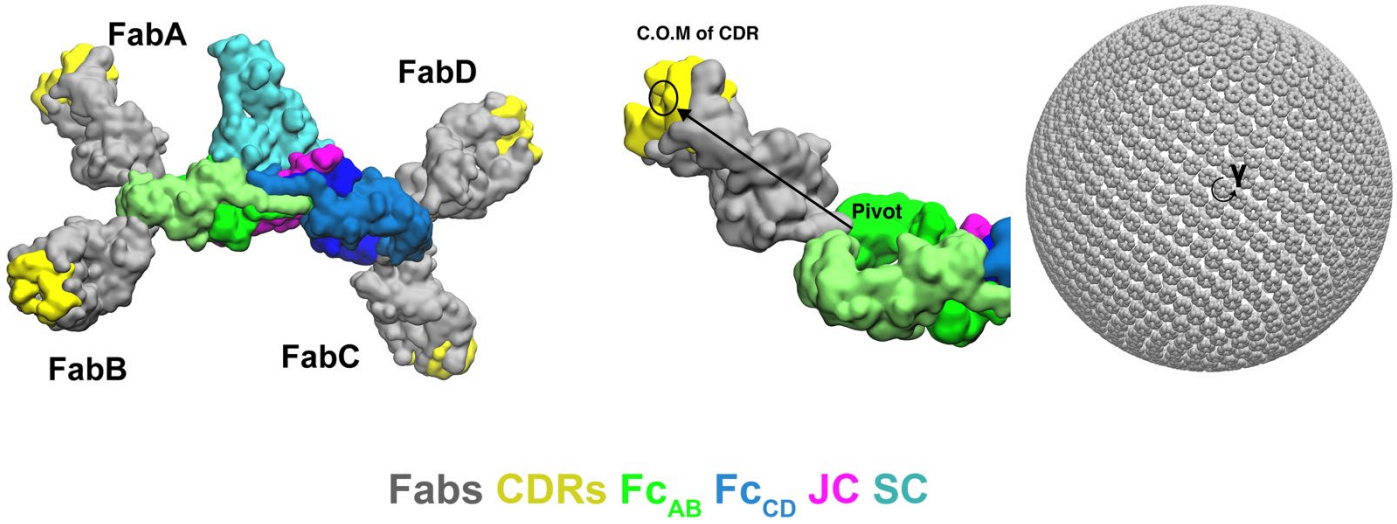

Figure S4. Modeling strategy for the computational search of SIgA Fab positions. (Left) SIgA structure shown with four modeled Fabs in one orientation. (Right) Schematic showing how Fab positions were modeled by mapping a vector, spanning from the N-terminal residue in the C_H2 domain (pivot) to the center of mass (C.O.M) of the CDRs (yellow), onto a Fibonacci spherical lattice (FSL) with one thousand points and γ angle intervals of 45^0.

Table S1. CryoEM data collection and refinement statistics.

Data Collection	**Secretory IgA**	**Dimeric IgA**
Magnification	130,000	105,000
Pixel size (Å)	0.652	0.836
Defocus Range (μ)	1.0-3.5	1.0-3.6
Defocus Mean (μ)	1.23	1.45
Frames per movie	40	40
Seconds per frame	0.03	0.05
Movies selected	1,266	2,258
Dosage (e/Å^2)	60.00	60.00
Initial particles	1,444,292	1,808,799
Final particles	184,737	288,823
Map sharpening B-factor	-85	n/a
Resolution at FSC=0.143 (Å)	3.70	3.30
Density and Fit		
Model/Data CC		
Mask	0.66	0.80
Box	0.75	0.89
Peaks	0.63	0.74
Volume	0.70	0.79
Model	**Secretory IgA**	**Dimeric IgA**
Refinement software	Rosetta cryo-EM	Rosetta cryo-EM
Validation software	Phenix	Phenix
Chains	6	5
Total atoms	24044	15811
Non-hydrogen atoms	12178	7965
Residues	1584	1046
Bond RMSD		
Length (Å)	0.035	0.020
Angle (°)	2.756	2.226
MolProbity score	2.47	1.02
Clash score	23.18	0.82
Ramachandran plot (%)		
Favored	93.62	96.12
Allowed	4.59	3.2
Outliers	1.79	0.68
Rotamer outliers (%)	1.80	0.54
Cβ outliers (%)	1.97	0.51
Peptide plane (%)		
Cis proline	2.9	6.4
Twisted proline	0	0
CaBLAM outliers (%)	4.48	2.95
EMRinger score	2.61	2.13