Structural Biology and Molecular Biophysics
Developmental Biology

Deep learning insights into the architecture of the mammalian egg-sperm fusion synapse

Reviewed Preprint

Revised by authors after peer review.

About eLife's process

Reviewed preprint version 2
April 10, 2024 (this version)

Reviewed preprint version 1
January 4, 2024

Sent for peer review
October 13, 2023

Posted to preprint server
September 27, 2023

Arne Elofsson, Ling Han, Enrica Bianchi, Gavin J. Wright, Luca Jovine

Science for Life Laboratory and Department of Biochemistry and Biophysics, Stockholm University, Box 1031, 171 21 Solna, Sweden • Department of Biosciences and Nutrition, Karolinska Institutet, 141 83 Huddinge, Sweden • Department of Biology, Hull York Medical School, York Biomedical Research Institute, University of York, YO10 5DD York, UK

Abstract

A crucial event in sexual reproduction is when haploid sperm and egg fuse to form a new diploid organism at fertilization. In mammals, direct interaction between egg JUNO and sperm IZUMO1 mediates gamete membrane adhesion, yet their role in fusion remains enigmatic. We used AlphaFold to predict the structure of other extracellular proteins essential for fertilization to determine if they could form a complex that may mediate fusion. We first identified TMEM81, whose gene is expressed by mouse and human spermatids, as a protein having structural homologies with both IZUMO1 and another sperm molecule essential for gamete fusion, SPACA6. Using a set of proteins known to be important for fertilization and TMEM81, we then systematically searched for predicted binary interactions using an unguided approach and identified a pentameric complex involving sperm IZUMO1, SPACA6, TMEM81 and egg JUNO, CD9. This complex is structurally consistent with both the expected topology on opposing gamete membranes and the location of predicted N-glycans not modeled by AlphaFold-Multimer, suggesting that its components could organize into a synapse-like assembly at the point of fusion. Finally, the structural modeling approach described here could be more generally useful to gain insights into transient protein complexes difficult to detect experimentally.

Impact statement

Structural modeling with AlphaFold-Multimer was used to investigate extracellular protein interactions involved in mammalian egg-sperm recognition, suggesting a putative pentameric complex that includes TMEM81, a sperm protein not previously involved in gamete recognition.

eLife assessment

This study offers **valuable** insights into the structural architecture of the mammalian egg-sperm fusion synapse, shedding light on the role of specific proteins in fertilization. The significance of the findings lies in the potential identification of a pentameric complex involved in gamete fusion by AlphaFold Multimer. The strength of evidence for the approach/methodology is **solid**, while the experimental validation is **incomplete** in supporting these interactions. This work will be of interest to biomedical researchers working on fertility and reproductive health.

Introduction

By merging the plasma membranes of egg and sperm and combining genetic material to initiate the development of a new individual, gamete fusion is the culmination of fertilization and a fundamental event in the life cycle of sexually reproducing species. Significant advances during the last twenty years have started to unravel the molecular basis of this phenomenon by identifying proteins essential for this process in organisms ranging from unicellular algae to mammals ***(Clark, 2018; Deneke and Pauli, 2021)***. In particular, recognition between egg glycosylphosphatidylinositol-anchored protein JUNO and sperm type I transmembrane protein IZUMO1 was found to be essential for the fusion of mouse gametes by mediating the juxtaposition of their plasma membranes ***(Bianchi et al., 2014; Inoue et al., 2005)***. In agreement with such a docking function, structural studies showed that, although the architecture of the ectodomain of IZUMO1 is reminiscent of *Plasmodium* invasion proteins, neither molecule resembles known fusogens ***(Aydin et al., 2016; Han et al., 2016; Kato et al., 2016; Nishimura et al., 2016; Ohto et al., 2016)***; at the same time, mouse IZUMO1 was recently reported to have fusogenic activity *in vitro* ***(Brukman et al., 2023)***, but whether this reflects a comparable function *in vivo* remains to be determined ***(Bianchi and Wright, 2023)***.

Despite the importance of the JUNO/IZUMO1 interaction for gamete fusion, gene ablation experiments in the mouse have identified several other egg and sperm molecules essential for this process. On the female side, these include two phylogenetically close tetraspanin membrane proteins, CD9 and CD81 ***(Miyado et al. 2000; Kaji et al. 2000; Miller et al. 2000; Rubinstein et al. 2006)***. CD9 concentrates to the gamete adhesion area concomitantly with IZUMO1 ***(Chalbi et al., 2014)*** and is thought to facilitate fusion by reshaping the oocyte's plasma membrane ***(Jégou et al., 2011; Umeda et al., 2020)***. CD81 is 44%-sequence identical to CD9 and can partially rescue the infertility of CD9-deficient mouse eggs ***(Kaji et al., 2002; Ohnami et al., 2012)***. On the male side, several surface-expressed molecules are required for mouse gamete fusion in addition to IZUMO1. These include sperm acrosome membrane-associated protein 6 (SPACA6) ***(Barbaux et al., 2020; Lamas-Toranzo et al., 2020; Lorenzetti et al., 2014; Noda et al., 2020)*** and transmembrane protein 95 (TMEM95) ***(Lamas-Toranzo et al., 2020)***, both of which are type I-transmembrane proteins with an IZUMO1-like ectodomain structure ***(Lamas-Toranzo et al., 2020; Nishimura et al., 2016; Vance et al., 2022)***. Sperm dendrocyte expressed seven transmembrane protein domain-containing proteins 1 and 2 (DCST1/2), which interact with each other ***(Noda et al., 2022)*** and are orthologues of molecules essential for fusion in worm (SPE-49/42) ***(Kroft et al., 2005; Wilson et al., 2018)*** and fly (SNEAKY/DCST-2) ***(Wilson et al., 2006)***, are required for fertility not only in the mouse but also in fish ***(Inoue et al., 2021; Noda et al., 2022)***. Finally, two other molecules necessary for mouse gamete fusion are fertilization influencing membrane protein (FIMP), the transmembrane domain-containing isoform of 4930451I11RIK ***(Fujihara et al., 2020)***, and sperm-oocyte fusion required 1 (SOF1) ***(Noda et al., 2020)***. In addition to this gene knockout-derived information, there is biochemical

evidence that IZUMO1 is part of rodent sperm multiprotein complexes that include structurally related molecules IZUMO2-4 (***Ellerman et al., 2009***). More recently, egg Fc receptor-like 3 (FCRL3/MAIA) was also suggested to be involved in human gamete adhesion and fusion by replacing JUNO as an IZUMO1-binding partner (***Vondrakova et al., 2022***), although others have not confirmed this (***Bianchi et al. 2024***).

The relatively large number of proteins that these studies collectively identified as required for mammalian egg-sperm fusion, together with the lack of conclusive evidence supporting a direct role of the JUNO/IZUMO1 complex in the fusion process itself, suggest that — in line with the concept of fertilization synapse (***Krauchunas et al., 2016***) — a larger macromolecular complex may orchestrate fusion. However, perhaps because such an assembly exists only transiently due to the need to prevent polyspermy, the identification of additional protein-protein interactions between the aforementioned factors has frustrated independent efforts by multiple laboratories.

Here, we show that, despite the clear centrality of the JUNO/IZUMO1 interaction, its mouse components have such a low affinity that, unlike their human homologs, they cannot be purified as a stable complex. Because the biochemical identification of other egg/sperm fusion factor complexes may be hindered by the fact that their binary affinities also vary significantly among different species, we attack the problem by taking advantage of the momentous advances in protein complex structure prediction using AlphaFold-Multimer (***Burke et al., 2023; Evans et al., 2021***). The rationale for using this approach lies in the fact that the availability of a significant number of sequences for the proteins of interest not only allows AlphaFold to predict possible complexes thereof highly accurately (***Jumper et al., 2021; Lee et al., 2023; Mirdita et al., 2022***), but also makes it largely insensitive to the species-specific affinity of a given protein-protein interaction.

Consistent with these considerations, the analysis of AlphaFold-Multimer predictions supports the suggestion that JUNO and IZUMO1 are part of a complex that includes additional fusion factors.

Results

Mouse JUNO and IZUMO1 do not form a biochemically stable complex

Whereas mammalian cell-expressed human JUNO and IZUMO1 ectodomains form a stable complex ($JUNO_E$/$IZUMO1_E$) that can be detected by size-exclusion chromatography (SEC), their murine homologs do not (***Figure 1 and Figure S1***). This is consistent with the low affinity of the interaction between the mouse proteins, whose 0.6-12 µM K_D is significantly higher than the ~50-90 nM K_D reported for the human $JUNO_E$/$IZUMO1_E$ complex expressed in insect cells (***Aydin et al., 2016; Bianchi et al., 2014; Nishimura et al., 2016; Ohto et al., 2016***). Notably, the K_D of wild-type mouse $JUNO_E$/$IZUMO1_E$ is also higher than the 360 nM K_D of the complex between human $IZUMO1_E$ and $JUNO_E$ W62A (***Aydin et al., 2016; Ohto et al., 2016***). The latter bears an interface mutation whose introduction into mouse JUNO abolishes its ability to rescue the sperm-fusion impairment of *Juno* null eggs, as well as halves its ability to support sperm binding to JUNO-expressing HEK293T cells (***Kato et al., 2016***). The low affinity of mouse $JUNO_E$/$IZUMO1_E$ could, in principle, be partially compensated by the avidity resulting from a high local concentration of receptors at the egg/sperm contact point. At the same time, consistent with the considerations made above, the binary interaction between JUNO and IZUMO1 may be stabilized within the context of a larger macromolecular complex.

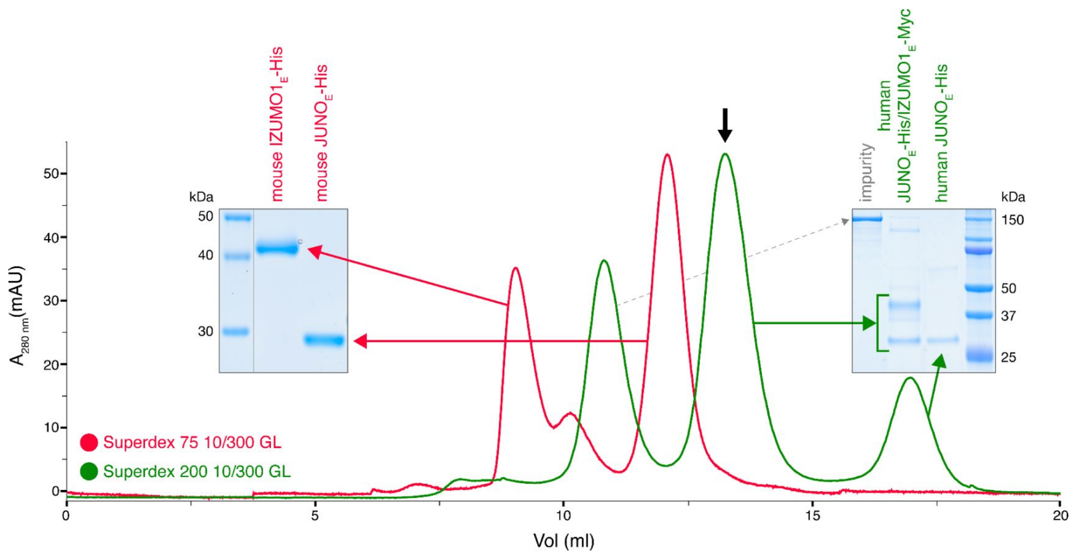


Figure 1.

Human but not mouse JUNO and IZUMO1 ectodomains form a stable complex in solution

The SEC elution profile of immobilized metal affinity chromatography (IMAC)-purified human $JUNO_E$-His/$IZUMO1_E$-Myc (green trace) shows a major peak that contains both proteins (black arrow), as well as a peak corresponding to unbound $JUNO_E$-His (SDS-PAGE analysis on the right). On the contrary, IMAC-purified mouse $JUNO_E$-His and mouse $IZUMO1_E$-His elute separately on gel filtration (red trace and SDS-PAGE analysis on the left). See also **Figure S1**.

AlphaFold-Multimer produces high-confidence predictions for both mouse and human JUNO/IZUMO1 ectodomain complexes

To assess whether the significant difference in affinity between the mouse and human complexes was reflected by the confidence of the corresponding AlphaFold-Multimer predictions, we compared the output of AlphaFold runs performed without using templates. This computational experiment showed that AlphaFold-Multimer not only generates a high-confidence model of human $JUNO_E$/$IZUMO1_E$ that accurately reproduces the corresponding crystal structure but also yields a model of mouse $JUNO_E$/$IZUMO1_E$ of comparable confidence (***Figure 2***). This is consistent with the expectation that, as long as a significant number of sequences can be aligned to those of a protein complex of interest and the interaction is evolutionarily conserved, the quality of the AlphaFold-Multimer predictions for this complex is not negatively affected by the low affinity that its components may have in a subset of species.

TMEM81 is a structural homolog of IZUMO1 and SPACA6

Considering that IZUMO1-4, SPACA6, and TMEM95 are part of a distinct superfamily of extracellular proteins implicated in gamete fusion (***Lamas-Toranzo et al., 2020***; ***Nishimura et al., 2016***; ***Vance et al., 2022***), we used Foldseek (***van Kempen et al., 2023***) to scan the AlphaFold/Swiss-Prot database for further proteins of similar structure. Despite insignificant sequence identities (16-27%), this search also identified transmembrane protein 81 (TMEM81) as a clear structural homolog of the conserved immunoglobulin (Ig)-like domain of IZUMO1 and SPACA6 (E-values 1.40e-8 - 1.39e-6) (***Figure 3A, B***). The TMEM81 hit was confirmed by the result of a search of the PDB database, carried out by generating an AlphaFold model of the protein's ectodomain and using it as input for Dali (***Holm, 2020***), which matched it to the crystal structure of human IZUMO1 (PDB 5JK9 (***Ohto et al., 2016***)) with a Z-score of 11.3 (significantly above the Z-score threshold of 8, which indicates very good structural superpositions (***Holm, 2020***)). Notably, TMEM81 is conserved in vertebrates (***NCBI, 2022***), and its gene is expressed in both mouse and human spermatids (***Jung et al., 2019***; ***Uhlén et al., 2015***; ***Yue et al., 2014***). Like IZUMO1-3, SPACA6, and TMEM95, TMEM81 is predicted to be a type I transmembrane protein with a large extracellular domain; moreover, it was previously anonymously suggested to be a β-sheet-rich molecule that may be structurally related to IZUMO1 (***Wikipedia, 2020***). Accordingly, the characteristic four-helix bundle (4HB) of IZUMO1 and SPACA6 is replaced by a three-stranded β-sheet in the AlphaFold model of TMEM81; however, the positioning of two invariant disulfide bonds that orient these highly different elements relative to the conserved Ig-like domain is remarkably similar in the three molecules (***Figure 3C***).

Prediction of interactions between human proteins associated with gamete fusion

To infer whether a larger macromolecular complex may be involved in gamete fusion without introducing a large number of possible false positives observed in attempts to perform a large-scale screening, we used AlphaFold-Multimer in template-free mode to examine all pairwise interactions of the human homologs of the 4 egg and 11 sperm proteins mentioned above, including TMEM81. Since we also considered the possibility that each of these 15 different molecules may also homodimerize, this amounted to a total of 120 unique combinations. Analysis of the corresponding predictions revealed a cluster of 7 possible interactions centered around IZUMO1, 5 of which were direct (JUNO, CD9, CD81, SPACA6, TMEM81) and 2 indirect (IZUMO4 (via JUNO), SOF1 (via SPACA6)). In addition, we detected isolated homodimeric interactions for IZUMO4 and DCST1, as well as heterodimeric interactions for IZUMO2/IZUMO3, TMEM95/FIMP and

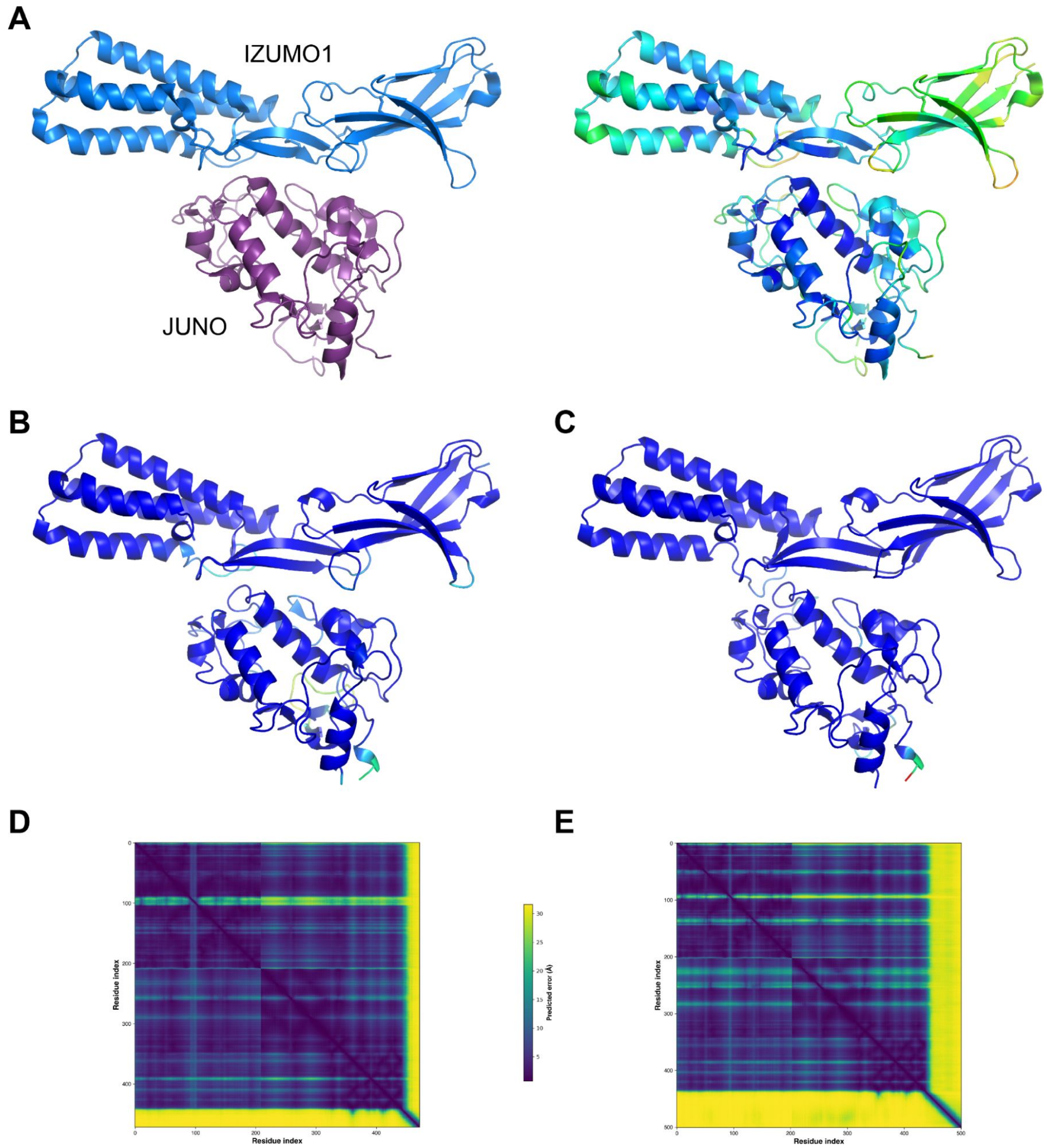


Figure 2.

Mouse JUNO-IZUMO1 complex structure prediction

(**A**) The crystal structure of the human $JUNO_E$/$IZUMO1_E$ complex (PDB 5F4E) ***(Aydin et al., 2016)***, shown in cartoon representation and colored by chain (left) or by B-factor (right).
(**B**) AlphaFold-Multimer template-free prediction of the structure of the human $JUNO_E$/$IZUMO1_E$ complex. The top-ranked model has a ranking confidence (rc = 0.8*predicted interface Template Modeling score (ipTM) + 0.2*predicted Template Modeling score (pTM)) of 0.87, and an average root mean square deviation (RMSD) from PDB 5F4E of 2.34 Å over 437 Cα (0.88 Å over 380 Cα after outlier rejection). Only the residues that match those resolved in the crystal structure are shown; the model is colored by prediction confidence from blue to red, according to a 100-(per-residue confidence (predicted local distance difference test, pLDDT) ***(Jumper et al., 2021)***) scale that ranges from 0 (blue; maximum confidence) to 100 (red; minimum confidence)), respectively.
(**C**) AlphaFold-Multimer top-ranked template-free prediction of the structure of the mouse $JUNO_E$/$IZUMO1_E$ complex (rc = 0.85; RMSD vs. 5F4E = 2.53 Å over 435 Cα (1.73 Å over 389 Cα after outlier rejection)), depicted and colored as in panel B.
(**D**) Predicted Aligned Error (PAE) plot for the human complex model shown in panel B. Residue indexes refer to the sequence of JUNO (amino acids G20-S228) followed by that of IZUMO1 (amino acids C22-Q284). The high PAE regions correspond to loop 2 of JUNO (residues V110-G123) and the C-terminal tail of $IZUMO1_E$ (residues K255-Q284), both of which have low pLDDT scores and are far away from the interface between the two proteins.
(**E**) PAE plot of the mouse complex shown in panel C, with residue indexes referring to JUNO (amino acids G20-G222) followed by IZUMO1 (amino acids C22-R319).

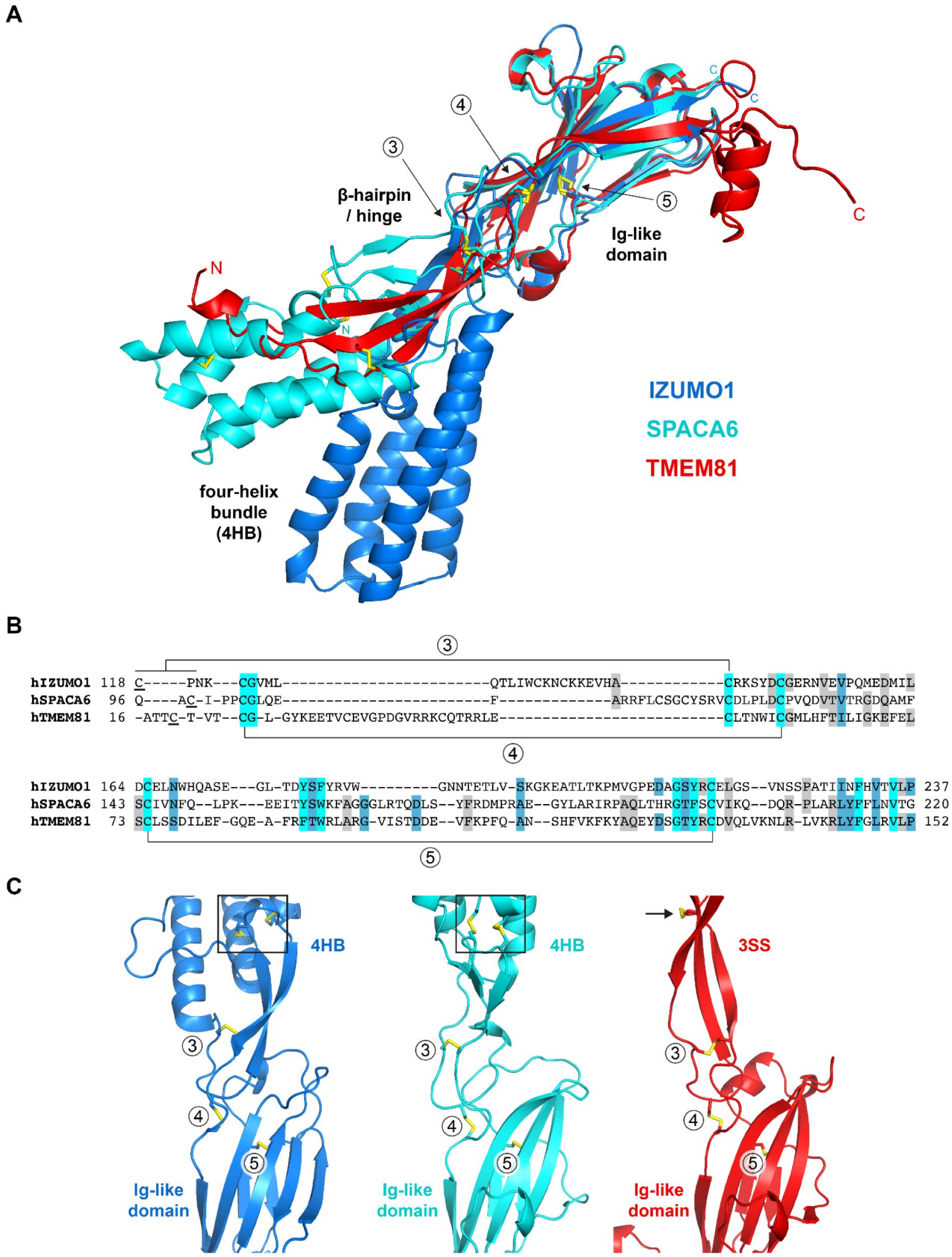


Figure 3.

Structural homology between IZUMO1, SPACA6 and TMEM81

(**A**) Structural superposition of the ectodomains of human IZUMO1 (residues C22-K255 of PDB 5JK9 chain A (***Aydin et al., 2016***)), human SPACA6 (residues C27-G246 of PDB 7TA2 (***Vance et al., 2022***)) and an AlphaFold model of the ectodomain of human TMEM81 (corresponding to residues I31-P218 of UniProt entry Q6P7N7). The three different regions of IZUMO1 and SPACA6 are indicated in black. Disulfide bonds are shown as yellow sticks, with arrows indicating disulfides 3-5 of IZUMO1 that are conserved in both SPACA6 and TMEM81. N- and C-termini are marked.

(**B**) Structure-based alignment of the sequence regions includes conserved disulfides 3 and 4, followed by the Ig-like domain harboring conserved disulfide 5.

(**C**) Partial grid view of the superposition shown in panel A, centered around the junction between the three molecules' variable (top) and conserved (bottom) domains. Note the strikingly similar relative arrangement of invariant disulfides 3, 4, and 5, and how an additional disulfide within the three-stranded sheet (3SS) of TMEM81 (black arrow) roughly matches the position of the double CXXC motifs of IZUMO1 and SPACA6 (black boxes).

DCST1/DCST2 (***Figure 4*** ***and Table S1***). Notably, the ~260 Å-long mace-shaped heterodimeric assembly predicted for DCST1/DCST2 is consistent with experimental evidence for interaction between the two proteins ***(Noda et al., 2022)***.

To assess the relative contribution of the components of the 7-interaction cluster, we used AlphaFold-Multimer to model the corresponding 8-protein complex. Analysis of the resulting predictions (***Figure 5A*** *and* ***Figure S2A***), as well as the predictions of the binary complexes IZUMO1/CD9 (***Figure S2B***) or IZUMO1/CD81 (***Figure S2C***), suggest that the two egg tetraspanins are interchangeable because they are predicted to bind to the same region of IZUMO1; moreover, in agreement with the observation that mouse fertility depends more on CD9 than CD81 ***(Kaji et al., 2002, 2000; Miller et al., 2000; Miyado et al., 2000; Ohnami et al., 2012; Rubinstein et al., 2006)***, IZUMO1 consistently interacts with the former when modeled together with both tetraspanins. The 8-protein complex predictions also indicate that IZUMO4 does not interact with the rest of the assembly (***Figure 5A*** *and* ***Figure S2A***), consistent with the observation that its predicted binary interaction with JUNO (***Figure 4***) is incompatible with the JUNO/IZUMO1 interface (***Figure S2D***). Finally, pDockQ and visual analysis of the predictions for the 8-protein complex indicate that SOF1 is mainly disordered and does not make significant contacts with other components (***Figure 5A*** *and* ***Figure S2A***). Taken together, these considerations leave egg JUNO and CD9 and sperm IZUMO1, SPACA6 and TMEM81 as subunits of a 5-protein complex that can be consistently modeled with acceptable ranking confidence and pDockQ scores (***Figure 5B, C***). Consistent with their central role in interfacing the egg and sperm plasma membranes, JUNO and IZUMO1 constitute the core of this putative assembly, where they interact in the same way that was observed crystallographically ***(Aydin et al., 2016; Ohto et al., 2016)*** and reproduced computationally ***(Figure 2)***. On the opposite side of the JUNO/IZUMO1 interface, the hinge region and 4HB of SPACA6 wrap around the 4HB of IZUMO1, generating a concave surface that interacts with the long extracellular loop (LEL) of CD9. Finally, TMEM81 adopts the same N-to-C orientation of IZUMO1 and SPACA6 and, by inserting its Ig-like module between the two proteins, links their C-terminal regions.

Discussion

In this study, we have taken advantage of the latest developments in protein structure and interaction prediction to model protein complex formation in the mammalian egg-sperm fusion synapse. We report the supramolecular organization of five cell surface proteins (three sperm and two egg) that form a core complex likely to be important for gamete recognition and fusion.

Because the only specific information about the target molecules that is used as input for AlphaFold-Multimer is their primary sequence, the neural network model does not incorporate any knowledge of data associated with the system's biology. As a result, biological information on egg-sperm fusion can be used as an independent criterion to validate the predictions. Firstly, because the majority of proteins involved in this process are either C-terminally membrane-anchored or transmembrane proteins, a basic feature expected in a gamete fusion synapse is that the C-termini of its egg subunits or the egg subunits themselves should all be located on the opposite side of the corresponding elements from sperm, relative to the gamete interface. This is true for the 5-component complex predictions (***Figure 5C***). On the egg plasma membrane side, JUNO and CD9 are positioned so that the GPI anchor attached to the C-terminus of the former (which is not modeled by AlphaFold, whose predictions are currently restricted to amino acids) would be located in correspondence with the transmembrane domains of CD9. Similarly, the general orientation and high flexibility of the juxtamembrane regions of IZUMO1, SPACA6, and TMEM81 are compatible with the fact that, in the context of the full-length proteins, these elements are connected to the single-spanning transmembrane helices that anchor the corresponding molecules to the sperm plasma membrane.

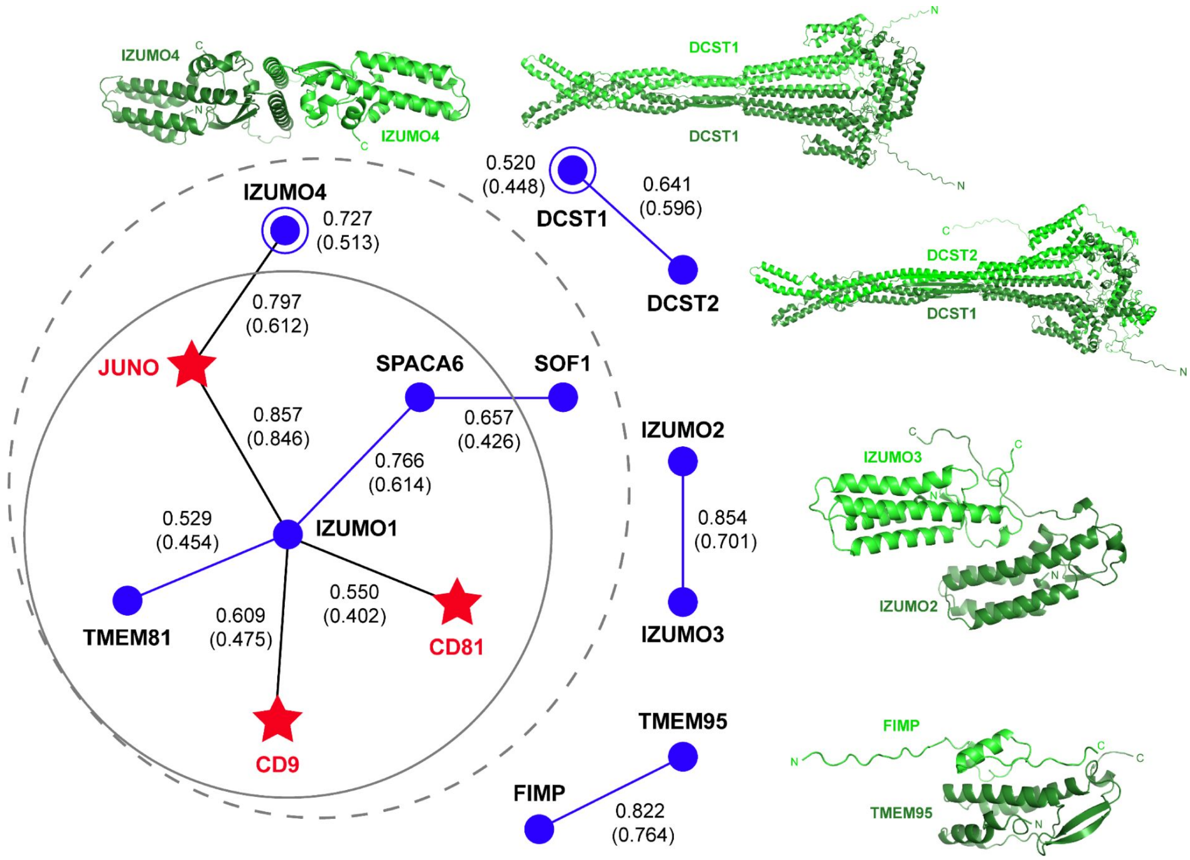


Figure 4.

AlphaFold-Multimer prediction of interactions between fusion-associated human gamete proteins

Egg and sperm proteins are indicated by red star and blue circle symbols, respectively. Interactions between egg and sperm proteins are shown as black lines connecting the respective symbols; homomeric and heteromeric interactions between sperm proteins are depicted as blue lines and open circles, respectively. For every interaction, the top-ranking model rc is reported, with the corresponding mean rc in parenthesis (for complete metrics, see Table S1). The gray dashed circle indicates a network of 7 interactions, identified using a mean rc cutoff of 0.4; the inner continuous circle highlights the 5 interactions within the network that involve sperm IZUMO1. Top-ranked predictions for the isolated binary interactions of other sperm subunits are shown in cartoon representation, with the two moieties of each complex colored dark and light green and the N- and C-termini of each chain indicated when possible.

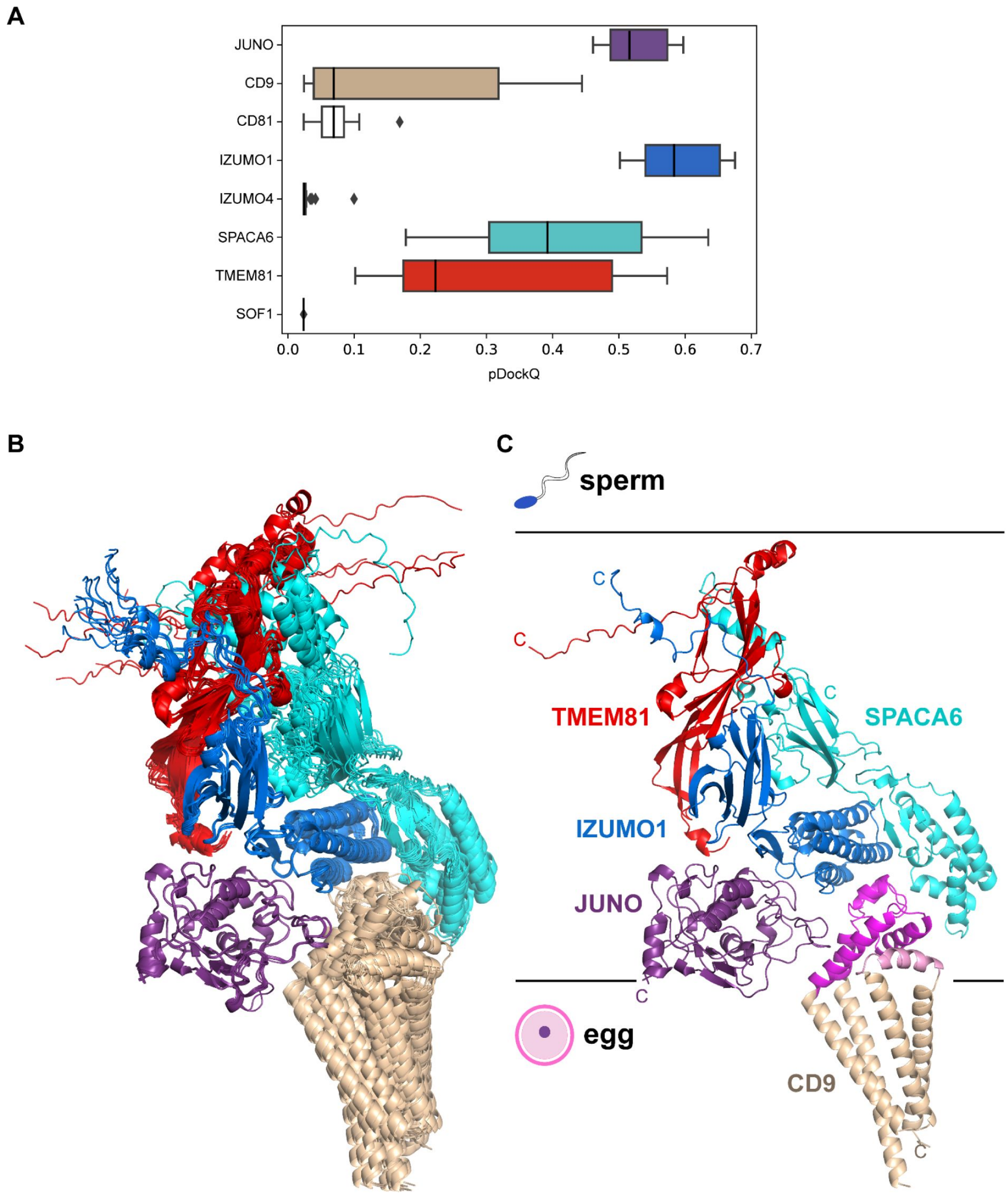


Figure 5.

A predicted five-subunit complex at the egg/sperm plasma membrane interface

(**A**) pDockQ analysis of 25 AlphaFold-Multimer predictions for a complex consisting of the 8 proteins enclosed by the dashed gray circle in **Figure 4**. The pDockQ score for each component of every prediction was calculated with respect to the rest of the corresponding complex, and the 25 scores for each chain were then plotted as a box plot.

(**B**) Superposition of the ten top-ranked AlphaFold predictions for a five-subunit complex consisting of egg CD9 and the ectodomains of egg JUNO and sperm IZUMO1, SPACA6 and TMEM81 (mean rc = 0.67, mean ipTM = 0.66). Proteins are shown in cartoon representation and colored by chain according to panel A.

(**C**) Top-ranked model from the ensemble in panel B (rc = 0.74, ipTM = 0.73). Subunits are colored as in the previous panels, except for CD9 whose short extracellular loop (SEL) and long extracellular loop (LEL) are highlighted in pink and magenta, respectively. Protein C-termini are marked, with horizontal lines representing the approximate surfaces of the gamete plasma membranes.

A second feature common to all the subunits in the modeled complexes is the presence of N-glycosylation sites. Because AlphaFold has no explicit knowledge of sequons and does not model carbohydrates, the N-glycans that decorate the native molecules could, in principle, interfere with predicted interfaces. As shown in ***Figure 6***, the predicted complex architecture is compatible with the location of all the possible N-glycosylation sites of JUNO, IZUMO1, SPACA6, and TMEM81, for both human and mouse homologs (amounting to a total of ten sites). One possible exception is a sequon within the short extracellular loop (SEL) of CD9, which is conserved in both species (corresponding to human N52 and mouse N50, respectively) but whose glycosylation remains to be experimentally verified. Interestingly, this site is located in relatively close proximity to where loop 3 of JUNO protrudes towards the region between the LEL and SEL of CD9 ***(Figure 7A)***. This suggests that if the conserved sequon of CD9 is glycosylated, this may interfere with the only minor contact that the protein makes with JUNO within the predicted complex. Notably, a fusion synapse architecture where CD9 makes little or no contact with JUNO but interacts with the 4HB of IZUMO would immediately explain the experimental observation that, in mouse oocytes, CD9 is recruited to the gamete fusion site only upon binding of JUNO to IZUMO1 ***(Chalbi et al., 2014)***. Moreover, the predicted CD9/IZUMO1 interface agrees with previous suggestions that the two proteins may interact, based on the observation that their sequences co-evolve ***(Claw et al., 2014; Vicens and Roldan, 2014)***.

Although there is a general agreement that the CD9 LEL plays an important role in gamete fusion, which of its residues are responsible for this is debated; in particular, an early suggestion that the 173-SFQ-175 motif of mouse CD9 LEL is required for fusion was recently challenged ***(Umeda et al., 2020; Zhu et al., 2002)***. Against this background, it is interesting to note that, in our predictions, the conserved CD9 Phe at the center of the SFQ tripeptide (175-TFT-177 in human) stacks against α-helix 2 of the IZUMO 4HB ***(Figure 7B)***. Not far from this interaction, the third α-helix of CD9 LEL makes hydrophobic contacts with IZUMO1 α2 and α4. These interactions are close to L115, a conserved IZUMO1 α4 residue thought to contribute to egg binding and fusion ***(Inoue et al., 2013)***, and directly involve W113, another conserved α4 amino acid that was recently implicated in fusion ***(Brukman et al., 2023)***. Notably, W113 bridges CD9 and SPACA6 by inserting between their LEL and double CXXC motif elements, respectively, while W88 — another IZUMO1 residue suggested to be important for fusion ***(Brukman et al., 2023)*** — also interacts hydrophobically with SPACA6 at the opposite side of IZUMO1's 4HB ***(Figure 7B)***.

Whereas all the data above is in good agreement with the structural predictions described in this manuscript, two aspects should be considered with caution. First, it remains unclear why, despite the fact that IZUMO1 complementation rescues the disappearance of SPACA6 from the mature sperm of IZUMO1 null mice ***(Inoue et al., 2021)***, attempts to biochemically identify a complex between IZUMO1 and SPACA6 have been met with limited success ***(Noda et al., 2020; Vance et al., 2022)***. Based on the predicted complex architecture, one obvious possibility would be that, in order to be stable, the interaction also requires the presence of TMEM81. One reason could be the low affinity of the interactions between these proteins, which is typical of extracellular receptor-ligand interactions and makes them difficult to detect experimentally ***(Wright and Bianchi, 2016)***. Also, to avoid any inappropriate membrane fusion events, the individual components of the complex may be purposefully spatially segregated until brought together at the moment of fusion, again making it difficult to detect this complex *in vivo*. The type of structural modeling approach described here could play a role in understanding the function of dynamic and transiently formed protein complexes in a range of biological processes that would otherwise be difficult to identify, although care should be taken as some complexes might be predicted due to interactions between homologs. Second, because it is difficult to assess the confidence of the interface between CD9 and IZUMO1+SPACA6 due to its relatively limited extent (combined interface area ~730 Å^2), it cannot be excluded that some protein other than CD9 may be the true counterpart of IZUMO1+SPACA6. In other words, it is, in principle, also possible that AlphaFold-

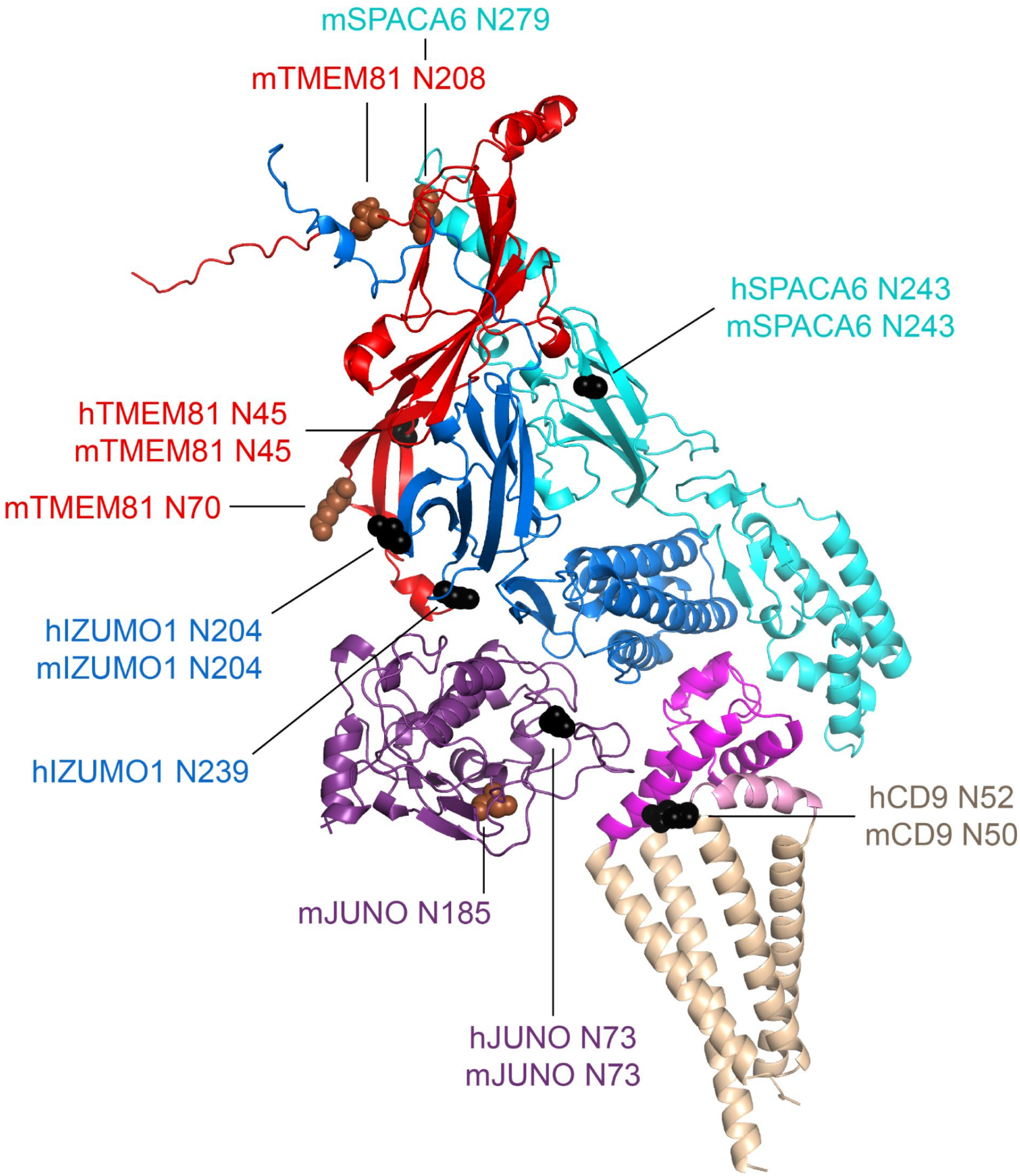


Figure 6.

Mapping of sequon positions onto the 5-subunit complex prediction

The positions of possible glycosylation sites are mapped onto the model of the 5-subunit assembly (depicted as in **Figure 5C**) by showing the corresponding Asn residues in sphere representation. Sequons found in human (h prefix) or in both human and mouse proteins are colored black, whereas sequons only found in mouse (m prefix) proteins are brown.

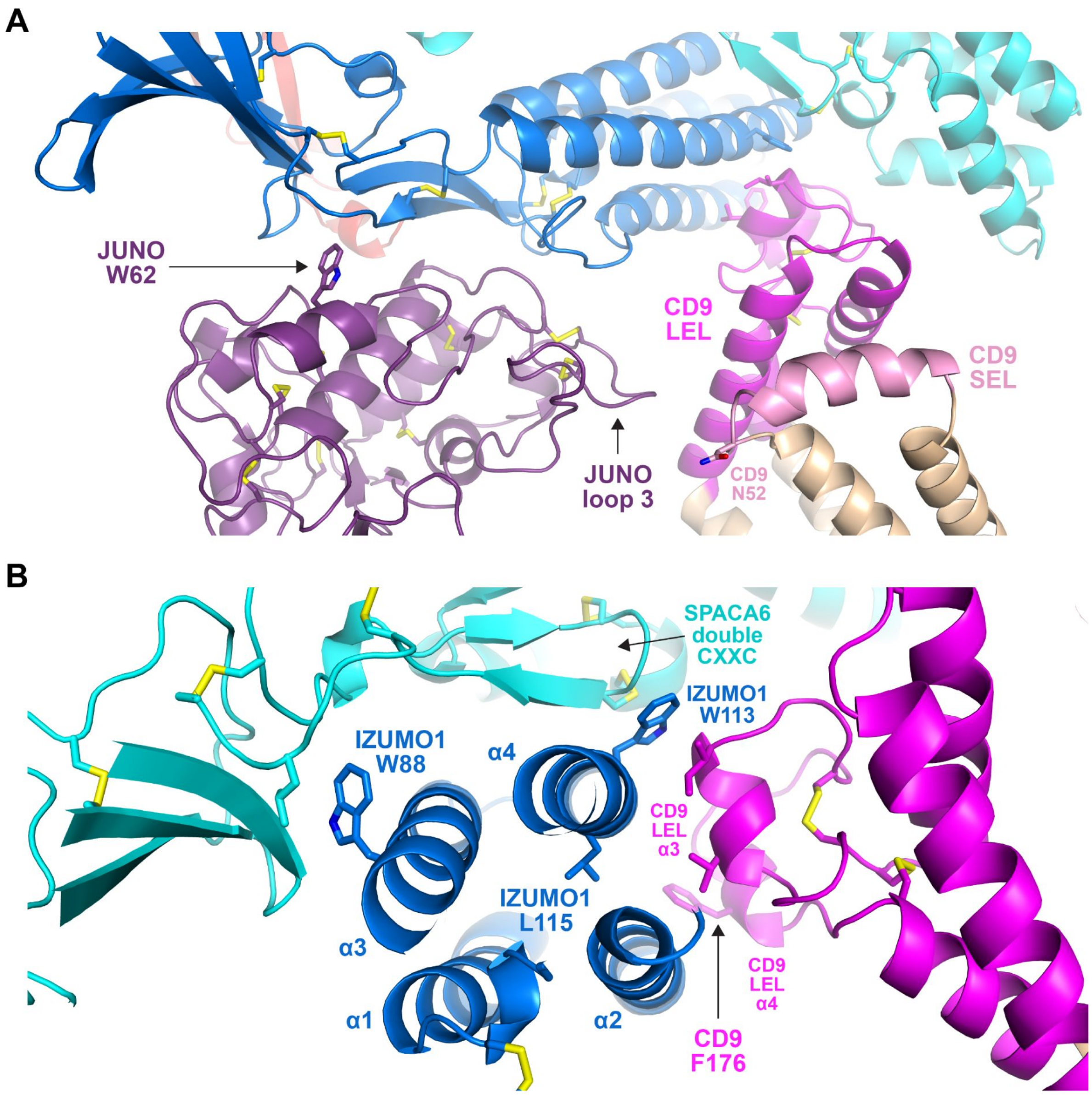


Figure 7.

Subunit interfaces of the predicted complex involve protein elements previously implicated in fusion.

(**A**) Detail of the prediction shown in **Figure 5C**, highlighting functionally important regions of JUNO and CD9, as well putatively N-glycosylated CD9 N52.

(**B**) Different view of the same complex prediction, centered around the IZUMO1 4HB.

Multimer simply recognizes that the concave surface generated by the combined 4HBs of IZUMO1 and SPACA6 is likely to engage in additional interactions and thus tries to fill it with another suitable input subunit.

Of direct relevance to these questions is an independent study, submitted back to back with the original preprint of the present manuscript, which provides experimental data for the existence of a trimeric IZUMO1/SPACA6/TMEM81 complex in zebrafish and suggests that this interacts with egg Bouncer (*Deneke et al., 2023*). Considering that the mammalian orthologue of Bouncer is expressed on sperm instead of the egg (*Fujihara et al., 2021*) and that role of CD9 in egg-sperm fusion is much more important in mammals than in fish (*Greaves et al., 2022*), the combination of our studies raises the intriguing possibility that, during the course of evolution, CD9 may have substituted Bouncer as a binding partner of the IZUMO1/SPACA6/TMEM81 complex.

Methods

Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Cell line (*Homo sapiens*)	HEK293T	Dr. Radu Aricescu and Dr. Yuguang Zhao (University of Oxford, UK)		
Cell line (*Homo sapiens*)	HEK293S GnTI⁻	ATCC	CRL-3022	
Antibody	Penta·His	QIAGEN	34660	Mouse igG1 mAb

Antibody	Anti-c-Myc	Sigma-Aldrich	M4439	Mouse igG1 mAb, clone 9E10
Antibody	Peroxidase AffiniPure Goat Anti-Mouse IgG (H+L)	Jackson ImmunoResearch Laboratories	115-035-003	
Recombinant DNA reagent	pHLsec3-hJUNO-$(GGGS)_2H_8$	This publication		
Recombinant DNA reagent	pHLsec3-hIZUMO1-Myc	This publication		
Recombinant DNA reagent	pHLsec3-mJuno-H_8	Han et al., 2016		
Recombinant DNA reagent	pHLsec3-mIzumo1-LEH_6	Nishimura et al., 2016		
Chemical compound, drug	25 kDa branched polyethyleneimine	Sigma-Aldrich	408727	
Chemical compound, drug	SimplyBlue SafeStain	Thermo Fisher Scientific	LC6060	
Software, algorithm	AlphaFold2	Jumper et al., 2021; Evans et al., 2021		
Software, algorithm	Belvu	Barson and Griffiths, 2016		
Software, algorithm	Dali	Holm, 2020		
Software, algorithm	Foldseek	van Kempen et al., 2023		
Software, algorithm	pDockQ	Bryant et al., 2022		
Software, algorithm	PyMOL	Schrödinger, LLC		
Software, algorithm	UCSF Chimera	Meng et al., 2006		

DNA constructs

For expression of human JUNO, a synthetic gene encoding the protein's ectodomain (residues G20-S228) followed by a 2x GGGS linker sequence (ATUM, Newark, CA, USA) was cloned into the *Age*I and *Xho*I restriction sites of mammalian expression vector pHLsec3 (***Raj et al., 2017***), in frame with 5' and 3' sequences encoding a CRYPα signal peptide/ETG tripeptide and an 8His-tag, respectively. pHLsec3 was also used to express a C-terminally Myc-tagged version of the ectodomain of human IZUMO1, preceded by its signal peptide (residues M1-L283). The ectodomains of mouse JUNO and IZUMO1 were expressed using previously described constructs (***Han et al., 2016***; ***Nishimura et al., 2016***).

Protein expression, purification and analysis

Polyethyleneimine-mediated transient transfection of HEK293 cells and protein purification by immobilized metal affinity chromatography (IMAC) and size-exclusion chromatography (SEC) was carried out following published protocols ***(Bokhove et al., 2016)***. While human $JUNO_E$-His and $IZUMO1_E$-Myc were always co-transfected, both individual transfection and co-transfection experiments were carried out in the case of mouse $JUNO_E$-His and $IZUMO1_E$-His. Samples separated on SDS-PAGE gels were detected with SimplyBlue SafeStain (Thermo Fisher Scientific) or subjected to immunoblotting with either Penta·His (1:1,000; QIAGEN) or Anti-c-Myc (1:5,000; Sigma-Aldrich) mouse monoclonal antibodies. Secondary antibody was horseradish peroxidase-conjugated goat anti-mouse IgG (1:10,000; Jackson ImmunoResearch Laboratories).

AlphaFold predictions

Predictions were generated with local copies of AlphaFold2 ***(Jumper et al., 2021; Evans et al., 2021)***, installed using versions 2.2-2.3.2 of the open-source code available at *https://github.com/deepmind/alphafold*, or by taking advantage of the Berzelius supercomputing resource (National Supercomputer Centre, Linköping University). All runs were performed using the full_dbs preset and excluding PDB templates. The human protein regions used for the binary interaction predictions whose network is shown in **Figure 4** were CD9 P2-V228 (UniProt P21926); CD81 M1-Y236 (P60033); DCST1 M1-G706 (Q5T197); DCST2 M1-K773 (Q5T1A1); FIMP A22-S77 (Q96LL3-2); JUNO G20-S228 (A6ND01); IZUMO1 C22-Q284 (Q8IYV9); IZUMO2 C21-P183 (Q6UXV1); IZUMO3 C21-D166 (Q5VZ72); IZUMO4 C18-H232 (Q1ZYL8); MAIA Q16-L580 (Q96P31); SOF1 S29-H122 (Q96L11); SPACA6 C27-T291 (W5XKT8); TMEM81 I31-P218 (Q6P7N7); TMEM95 C17-D140 (Q3KNT9). Additional prediction runs were performed using full-length sequences (excluding N-terminal signal peptide regions) also for sperm type I-transmembrane proteins and sequences that lacked disordered protein regions. The pDockQ confidence score of multi-chain predictions was calculated as described earlier ***(Bryant et al., 2022)***.

Structure analysis and comparison

Model coordinates were visualized, inspected and superimposed with PyMOL (Schrödinger, LLC), which was also used to generate all structural figures. Database searches were carried out using Dali ***(Holm, 2020)*** and Foldseek ***(van Kempen et al., 2023)***; structure-based alignments were generated with UCSF Chimera ***(Meng et al., 2006)*** and manually edited with Belvu ***(Barson and Griffiths, 2016)***.

Article and author information

Contribution: Setting up the AlphaFold runs, providing computational evaluation of potential complexes, discussion with the other authors, proofreading the manuscript.

Competing interests: No competing interests declared

Contribution: Investigation, Validation

Competing interests: No competing interests declared

Contribution: Writing–review & editing: critical review, commentary, or revision

Competing interests: No competing interests declared

Contribution: Writing–review & editing: critical review, commentary, or revision

Competing interests: No competing interests declared

Contribution: Conceptualization, Validation, Formal analysis, Resources, Writing–original draft, Writing–review & editing, Visualization, Supervision, Project administration, Funding acquisition

Competing interests: No competing interests declared

Funding

Knut and Alice Wallenberg Foundation (2018.0042)

- Luca Jovine

Swedish Research Council (2020-04936, 2021-03979)

- Arne Elofsson
- Luca Jovine

Biotechnology and Biological Sciences Research Council (BB/T006390/1)

- Enrica Bianchi
- Gavin J. Wright

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

Acknowledgements

Computations and data handling were enabled by the supercomputing resource Berzelius provided by the National Supercomputer Centre at Linköping University, the Knut and Alice Wallenberg Foundation, and SNIC (grants Berzelius-2021-29 and SNIC 2021/5-297). We thank Andrea Pauli (IMP, Vienna) for sharing her preprint before submission to *bioRxiv*.

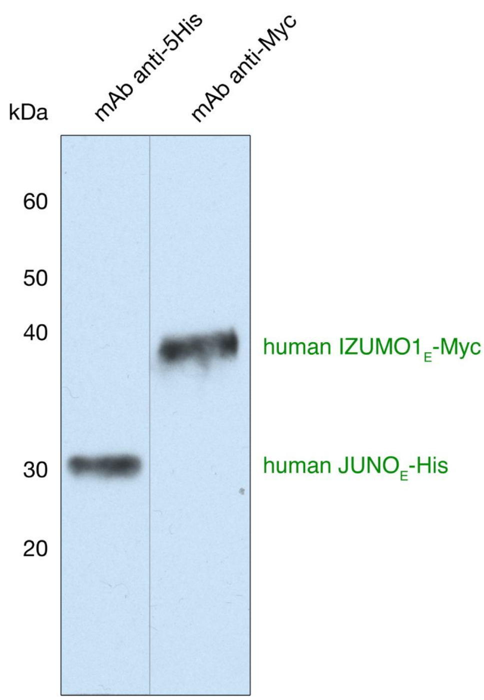


Figure S1.

Immunoblot analysis of the SEC peak corresponding to the human JUNO/IZUMO1 ectodomain complex. The peak indicated by a black arrow in **Figure 1** was separated in parallel on two separate SDS-PAGE gels under reducing conditions and then probed by immunoblot using the indicated monoclonal antibodies (mAbs).

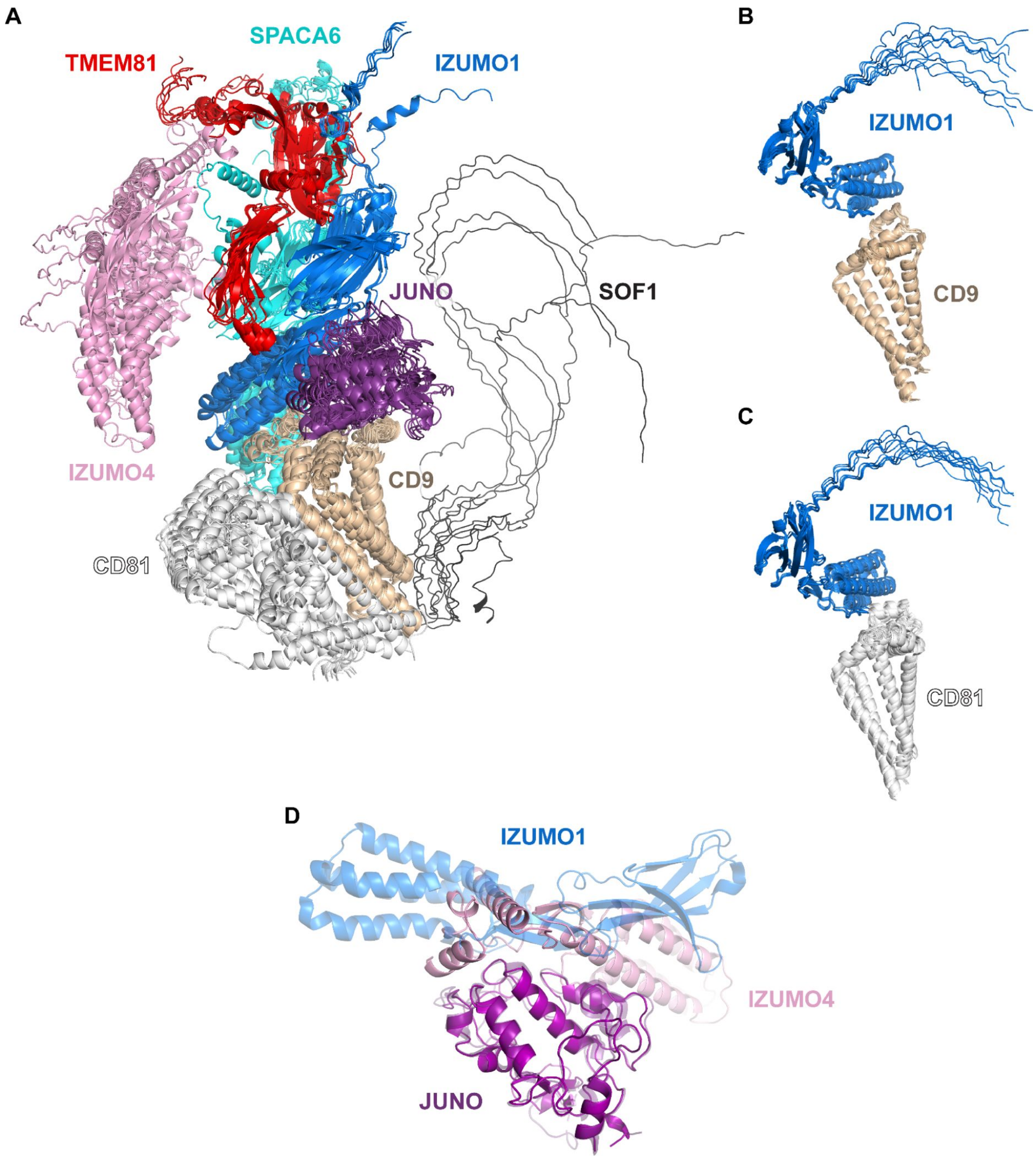


Figure S2.

Modeling of the 8-protein network and binary subcomplexes thereof.
(**A**) Highest-scoring cluster of predictions for a complex that includes the 8 proteins enclosed by the dashed gray circle in **Figure 4** (mean rc = 0.49, top rc = 0.51).
(**B**) Superposition of the ten top-ranked predictions for the IZUMO1/CD9 binary interaction (mean rc = 0.57, top rc = 0.61).
(**C**) Superposition of the ten top-ranked predictions for the IZUMO1/CD81 binary interaction (mean rc = 0.54, top rc = 0.55).
(**D**) Comparison of the human $JUNO_E$/$IZUMO1_E$ complex crystal structure (PDB 5F4E) ***(Aydin et al., 2016)*** and the top-ranked prediction for a binary complex consisting of $JUNO_E$ and IZUMO4 (rc = 0.80). The models have been superimposed over $JUNO_E$ (RMSD 2.5 Å over 204 Cα atoms, or 0.6 Å over 166 Cα atoms after outlier rejection) and colored as in **Figure 2A** and **Figure S2A**, respectively, with the experimental structure cartoon shown semi-transparent.

Uncropped gel scans for **Figure 1**.

Figure S1, left

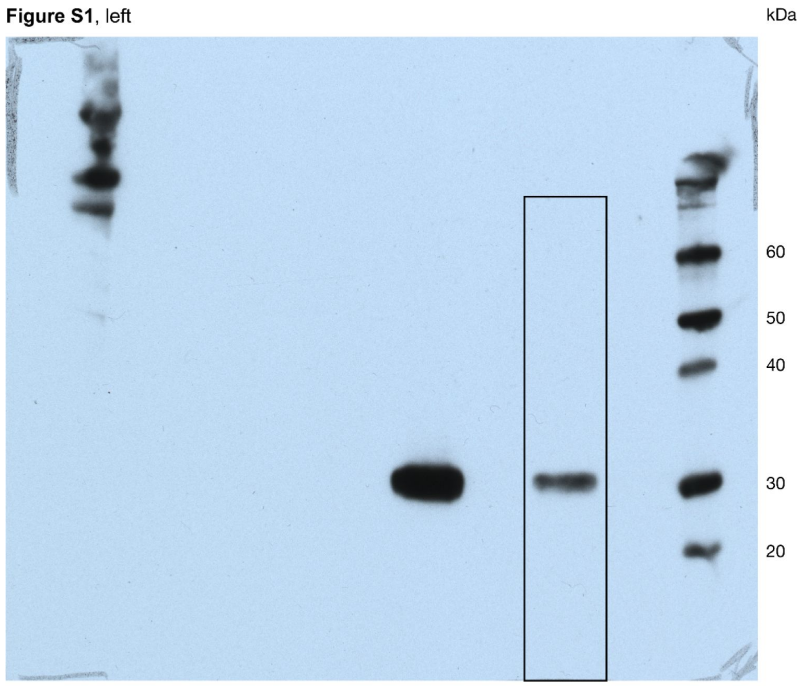
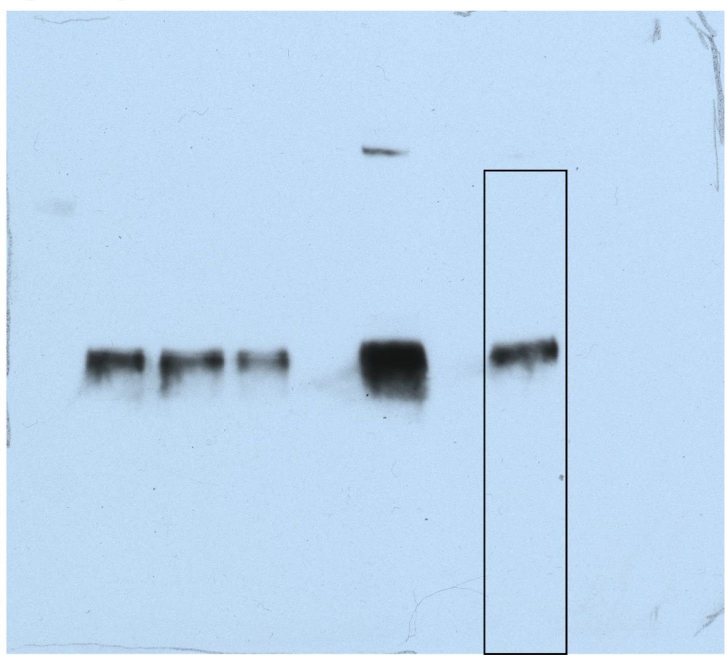


Figure S1, right



Uncropped blot scans for **Figure S1**.

Table S1. Evaluation metrics for protein complex predictions

References

Aydin H, Sultana A, Li S, Thavalingam A, Lee JE (2016) **Molecular architecture of the human sperm IZUMO1 and egg JUNO fertilization complex** *Nature* **534**:562–565 https://doi.org/10.1038/nature18595

Barbaux S, Ialy-Radio C, Chalbi M, Dybal E, Homps-Legrand M, Do Cruzeiro M, Vaiman D, Wolf J-P, Ziyyat A (2020) **Sperm SPACA6 protein is required for mammalian Sperm-Egg Adhesion/Fusion** *Scientific Reports* **10** https://doi.org/10.1038/s41598-020-62091-y

Barson G, Griffiths E (2016) **SeqTools: visual tools for manual analysis of sequence alignments** *BMC Research Notes* **9** https://doi.org/10.1186/s13104-016-1847-3

Bianchi E, Doe B, Goulding D, Wright GJ (2014) **Juno is the egg Izumo receptor and is essential for mammalian fertilization** *Nature* **508**:483–487 https://doi.org/10.1038/nature13203

Bianchi E, Wright GJ (2023) **Mammalian fertilization: Does sperm IZUMO1 mediate fusion as well as adhesion?** *The Journal Of Cell Biology* **222** https://doi.org/10.1083/jcb.202301035

Bianchi E, Jiménez-Movilla M, Cots-Rodríguez P, Viola C., Wright GJ (2024) **No evidence for a direct extracellular interaction between human Fc receptor–like 3 (MAIA) and the sperm ligand IZUMO1** *Science Advances* **10** https://doi.org/10.1126/sciadv.adk6352

Bokhove M, Al Hosseini H Sadat, Saito T, Dioguardi E, Gegenschatz-Schmid K, Nishimura K, Raj I, de Sanctis D, Han L, Jovine L. (2016) **Easy mammalian expression and crystallography of maltose-binding protein-fused human proteins** *Journal Of Structural Biology* **194**:1–7 https://doi.org/10.1016/j.jsb.2016.01.016

Brukman NG, Nakajima KP, Valansi C, Flyak K, Li X, Higashiyama T, Podbilewicz B (2023) **A novel function for the sperm adhesion protein IZUMO1 in cell-cell fusion** *The Journal Of Cell Biology* **222** https://doi.org/10.1083/jcb.202207147

Bryant P, Pozzati G, Elofsson A (2022) **Improved prediction of protein-protein interactions using AlphaFold2** *Nature Communications* **13** https://doi.org/10.1038/s41467-022-28865-w

Burke DF *et al.* (2023) **Towards a structurally resolved human protein interaction network** *Nature Structural & Molecular Biology* **30**:216–225 https://doi.org/10.1038/s41594-022-00910-8

Chalbi M *et al.* (2014) **Binding of sperm protein Izumo1 and its egg receptor Juno drives Cd9 accumulation in the intercellular contact area prior to fusion during mammalian fertilization** *Development* **141**:3732–3739 https://doi.org/10.1242/dev.111534

Clark T (2018) **HAP2/GCS1: Mounting evidence of our true biological EVE?** *PLoS Biology* **16** https://doi.org/10.1371/journal.pbio.3000007

Claw KG, George RD, Swanson WJ (2014) **Detecting coevolution in mammalian sperm-egg fusion proteins** *Molecular Reproduction And Development* **81**:531–538 https://doi.org/10.1002/mrd.22321

Deneke VE *et al.* (2023) **A conserved fertilization complex of Izumo1, Spaca6, and Tmem81 mediates sperm-egg interaction in vertebrates** *bioRxiv* https://doi.org/10.1101/2023.07.27.550750

Deneke VE, Pauli A (2021) **The Fertilization Enigma: How Sperm and Egg Fuse** *Annual Review Of Cell And Developmental Biology* **37**:391–414 https://doi.org/10.1146/annurev-cellbio-120219-021751

Ellerman DA, Pei J, Gupta S, Snell WJ, Myles D, Primakoff P (2009) **Izumo is part of a multiprotein family whose members form large complexes on mammalian sperm** *Molecular Reproduction And Development* **76**:1188–1199 https://doi.org/10.1002/mrd.21092

Evans R *et al.* (2021) **Protein complex prediction with AlphaFold-Multimer** *bioRxiv* https://doi.org/10.1101/2021.10.04.463034

Fujihara Y *et al.* (2021) **The conserved fertility factor SPACA4/Bouncer has divergent modes of action in vertebrate fertilization** *Proceedings Of The National Academy Of Sciences Of The United States Of America* **118** https://doi.org/10.1073/pnas.2108777118

Fujihara Y, Lu Y, Noda T, Oji A, Larasati T, Kojima-Kita K, Yu Z, Matzuk RM, Matzuk MM, Ikawa M (2020) **Spermatozoa lacking Fertilization Influencing Membrane Protein (FIMP) fail to fuse with oocytes in mice** *Proceedings Of The National Academy Of Sciences Of The United States Of America* **117**:9393–9400 https://doi.org/10.1073/pnas.1917060117

Greaves S, Marsay KS, Monk PN, Roehl H, Partridge LJ (2022) **Tetraspanin Cd9b plays a role in fertility in zebrafish** *PloS One* **17** https://doi.org/10.1371/journal.pone.0277274

Han L, Nishimura K, Sadat Al Hosseini H, Bianchi E, Wright GJ, Jovine L. (2016) **Divergent evolution of vitamin B9 binding underlies Juno-mediated adhesion of mammalian gametes** *Current Biology* **26**:R100–R101 https://doi.org/10.1016/j.cub.2015.12.034

Holm L (2020) **Using Dali for Protein Structure Comparison** *Methods In Molecular Biology* **2112**:29–42 https://doi.org/10.1007/978-1-0716-0270-6_3

Inoue N, Hagihara Y, Wada I (2021) **Evolutionarily conserved sperm factors, DCST1 and DCST2, are required for gamete fusion** *eLife* **10** https://doi.org/10.7554/eLife.66313

Inoue N, Hamada D, Kamikubo H, Hirata K, Kataoka M, Yamamoto M, Ikawa M, Okabe M, Hagihara Y (2013) **Molecular dissection of IZUMO1, a sperm protein essential for sperm-egg fusion** *Development* **140**:3221–3229 https://doi.org/10.1242/dev.094854

Inoue N, Ikawa M, Isotani A, Okabe M (2005) **The immunoglobulin superfamily protein Izumo is required for sperm to fuse with eggs** *Nature* **434**:234–238 https://doi.org/10.1038/nature03362

Jégou A, Ziyyat A, Barraud-Lange V, Perez E, Wolf JP, Pincet F, Gourier C (2011) **CD9 tetraspanin generates fusion competent sites on the egg membrane for mammalian fertilization** *Proceedings Of The National Academy Of Sciences Of The United States Of America* **108**:10946–10951 https://doi.org/10.1073/pnas.1017400108

Jumper J *et al.* (2021) **Highly accurate protein structure prediction with AlphaFold** *Nature* **596**:583–589 https://doi.org/10.1038/s41586-021-03819-2

Jung M, Wells D, Rusch J, Ahmad S, Marchini J, Myers SR, Conrad DF (2019) **Unified single-cell analysis of testis gene regulation and pathology in five mouse strains** *eLife* **8** https://doi.org/10.7554/eLife.43966

Kaji K, Oda S, Miyazaki S, Kudo A (2002) **Infertility of CD9-deficient mouse eggs is reversed by mouse CD9, human CD9, or mouse CD81; polyadenylated mRNA injection developed for molecular analysis of sperm-egg fusion** *Developmental Biology* **247**:327–334 https://doi.org/10.1006/dbio.2002.0694

Kaji K, Oda S, Shikano T, Ohnuki T, Uematsu Y, Sakagami J, Tada N, Miyazaki S, Kudo A (2000) **The gamete fusion process is defective in eggs of Cd9-deficient mice** *Nature Genetics* **24**:279–282 https://doi.org/10.1038/73502

Kato K, Satouh Y, Nishimasu H, Kurabayashi A, Morita J, Fujihara Y, Oji A, Ishitani R, Ikawa M, Nureki O (2016) **Structural and functional insights into IZUMO1 recognition by JUNO in mammalian fertilization** *Nature Communications* **7** https://doi.org/10.1038/ncomms12198

Krauchunas AR, Marcello MR, Singson A (2016) **The molecular complexity of fertilization: Introducing the concept of a fertilization synapse** *Molecular Reproduction And Development* **83**:376–386 https://doi.org/10.1002/mrd.22634

Kroft TL, Gleason EJ, L'Hernault SW (2005) **The spe-42 gene is required for sperm-egg interactions during C. elegans fertilization and encodes a sperm-specific transmembrane protein** *Developmental Biology* **286**:169–181 https://doi.org/10.1016/j.ydbio.2005.07.020

Lamas-Toranzo I *et al.* (2020) **TMEM95 is a sperm membrane protein essential for mammalian fertilization** *eLife* **9** https://doi.org/10.7554/eLife.53913

Lee S, Kim G, Karin EL, Mirdita M, Park S, Chikhi R, Babaian A, Kryshtafovych A, Steinegger M (2023) **Petascale Homology Search for Structure Prediction** *bioRxiv* https://doi.org/10.1101/2023.07.10.548308

Lorenzetti D, Poirier C, Zhao M, Overbeek PA, Harrison W, Bishop CE (2014) **A transgenic insertion on mouse chromosome 17 inactivates a novel immunoglobulin superfamily gene potentially involved in sperm-egg fusion** *Mammalian Genome* **25**:141–148 https://doi.org/10.1007/s00335-013-9491-x

Meng EC, Pettersen EF, Couch GS, Huang CC, Ferrin TE (2006) **Tools for integrated sequence-structure analysis with UCSF Chimera** *BMC Bioinformatics* **7** https://doi.org/10.1186/1471-2105-7-339

Miller BJ, Georges-Labouesse E, Primakoff P, Myles DG (2000) **Normal fertilization occurs with eggs lacking the integrin α6β1 and is CD9-dependent** *The Journal Of Cell Biology* **149**:1289–1296 https://doi.org/10.1083/jcb.149.6.1289

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M (2022) **ColabFold: making protein folding accessible to all** *Nature Methods* **19**:679–682 https://doi.org/10.1038/s41592-022-01488-1

Miyado K *et al.* (2000) **Requirement of CD9 on the egg plasma membrane for fertilization** *Science* **287**:321–324 https://doi.org/10.1126/science.287.5451.32

NCBI (2022) **NCBI. 2022. TMEM81 orthologs. https://www.ncbi.nlm.nih.gov/gene/388730/ortholog/?scope=7776&term=TMEM 81**

Nishimura K, Han L, Bianchi E, Wright GJ, de Sanctis D, Jovine L. (2016) **The structure of sperm Izumo1 reveals unexpected similarities with Plasmodium invasion proteins** *Current Biology* **26**:R661–R662 https://doi.org/10.1016/j.cub.2016.06.028

Noda T *et al.* (2022) **Sperm membrane proteins DCST1 and DCST2 are required for sperm-egg interaction in mice and fish** *Communications Biology* **5** https://doi.org/10.1038/s42003-022-03289-w

Noda T, Lu Y, Fujihara Y, Oura S, Koyano T, Kobayashi S, Matzuk MM, Ikawa M (2020) **Sperm proteins SOF1, TMEM95, and SPACA6 are required for sperm-oocyte fusion in mice** *Proceedings Of The National Academy Of Sciences Of The United States Of America* **117**:11493–11502 https://doi.org/10.1073/pnas.1922650117

Ohnami N *et al.* (2012) **CD81 and CD9 work independently as extracellular components upon fusion of sperm and oocyte** *Biology Open* **1**:640–647 https://doi.org/10.1242/bio.20121420

Ohto U, Ishida H, Krayukhina E, Uchiyama S, Inoue N, Shimizu T (2016) **Structure of IZUMO1–JUNO reveals sperm–oocyte recognition during mammalian fertilization** *Nature* **534**:566–569 https://doi.org/10.1038/nature18596

Raj I, Al Hosseini H Sadat, Dioguardi E, Nishimura K, Han L, Villa A, de Sanctis D, Jovine L. (2017) **Structural Basis of Egg Coat-Sperm Recognition at Fertilization** *Cell* **169**:1315–1326 https://doi.org/10.1016/j.cell.2017.05.033

Rubinstein E, Ziyyat A, Prenant M, Wrobel E, Wolf J-P, Levy S, Le Naour F, Boucheix C. (2006) **Reduced fertility of female mice lacking CD81** *Developmental Biology* **290**:351–358 https://doi.org/10.1016/j.ydbio.2005.11.031

Uhlén M *et al.* (2015) **Proteomics. Tissue-based map of the human proteome** *Science* **347** https://doi.org/10.1126/science.1260419

Umeda R *et al.* (2020) **Structural insights into tetraspanin CD9 function** *Nature Communications* **11** https://doi.org/10.1038/s41467-020-15459-7

Vance TDR, Yip P, Jiménez E, Li S, Gawol D, Byrnes J, Usón I, Ziyyat A, Lee JE (2022) **SPACA6 ectodomain structure reveals a conserved superfamily of gamete fusion-associated proteins** *Communications Biology* **5** https://doi.org/10.1038/s42003-022-03883-y

van Kempen M, Kim SS, Tumescheit C, Mirdita M, Lee J, Gilchrist CLM, Söding J, Steinegger M. (2023) **Fast and accurate protein structure search with Foldseek** *Nature Biotechnology* :1–4 https://doi.org/10.1038/s41587-023-01773-0

Vicens A, Roldan ERS (2014) **Coevolution of positively selected IZUMO1 and CD9 in rodents: evidence of interaction between gamete fusion proteins?** *Biology Of Reproduction* **90** https://doi.org/10.1095/biolreprod.113.116871

Vondrakova J *et al.* (2022) **MAIA, Fc receptor-like 3, supersedes JUNO as IZUMO1 receptor during human fertilization** *Science Advances* **8** https://doi.org/10.1126/sciadv.abn0047

Wikipedia (2020) **Wikipedia. 2020. TMEM81. https://en.wikipedia.org/w/index.php?title=TMEM81&oldid=954520237**

Wilson KL, Fitch KR, Bafus BT, Wakimoto BT (2006) **Sperm plasma membrane breakdown during Drosophila fertilization requires sneaky, an acrosomal membrane protein** *Development* **133**:4871–4879 https://doi.org/10.1242/dev.02671

Wilson LD, Obakpolor OA, Jones AM, Richie AL, Mieczkowski BD, Fall GT, Hall RW, Rumbley JN, Kroft TL (2018) **The Caenorhabditis elegans spe-49 gene is required for fertilization and encodes a sperm-specific transmembrane protein homologous to SPE-42** *Molecular Reproduction And Development* **85**:563–578 https://doi.org/10.1002/mrd.22992

Wright GJ, Bianchi E (2016) **The challenges involved in elucidating the molecular basis of sperm-egg recognition in mammals and approaches to overcome them** *Cell And Tissue Research* **363**:227–235 https://doi.org/10.1007/s00441-015-2243-3

Yue F *et al.* (2014) **A comparative encyclopedia of DNA elements in the mouse genome** *Nature* **515**:355–364 https://doi.org/10.1038/nature13992

Zhu G-Z, Miller BJ, Boucheix C, Rubinstein E, Liu CC, Hynes RO, Myles DG, Primakoff P (2002) **Residues SFQ (173-175) in the large extracellular loop of CD9 are required for gamete fusion** *Development* **129**:1995–2002 https://doi.org/10.1242/dev.129.8.1995

Article and author information

Arne Elofsson
Science for Life Laboratory and Department of Biochemistry and Biophysics, Stockholm University, Box 1031, 171 21 Solna, Sweden
For correspondence: arne@bioinfo.se
ORCID iD: 0000-0002-7115-9751

Ling Han
Department of Biosciences and Nutrition, Karolinska Institutet, 141 83 Huddinge, Sweden
ORCID iD: 0000-0001-9310-4789

Enrica Bianchi
Department of Biology, Hull York Medical School, York Biomedical Research Institute, University of York, YO10 5DD York, UK
ORCID iD: 0000-0001-8124-7328

Gavin J. Wright
Department of Biology, Hull York Medical School, York Biomedical Research Institute, University of York, YO10 5DD York, UK
ORCID iD: 0000-0003-0537-0863

Luca Jovine
Department of Biosciences and Nutrition, Karolinska Institutet, 141 83 Huddinge, Sweden
For correspondence: luca.jovine@ki.se
ORCID iD: 0000-0002-2679-6946

Copyright

Editors

Reviewing Editor
Jean-Ju Chung
Yale University, New Haven, United States of America

Senior Editor
Wei Yan
The Lundquist Institute, Torrance, United States of America

Reviewer #2 (Public Review):

Summary:

Fertilization is a crucial event in sexual reproduction, but the molecular mechanisms underlying egg-sperm fusion remain elusive. Elofsson A et al. used AlphaFold to explore possible synapse-like assemblies between sperm and egg membrane proteins during fertilization. Using a systematic search of protein-protein interactions, the authors proposed a pentameric complex of three sperm (IZUMO1, SPACA6, and TMEM81) and two egg (JUNO and CD9) proteins, providing a new structural model to be used in future structure-function studies.

Strengths:

(1) The study uses the AlphaFold algorithm to predict higher-order assemblies. This approach could offer insights into a highly transient protein complex, which are challenging to detect experimentally.
(2) The article predicts a pentameric complex between proteins involved in fertilization, shedding light on the architectural aspects of the egg-sperm fusion synapse.

Weaknesses:

The proposed model, which is a prediction from a modeling algorithm, lacks experimental validation of the identity of the components and the predicted contacts.

It is noteworthy that in an independent study, Deneke et al. provides experimental evidence of the interaction between IZUMO1/SPACA6/TMEM81 in zebrafish. This is an important element that supports the findings presented in this manuscript

Regarding the authors response on the question of a global search:
I understand that a global search might be difficult to interpret because a large number of putative false positives. But it is this type of information that is needed to assess the validity of the model and the scoring power in the absence of any experimental validation. At minimum, the search should include a negative control set of proteins known to be unrelated to sperm fertilization or homologous egg-sperm fusion complexes from incompatible species to account for species-specific interactions.

I acknowledge that experimentally validating highly transient complexes presents technical hurdles. However, a high-confidence structural model could enable the design of point mutations specifically disrupting the predicted interactions. Subsequent rescue experiments could then validate the directionality of these interactions. Ultimately, such experiments are crucial for robust model validation.

https://doi.org/10.7554/eLife.93131.2.sa1

Reviewer #3 (Public Review):

Summary:

Sperm-egg fusion is a critical step in successful fertilization. Although several proteins have been identified in mammals that are required for sperm-egg adhesion and fusion, it is still unclear whether there are other proteins involved in this process and how the reported proteins complex co-operate to complete the fusion process. In this study, the authors first identified TMEM81 as a structural homologue of IZUMO1 and SPACA6, and predicted the interactions with a pool of human proteins associated with gamete fusion, using AlphaFold-Multimer, a recent advance in protein complex structure prediction. The prediction is compelling and well discussed, and the experimental evidence to verify this interaction is lacking in this study but supported by a complementary and independent study by another group.

Strengths:

The authors present a pentameric complex formation of four previously reported proteins involved in egg/sperm interaction together with TMEM181 using a deep learning tool, AlphaFold-Multimer.

Weaknesses:

It is intriguing to see that some of the proteins involved in sperm-egg interaction are successfully predicted to be assembled into a single multimeric structure by AlphaFold-Multimer. The experimental validation of the interactions is not directly supported in this study. As there are more candidate proteins in the process, testing other possible protein interactions more comprehensively will provide more rationale for the current 3D multi-protein modeling.

https://doi.org/10.7554/eLife.93131.2.sa0

Author response:

The following is the authors' response to the original reviews.

> ***Reviewer #1***
>
> *The authors should include experiments such as Cryo-EM and genetically modified animals to demonstrate the physiological importance of the TMEM81 complex.*

While we intend to pursue cryo-EM studies of the putative complex (or subcomplexes thereof), this is clearly not a straightforward endeavor and goes beyond the scope of the present manuscript. Concerning the generation of genetically modified animals, we would like to underline that the majority of the proteins that we used for AlphaFold-Multimer complex predictions were precisely chosen based on the fact that - as detailed in the publications referenced in the Introduction - ablation of the respective genes caused sex-specific infertility due to defects in gamete fusion (the other criterion used for inclusion being structural similarity to IZUMO1 coupled with expression in the testis (IZUMO2-4 and TMEM81), or evidence from other kinds of experiments in the case of human-specific MAIA). Concerning TMEM81, experimental evidence for a direct involvement in gamete fusion is described in the referenced preprint by Daneke et al., which was submitted to bioRxiv concomitantly with the present work.

> ***Reviewer #2***
>
> *I believe that the manuscript would benefit from the authors providing more information about the systematic search (Figure 4). For example, by indicating for each pair tested the average pDock score in a 2D plot (or table) and as raw data in the supplementary information.*

Figure 4 has been modified to report both the top and the mean ranking scores for every interaction. Furthermore, additional metrics for the systematic search summarized in Figure 4, including pDockQ scores, are provided in this manuscript revision as supplementary Table S1.

> *A global search, such as including all membrane proteins expressed in eggs or sperm, could not only be more informative but could also allow the reader to understand the pDock score discrimination power for this particular subset.*

The possibility of carrying out a global search was evaluated by performing preliminary computational experiments on an extended ensemble of sperm and egg proteins. In order to do so, we compiled a list of sperm membrane proteins by referring to 4 proteomic datasets (PMIDs 36384108, 36896575, 31824947, 24082039) and identifying ~600 proteins that were found in at least two of them; among these, 250 were single-pass type I or type II membrane proteins, or GPI-anchored proteins. Similarly, a list of 160 egg surface membrane proteins, excluding multipass and secreted ones, was obtained by comparing oocyte cDNA library NIH_MGC_257_N (Express Genomics, USA) with 4 proteomic datasets (PMIDs 35809850, 36042231, 29025019, 27215607). As we briefly commented at the beginning of the section “Prediction of interactions between human proteins associated with gamete fusion” of the revised manuscript, the tests carried out using the resulting list of sperm and egg proteins suggested that interpreting the results of a global search would be severely complicated by a relatively large number of putative false positives. Moreover, the tests showed that performing a complete systematic search would be beyond our current access to computing power. Based on these observations, we preferred to maintain the present study limited to proteins that had been previously clearly implicated in gamete fusion and/or matched specific structural features of IZUMO1.

> *Figure 5 could be improved in clarity by schematically indicating to which cell each protein is anchored.*

This has been done in the revised version of the manuscript.

> ***Reviewer #3***
>
> *Major comments*
>
> *(1) In Figure 1, how the protein of mouse/human IZUMO1 and JUNO is purified is not mentioned in the main text nor in the Methods. Are the mouse IZUMO1-His and mouse JUNO-His transfected together or separately? Are human JUNO-His and human IZUMO1-Myc transfected together into HEK293 cells? And purified by IMAC?*

Transfection information has been included in the Methods section “Protein expression, purification and analysis” (previously “Protein expression and purification”). Concerning the purification procedure, we had already stated in the legend of Figure 1 that human JUNOE-His/IZUMO1E-Myc had been purified by IMAC before SEC, and have now done the same for mouse JUNOE-His and IZUMO1E-His.

> *(2) It would be easier to understand the figure if the author could run a WB to indicate which band above JUNO is specifically IZUMO1-Myc in Figure 1.*

This has been done and reported in a new Figure S1 (with the original Figure S1 having now become Figure S2). Details about the antibodies used for immunoblot have been included in both Methods section "Protein expression, purification and analysis" and the Key Resources Table.

> *(3) Figure 4: Analysis of more proteins that have been suggested as possible candidates for sperm-egg interaction will help to highlight the following results. Also, providing a score for the possibility of interaction might help in selecting those proteins in Figures 5 and 6.*

Please refer to the answer to the first question of Reviewer #2.

> *(4) Figure 7: The authors take advantage of the latest developments in protein structure and interaction to model protein complex formation. However, some experimental experiments such as Co-IP, pull down to support the prediction to verify some of this predicated interaction is necessary.*

We agree with the reviewer; however, for the reasons we discussed during our comparison of the biochemical properties of the JUNO/IZUMO1 interaction between mouse and human, pursuing this line of inquiry will likely necessitate an extensive set of parallel experiments using proteins from different species. This work is being planned and will be the focus of future studies. However, as we mentioned at the end of the Abstract, one should also consider that some of these complexes are likely to be highly transient. Because of this, while they may have important regulated roles in vivo (function at a specific time and place), they could be very challenging to detect using standard approaches in vitro. We thus see this as a significant advance that structural modeling could contribute to the identification of such functionally important but transient interactions.

Minor points

> *(1) In the abstract, "three sperm (IZUMO1, SPACA6 and TMEM81) "should be "three sperm proteins."*

The Abstract has been condensed to fit within the suggested 200-word limit and, as part of this, the sentence has been changed to "complex involving sperm IZUMO1, SPACA6, TMEM81 and egg JUNO, CD9".

> *(2) How do the predictions of the binary complex IZUMO1/CD9 (Figure S1B) or IZUMO1/CD81 (Figure S1C) suggest "the two egg tetraspanins are interchangeable"? Was it because they are quite similar? Please provide more explanation for this speculation. Interchangeable by function or for complex formation? To support the conclusion, biochemical data is required. Otherwise, it needs to be toned down.*

This is because, in the AlphaFold-Multimer predictions of the pentameric complex, CD9 and CD81 are placed in essentially the same way relative to the other subunits.

We have now clarified this at the end of page 6:

"(...) suggest that the two egg tetraspanins are interchangeable because they are predicted to bind to the same region of IZUMO1; (...)"

> *(3) It would be more reader-friendly if the author could label the name of each protein in the figure in Figure S1, especially when the name is not written in the figure legend.*

This has been done in Figure S2 of the revised manuscript (corresponding to original Figure S1).