An avidity-driven mechanism of extracellular BMP regulation by Twisted gastrulation

Gareth Moore[1], Lauren Forbes-Beadle[1], Holly Birchenough[2], Clair Baldock[3]* and Hilary L Ashe[1]*

[1]Faculty of Biology, Medicine and Health, University of Manchester, Manchester M13 9PT, UK.
[2]Biomolecular Analysis Facility, University of Manchester, Manchester M13 9PT, UK
[3]Wellcome Centre for Cell-Matrix Research, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, University of Manchester, Manchester, M13 9PT, UK.
*Authors for correspondence (hilary.ashe@manchester.ac.uk, clair.baldock@manchester.ac.uk)

Abstract

Bone Morphogenetic Protein (BMP) signalling is tightly regulated extracellularly by specific protein-protein interactions. During dorsoventral patterning of vertebrate and invertebrate embryos, the conserved regulator Twisted gastrulation (Tsg) precisely modulates BMP signalling by binding Chordin/Short gastrulation (Sog) to promote formation of the inhibitory Tsg-Sog/Chordin-BMP ternary complex. Here we elucidate the mechanism by which Tsg interacts with Sog/Chordin to modulate BMP signalling extracellularly. Using AlphaFold predictions, we identify a Chordin binding epitope in the Tsg C-terminal domain, which we validate using *in vitro* binding studies with targeted point mutants. Introduction of the equivalent point mutation into *Drosophila* Tsg, to disrupt Tsg-Sog interaction, results in an unexpectedly mild perturbation to embryonic dorsoventral patterning *in vivo*, in the form of a shallower BMP gradient. Using binding assays, we provide a molecular explanation for this mild phenotype by showing that the BMP ligand can partially rescue ternary complex formation when the Tsg-Sog interaction is disrupted. Additionally, we show that an evolutionary divergent Tsg C-terminal extension is essential for full Tsg function in *Drosophila* embryos. Based on these findings we propose that Tsg promotes formation of a Tsg-Sog/Chordin-BMP complex by an avidity-driven mechanism, which will be relevant to a broad range of developmental contexts.

Introduction

The Bone Morphogenetic Proteins (BMPs) are a major family of signalling molecules, belonging to the TGF-β superfamily. BMPs are involved in the development and homeostasis of many organs including the skeletal (Salazar et al., 2016) and cardiovascular (Cai et al., 2012) systems. Preceding this, BMP signalling plays a conserved role in embryonic dorsal-ventral axis patterning in Bilaterians (Mörsdorf et al., 2024).

In the *Drosophila* embryo, a heterodimer of the BMP ligands Decapentaplegic (Dpp) and Screw (Scw) patterns the dorsal ectoderm (Arora et al., 1994; Shimmi et al., 2005; St Johnston and Gelbart, 1987). *dpp* is uniformly expressed in the dorsal ectoderm (Irish and Gelbart, 1987; Jackson and Hoffmann, 1994) and *scw* is ubiquitously expressed (Arora et al., 1994). However, the Dpp-Scw protein distribution is rapidly redistributed forming a BMP activity gradient in dorsal ectoderm, peaking at the dorsal midline where the amnioserosa tissue is specified (Ferguson and Anderson, 1992; Shimmi et al., 2005; Wang and Ferguson, 2005). Distinct thresholds of Dpp/Scw activity have been identified (Ashe et al., 2000) and the gradient is decoded at the cellular level by controlling the rate of promoter activation of BMP target genes (Hoppe et al., 2020).

This BMP activity gradient is established through the concerted action of a set of extracellular regulators. The antagonist *short gastrulation* (*sog*) is expressed in ventrolateral regions and diffuses dorsally where it binds Dpp/Scw, facilitated by collagen IV which acts as a scaffold for complex assembly (Wang et al., 2008). The Sog-Dpp/Scw complex is released from collagen IV upon the binding of Twisted gastrulation (Tsg) and this ternary Tsg-Sog-Dpp/Scw complex diffuses dorsally (Sawala et al., 2012; Wang et al., 2008). In this ternary complex, Dpp/Scw is prevented from binding its cognate BMP receptors by Tsg and Sog. Both Tsg and the Type I receptor directly bind to BMP's 'wrist' epitope (Kirsch et al., 2000; Malinauskas et al., 2024), whereas Sog and the Type II BMP receptor bind to the BMP 'knuckle' epitope (Marqués et al., 1997; Piccolo et al., 1996; Zhang et al., 2007). The Dpp/Scw ligand is released from the ternary complex by the metalloprotease Tolloid (Tld), which cleaves Sog (Marqués et al., 1997; Winstanley et al., 2015). In dorsolateral regions, where Sog levels are higher, the liberated ligand is rebound by Sog and Tsg whereas at the dorsal midline, where Sog levels are lowest, the ligand binds receptors to initiate signalling (Holley et al., 1996; Sawala et al., 2012). Through multiple rounds of binding and release, Dpp/Scw is therefore redistributed to an activity gradient peaking at the dorsal midline.

An alternative source-sink mechanism has been proposed in Zebrafish, whereby BMPs diffuse dorsally towards a Chordin (vertebrate homologue of Sog) sink which is spatially restricted by

Tld (Tuazon et al., 2020; Zinski et al., 2017). Despite differing mechanisms, the presence and action of Chordin/Sog as a BMP antagonist (François et al., 1994; Marqués et al., 1997; Oelgeschläger et al., 2003a) and Tld as a BMP agonist (Ferguson and Anderson, 1992; Piccolo et al., 1997; Shimell et al., 1991; Xie and Fisher, 2005) are conserved between vertebrates and invertebrates. Through shuttling, Sog also acts as BMP agonist to achieve peak signalling levels at long range (Ashe and Levine, 1999). Tsg function is also complicated with both agonistic and antagonistic actions. Tsg acts as an antagonist by increasing Chordin/Sog binding of BMP through the formation of the ternary inhibitory complex (Chang et al., 2001; Larraín et al., 2001; Ross et al., 2001; Shimmi et al., 2005). However, Tsg also increases the rate of Chordin/Sog cleavage by Tld to subsequently release the ligand from inhibitory complexes and therefore acts as an agonist (Larraín et al., 2001; Scott et al., 2001; Shimmi and O'Connor, 2003; Xie and Fisher, 2005).

The recent crystal structure of Tsg reveals two discretely folded domains joined by a short, flexible linker (Malinauskas et al., 2024). Tsg binds BMP ligands via its N-terminal domain (NTD) and abrogation of this interaction *in vivo* ablates BMP gradient formation in the *Drosophila* embryo (Malinauskas et al., 2024; Oelgeschläger et al., 2000). Tsg binds Chordin/Sog via its C-terminal domain (CTD) (Oelgeschläger et al., 2003b). This interaction is mediated through the von Willebrand Factor type C (vWC) domains of the Chordin family of antagonists; the Chordin vWC1, vWC3, and vWC4 and Chordin-like 2 (CHRDL2) vWC1 and vWC3 domains (Fujisawa et al., 2009; Troilo et al., 2016; Zhang et al., 2007). However, the molecular mechanism by which Tsg binds Chordin/Sog and promotes ternary complex formation is so far undetermined.

Here, we use structural predictions with AlphaFold-Multimer (Evans et al., 2022; Jumper et al., 2021) to identify the Chordin family binding site in the Tsg CTD which suggests a common mode of interaction centred around a conserved hydrophobic patch in the Tsg CTD. *In vitro* binding studies with wildtype and mutant Tsg support this proposed mode of interaction, identifying a key leucine residue, Leu103, as important for binding. Introducing the equivalent mutation in *Drosophila* led to mild perturbation of BMP activity gradient formation, due to weaker Tsg-Sog-Dpp/Scw complex formation. Based on these data, we propose that Tsg promotes ternary complex formation through an avidity effect. Finally, phylogenetic analysis uncovers an evolutionary divergent C-terminal sequence, conserved amongst protostome but not deuterostome Tsg sequences. Truncation of this C-terminal extension *in vivo* results in a lethal perturbation of BMP activity gradient establishment, revealing the sequence is essential for Tsg function.

Results

AlphaFold predicts a conserved Tsg-Chordin family binding mode

To identify the potential Chordin family binding surface in the Tsg CTD, we used AlphaFold-Multimer (Evans et al., 2022) to model the human Tsg-Chordin complex. AlphaFold confidently predicts the four vWC domains in addition to the four CHRD domains of Chordin (Troilo et al., 2014) and the two domains of Tsg (Malinauskas et al., 2024), with the CTD of Tsg positioned to interact with the vWC4 domain of Chordin (Fig. 1A). The interface predicted Template Modelling (ipTM) score gives a global model confidence score, weighted towards the interface residues (Evans et al., 2022; Jumper et al., 2021). Overall, the highest confidence model scored an ipTM of only 0.44, although inspection of the predicted aligned error (PAE) plot suggests a higher confidence between the relative positions of the Tsg CTD and Chordin vWC4 (Fig. 1A). Given the high degree of uncertainty in the Chordin model due to long linker regions, Tsg was modelled with Chordin vWC4 alone giving a more confident prediction (ipTM of 0.65) (Fig. 1B). This putative interface is in agreement with binding studies showing Tsg can bind to Chordin vWC4 (Zhang et al., 2007).

Inspection of the Tsg-Chordin vWC4 model shows the Tsg CTD interface is composed of two α-helices (α4 and α5) and the central β-sheet (Malinauskas et al., 2024) (Fig. 1B). This surface faces the Chordin vWC4 β-hairpin and subdomain 1 (SD1), composed of an extended loop and a two-stranded antiparallel β-sheet (Fig.1B). The loop forms an extended surface that cups the face of the Tsg CTD formed by α5 and the β-sheet, while the β2β3-strand of vWC4 sits adjacent and almost parallel to α4 of the Tsg CTD, with the β-hairpin reaching down to the intersection of Tsg CTD α4 and 5 (Fig. 1B and C). This creates a buried surface area of 823 $Å^2$ between the two proteins. To gain insight into the nature of the Tsg-Chordin interaction, the surface hydrophobicity was plotted (Fig. 1D). This shows Chordin vWC4 has a central hydrophobic cavity formed by the β-hairpin (Phe874, Ala875), β1 loop (Pro887, Val889, Pro890, Pro891, Phe892, Met895) and β2β3-sheet (Ile898, Cys900, Val906, Pro907, Cys909). This faces a reciprocal hydrophobic patch on the Tsg CTD formed by α4 (Ile100, Leu103, Leu107), β-sheet (Leu114, Trp116, Phe172, Met176) and α5 (Met188).

To investigate whether Tsg could employ a common mode of interaction with the Chordin family of antagonists, human CHRDL2 and *Drosophila* Sog - Tsg complexes were modelled in AlphaFold Multimer. CHRDL2 has three vWC domains and AlphaFold models the Tsg CTD bound to vWC3 of CHRDL2 with an ipTM score of 0.48 (Fig. S1A). Modelling Tsg-CHRDL2 vWC3 specifically yields an ipTM of 0.83 and this model shows the same interface as for the Tsg-Chordin complex with the β-sheet of CHRDL2 vWC3 adjacent to Tsg CTD α4 and the loop

of vWC3 cupping the Tsg CTD (Fig. S1B). AlphaFold also predicts the same mode of binding for the *Drosophila* Tsg-Sog complex (ipTM 0.5), with *Drosophila* Tsg CTD bound to vWC4 of Sog (ipTM 0.81) via the same interface (Fig. S1.C and D). Furthermore, modelling Tsg in complex with all the individual vWC domains of Chordin, CHRDL2 and Sog supports a common mode of interaction between Tsg and the Chordin family vWC domains. This suggests a conserved interaction between Tsg and the Chordin family of antagonists, from flies to humans.

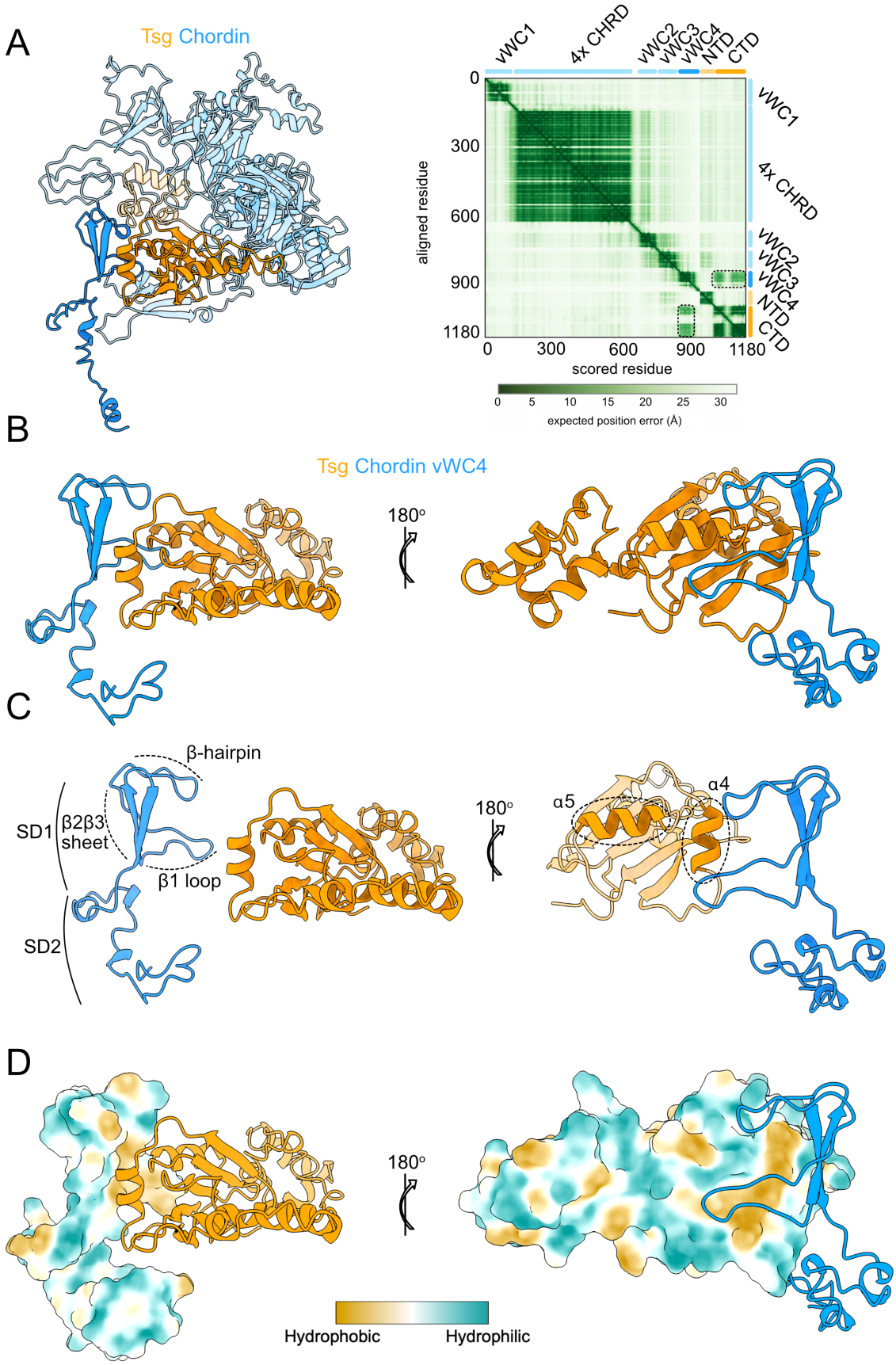

Figure 1. AlphaFold identifies a putative Chordin binding epitope in the Tsg CTD.
(A) AlphaFold model (left) and corresponding PAE plot (right) of the human Tsg (orange) - Chordin (blue) complex with a putative interface between Tsg CTD (dark orange) and Chordin vWC4 (dark blue) as indicated by the dashed regions on the PAE plot. (B) AlphaFold model of the Tsg-Chordin vWC4 complex. (C) Model views shown in (B) but with chains separated for visibility of labelled secondary structure elements. SD1, subdomain 1; SD2, subdomain 2 (D) Tsg-Chordin vWC4 complex as in (B) with surface hydrophobicity plotted.

Conserved leucine residues mediate the Tsg-Chordin interaction

As modelling Tsg in complex with various Chordin family members suggested a common mode of interaction, we next addressed the extent to which the Tsg CTD hydrophobic patch is conserved. Alignment of the Tsg CTD from a range of species showed conservation of the hydrophobic surface centred around two highly conserved leucine residues in α4, as identified by sequence and structural alignment (Fig. 2A and Fig. S2A). As these residues, Leu103 and Leu107 in human Tsg, are highly conserved and placed at the centre of the putative Tsg-Chordin interface they were selected as candidates for mutagenesis (Fig. 2B). Each leucine residue was mutated to alanine and the effect on binding to different human Chordin family members was assessed by SPR. These include ΔN-Chordin, a truncated form of Chordin which is missing the N-terminal vWC1 domain but retains the C-terminal vWC2-4 domains to which Tsg binds more strongly (Troilo et al., 2016; Zhang et al., 2007), and CHRDL2 consisting of 3 vWC domains (Fig. 2C).

Human ΔN-Chordin bound immobilised human Tsg with high affinity (K_D = 14 nM), consistent with previous data (Troilo et al., 2016), but showed no interaction with Tsg^{L103A} and a weaker interaction with Tsg^{L107A} (K_D = 60 nM) (Fig. 2D). Similarly, CHRDL2 bound the wildtype Tsg protein tightly (K_D= 5 nm) but showed no binding to Tsg^{L103A} and a weakened interaction with Tsg^{L107A} (K_D = 11 nM) (Fig. 2E). To confirm this result, we reversed the orientation of the interaction, immobilising CHRDL2 and flowing Tsg, Tsg^{L103A}, or Tsg^{L107A} as the analyte. Again, in this orientation no binding of Tsg^{L103A} to CHRDL2 was observed and the affinity of Tsg^{L107A} to CHRDL2 was weaker relative to wild type Tsg (Fig. S2B). We observed lower affinity interactions in this reversed orientation, consistent with previous binding studies of Tsg and Chordin family members (Fujisawa et al., 2009; Troilo et al., 2016; Zhang et al., 2007). To ensure the disrupted binding was not due to a destabilised Tsg CTD we assessed the thermal stability of wildtype and mutant Tsg by static light scattering (SLS). All three proteins showed remarkable thermal stability with no melting and aggregation detected from 20-95°C (Fig. S2C). This stability could be due to the high number of disulphide bonds, with 12 intramolecular disulphide bonds in the 225 residue protein (Malinauskas et al., 2024). Furthermore, neither mutation disrupted the protein secondary structure composition as assessed by circular dichroism (Fig. S2D and Table S1). Overall, these data support the proposed binding interface identified by the structural model of Tsg-Chordin and show that the conserved leucine residues are important to binding, with Leu103 particularly important in mediating the interaction. Given the consistency of this pattern in the interaction between Tsg and different Chordin family members, it suggests a common mode of binding at the molecular level.

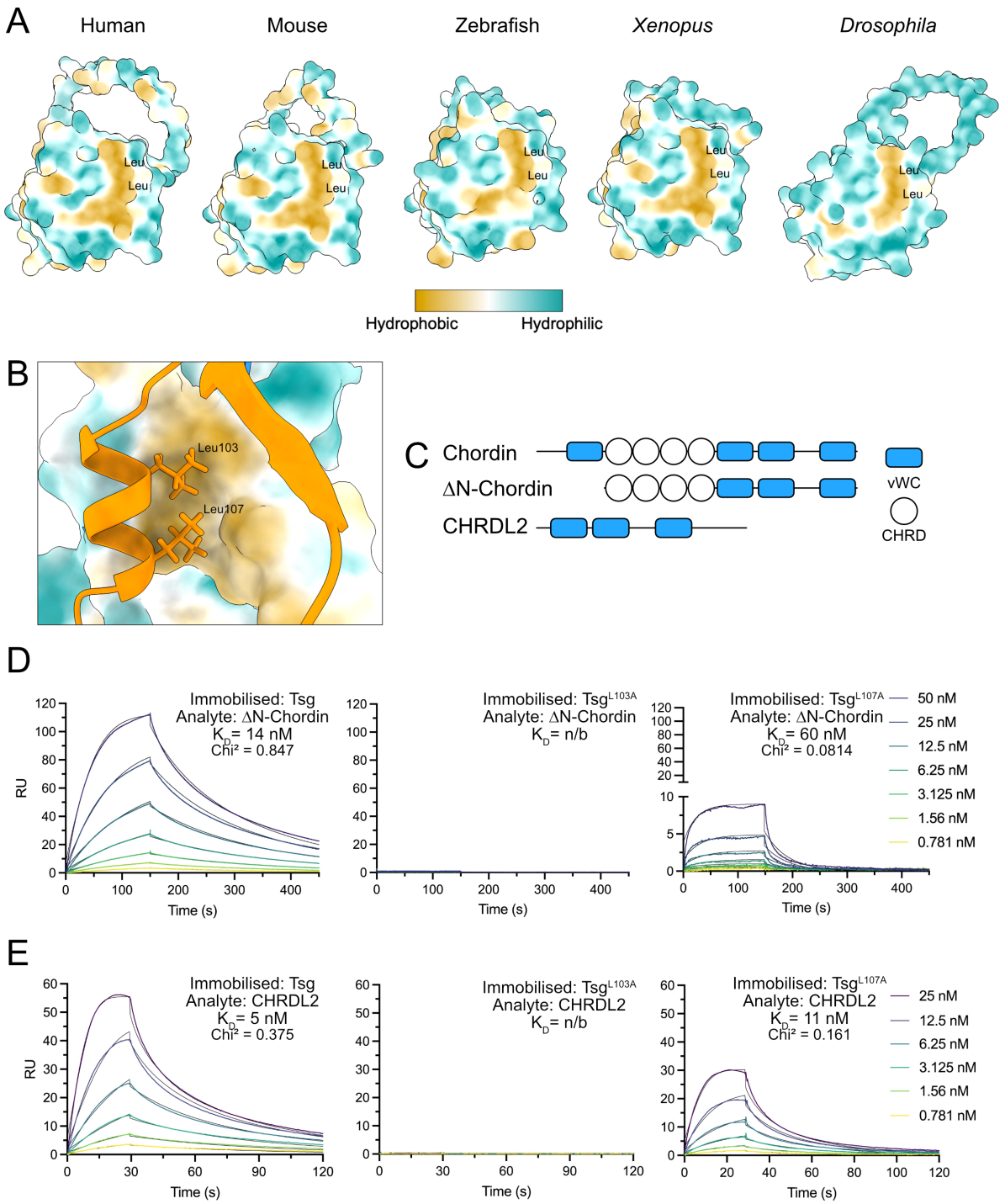

Figure 2. Tsg binds the Chordin family through conserved leucine residues.
(A) AlphaFold models of the Tsg CTD from different species with surface hydrophobicity plotted shows a conserved hydrophobic patch centred around two conserved leucine residues. (B) Close view of the human Tsg CTD α4 helix showing positioning relative to Chordin vWC4. Conserved Leu103 and Leu107 highlighted. (C) Domain organisation of Chordin family members. (D, E) SPR sensorgrams of ΔN-Chordin (D) and CHRDL2 (E) binding to immobilised Tsg, Tsg^{L103A}, or Tsg^{L107A}. Analyte was flowed over at a range of concentrations and binding kinetics calculated by fitting curves to a Langmuir 1:1 binding model. All experiments were performed in triplicate, average curves plotted with model fit in black, no binding (n/b).

Embryonic Dpp/Scw gradient formation is robust to a perturbed Tsg-Sog interaction

Given the high degree of sequence conservation of Tsg (Vilmos et al., 2001), including conservation of the key leucine residues studied above (Fig. S2A), we next tested the impact of the equivalent Leu103Ala mutation on the *Drosophila* Tsg-Sog interaction (Leu100Ala in *Drosophila* Tsg). To test the effect of the Tsg Leu100Ala mutation in vivo, we introduced *tsg^{L100A}:ALFA* sequences at the *Drosophila tsg* locus by ΦC31-mediated transgenesis in a *tsg^{attP}* line (Malinauskas et al., 2024) (Fig. 3A). As the *tsg* gene is located on the X chromosome, *tsg* null males are not viable (Mason et al., 1994). Unexpectedly, the *tsg^{L100A}* mutation did not confer male lethality. As the *tsg^{L100A}* mutation was sublethal *in vivo*, we next investigated the impact on embryonic Dpp-Scw activity gradient formation. The expression domain of the BMP target gene *u-shaped* (*ush*) responds to intermediate levels of BMP activity, thereby providing a sensitive readout of gradient formation (Ashe et al., 2000) (Fig. 3B). Using single molecule fluorescence *in situ* hybridisation (smFISH) we visualised the expression domain of *ush* (Fig. 3C). Null *tsg^{attP}* embryos fail to establish a BMP activity gradient and therefore display an expanded *ush* expression, while reintegration of *tsg:ALFA* sequences rescues gradient formation to establish the wildtype *ush* pattern (Malinauskas et al., 2024) (Fig. 3C). Mutant *tsg^{L100A}:ALFA* embryos display a broader *ush* expression domain than *tsg:ALFA* embryos, however this pattern is not as broad as in null *tsg^{attP}* embryos (Fig. 3C). Quantification of the *ush* expression domain width shows that *tsg^{L100A}:ALFA* embryos have an expression domain of 22-24 nuclei, significantly expanded from the *tsg:ALFA* embryos but not to the extent of full loss-of-function *tsg^{attP}* embryos, where gradient formation is absent (Fig. 3D). This suggests that limited gradient formation is established in the dorsal ectoderm of *tsg^{L100A}:ALFA* embryos, but the shallower gradient is sufficient to permit full development.

To understand how abrogation of the Tsg-Sog interaction could result in a relatively mild phenotype, we next directly tested the impact of the Leu100Ala mutation on the binding of *Drosophila* Tsg to Sog. To test the effect on binding, we performed a co-immunoprecipitation assay by mixing Tsg:ALFA or Tsg^{L100A}:ALFA conditioned media with Sog conditioned media in the presence and absence of Dpp:HA-Scw:Flag conditioned media. Tsg was immobilised on anti-ALFA resin and the amount of Sog and Dpp bound was assessed by Western blot (Fig. 3E). Tsg binds Sog and this interaction is increased in the presence of Dpp-Scw (Ross et al., 2001; Shimmi et al., 2005) (Fig. 3E). Tsg^{L100A} did not bind Sog (Fig. 3E) consistent with the findings for the human proteins (Fig. 2) and further supports the conserved binding mode between Tsg and the Chordin family of antagonists, as suggested by the AlphaFold modelling (Fig. S1C). However, we observed a partial rescue of the Tsg-Sog interaction, albeit weaker

binding than the wildtype, in the presence of Dpp-Scw (Fig. 3E). This suggests that addition of Dpp-Scw to form the ternary complex can partially overcome the disruption to the Tsg-Sog interaction. However, ternary complex formation was weaker with Tsg^{L100A} than wildtype Tsg, as seen by the reduction in Dpp binding (Fig. 3E).

This finding explains the mild perturbation to gradient formation observed in tsg^{L100A} embryos. The presence of Dpp-Scw can partially compensate the disruption to the Tsg-Sog interaction, although the weaker ternary Tsg-Sog-BMP complex formation leads to a shallower BMP gradient and partial expansion of the *ush* expression domain (Fig. 3C). Overall, these results suggest that BMP activity gradient establishment is robust to perturbation of the Tsg-Sog interaction, likely due to the synergistic action of all three components of the shuttling complex *in vivo* (Ross et al., 2001; Shimmi et al., 2005).

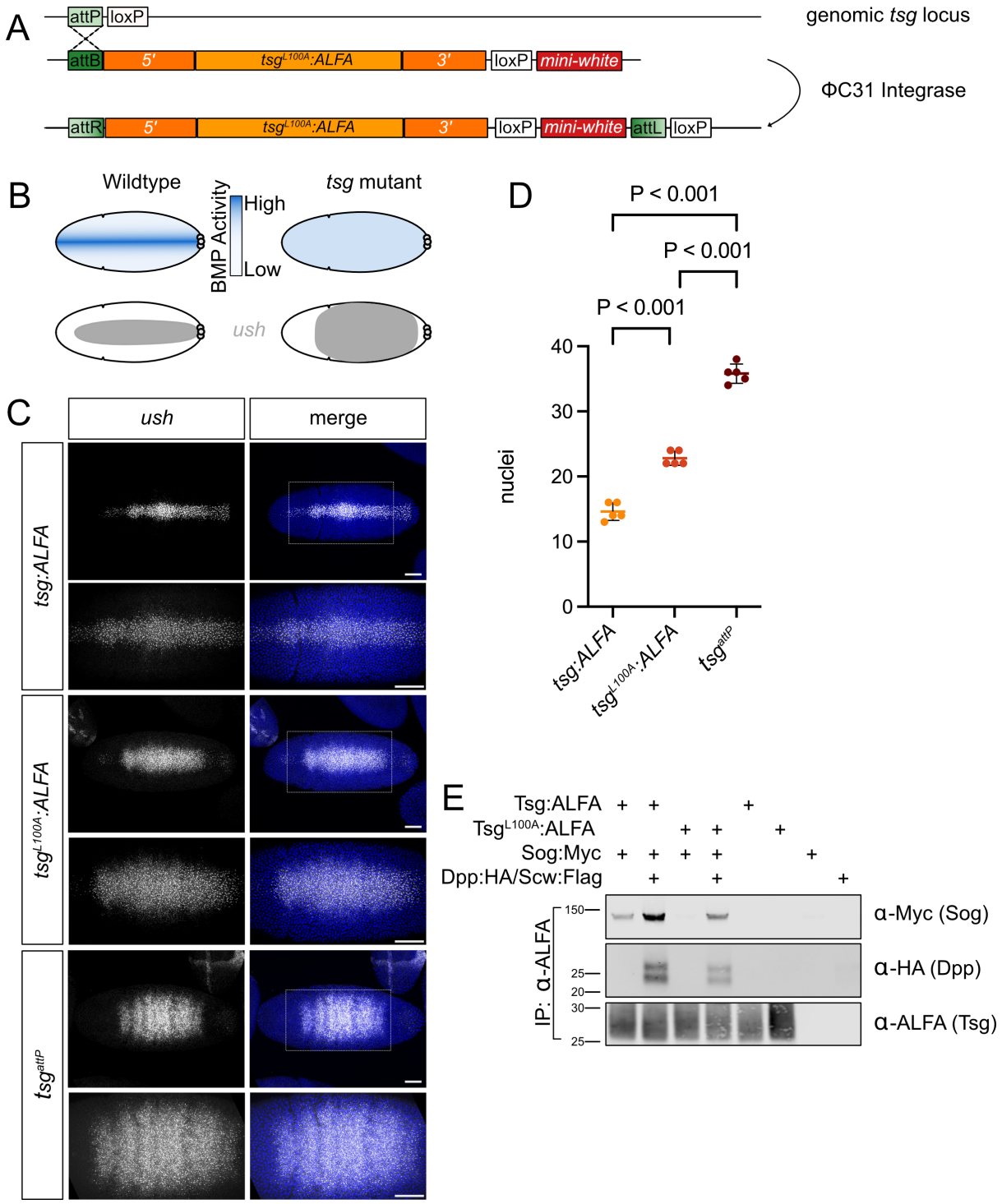

Figure 3. Disrupting the Tsg-Sog interaction perturbs Dpp/Scw gradient formation.
(A) Schematic of strategy to introduce *tsg*^{L100A} sequences at the *tsg* locus in *tsg*^{attP} flies. (B) Cartoon of *Drosophila* embryo showing (top) the BMP activity gradient and (bottom) *ush* expression domain in wild type and mutant *tsg* embryos. (C) smFISH of the *ush* expression domain at 20x (top) and 40x (bottom) in *tsg:ALFA* or *tsg* mutant embryos (dorsal view). Nuclei are stained with DAPI. Scale bar = 50 μm. (D) Quantification of the *ush* expression domain width in nuclei. One-way ANOVA with Tukey's multiple comparisons test. n=5, error bars represent standard deviation. (E) Western blot showing the amount of Sog and Dpp co-immunoprecipitated with Tsg:ALFA, immobilised on anti-ALFA resin. Protein combinations are as shown in the key.

Protostomes conserve a C-terminal Tsg tail sequence

In a previous study, comparison of Tsg sequences from different *Drosophila* species with vertebrate species identified a C-terminal extension of 40 residues in the *Drosophila* sequences relative to the vertebrate orthologues, termed a tail Vilmos *et. al* (2001). As the function of this C-terminal Tsg tail is unknown, we first studied its conservation. To this end, a reference proteome was queried for *tsg* sequences yielding a list of 60 sequences across a broad taxonomic diversity. These sequences were then aligned to reveal a dichotomy in the C-termini of different *tsg* sequences; 32/60 species had a consistent C-terminal end two residues after the final conserved cysteine, whereas the remaining 28/60 species showed some degree of C-terminal extension (Fig. 4).

Phylogenetic analysis revealed the presence or absence of a tail sequence traced back some 600 million years ago at the protostome-deuterostome split. All 28 species with an extension were protostomes and all 32 species lacking Tsg tail sequences were deuterostomes. While clearly present at the C-terminus of protostome *tsg* sequences, there is little-to-no conservation at the sequence level or indeed the length of the extension. Furthermore, AlphaFold predicts that the *Drosophila* Tsg tail lacks any significant secondary structure, and this prediction was supported by other *in silico* tools that score the Tsg tail region as likely to be an intrinsically disordered region (IDR) (Fig. S3).

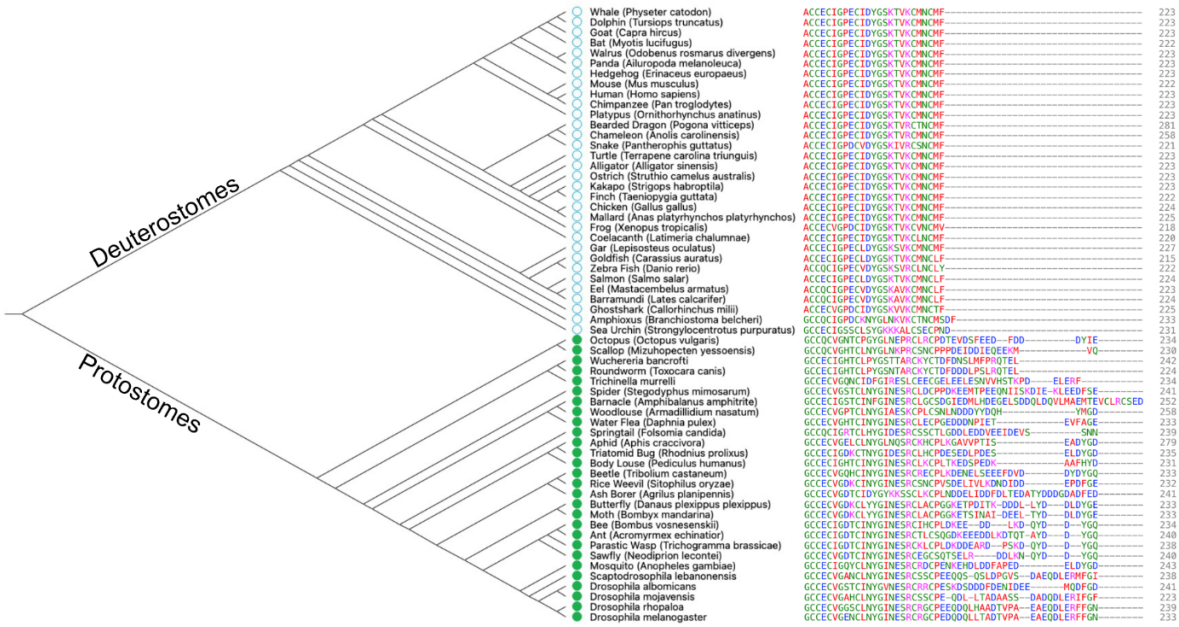

Figure 4. Protostomes but not deuterostomes have a conserved C-terminal Tsg tail sequence.
Straight cladogram and accompanying sequence alignment showing the evolutionary divergence of the Tsg tail, with deuterostomes lacking (blue circle) and protostomes conserving (green dot) a Tsg tail. Cladogram rooted and not to scale.

The Tsg tail is essential to BMP activity gradient formation

As the Tsg tail is not universally conserved and shows both length and sequence diversity, we next asked how important the tail is to Tsg function. We reintegrated truncated *Drosophila tsg* sequences (Met1-Glu208), which removes the tail sequence (Asp209-Ser249) two residues after the final cysteine as in deuterostome sequences, at the *tsg* locus in tsg^{attP} flies (Fig. 5A). Reintegration of $tsg^{\Delta tail}$*:ALFA* did not rescue function and resulted in male lethality. This indicates that the tail performs a specific function essential to full Tsg action and viability of the animal. To determine whether the tail plays a role during BMP gradient formation we assessed the *ush* expression domain by smFISH (Fig. 5B). Embryos containing the truncated *tsg* sequences show a broad expansion of the *ush* expression domain as in null tsg^{attP} embryos. This demonstrates that the Tsg tail contributes to Tsg function during BMP gradient formation (Fig. 5B). The lack of a BMP gradient explains the loss of viability in $tsg^{\Delta tail}$*:ALFA* males. Together, these results show that despite a lack of sequence or structural conservation, the protostome-specific Tsg tail is necessary for the full action of the protein during dorsoventral patterning in *Drosophila* embryogenesis.

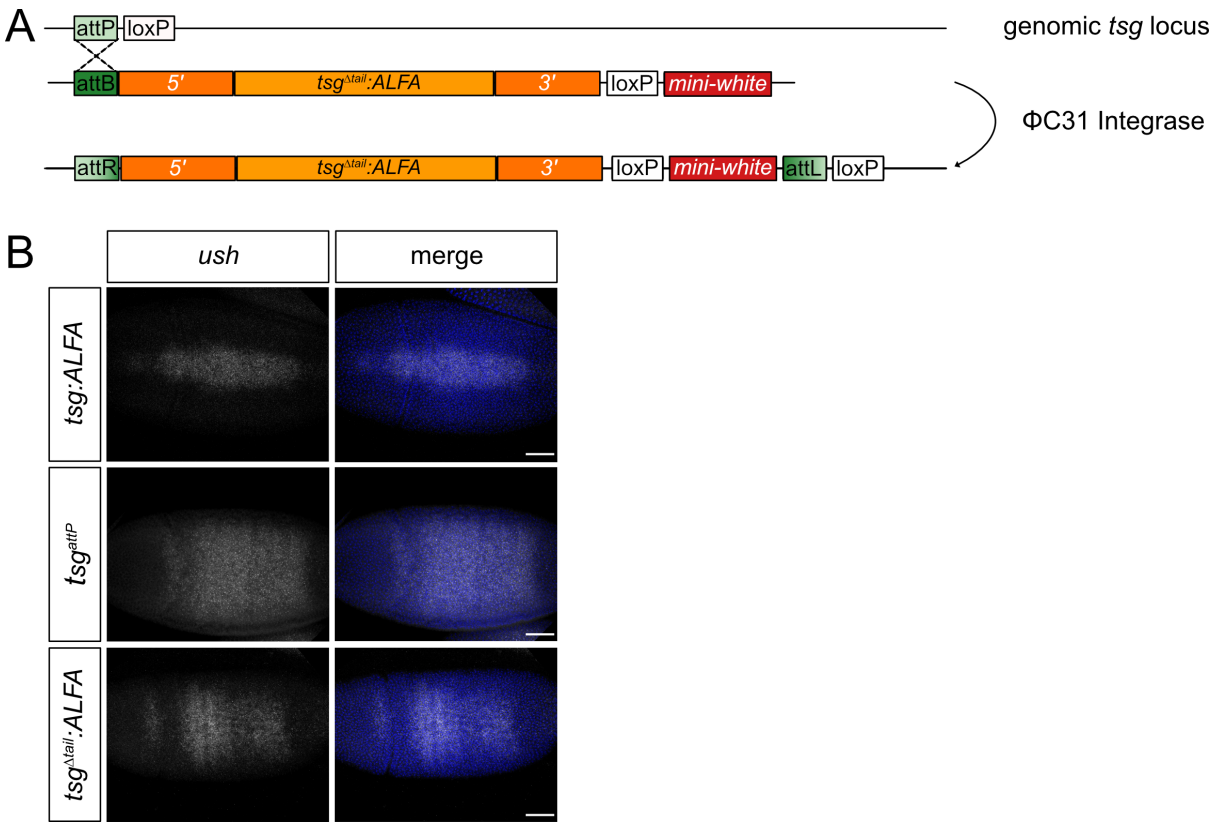

Figure 5. The C-terminal Tsg tail sequence is essential to Tsg function.
(A) Strategy to reintegrate truncated *tsg* sequences, lacking the Tsg tail, at the *tsg* locus in tsg^{attP} flies. (B) smFISH of *ush* expression domain in *tsg:ALFA*, tsg^{attP}, and $tsg^{\Delta tail}$*:ALFA* embryos, dorsal view with nuclei stained with DAPI. Scale bars = 50 μm.

Discussion

In this study we used *in silico* structural modelling with AlphaFold to identify a binding site in Tsg for Chordin family members, which is conserved from flies to humans. Modelling proposed a conserved binding mode between the Tsg CTD and Chordin family vWC domains, consistent with previous binding studies showing that Tsg binds the Chordin family via its CTD (Malinauskas et al., 2024; Oelgeschläger et al., 2003b). The newly identified interface has a hydrophobic core centred around a pair of conserved leucine residues, Leu103 and Leu107 in human Tsg. The reciprocal hydrophobic surface in the Chordin vWC domain is formed by an extended loop and β-sheet that comprises SD1 of the vWC domain. Mutation to residues in this putative interface in CHRDL2 vWC domains reduced their affinity for Tsg (Fujisawa et al., 2009), consistent with the proposed binding interface identified here.

Chordin family vWC domains have a long insertion in the SD1 loop region which has been proposed to be the major determinant of binding BMPs and Tsg (Fujisawa et al., 2009; Xu et al., 2017). This prediction is consistent with binding studies of CHRDL2 vWC SD1 mutants (Fujisawa et al., 2009) and the AlphaFold models of human Tsg-Chordin, human Tsg-CHRDL2, and *Drosophila* Tsg-Sog shown here. Analysis of CHRDL2's three vWC domains revealed that vWC1 and vWC3, which contain an extended loop region in SD1, bind to both Tsg and BMP (Fujisawa et al., 2009; Zhang et al., 2007). In contrast, vWC2 of CHRDL2 has a short linker between the β-hairpin and β-sheet and does not bind Tsg or BMP (Fujisawa et al., 2009; Zhang et al., 2007). Furthermore, Chordin and Sog have four vWC domains and the same pattern was observed, with vWC2 lacking the extended loop while the remaining vWC 1, 3 and 4 domains have an extended loop and show binding to Tsg and BMP (Troilo et al., 2014; Zhang et al., 2007). This supports the proposal that the SD1 loop mediates binding of Chordin family vWC domains for a general model of binding to Tsg. Domain specific sequence differences in this SD1 loop then could explain the differential binding ability of Chordin family vWC domains. While AlphaFold modelling can be used to generate hypotheses of binding modes, is does not necessarily provide accurate positioning of amino acid side chains, which is of particular importance when studying binding interfaces in protein-protein interactions (Terwilliger et al., 2023). Therefore, it is important to experimentally validate *in silico* predictions, which we have done here with solid state and in solution binding assays, to test the AlphaFold models of the Tsg-Chordin family interaction.

Phylogenetic analysis here uncovered an evolutionary divergence at the Tsg C-terminus. A previously identified Tsg tail sequence (Vilmos et al., 2001) was found to be protostome-specific and shown to be required full Tsg action in patterning the embryonic dorsal ectoderm.

In silico characterisation suggests the tail to be an intrinsically disordered region (IDR). Some IDRs adopt secondary structure in the presence of a binding partner suggesting a possible role of the tail in binding another extracellular factor. Alternatively, the tail may be needed for movement of Tsg, which is highly diffusive(Mason et al., 1997). Future studies of Tsg function and fate with and without the tail sequence are necessary to uncover the action of the Tsg tail.

Proteins can display cooperative binding, where interaction of one protein promotes subsequent interaction. This is true of Tsg, which increases Chordin/Sog interaction with BMP ligands by promoting formation of ternary Tsg-Sog/Chordin-BMP complexes (Chang et al., 2001; Larraín et al., 2001; Ross et al., 2001; Scott et al., 2001). Here, the presence of Dpp-Scw partially recovered the perturbed Tsg^{L100A}-Sog interaction via ternary complex formation, albeit less potently than with wild type Tsg. Based on this finding, we propose the simplest hypothesis for Tsg promoting ternary complex formation is through an avidity mechanism: where binding of one epitope makes a second interaction more likely, for example by bringing a second epitope in closer proximity for interaction. Chordin family vWC domains bind BMP ligands at the type II receptor knuckle epitope (Zhang et al., 2007) so, despite containing four vWC domains, only two vWC domains of Sog/Chordin can bind one BMP dimer at once. Tsg binds the BMP ligand at the wrist (type I receptor) epitope through its N-terminal domain (NTD) (Malinauskas et al., 2024; Zhang et al., 2007). By simultaneously binding the BMP wrist epitope, through its NTD, and Sog/Chordin, through its CTD, Tsg provides an additional link between Sog/Chordin and the BMP ligand. Interaction at one binding site could then aid subsequent interaction by bringing the remaining epitopes closer, thus promoting interaction via an avidity mechanism. Such a mechanism of Tsg action is supported by the observation that the linker region is of conserved length, suggesting the spacing of the NTD and CTD is key to function (Malinauskas et al., 2024).

Avidity of Tsg-Sog/Chordin binding the BMP ligand would mirror BMP receptor complex formation, which is thought to assemble through avidity-driven complexing (Groppe et al., 2008). In the context of BMP gradient formation in the early *Drosophila* embryo, such avidity of Tsg-Sog/Chordin binding BMP would be important. Quantitative models of Dpp-Scw shuttling require Sog to the bind the ligand in the low pM range (Eldar et al., 2002; Mizutani et al., 2005; Umulis et al., 2006), while binding studies with the vertebrate homologues suggest the interaction is in low nM range (Zhang et al., 2007). Both Sog and Dpp bind the *Drosophila* collagen IV proteins Viking and Dcg1 (Wang et al., 2008) and modelling demonstrates that this could increase the affinity of Sog for Dpp-Scw by 50- to 100 fold (Umulis et al., 2009). Tsg releases Sog and Dpp from collagen IV (Wang et al., 2008); however it is that possible that,

prior to release, the avidity-driven effects of Tsg are needed to achieve the ~1000 fold increase in affinity of Sog for Dpp-Scw to achieve shuttling *in vivo*.

The partial rescue of the Tsg^{L100A}-Sog interaction by Dpp-Scw through avidity, when all three components are present, provides an explanation for the viability of the tsg^{L100A} line. In contrast, abrogation of the Tsg-Dpp-Scw interaction leads to a complete loss of ternary complex formation and gradient establishment in the *Drosophila* embryo (Malinauskas et al., 2024), suggesting the Tsg-BMP interaction is the key interaction in driving complex formation. This is consistent with a proposed mechanism of ternary complex assembly where Dpp-Scw and Sog are immobilised on collagen IV, and Tsg binds both proteins to release Tsg-Sog-Dpp-Scw complexes from the extracellular matrix (Sawala et al., 2012). Furthermore, increasing the affinity of each component through cooperative binding could confer robustness to the system, overcoming fluctuations in the level of an individual component by promoting its binding through the other two components in the ternary complex. BMP gradient formation is robust to gene dosage of several key components, including *tsg* (Eldar et al., 2002; Mason et al., 1997), and different aspects of the system are proposed to mediate this. For example, the use of Dpp-Scw heterodimers as the most potent ligand aids in buffering against fluctuations in individual ligand levels (Shimmi et al., 2005). The ability of ternary complex formation to partially overcome disruption of the Tsg-Sog interaction provides an explanation for the robustness of the system to fluctuation in Tsg levels.

In conclusion, we identify the Chordin family binding site in Tsg and combine *in vitro* binding studies and *in vivo* genetics to propose a mechanism by which Tsg promotes the formation of ternary Tsg-Sog/Chordin-BMP complexes. This furthers our understanding of this key regulatory complex across vertebrate and invertebrate development.

Methods

AlphaFold Multimer

To perform *in silico* protein structure prediction, AlphaFold-Multimer (V2.1.1) (Evans et al., 2022; Jumper et al., 2021) was implemented on the High-Performance Computing (HPC) cluster through the Computational Shared Facility 3 (CSF3), University of Manchester. Jobs were run using 1 Graphics Processing Units (GPU) with 8 cores. The model was trained on Protein Data Bank (PDB) files deposited before 01/11/2021 and used the ‘full genetic database’ for multiple sequence alignment (MSA) configuration. Each prediction was performed with the default number of recycles to generate five relaxed predictions which were

ranked based on predicted template alignment (pTM) for monomer of the pTM and interface predicted template alignment (ipTM) score defined as: $0.2 * pTM + 0.8 * ipTM$. Predicted alignment error (PAE) plots were generated for each model through a custom python script implemented on the CSF3. All five model outputs were analysed, and the highest confidence prediction presented.

Cloning, expression and purification of human CHRDL2 and wild type and mutant human Tsg

A synthetic sequence of human Twisted gastrulation (residues Cys26-Phe223) codon optimised for mammalian expression (ThermoFisher Scientific) with a C-terminal TwinStrep-tag (IBA LifeScience) joined by a 3xGlyGlyGlySer linker was inserted into the pHLsec vector (Addgene_99845) between AgeI and KpnI restriction sites. Site directed mutagenesis to introduce single point mutations was performed by PCR and mutated PCR products annealed using In-Fusion HD cloning (Takara Bioscience). The human CHRDL2 sequence (Arg24-Thr429) with a C-terminal TwinStrep-tag (IBA LifeScience) joined by a 3xGlyGlyGlySer linker was also inserted into the pHLsec vector (Addgene_99845) between AgeI and KpnI restriction sites. pHLsec-Tsg-TwinStrep and pHLsec-CHRDL2-TwinStrep plasmids were transfected to Expi293-F cells (ThermoFisher Scientific) using the ExpiFectamine 293 Transfection Kit (ThermoFisher Scientific). Cells were cultured for 4 days before conditioned media was collected, centrifuged at 4500 x g at 4°C for 15 mins and filtered using 11 μm filter paper (Whatman) and then 0.65 μM PVDF filter membrane (Millipore). Recombinantly expressed protein was purified by affinity chromatography using Strep-Tactin XT 4Flow High-Capacity resin (IBA LifeSciences) and protein eluted in elution buffer (20 mM HEPES, pH 7.4, 500 mM NaCl, 75 mM Biotin). Eluted protein was further purified by size exclusion chromatography (SEC) using a Superdex 200 Increase GL column (Cytiva) on an Äkta Purifier FPLC (Cytiva) with SEC buffer (20 mM HEPES, 150 mM NaCl, pH 7.4). Protein was resolved at 0.5 ml/min, at room temperature. Protein elution from the column was monitored using UV absorbance at 280 nm and eluted in 0.5 ml fractions. Fractions of interest were assessed for protein content by SDS-PAGE and stored at -80°C. Purified, recombinantly expressed ΔN-Chordin was prepared as described previously (Troilo et al., 2014).

Surface Plasmon Resonance

To determine binding kinetics and affinities, surface plasmon resonance (SPR) was performed on a Biacore T200 (Cytiva). Ligand was diluted to a concentration of 30 μg/ml in sodium acetate buffer pH 5.5 and immobilised onto CM5 Series S Sensor Chips (Cytiva) via amine coupling using NHS/EDC cross linking, with exposed activated amines subsequently blocked

with ethanolamine. Assays were performed in running buffer (20 mM HEPES, 150 mM NaCl, 0.05% Tween20, pH 7.4) at 25°C, and all analytes were diluted in running buffer. The surface was regenerated between cycles with 0.1 M glycine pH 2 followed by 1 M NaCl pH 7.4, each at 30 μl/min for 30 secs. For kinetic assays, analytes were diluted in a two-fold concentration series and flowed over immobilised ligand at 30 μl/min for 200 secs, followed by a 300 sec dissociation phase. Binding parameters were calculated using non-linear fitting of a Langmuir 1:1 binding model performed using the Biacore T200 Evaluation Software (V2). Purified, recombinantly expressed ΔN-Chordin for binding studies was provided by Dr Rana Dajani (University of Manchester).

Circular dichroism

To determine secondary structure composition of wild type and mutant human Tsg, circular dichroism (CD) was performed on a J810 spectrophotometer (JASCO). To minimise background signal from buffer, all proteins were buffer exchanged to 5 mM Tris, 150 mM NaCl pH 7.4 buffer, using a 5 ml HiTrap Desalting Sephafex G-25 column (Cytiva). Following buffer exchange, protein concentration was determined by absorbance at 280 nm on a Nanodrop-1000 spectrophotomer (ThermoFisher Scientific) and adjusted to 0.3-0.4 mg/ml. Spectra from 260-190 nm range were collected in 0.2 nm steps, with a 0.5 sec exposure time to a total of 10 accumulations which were averaged. The same measurements were taken with a blank buffer sample which was subtracted from experimental readings. Data analysis was performed with the online DichroWeb suite using the CDSSTR ((Manavalan and Johnson, 1987; Sreerama and Woody, 2000) , Contin (van Stokkum et al., 1990), K2D (Andrade et al., 1993) and Selcon3 (Sreerama et al., 1999; Sreerama and Woody, 1993) algorithms with the reference set SP175t (Lees et al., 2006).

Static Light Scattering

Protein thermal stability was assessed by static light scattering using an Uncle (Unchained Labs). Samples were diluted to 200 μg/ml in 20 mM HEPES, 150 mM NaCl, pH 7.4. Sample aggregation temperature (T_{agg}) was assessed by static light scattering recorded at 473 nm through a temperature ramp from 20-95°C, in 1°C steps per min. All measurements were recorded in triplicate.

S2 cell co-immunoprecipitation assay

S2R+ cells were cultured in Schneider's Drosophila media (Gibco) supplemented with 10% v/v FBS, 1% v/v penicillin/streptomycin at 25°C. Cells were transfected using Effectene (Qiagen) in 12-well plates. A pMT-BiP-Tsg:ALFA plasmid was generated by PCR amplification

of the Tsg coding sequence from pMT-BiP-Tsg:His (Winstanley et al., 2015) to include C-terminal ALFA tag (Götzke et al., 2019) and inserted in pMT-BiP digested with EcoRI and AgeI. A single Leu100Ala point mutation was introduced by In-Fusion cloning (Takara) to generate a pMT-BiP-Tsg^{L100A}:ALFA plasmid. Sog, Tsg, or Tsg^{L100A} conditioned media was produced by transfecting cells with 1 μg of pMT-Sog:Myc (Sawala et al., 2012), pMT-BiP-Tsg:ALFA, or pMT-BiP-Tsg^{L100A}:ALFA. Dpp-Scw conditioned media was generated as in Sawala et al., (2012) by transfecting cells in a T75 flask (Corning) with 10 μg pMT-Dpp:HA and 4 μg pMT-Scw:Flag. Protein expression was induced 24 hours after transfection by supplementing media with 0.5 mM $CuSO_4$. Conditioned media was collected a further 72 hours after induction and centrifuged at 5000 x g, 15 mins at room temperature. Conditioned media was then mixed and incubated at room temperature for 3 hours before binding anti-ALFA Selector PE resin (NanoTag Biotechnologies) at 18°C for 1 hour. Matrix was collected and washed four times with buffer (20 mM HEPES, pH 7.4, 150 mM NaCl) and samples eluted by boiling in NuPAGE LDS Sample Buffer (ThernoFisher Scientific) + 20% β-mercaptoethanol. Samples were analysed by Western blot using Mouse anti-Myc (1:1000, Millipore, RRID: AB_11211891), Rabbit anti-ALFA (1:1000, NanoTag Biotechnologies, RRID:AB_3075998) and Mouse anti-HA (1:1000, Roche, RRID: AB_514505) primary antibodies, with IRDye® 680RD donkey anti-Rabbit IgG (1:10,000, LI-COR, RRID: AB_2716687) and IRDye 800CW donkey anti-Mouse IgG (1:10,000, LI-COR, RRID: AB_621847) secondary antibodies. Western blots were visualised on a LI-COR Odyssey CLx.

Drosophila stocks and maintenance

Fly stocks were maintained on standard fly food media (yeast 50 g/l, glucose 78 g/l, maize 72 g/l, agar 8 g/l, 10% nipagen in EtOH 27 ml/l and propionic acid 3 ml/l). *D. melanogaster* stocks used in this work were: *tsg^{attP}/FM7c-ftz-lacZ*; *tsg:ALFA* (Malinauskas et al., 2024); *tsg^{L100A}:ALFA* and *tsg$^{\Delta tail}$:ALFA/FM7c-ftz-lacZ* (see below). The *tsg^{attP}* and *tsg$^{\Delta tail}$:ALFA* stocks were balanced with *FM7c-ftz-lacZ* allowing visualisation of mutant embryos based on the absence of *lacZ* staining.

Generation of mutant *tsg Drosophila* stocks

Both the *tsg^{L100A}:ALFA* *tsg$^{\Delta tail}$:ALFA* fly stocks were generated using ΦC31-mediated transgenesis in a *tsg^{attP}* stock, as outlined in Malinauskas et al., (2024). Briefly, the *tsg^{attP}* stock replaces the *tsg* locus with an attP landing site. Reintegration using a RIV^{white}-Tsg:ALFA plasmid repairs the *tsg* locus to the wildtype sequence other than the addition of an inert C-terminal ALFA tag to the *tsg* coding sequence (Malinauskas et al., 2024). To generate the point mutant reintegration vector, Leu100 was mutated to Ala in the RIV^{white}-Tsg:ALFA plasmid using

In-Fusion cloning. To generate the truncated *tsg* sequence, the tail sequence (Asp209-Ser249) was excised from RIV^{white}-Tsg:ALFA, leaving a truncated coding sequence (Met1-Glu208). Reintegration plasmids were co-injected with a ΦC31 encoding plasmid into *tsg^{attP}* embryos and survivors crossed back to the balanced *tsg^{attP}/FM7c-ftz-lacZ* stock. Successful recombinants were screened for using the *mini-white* marker and confirmed by sequencing genomic DNA. Male *tsg$^{\Delta tail}$:ALFA* flies were not viable and so the line remained balanced over the *FM7c-ftz-lacZ* chromosome. Male *tsg^{L100A}:ALFA* flies were viable and so the balancer chromosome was crossed out, to generate a homozygous *tsg^{L100A}:ALFA* stock.

smFISH and imaging of *Drosophila* embryos

Drosophila embryos (2-4 hours) were fixed as previously described (Hoppe et al., 2020) and stored in methanol at -20°C long term. For smFISH, embryos were stained with *ush* Stellaris and *lacZ* Stellaris probes (Frampton et al., 2022) and nuclei were stained with DAPI (1:1000, NEB 4083). Samples were mounted in ProLong Diamond antifade mountant (Thermo Fisher) and set overnight before long term storage at -20°C.

Embryos shown in Fig. 3 were imaged on an Andor Dragonfly200 spinning disk inverted confocal microscope with both a 20x/0.75 MImm Plan Fluor and 40x/1.3 Super Fluor objective. Samples were excited using 405 nm (5%, 100ms exposure), 561 nm (10%, 100ms exposure) and 637 nm (10%, 150ms exposure) diode lasers and collected through blue (450nm), red (600nm) or far red (700nm) filters. Each channel was collected sequentially with 4x averaging on the iXon EMCCD (1024 X 1024) camera and multiple Z slices were obtained at system optimised spacing. Images were deconvolved using inbuilt Andor deconvolution software.

Embryos shown in Fig. 5 were imaged on a Leica TCS SP8 AOBS confocal microscope with a HC PL APO CS2 40x/1.30 oil objective. Images were collected at 0.75 zoom, pinhole 1 airy unit, scan speed 600 Hz bidirectional, 2048×2048 format, at 8-bit, 4x Line Averaging, 0.35 µm Z step size, using the white light laser with 405 nm (5%), 548 nm (20%) and 647 nm (20%) and hybrid detectors. Raw images were deconvolved with Huygens Remote Manager software v 3.7.1 (SVI). *tsg$^{\Delta tail}$:ALFA* hemizygous males were identified by a lack of a *lacZ* signal in the 548 nm channel.

The *ush* expression domain width was quantified in deconvolved images in Fiji (Schindelin et al., 2012). Maximum intensity projections of imaging stacks were made, and nuclei counted at 50% embryo length across the *ush* domain in a line perpendicular to the anterior-posterior axis. Statistical analysis was performed in GraphPad Prism 10 (v.10.2.2).

Phylogenetic analysis of *tsg* sequences

To determine the evolutionary conservation of the *Drosophila tsg* C-terminal extension, a reference proteome was used. This reference proteome is curated by UniProt to landmark proteome space and has been selected manually and algorithmically to provide a representative overview of taxonomic diversity. The text search function was used on UniProt, queried with 'Twisted Gastrulation and Reference Proteome'. This returned 74 entries which was manually curated to remove closely related species and poorly annotated sequences, yielding a list of 60 species sequences. This condensed list was submitted to the NCBI Taxonomy Tree browser (Schoch et al., 2020) and the results exported as a .phylip file which was then visualised and annotated in Mega11 (Stecher et al., 2020; Tamura et al., 2021).

Acknowledgments

We thank Catherine Sutcliffe, Sanjai Patel, and the Manchester Fly Facility for assistance in generating the *tsg* mutant fly lines, Dr Thomas Jowitt (University of Manchester Biomolecular Analysis Facility) for advice on the circular dichroism and SLS, Dr Rana Dajani for providing ΔN-chordin protein, and the University of Manchester Bioimaging facility. Work was supported by the Biotechnology and Biological Sciences Research Council (BB/ V008099/1 to H.L.A. and C.B). G.M. was supported by a BBSRC DTP PhD studentship (BB/M011208/1). The Wellcome Centre for Cell-Matrix Research is funded by Wellcome Trust core award 203128/Z/16/Z.

Author contributions

Conceptualisation, G.M., C.B. and H.L.A.; investigation, G.M., L.F.B., H.B.; writing, G.M., C.B., and H.L.A.; supervision, C.B. and H.L.A.; funding acquisition, C.B. and H.L.A.

Competing interests

The authors declare no competing interests.

Supplementary information

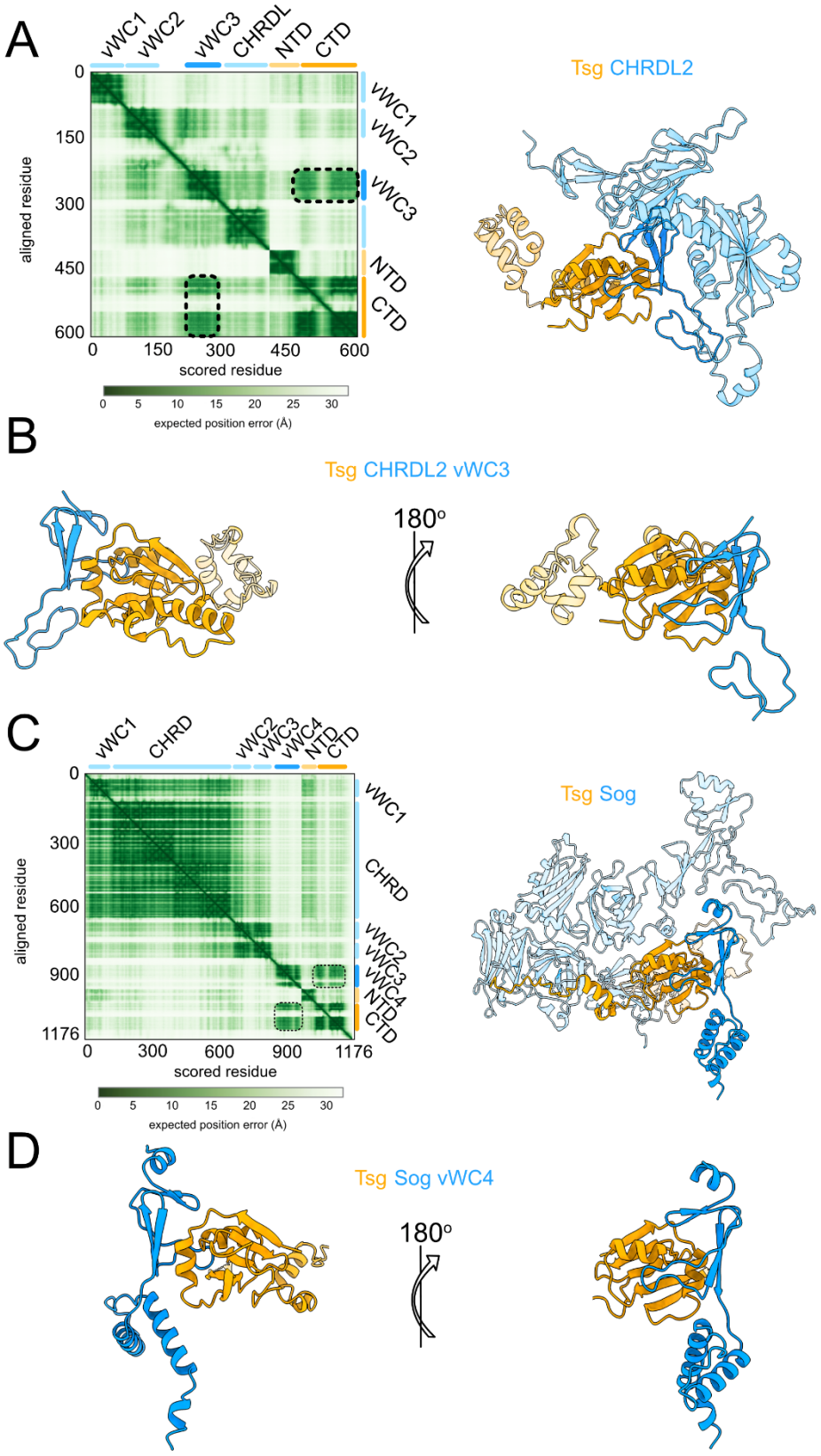

Figure S1. AlphaFold models of Tsg and Chordin family antagonists.
(A) AlphaFold model (right) and corresponding PAE plot (left) of human Tsg (orange) – CHRDL2 (blue) complex with a putative interface between Tsg CTD (dark orange) and CHRDL2 vWC3 (dark blue) as indicated by the dashed regions on the PAE plot. (B) AlphaFold model of Tsg-CHRDL2 vWC3 complex, surface hydrophobicity plotted. (C) and (D) as above but for *Drosophila* Tsg-Sog complex.

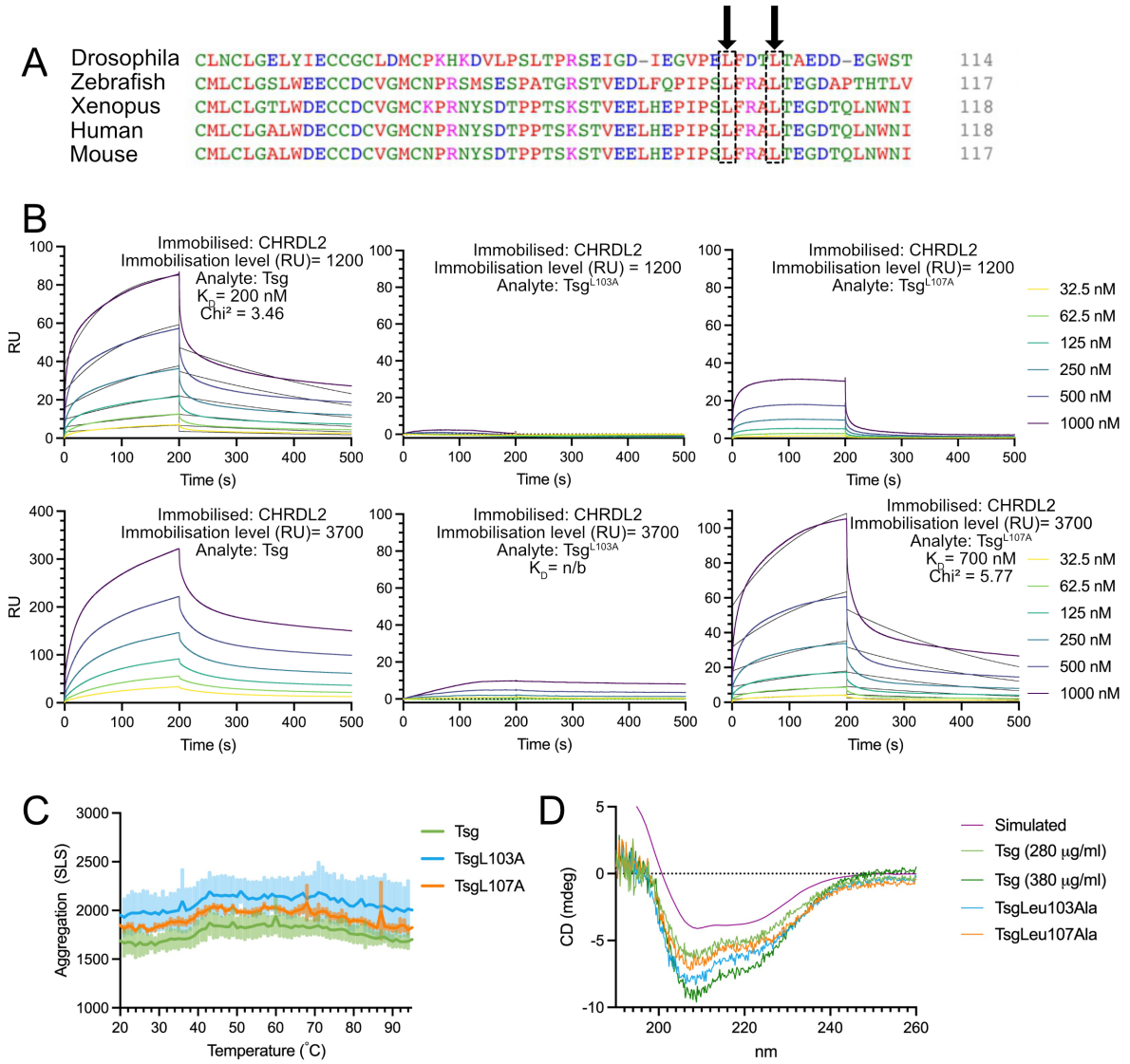

Figure S2. Binding, stability and foldness of Tsg point mutants

(A) Sequence alignment of Tsg sequences from five species showing conservation of identified Leu residues (black arrows). (B) SPR sensorgrams of Tsg, Tsg^{L103A}, or Tsg^{L107A} binding immobilised CHRDL2. CHRDL2 was immobilised at two surface densities for optimal signal detection of both wild type and mutant Tsg binding. As such, binding kinetics for wild type Tsg were calculated from the low surface density measurements, and mutant Tsg kinetics calculated from the high surface density measurements. Binding kinetics were calculated by fitting curves to a Langmuir 1:1 binding model. All experiments were performed in triplicate, average curves plotted with model fit in black. n/b = no binding. (C) Static light scattering (SLS) traces of wild type Tsg (green), Tsg^{L103A} (blue), and Tsg^{L107A} (orange) recorded at 473 nm from 25-95°C. Mean as solid line of triplicate samples with standard deviation from the mean. (D) Buffer subtracted circular dichroism spectra of wild type Tsg at 280 µg/ml (light green) and 380 µg/ml (dark green), TsgLeu103Ala at 330 µg/ml (blue) and TsgLeu107Ala at 290 µg/ml (orange) with a simulated spectrum based on AlphaFold model of Tsg (purple). The theoretical spectrum was generated using PDB2CD webserver from the AlphaFold model of Tsg. For all samples, 10 spectra were recorded and averaged.

Table 1 Estimated secondary structural element proportion in Tsg wild type and mutants

Tsg (280 μg/ml)

Method	α-Helix	β-strand	Turns	Unordered	NMRSD
Contin	0.055	0.182	0.159	0.467	0.165
CDSSTR	0.07	0.24	0.14	0.41	0.094
K2D	0.14	0.33	n/a	0.53	0.627
Selcon3	0.051	0.169	0.13	0.434	0.278
Mean	**0.079**	**0.23025**	**0.143**	**0.46025**	**n/a**

Tsg (380 μg/ml)

Method	α-Helix	β-strand	Turns	Unordered	NMRSD
Contin	0.052	0.19	0.157	0.46	0.099
CDSSTR	0.07	0.23	0.13	0.42	0.059
K2D	0.14	0.33	n/a	0.54	0.577
Selcon3	0.053	0.113	0.137	0.395	0.132
Mean	**0.07875**	**0.21575**	**0.14133333**	**0.45375**	**n/a**

Tsg^{L103A}

Method	α-Helix	β-strand	Turns	Unordered	NMRSD
Contin	0.066	0.169	0.148	0.499	0.165
CDSSTR	0.07	0.23	0.14	0.42	0.061
K2D	0.13	0.33	n/a	0.54	0.579
Selcon3	0.066	0.155	0.14	0.426	0.262
Mean	**0.083**	**0.221**	**0.14266667**	**0.47125**	**n/a**

Tsg^{L107A}

Method	α-Helix	β-strand	Turns	Unordered	NMRSD
Contin	0.062	0.169	0.161	0.47	0.144
CDSSTR	0.07	0.23	0.14	0.41	0.07
K2D	0.14	0.32	n/a	0.55	0.544
Selcon3	0.056	0.17	0.136	0.441	0.224
Mean	**0.082**	**0.22225**	**0.14566667**	**0.46775**	**n/a**

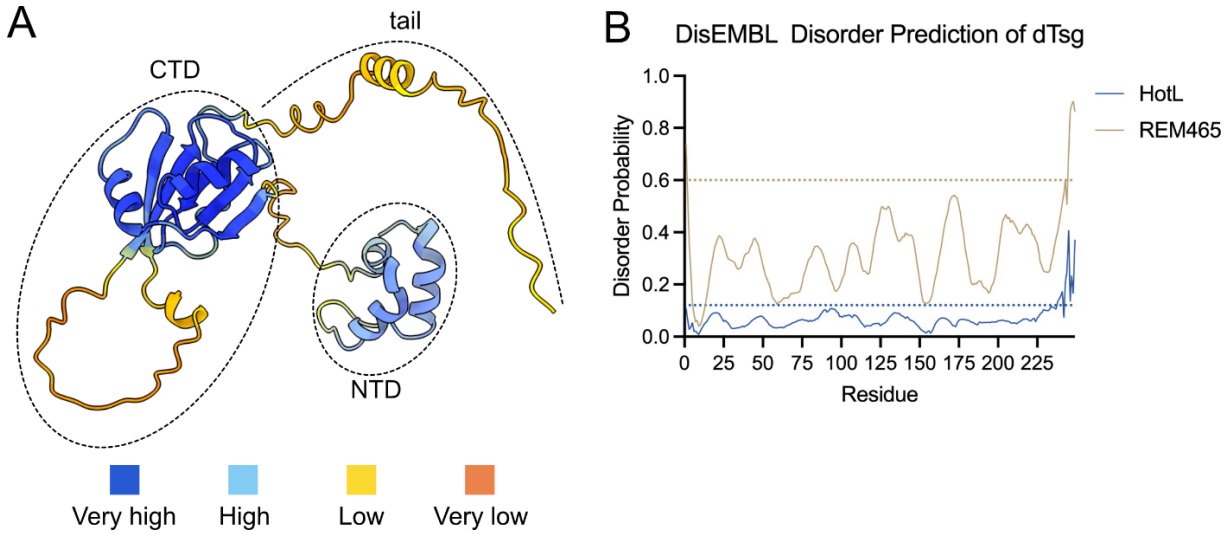

Figure S3. *In silico* tools predict a disordered tail.
(A) AlphaFold model of *Drosophila* Tsg predicts the N-terminal domain (NTD) and C-terminal domain (CTD) with very high (pLDDT >90) or high (90>pLDDT>70) confidence but the tail sequence with low (70>pLDDT>50) or very low (pLDDT<50) confidence. pLDDT, predicted Local Distance Difference Test, plotted per colour key shown below. AlphaFold identifier: AF-P54356-F1. (B) DisEMBL disorder prediction algorithms HotLoops and REMARK-465 for *Drosophila tsg* sequence, dotted line represents model random expectation threshold.

References

Andrade, M.A., Chacón, P., Merelo, J.J., Morán, F., 1993. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. Protein Eng. 6, 383–390. https://doi.org/10.1093/protein/6.4.383

Arora, K., Levine, M.S., O'Connor, M.B., 1994. The screw gene encodes a ubiquitously expressed member of the TGF-β family required for specification of dorsal cell fates in the Drosophila embryo. Genes Dev. 8, 2588–2601. https://doi.org/10.1101/gad.8.21.2588

Ashe, H.L., Levine, M., 1999. Local inhibition and long-range enhancement of Dpp signal transduction by Sog. Nature 398, 427–431. https://doi.org/10.1038/18892

Ashe, H.L., Mannervik, M., Levine, M., 2000. Dpp signaling thresholds in the dorsal ectoderm of the Drosophila embryo. Development 127, 3305–3312. https://doi.org/10.1242/dev.127.15.3305

Cai, J., Pardali, E., Sánchez-Duffhues, G., ten Dijke, P., 2012. BMP signaling in vascular diseases. FEBS Lett. 586, 1993–2002. https://doi.org/10.1016/j.febslet.2012.04.030

Chang, C., Holtzman, D.A., Chau, S., Chickering, T., Woolf, E.A., Holmgren, L.M., Bodorova, J., Gearing, D.P., Holmes, W.E., Brivanlou, A.H., 2001. Twisted gastrulation can function as a BMP antagonist. Nature 410, 483–487. https://doi.org/10.1038/35068583

Eldar, A., Dorfman, R., Weiss, D., Ashe, H., Shilo, D.Z., Barkal, N., 2002. Robustness of the BMP morphogen gradient in Drosophila embryonic patterning. Nature 419, 304–308. https://doi.org/10.1038/nature01061

Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., Ronneberger, O., Bodenstein, S., Zielinski, M., Bridgland, A., Potapenko, A., Cowie, A., Tunyasuvunakool, K., Jain, R., Clancy, E., Kohli, P., Jumper, J., Hassabis, D., 2022. Protein complex prediction with AlphaFold-Multimer. https://doi.org/10.1101/2021.10.04.463034

Ferguson, Anderson, 1992. Decapentaplegic acts as a morphogen to organize dorsal-ventral pattern in the Drosophila embryo. Cell 71, 451–461. https://doi.org/10.1016/0092-8674(92)90514-d

Frampton, S.L., Sutcliffe, C., Baldock, C., Ashe, H.L., 2022. Modelling the structure of Short Gastrulation and generation of a toolkit for studying its function in Drosophila. Biol. Open 11, bio059199. https://doi.org/10.1242/bio.059199

François, V., Solloway, M., O'Neill, J.W., Emery, J., Bier, E., 1994. Dorsal-ventral patterning of the Drosophila embryo depends on a putative negative growth factor encoded by the short gastrulation gene. Genes Dev. 8, 2602–2616. https://doi.org/10.1101/gad.8.21.2602

Fujisawa, T., Huang, Y., Sebald, W., Zhang, J.L., 2009. The binding of von Willebrand factor type C domains of Chordin family proteins to BMP-2 and Tsg is mediated by their SD1 subdomain. Biochem. Biophys. Res. Commun. 385, 215–219. https://doi.org/10.1016/j.bbrc.2009.05.041

Götzke, H., Kilisch, M., Martínez-Carranza, M., Sograte-Idrissi, S., Rajavel, A., Schlichthaerle, T., Engels, N., Jungmann, R., Stenmark, P., Opazo, F., Frey, S., 2019. The ALFA-tag is a highly versatile tool for nanobody-based bioscience applications. Nat. Commun. 10, 4403. https://doi.org/10.1038/s41467-019-12301-7

Groppe, J., Hinck, C.S., Samavarchi-Tehrani, P., Zubieta, C., Schuermann, J.P., Taylor, A.B., Schwarz, P.M., Wrana, J.L., Hinck, A.P., 2008. Cooperative Assembly of TGF-β Superfamily Signaling Complexes Is Mediated by Two Disparate Mechanisms and Distinct Modes of Receptor Binding. Mol. Cell 29, 157–168. https://doi.org/10.1016/j.molcel.2007.11.039

Holley, S.A., Neul, J.L., Attisano, L., Wrana, J.L., Sasai, Y., O'Connor, M.B., De Robertis, E.M., Ferguson, E.L., 1996. The Xenopus dorsalizing factor noggin ventralizes Drosophila embryos by preventing DPP from activating its receptor. Cell 86, 607–617. https://doi.org/10.1016/S0092-8674(00)80134-8

Hoppe, C., Bowles, J.R., Minchington, T.G., Sutcliffe, C., Upadhyai, P., Rattray, M., Ashe, H.L., 2020. Modulation of the Promoter Activation Rate Dictates the Transcriptional Response to Graded BMP Signaling Levels in the Drosophila Embryo. Dev. Cell 54, 727-741.e7. https://doi.org/10.1016/j.devcel.2020.07.007

Irish, V.F., Gelbart, W.M., 1987. The decapentaplegic gene is required for dorsal-ventral patterning of the Drosophila embryo. Genes Dev. 1, 868–879. https://doi.org/10.1101/gad.1.8.868

Jackson, P.D., Hoffmann, F.M., 1994. Embryonic expression patterns of the drosophila decapentaplegic gene: Separate regulatory elements control blastoderm expression and lateral ectodermal expression. Dev. Dyn. 199, 28–44. https://doi.org/10.1002/aja.1001990104

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S.A.A., Ballard, A.J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A.W., Kavukcuoglu, K., Kohli, P., Hassabis, D., 2021. Highly accurate protein structure prediction with AlphaFold. Nature 596, 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kirsch, T., Sebald, W., Dreyer, M.K., 2000. Crystal structure of the BMP-2–BRIA ectodomain complex. Nat. Struct. Biol. 7, 492–496. https://doi.org/10.1038/75903

Larraín, J., Oelgeschläger, M., Ketpura, N.I., Reversade, B., Zakin, L., De Robertis, E.M., 2001. Proteolytic cleavage of Chordin as a switch for the dual activities of twisted gastrulation in BMP signaling. Development 128, 4439–4447.

Lees, J.G., Miles, A.J., Wien, F., Wallace, B.A., 2006. A reference database for circular dichroism spectroscopy covering fold and secondary structure space. Bioinforma. Oxf. Engl. 22, 1955–1962. https://doi.org/10.1093/bioinformatics/btl327

Malinauskas, T., Moore, G., Rudolf, A.F., Eggington, H., Belnoue-Davis, H.L., El Omari, K., Griffiths, S.C., Woolley, R.E., Duman, R., Wagner, A., Leedham, S.J., Baldock, C., Ashe, H.L., Siebold, C., 2024. Molecular mechanism of BMP signal control by Twisted gastrulation. Nat. Commun. 15, 4976. https://doi.org/10.1038/s41467-024-49065-8

Manavalan, P., Johnson, W.C., 1987. Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. Anal. Biochem. 167, 76–85. https://doi.org/10.1016/0003-2697(87)90135-7

Marqués, G., Musacchio, M., Shimell, M.J., Wünnenberg-Stapleton, K., Cho, K.W.Y., O'Connor, M.B., 1997. Production of a DPP activity gradient in the early drosophila embryo through the opposing actions of the SOG and TLD proteins. Cell 91, 417–426. https://doi.org/10.1016/S0092-8674(00)80425-0

Mason, E.D., Konrad, K.D., Webb, C.D., Lawrence Marsh, J., 1994. Dorsal midline fate in Drosophila embryos requires twisted gastrulation, a gene encoding a secreted protein related to human connective tissue growth factor. Genes Dev. 8, 1489–1501. https://doi.org/10.1101/gad.8.13.1489

Mason, E.D., Williams, S., Grotendorst, G.R., Marsh, J.L., 1997. Combinatorial signaling by twisted gastrulation and decapentaplegic. Mech. Dev. 64, 61–75. https://doi.org/10.1016/S0925-4773(97)00049-X

Mizutani, C.M., Nie, Q., Wan, F.Y.M., Zhang, Y.T., Vilmos, P., Sousa-Neves, R., Bier, E., Marsh, J.L., Lander, A.D., 2005. Formation of the BMP activity gradient in the drosophila embryo. Dev. Cell 8, 915–924. https://doi.org/10.1016/j.devcel.2005.04.009

Mörsdorf, D., Knabl, P., Genikhovich, G., 2024. Highly conserved and extremely evolvable: BMP signalling in secondary axis patterning of Cnidaria and Bilateria. Dev. Genes Evol. https://doi.org/10.1007/s00427-024-00714-4

Oelgeschläger, M., Kuroda, H., Reversade, B., Robertis, E.M.D., 2003a. Chordin Is Required for the Spemann Organizer Transplantation Phenomenon in Xenopus Embryos. Dev. Cell 4, 219–230. https://doi.org/10.1016/S1534-5807(02)00404-5

Oelgeschläger, M., Larrain, J., Geissert, D., De Robertis, E.M., 2000. The evolutionarily conserved BMP-binding protein Twisted gastrulation promotes BMP signalling. Nature 405, 757–763. https://doi.org/10.1038/35015500

Oelgeschläger, M., Reversade, B., Larraín, J., Little, S., Mullins, M.C., De Robertis, E.M., 2003b. The pro-BMP activity of Twisted Gastrulation is independent of BMP binding. Development 130, 4047–4056. https://doi.org/10.1242/dev.00633

Piccolo, S., Agius, E., Lu, B., Goodman, S., Dale, L., De Robertis, E.M., 1997. Cleavage of Chordin by Xolloid Metalloprotease Suggests a Role for Proteolytic Processing in the Regulation of Spemann Organizer Activity. Cell 91, 407–416.

Piccolo, S., Sasai, Y., Lu, B., De Robertis, E.M., 1996. Dorsoventral patterning in Xenopus: Inhibition of ventral signals by direct binding of chordin to BMP-4. Cell 86, 589–598. https://doi.org/10.1016/S0092-8674(00)80132-4

Ross, J.J., Shimmi, O., Vilmos, P., Petryk, A., Kim, H., Gaudenz, K., Hermanson, S., Ekker, S.C., O'connor, M.B., Marsh, J.L., 2001. Twisted is a conserved extracellular BMP antagonist. Nature 410, 479–483. https://doi.org/10.1038/35068578

Salazar, V.S., Gamer, L.W., Rosen, V., 2016. BMP signalling in skeletal development, disease and repair. Nat. Rev. Endocrinol. 12, 203–221. https://doi.org/10.1038/nrendo.2016.12

Sawala, A., Sutcliffe, C., Ashe, H.L., 2012. Multistep molecular mechanism for bone morphogenetic protein extracellular transport in the Drosophila embryo. Proc. Natl. Acad. Sci. U. S. A. 109, 11222–11227. https://doi.org/10.1073/pnas.1202781109

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.-Y., White, D.J., Hartenstein, V., Eliceiri, K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image analysis. Nat. Methods 9, 676–682. https://doi.org/10.1038/nmeth.2019

Schoch, C.L., Ciufo, S., Domrachev, M., Hotton, C.L., Kannan, S., Khovanskaya, R., Leipe, D., Mcveigh, R., O'Neill, K., Robbertse, B., Sharma, S., Soussov, V., Sullivan, J.P., Sun, L., Turner, S., Karsch-Mizrachi, I., 2020. NCBI Taxonomy: a comprehensive update on curation, resources and tools. Database 2020, baaa062. https://doi.org/10.1093/database/baaa062

Scott, I.C., Blitz, I.L., Pappano, W.N., Maas, S.A., Cho, K.W.Y., Greenspan, D.S., 2001. Homologues of twisted gastrulation am extracellular cofactors in antagonism of BMP signalling. Nature 410, 475–478. https://doi.org/10.1038/35068572

Shimell, M.J., Ferguson, E.L., Childs, S.R., O'Connor, M.B., 1991. The Drosophila dorsal-ventral patterning gene tolloid is related to human bone morphogenetic protein 1. Cell 67, 469–481. https://doi.org/10.1016/0092-8674(91)90522-Z

Shimmi, O., O'Connor, M.B., 2003. Physical properties of Tld, Sog, Tsg and Dpp protein interactions are predicted to help create a sharp boundary in Bmp signals during dorsoventral patterning of the Drosophila embryo. Development 130, 4673–4682. https://doi.org/10.1242/dev.00684

Shimmi, O., Umulis, D., Othmer, H., O'Connor, M.B., 2005. Facilitated transport of a Dpp/Scw heterodimer by Sog/Tsg leads to robust patterning of the Drosophila blastoderm embryo. Cell 120, 873–886. https://doi.org/10.1016/j.cell.2005.02.009

Sreerama, N., Venyaminov, S.Y., Woody, R.W., 1999. Estimation of the number of alpha-helical and beta-strand segments in proteins using circular dichroism spectroscopy. Protein Sci. Publ. Protein Soc. 8, 370–380.

Sreerama, N., Woody, R.W., 2000. Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. Anal. Biochem. 287, 252–260. https://doi.org/10.1006/abio.2000.4880

Sreerama, N., Woody, R.W., 1993. A self-consistent method for the analysis of protein secondary structure from circular dichroism. Anal. Biochem. 209, 32–44. https://doi.org/10.1006/abio.1993.1079

St Johnston, R.D., Gelbart, W.M., 1987. Decapentaplegic transcripts are localized along the dorsal-ventral axis of the Drosophila embryo. EMBO J. 6, 2785–2791. https://doi.org/10.1002/j.1460-2075.1987.tb02574.x

Stecher, G., Tamura, K., Kumar, S., 2020. Molecular Evolutionary Genetics Analysis (MEGA) for macOS. Mol. Biol. Evol. 37, 1237–1239. https://doi.org/10.1093/molbev/msz312

Tamura, K., Stecher, G., Kumar, S., 2021. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Mol. Biol. Evol. 38, 3022–3027. https://doi.org/10.1093/molbev/msab120

Terwilliger, T.C., Liebschner, D., Croll, T.I., Williams, C.J., McCoy, A.J., Poon, B.K., Afonine, P.V., Oeffner, R.D., Richardson, J.S., Read, R.J., Adams, P.D., 2023. AlphaFold predictions are valuable hypotheses and accelerate but do not replace experimental structure determination. Nat. Methods 1–7. https://doi.org/10.1038/s41592-023-02087-4

Troilo, H., Barrett, A.L., Zuk, A.V., Lockhart-Cairns, M.P., Wohl, A.P., Bayley, C.P., Dajani, R., Tunnicliffe, R.B., Green, L., Jowitt, T.A., Sengle, G., Baldock, C., 2016. Structural characterization of twisted gastrulation provides insights into opposing functions on the BMP signalling pathway. Matrix Biol. 55, 49–62. https://doi.org/10.1016/j.matbio.2016.01.019

Troilo, H., Zuk, A.V., Tunnicliffe, R.B., Wohl, A.P., Berry, R., Collins, R.F., Jowitt, T.A., Sengle, G., Baldock, C., 2014. Nanoscale structure of the BMP antagonist chordin supports cooperative BMP binding. Proc. Natl. Acad. Sci. U. S. A. 111, 13063–13068. https://doi.org/10.1073/pnas.1404166111

Tuazon, F.B., Wang, X., Andrade, J.L., Umulis, D., Mullins, M.C., 2020. Proteolytic Restriction of Chordin Range Underlies BMP Gradient Formation. Cell Rep. 32, 108039. https://doi.org/10.1016/j.celrep.2020.108039

Umulis, D., O'Connor, M.B., Blair, S.S., 2009. The extracellular regulation of bone morphogenetic protein signaling. Development 136, 3715–3728. https://doi.org/10.1242/dev.031534

Umulis, D.M., Serpe, M., O'Connor, M.B., Othmer, H.G., 2006. Robust, bistable patterning of the dorsal surface of the Drosophila embryo. Proc. Natl. Acad. Sci. U. S. A. 103, 11613–11618. https://doi.org/10.1073/pnas.0510398103

van Stokkum, I.H., Spoelder, H.J., Bloemendal, M., van Grondelle, R., Groen, F.C., 1990. Estimation of protein secondary structure and error analysis from circular dichroism spectra. Anal. Biochem. 191, 110–118. https://doi.org/10.1016/0003-2697(90)90396-q

Vilmos, P., Gaudenz, K., Hegedus, Z., Marsh, J.L., 2001. The twisted gastrulation family of proteins, together with the IGFBP and CCN families, comprise the TIC superfamily of cysteine rich secreted factors. J. Clin. Pathol. - Mol. Pathol. 54, 317–323. https://doi.org/10.1136/mp.54.5.317

Wang, X., Harris, R.E., Bayston, L.J., Ashe, H.L., 2008. Type IV collagens regulate BMP signalling in Drosophila. Nature 455, 72–77. https://doi.org/10.1038/nature07214

Wang, Y.C., Ferguson, E.L., 2005. Spatial bistability of Dpp-receptor interactions during Drosophila dorsal-ventral patterning. Nature 434, 229–234. https://doi.org/10.1038/nature03318

Winstanley, J., Sawala, A., Baldock, C., Ashe, H.L., 2015. Synthetic enzyme-substrate tethering obviates the Tolloid-ECM interaction during Drosophila BMP gradient formation. eLife 4, e05508. https://doi.org/10.7554/eLife.05508

Xie, J., Fisher, S., 2005. Twisted gastrulation enhances BMP signaling through chordin dependent and independent mechanisms. Development 132, 383–391. https://doi.org/10.1242/dev.01577

Xu, E.-R., Blythe, E.E., Fischer, G., Hyvönen, M., 2017. Structural analyses of von Willebrand factor C domains of collagen 2A and CCN3 reveal an alternative mode of binding to bone morphogenetic protein-2. J. Biol. Chem. 292, 12516–12527. https://doi.org/10.1074/jbc.M117.788992

Zhang, J.L., Huang, Y., Qiu, L.Y., Nickel, J., Sebald, W., 2007. Von Willebrand factor type C domain-containing proteins regulate bone morphogenetic protein signaling through different recognition mechanisms. J. Biol. Chem. 282, 20002–20014. https://doi.org/10.1074/jbc.M700456200

Zinski, J., Bu, Y., Wang, X., Dou, W., Umulis, D., Mullins, M.C., 2017. Systems biology derived source-sink mechanism of BMP gradient formation. https://doi.org/10.7554/eLife.22199.001