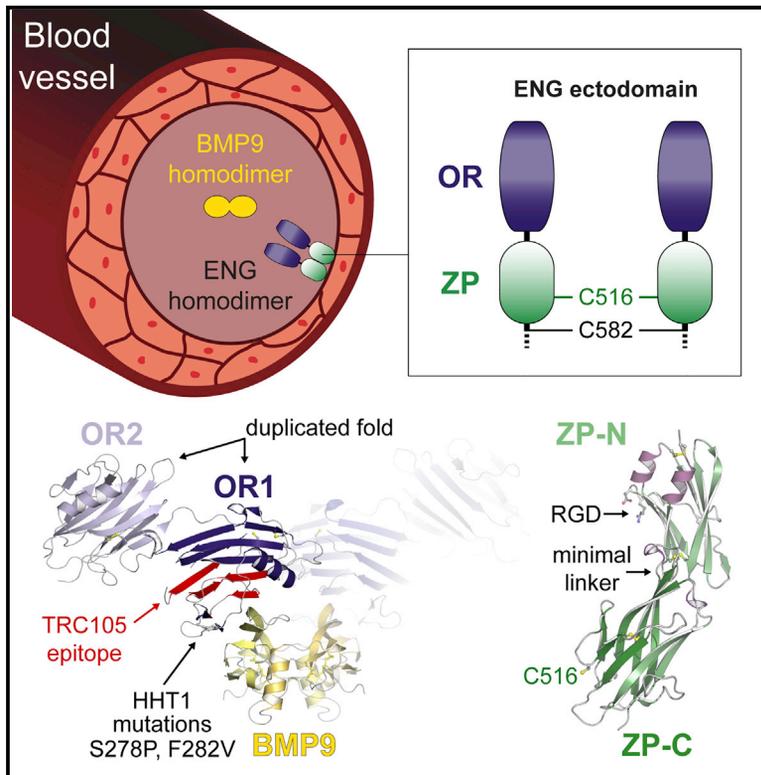


Structural Basis of the Human Endoglin-BMP9 Interaction: Insights into BMP Signaling and HHT1

Graphical Abstract



Authors

Takako Saito, Marcel Bokhove, Romina Croci, ..., Michelle Letarte, Daniele de Sanctis, Luca Jovine

Correspondence

luca.jovine@ki.se

In Brief

Endoglin (ENG)/CD105, a key player in angiogenesis and vascular homeostasis, is mutated in the genetic disorder HHT1 and implicated in tumor angiogenesis and preeclampsia. Saito et al. determine structures of human ENG alone and in complex with the physiological ligand BMP9, shedding light onto the molecular basis of BMP signaling.

Highlights

- Crystal structures of human ENG and its complex with BMP9 were determined
- The orphan domain of ENG adopts a fold that explains the effect of HHT1 mutations
- ZP module-mediated dimerization of ENG creates a clamp that secures homodimeric BMP9
- ENG-bound BMP9 can interact with the ALK1 receptor but not the ActRIIB receptor

Accession Numbers

5I04
5I05
5HZW
5HZV



Structural Basis of the Human Endoglin-BMP9 Interaction: Insights into BMP Signaling and HHT1

Takako Saito,^{1,5,6} Marcel Bokhove,^{1,6} Romina Croci,¹ Sara Zamora-Caballero,¹ Ling Han,¹ Michelle Letarte,^{2,3} Daniele de Sanctis,⁴ and Luca Jovine^{1,7,*}

¹Department of Biosciences and Nutrition and Center for Innovative Medicine, Karolinska Institutet, Huddinge 14183, Sweden

²Molecular Medicine, Hospital for Sick Children, Peter Gilgan Centre for Research and Learning, University of Toronto, Toronto, ON M5G 0A4, Canada

³Department of Immunology, University of Toronto, Toronto, ON M5G 0A4, Canada

⁴ESRF—The European Synchrotron, Grenoble 38000, France

⁵Present address: Department of Cell Science, Institute of Biomedical Sciences, School of Medicine, Fukushima Medical University, Fukushima 9601295, Japan

⁶These authors contributed equally

⁷Lead Contact

*Correspondence: luca.jovine@ki.se

<http://dx.doi.org/10.1016/j.celrep.2017.05.011>

SUMMARY

Endoglin (ENG)/CD105 is an essential endothelial cell co-receptor of the transforming growth factor β (TGF- β) superfamily, mutated in hereditary hemorrhagic telangiectasia type 1 (HHT1) and involved in tumor angiogenesis and preeclampsia. Here, we present crystal structures of the ectodomain of human ENG and its complex with the ligand bone morphogenetic protein 9 (BMP9). BMP9 interacts with a hydrophobic surface of the N-terminal orphan domain of ENG, which adopts a new duplicated fold generated by circular permutation. The interface involves residues mutated in HHT1 and overlaps with the epitope of tumor-suppressing anti-ENG monoclonal TRC105. The structure of the C-terminal zona pellucida module suggests how two copies of ENG embrace homodimeric BMP9, whose binding is compatible with ligand recognition by type I but not type II receptors. These findings shed light on the molecular basis of the BMP signaling cascade, with implications for future therapeutic interventions in this fundamental pathway.

INTRODUCTION

The transforming growth factor β (TGF- β) superfamily consists of type I and II serine/threonine kinase receptors that interact with structurally related ligands, including TGF- β s and bone morphogenetic proteins (BMPs). Ligand-induced receptor complex formation results in phosphorylation of a type I receptor by a type II receptor, which, in turn, leads to phosphorylation of one of several downstream Smad proteins and signal transduction (Horbelt et al., 2012).

ENG is an abundant type I transmembrane glycoprotein of the vascular endothelium that was initially recognized as a co-recep-

tor for TGF- β 1 and TGF- β 3, associating with TGF- β type I and II receptors (Cheifetz et al., 1992; Gougos and Letarte, 1988a, 1990; Quackenbush and Letarte, 1985). In most cells, the TGF- β type I receptor is activin receptor-like kinase 5 (ALK5), whereas, in endothelial cells, there is also ALK1. ENG and ALK1 were subsequently shown to bind with much higher affinity to BMP9 (also known as growth differentiation factor 2 [GDF2]) than to TGF- β 1 and mediate its proliferation signal (Brown et al., 2005; Nolan-Stevaux et al., 2012; Scharpfenecker et al., 2007). Post-translationally processed homodimeric BMP9 mediates signal transduction via a set of interactions involving ENG, ALK1, and a BMP type II receptor (BMPRII or ActRIIB) (Bidart et al., 2012; Brown et al., 2005; David et al., 2007; Kienast et al., 2016; Scharpfenecker et al., 2007; Upton et al., 2009). Like ALK1, ENG is essential for heart development and vascular homeostasis; *Eng*^{-/-} mice die at mid-gestation because of defective cardiovascular development and angiogenesis (Bourdeau et al., 1999; Li et al., 1999). ENG is overexpressed in the vascular endothelium of solid tumors (Bernabeu et al., 2009; Miller et al., 1999), and functionally blocking antibodies to ENG can inhibit tumor growth via the ENG/BMP9/ALK1 pathway (Lebrin et al., 2004). Mutations in the *ENG* gene cause the autosomal dominant disorder hereditary hemorrhagic telangiectasia type 1 (HHT1), a vascular dysplasia with life-threatening consequences that affects 1 in 10,000 individuals and is characterized by arteriovenous malformations, recurrent hemorrhages, and numerous sequelae (Abdalla and Letarte, 2006; Bourdeau et al., 1999; McAllister et al., 1994). Over 500 human *ENG* gene variants linked to HHT1 are currently reported in the University of Utah Department of Pathology HHT and ENG Database (http://www.arup.utah.edu/database/ENG/ENG_welcome.php; Stenson et al., 2014). Moreover, a circulating form of the ENG ectodomain, suggested to be released upon C-terminal cleavage of ENG by matrix metalloproteinase 14 (MMP-14) (Hawinkels et al., 2010), is increased at early stages of preeclampsia and serves as a predictive marker (Gregory et al., 2014; Venkatesha et al., 2006). Despite the important role of ENG in the vascular system, its association with several diseases and potential use



as a target of anti-angiogenic cancer therapy (Bernabeu et al., 2009; Paauwe et al., 2016), no high-resolution structural information is available for the protein and how it interacts with the key ligand BMP9 (Castonguay et al., 2011).

ENG is predominantly expressed as an ~180 kDa homodimer linked by intermolecular disulfide bonds (Gougos and Letarte, 1988b, 1990). It is composed of an N-terminal orphan region (OR), a C-terminal bipartite zona pellucida module (ZP), a single transmembrane domain, and a short cytoplasmic peptide that differs in the L and S isoforms of the protein (Figure S1A; Bellón et al., 1993; Bernabeu et al., 2009; Gregory et al., 2014; Litscher and Wassarman, 2015). Extensive attempts to crystallize the full-length ectodomain of ENG (ECTO; residues E26–G586), using different strategies to reduce N-glycan heterogeneity, were not successful. We reasoned that this may be caused by flexibility of the OR/ZP linker (Figure S1A, arrow), which could underlie the different conformations observed by small-angle X-ray scattering (Alt et al., 2012; Van Le et al., 2009) and negative stain electron microscopy (Llorca et al., 2007). To overcome this obstacle, we combined a divide and conquer strategy with a recently developed eukaryotic expression system where proteins of interest are fused to a mammalianized version of bacterial maltose-binding protein (mMBP) (Bokhove et al., 2016a). Using this approach, we first elucidated the crystal structure of ENG OR, which plays a key role in signaling by binding BMP9 (Castonguay et al., 2011) and does not show homology to proteins of known structure (Alt et al., 2012; Llorca et al., 2007); subsequently, we determined the structure of the OR-BMP9 complex as well as the structure of ENG ZP.

The structure of ENG OR reveals a new fold wherein a β helix-like domain with an unprecedented topology has been duplicated, followed by circular permutation. The electron density map of ENG OR in complex with BMP9 shows how the knuckle region of the ligand interacts hydrophobically with a ridge at the edge of OR, including residues previously implicated in BMP9 binding (Mallet et al., 2015). Finally, the crystal structure of ENG ZP reveals a minimal fold with the short ZP-N/ZP-C interdomain linker signature of non-polymerizing ZP modules (Bokhove et al., 2016b). Based on the structures of OR and ZP, a model can be presented of the ENG disulfide-bonded dimer in complex with BMP9 and the type I receptor ALK1.

RESULTS AND DISCUSSION

The Crystal Structure of OR Reveals Two Intertwined Domains

The structure of OR (ENG E26–S337) fused to mMBP (^MOR; Figures S1B and S2A–S2C) was solved at 2.4 Å resolution by molecular replacement (MR) with MBP (Table S1). The electron density map (Figure S2B) reveals that OR is a monomer and consists of two domains that share a common new fold (OR1/OR2; Figure 1). Each domain consists of 12 β strands, which form a parallel β helix-like structure, and a single α helix. Interestingly, OR2 (β 3– β 14 and α 1) is inserted C-terminally to OR1 β 1– β 2, after which OR1 (β 15– β 24 and α 2) is completed (Figure 1B). Both domains contain a conserved disulfide bond involving C30–C207 and C53–C182, respectively (Figures 1A–1C; Figure S2D). Two additional Cys in OR1, C242 and C330 (Figure S1A), are likely to form a third disul-

fide that is not resolved because of structural disorder of residues S329–S337. However, an ^MOR variant lacking both Cys is secreted as efficiently as wild-type ^MOR and retains BMP9 binding activity, indicating that C242 and C330 are not crucial for folding and interaction with the ligand (Figure S1C).

OR Is a Result of Gene Duplication Followed by Circular Permutation

OR1 and OR2 are only 18% identical in sequence but can be superimposed with a C α root-mean-square deviation of 2.2 Å for 102 overlapping residues (Figure 1C; Figure S3A). Although OR2 does not require OR1 for folding and secretion (Figure S3B), an OR1-only construct was not secreted by itself (Figure S3C). Notably, the relative sequence position of the Cys moieties of the conserved disulfide is reversed in OR1 compared with OR2 so that C30 and C207 correspond to C182 and C53, respectively. Moreover, the N-terminal sequence of OR1 (including strands β 1 and β 2) aligns with the C-terminal part of its OR2 domain (Figure S3A). Taken together, these observations suggest that OR1 was generated by a duplication of an ancestral OR2 gene, followed by circular permutation (Figure 1D). Considering that ENG is only found in amniotes, a possible evolutionary relationship for OR2 may be, in turn, indicated by the fact that its structure distantly resembles tubulin cofactor C (Figure 1E). However, this β -helical molecular chaperone has a markedly different topology with a C-terminal α helix (Nithianantham et al., 2015).

BMP9 Binds to a Conserved Hydrophobic Groove of OR1

To gain insights into how ENG modulates signaling (including cell proliferation) by interacting with ligands, we co-expressed ^MOR and BMP9 in mammalian cells and crystallized the resulting complex (Figure S4A). Although the crystals only diffracted to ~4.5 Å, the data could be phased by MR using mMBP, OR1, and an independently determined structure of BMP9 (Figure S4B; Table S1). In the complex crystals, the disulfide-bonded BMP9 homodimer spans two adjacent asymmetric units by lying on a symmetry axis (Figure S4C), and the knuckle of each BMP9 molecule binds to a copy of OR1 (Figure S4D), resulting in a “dimer of heterodimers” (Figure 2A). This arrangement is consistent with co-expression experiments of His- and Myc-tagged ^MOR with BMP9 followed by His pull-down. Immunoblot analysis shows that Myc-^MOR co-elutes with His-^MOR only in the presence of BMP9, indicating that the dimeric ligand binds two copies of OR that do not interact with each other (Figure 2B). In agreement with the fact that mMBP and BMP9 are positioned on opposite sides of OR1 in the crystals (Figure S4C), mMBP does not inhibit ENG-BMP9 binding (Figures 2A and 2B), nor does it have the ability to bind BMP9 itself (Figure S4E, lane 11). Moreover, the OR1-BMP9 interaction is consistent with the observation that OR2—which appears to have some degree of flexibility relative to OR1 (Figure S4F)—does not bind BMP9 when expressed by itself (Figure S3B).

The structure of the complex reveals that the ENG-BMP9 interface is largely hydrophobic (Figure 2C). It comprises OR1 residues that contribute to ENG strand β 20 or immediately precede it (I271 and Q270 or N268, respectively; Figure 2D; Figure S3A), which are close to BMP9 residues P412, T413, and Y416 (Figure 2D). Consistent with a nanomolar binding affinity

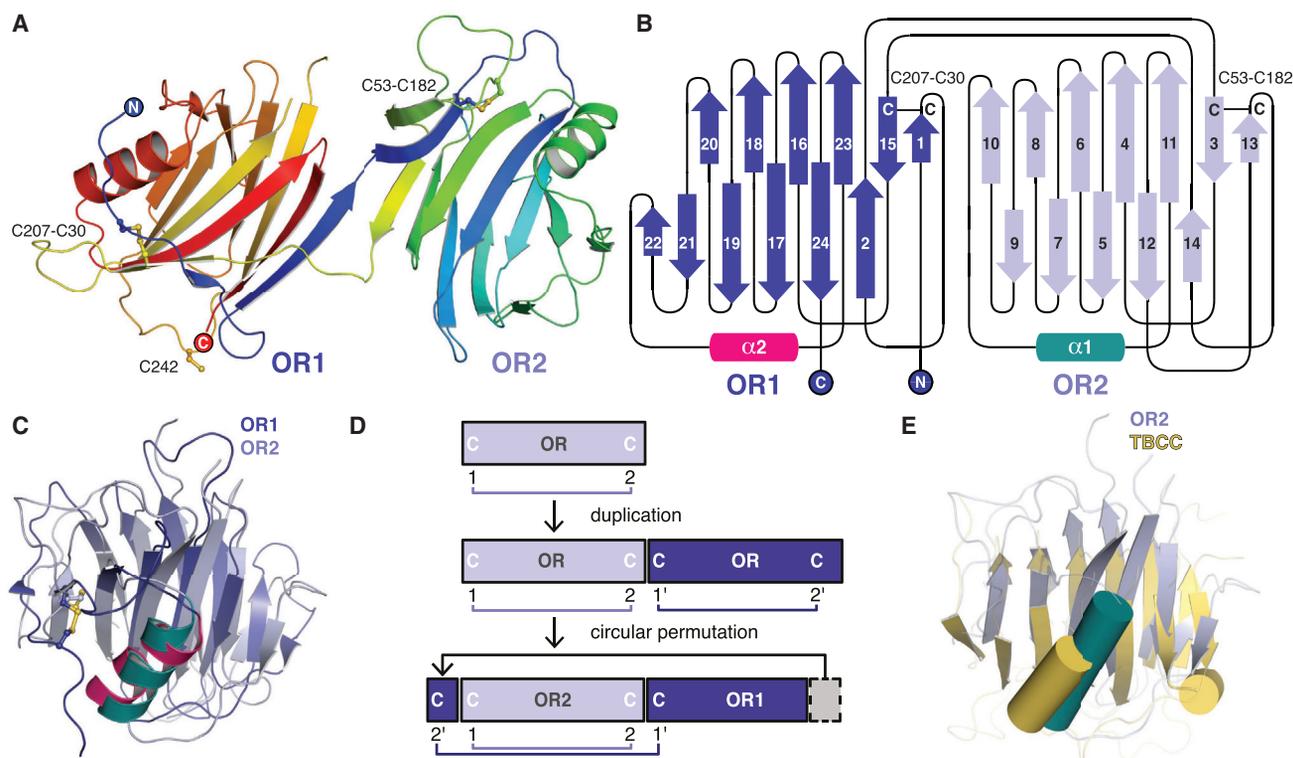


Figure 1. ENG OR Consists of a Duplicated and Circularly Permuted Domain

(A) Cartoon representation of the OR crystal structure, consisting of two domains (OR1/OR2) and rainbow-colored from blue (N terminus) to red (C terminus). Cys residues are shown in ball-and-stick representation.

(B) OR topology with OR1 and OR2 β helices colored dark and light blue, respectively. β Strands (arrows) are numbered.

(C) Superposition of OR1 and OR2, colored as in (B).

(D) Suggested evolution of ENG OR by gene duplication and circular permutation.

(E) Superposition of ENG OR2 and tubulin cofactor C (TBCC).

See also [Figures S1–S3](#) and [S6](#) and [Table S1](#).

for BMP9 ([Table 1](#); [Alt et al., 2012](#); [Castonguay et al., 2011](#); [Gregory et al., 2014](#)), individual mutations of ENG residues N268, Q270, and I271 are not sufficient to disrupt the interaction between unfused OR and BMP9 ([Figure S4E](#), lanes 2, 4, and 5). However, double and triple combinations of these mutations show that OR1 residues Q270 and I271 ([Figure 2D](#); [Figure S3A](#)), which are highly conserved in mammals, are essential for BMP9 binding ([Figure 2E](#)). Importantly, neither the aforementioned individual mutations nor their combinations affect OR folding because the corresponding proteins are secreted comparably with wild-type OR ([Figure S4E](#), lanes 2, 4, 5, and 10; [Figure 2E](#)). Similarly, alanine mutants of D246, Y277, and F290 are secreted normally, although, in this case, the mutant proteins are still able to bind BMP9 ([Figure S4E](#), lanes 1, 6, and 9). In contrast, mutation of the buried hydrophobic residue M269 severely affects OR folding and secretion; however, the fraction of protein that is correctly folded and secreted retains BMP9-binding activity ([Figure S4E](#), lane 3). Interestingly, 9 of 11 residues located in the BMP9 $\beta 6$ – $\beta 7$ region that contacts OR1 are conserved in BMP10, the other high-affinity ligand of ENG ([Castonguay et al., 2011](#)). On the contrary, these residues differ significantly in BMPs that either do not bind ENG or interact only weakly

with it ([Castonguay et al., 2011](#); [Gregory et al., 2014](#); [Figure S4G](#)). Most importantly, the interface region of ENG also involves conserved S278 and invariant F282, whose mutation in HHT1 patients disrupts BMP9-dependent signaling ([Mallet et al., 2015](#)) as well as in vitro binding between OR and BMP9 ([Figure 2E](#); [Figure S4E](#), lanes 7 and 8). Further supporting the cell signaling studies ([Mallet et al., 2015](#)) and our pull-down experiments ([Figure 2E](#)), quantitative analysis of protein-protein interaction by biolayer interferometry (BLI) shows that the mutations Q270A/I271A, S278P, and F282V completely abolish the binding of unfused OR to BMP9. Moreover, BLI confirms that mMBP fusion does not interfere with the interaction between wild-type OR and BMP9 ([Table 1](#); [Figure S8](#)).

Interestingly, the epitope of TRC105/SN6j, a monoclonal antibody to ENG used in clinical trials and acting as a tumor suppressor ([She and Seon, 2001](#); [Bernabeu et al., 2009](#); [Nolan-Stevaux et al., 2012](#); [Paauwe et al., 2016](#)), overlaps with the BMP9-binding region of OR1 ([Figure 2A](#), red). This explains why TRC105, but not SN6h, whose target epitope is elsewhere on ENG ([Figure 2A](#), cyan), inhibits BMP9-dependent signaling ([Nolan-Stevaux et al., 2012](#)). Similarly, overlap of the BMP9 binding site with the prodomain region of pro-BMP9 ([Mi et al., 2015](#))

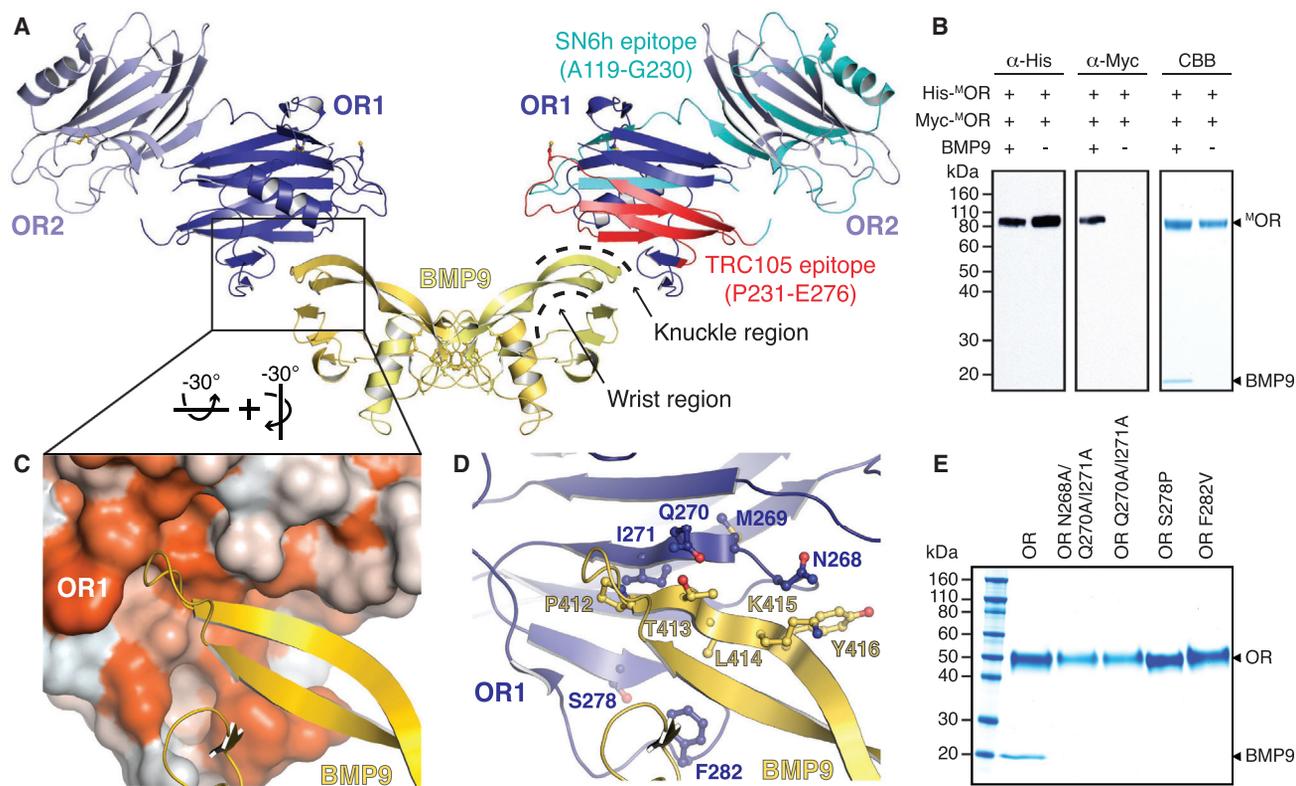


Figure 2. A Hydrophobic Surface of ENG OR1 Binds to the Knuckle Region of BMP9

(A) Crystal structure of the OR-BMP9 complex, showing the dimer generated by crystal symmetry. OR is colored as in Figures 1B and 1C, and BMP9 chains are shown in yellow.

(B) Non-reducing immunoblot and Coomassie (CBB) analysis of His pull-down of ^MOR/BMP9 co-expression experiments. Two copies of OR bind dimeric BMP9 but do not interact with each other.

(C) Detail of the OR1-BMP9 interface, with the surface of OR1 colored by hydrophobicity (orange, most hydrophobic; white, least hydrophobic).

(D) Alternative view of (C), with both proteins in cartoon representation and interface residues shown in ball-and-stick representation.

(E) BMP9 is pulled down by non-mMBP fused wild-type OR but not by double or triple mutant ORs, indicating that the residues shown in (D) are essential for binding.

See also Figures S4 and S8 and Table S1.

(Figure S4H) explains why the prodomain is displaced upon interaction with ENG (Kienast et al., 2016).

The ZP Module of ENG Adopts a Minimal Fold with a Short ZP-N/ZP-C Linker

Whereas the BMP9-binding activity of ENG depends on OR, the ZP moiety mediates its dimerization and spatial organization within the complex. Whereas an unfused ZP construct was barely secreted (Bokhove et al., 2016a), mMBP-fused ENG P338–G581 (^MZP; Figure S1B) could be efficiently expressed as a mixture of monomer and intermolecularly disulfide-bonded homodimer (Figure S5A, left). Although the relative ratio of the latter species significantly increased upon sample concentration (Figure S5A, right), crystallization selected the monomer (Figure S5B). The 2.7 Å structure of ^MZP displays the hallmarks of a typical ZP module composed of ZP-N and ZP-C domains with C₁-C₄, C₂-C₃ and C₅-C₇ disulfide connectivity, respectively (Bokhove et al., 2016b; Han et al., 2010; Litscher and Wassarman, 2015; Figures 3A and 4A; Figures S5C, S5D, and S6); however, it also contains several specific features. Consistent with a role of ENG in integrin-mediated

cell adhesion (Gougos and Letarte, 1990; Rossi et al., 2015), the RGD motif of the protein constitutes a short exposed loop of ZP-N (Figure 3A, left). Moreover, following the disordered OR/ZP linker residues P338–D348, C350 before α A forms a disulfide with C382 located after α BC (Figure 3A; Figure S5D). This is reminiscent of a comparable disulfide in uromodulin (UMOD), another ZP module-containing protein (Bokhove et al., 2016b). In agreement with an important structural role of this disulfide, loss of C382 is associated with HHT1 (Stenson et al., 2014; University of Utah Department of Pathology HHT and ENG Database, http://www.arup.utah.edu/database/ENG/ENG_welcome.php), and mutation of C350 and/or C382 disrupts ENG folding and secretion (Figures 4B and 4C).

Although the structure of ZP does not support a role for C350 in ENG dimerization (Alt et al., 2012), Cys mutagenesis reveals that, in addition to a C582-C582 disulfide (Guerrero-Estero et al., 2002), dimerization involves C516 (Figures 4A–4C). This agrees with the observation that, in the monomeric ZP structure, C516 is exposed on the surface of ZP-C (Figure 3A). By consisting of only 133 residues and containing a single intramolecular

Table 1. Kinetics Analysis of ENG-BMP9 Interaction by BLI

	K_D (M)	k_{on} ($M^{-1}s^{-1}$)	k_{off} (s^{-1})
ECTO	9.81×10^{-9}	$2.0 \times 10^5 \pm 1.69 \times 10^3$	$1.98 \times 10^{-3} \pm 1.15 \times 10^{-5}$
OR	2.95×10^{-7}	$2.95 \times 10^5 \pm 8.96 \times 10^3$	$8.71 \times 10^{-2} \pm 9.08 \times 10^{-4}$
^M OR	1.79×10^{-7}	$3.49 \times 10^5 \pm 1.3 \times 10^4$	$6.22 \times 10^{-2} \pm 7.76 \times 10^{-4}$
OR F282V	NBD		
OR S278P	NBD		
OR Q270A/I271A	NBD		
mMBP	NBD		

NBD, no binding detected under the tested experimental conditions. See also [Figure S8](#).

disulfide bond, the latter represents a simplified version of the ZP-C domain found in other ZP module proteins ([Bokhove et al., 2016b](#); [Han et al., 2010](#)) ([Figure 3B](#)), including closely related betaglycan (BG)/TGF- β receptor III ([Bernabeu et al., 2009](#); [Diestel et al., 2013](#); [Lin et al., 2011](#)) ([Figure S6](#)).

In the ZP module of ENG, a minimal interdomain linker forces a 3_{10} helix within the FG loop of ZP-C to pack against the β A- β G surface of ZP-N; remarkably, the corresponding surface of UMOD mediates ZP-N/ZP-N homodimerization and is essential for protein polymerization ([Bokhove et al., 2016b](#)) ([Figures 3A and 3B](#)). Taken together, these observations support the idea that, in addition to the lack of a consensus cleavage site within the FG loop of ZP-C ([Lin et al., 2011](#)), close spacing of the ZP-N and ZP-C domains of ENG and BG (which also contains a helix within its ZP-C FG loop; [Figure S6](#); [Lin et al., 2011](#)) underlies their inability to polymerize like other ZP module proteins ([Bokhove et al., 2016b](#)) ([Figure 3B](#)).

ZP-Mediated Dimerization of ENG Clamps BMP9 within the Signaling Complex

To investigate the solution stoichiometry of the ENG-BMP9 complex, we co-expressed His-^MECTO (~ 120 kDa), ECTO-Myc (~ 80 kDa), and BMP9 (2×12.1 kDa) ([Figure 4D](#)). This generates five distinct species, shown in colored boxes: His-^MECTO homodimer (~ 240 kDa, blue), His-^MECTO/ECTO-Myc heterodimer (~ 200 kDa, green), ECTO-Myc homodimer (~ 160 kDa, red), His-^MECTO monomer (cyan), and ECTO-Myc monomer (magenta). His pull-down followed by anti-Myc immunoblot detects disulfide-bonded His-^MECTO/ECTO-Myc heterodimers bound to BMP9 ([Figure 4D](#), lane 1). Because no monomeric ECTO-Myc is present in lane 1, BMP9 is preferentially bound by disulfide-bonded dimeric ENG. Furthermore, the same lane shows lack of homodimeric ECTO-Myc among the material pulled down by immobilized metal affinity chromatography (IMAC), excluding that BMP9 binds more than one ENG dimer or higher oligomeric forms of the protein. Analysis of a purified ECTO-BMP9 complex by size exclusion chromatography with multi-angle light scattering (SEC-MALS) confirmed these results by detecting a molecular mass of ~ 150 kDa ([Figure 4E](#)).

Based on the combination of our crystallographic and biochemical data ([Figures 1, 2, 3, and 4A–4E](#)), we have modeled

the BMP9-bound state of homodimeric ENG ([Figure 4F](#)). As a result of the intermolecular cross-linking provided by ZP disulfide C516-C516 ([Figures 4B and 4C](#)) as well as membrane-proximal C582-C582 ([Guerrero-Esteo et al., 2002](#)), the two ENG molecules generate a clamp-like structure whose OR arms protrude into the extracellular space to secure the knuckles of the BMP9 homodimer. Consistent with this proposed architecture, which is compatible with MMP-14 cleavage of ENG at G586/L587 ([Hawinkels et al., 2010](#)), pull-down experiments suggest that ENG ZP does not interact directly with BMP9 ([Figure S5E](#)). However, the ZP must indirectly lower the dissociation of BMP9 from OR, because, as previously observed by surface plasmon resonance (SPR) using an OR construct truncated at P338 ([Alt et al., 2012](#)), the dissociation rate constant (k_{off}) of the OR-BMP9 complex is significantly higher than that of ECTO-BMP9 ([Table 1](#)). Together with the structure and SEC-MALS-derived stoichiometry ([Figure 4E](#)), this observation suggests that, by covalently cross-linking two protein chains ([Figure 4B](#)), ZP confers an antibody-like architecture to ENG. This has the effect of increasing the avidity of the ENG-BMP9 interaction by presenting two closely spaced OR1 binding sites that both match the homodimeric structure of the ligand and counteract its dissociation from a single OR1 site ([Figure 4F](#)). Consistent with this idea, longer OR constructs that are non-physiologically cross-linked into dimers by pairs of C350 (which normally makes an intramolecular disulfide with C382; [Figure 3A](#); [Figure S5D](#)) dissociate much slower from BMP9 than monomeric OR ([Alt et al., 2012](#); [Castonguay et al., 2011](#)).

Taken together, these considerations suggest that ZP not only provides a structural scaffold that spaces OR from the plasma membrane, positioning BMP9 for interaction with type I and II receptors such as ALK1 and ActRIIB, but also affects signaling by indirectly increasing ligand retention.

The ENG-BMP9 Interface Is Compatible with Binding of ALK1, but not ActRIIB, to BMP9

BMP9-dependent signaling in endothelial cells is thought to be regulated by the association of the ligand with ENG and the type I receptor ALK1 ([Scharpfenecker et al., 2007](#)). By merging the model of ENG-BMP9 ([Figure 4F](#)) with the structure of ALK1-BMP9 in crystals of ALK1-BMP9-ActRIIB ([Townson et al., 2012](#)), the extracellular part of an ENG-BMP9-ALK1 ternary complex can be straightforwardly modeled ([Figures 5A and 5B](#)). The resulting assembly suggests that, whereas ALK1 binds to the wrist of BMP9 ([Townson et al., 2012](#)), the knuckle region of the ligand is recognized non-competitively by ENG OR1 ([Figures 2A–2D](#)). Also considering that ENG and ALK1 are physically spaced from each other by the finger/knuckle region of the ligand ([Figure 5A](#)), this agrees with SPR binding studies suggesting that ENG and ALK1 bind independently and non-cooperatively to BMP9 ([Alt et al., 2012](#); [Castonguay et al., 2011](#)). Accordingly, a soluble triple complex could be detected upon co-expression of BMP9, ECTO, and the extracellular part of ALK1 ([Figure 5C](#)). On the other hand, comparison of the ternary complex model ([Figure 5B](#)) with the structure of ALK1-BMP9-ActRIIB ([Townson et al., 2012](#); [Figure 5D](#)) indicates that OR1 and the type II receptor ActRIIB compete for the same region of BMP9. This is consistent with SPR-based evidence that ENG and ActRIIB bind to BMP9 in a mutually exclusive way ([Castonguay et al., 2011](#)) and likely also

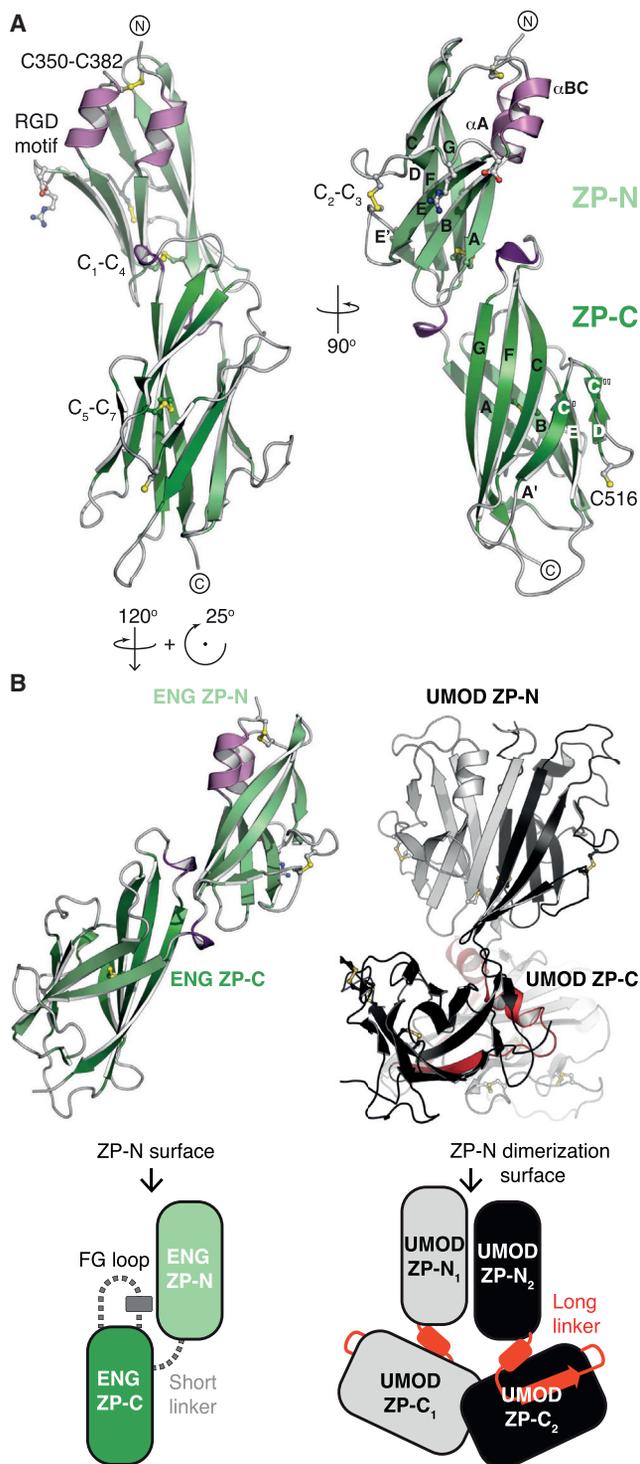


Figure 3. ENG Contains a Minimal ZP Module with Closely Interacting ZP-N and ZP-C Domains

(A) Crystal structure of ENG ZP, colored according to secondary structure. Canonical disulfide bonds and the RGD motif are shown in ball-and-stick representation. The extra ZP-N disulfide (C350–C382) and the free ZP-C Cys involved in dimerization (C516) are also indicated.

(B) Side-by-side comparison of the ENG and UMOD ZP modules. The secondary structure elements of ENG ZP-N and ZP-C are only spaced by five

applies to BMPRII because previous structural studies showed that type II receptors always interact with the knuckle region of growth factors (Hinck et al., 2016).

Together with the observation that ENG binds BMP9 with comparable affinity to ALK1 (Alt et al., 2012) but is significantly more abundant than the type I receptor in endothelial cells (Abdalla et al., 2000; Cheifetz et al., 1992; Ohta et al., 1987; Quack-enbush and Letarte, 1985), these findings support a model where homodimeric BMP9 is first efficiently captured and retained on the surface of endothelial cells by the bivalent architecture of ENG, and a ternary complex is then formed upon non-competitive recruitment of ALK1 (Figures 5A–5C). Presentation of the ligand to a type II receptor, a step requiring BMP9 release from its binding site on ENG, would finally initiate signal transduction. Mirroring the case of antagonistic TGF- β family ligands (Aykul and Martinez-Hackert, 2016), this model would corroborate the idea that different receptors can compete for the same ligand to modulate signaling (Mahlawat et al., 2012). Although it remains to be established how captured BMP9 becomes available to ActRIIB/BMPRII, the fact that the ligand has a relatively high k_{off} for a single OR1 site (Table 1) suggests that this interaction may be triggered by transient dissociation of homodimeric BMP9 from one of its two binding sites on ENG. This would temporarily free one of the knuckle regions of BMP9 for recognition by a type II receptor located close to the ternary ENG-BMP9-ALK1 complex on the plasma membrane. Considering that ActRIIB has a lower k_{off} than OR (Table 1; Alt et al., 2012; Aykul and Martinez-Hackert, 2016; Townson et al., 2012), subsequent dissociation of the second knuckle of BMP9 from ENG would effectively result in transfer of the ALK1-bound ligand to the type II receptor. Elucidation of the details of the signal transduction mechanism will require further structural studies of the intact ternary complex. This may adopt a range of conformations not recapitulated by our crystal structure-based model as a result of flexibility at the ENG OR/ZP junction—possibly underlying the two ECTO-BMP9 states observed by SEC-MALS (Figure 4E)—as well as membrane anchoring.

Structural Basis of HHT1

Mapping of almost 100 gene variants linked to HHT1 (Mallet et al., 2015; Paquet et al., 2001; Stenson et al., 2014; University of Utah Department of Pathology HHT and ENG database, http://www.arup.utah.edu/database/ENG/ENG_welcome.php) onto the structure of ENG indicates that, with the exception of previously discussed OR1 residues S278 and F282 (Figures 2D and 2E), the majority of HHT1 missense mutations affect buried hydrophobic residues (Figure S7).

Fifteen representative mutation sites, ten in OR and five in ZP, are highlighted in Figures 6A and 6B, respectively. The structure of ENG reveals that mutations L32R/H, G52D/V, L82H, V125D, L162R, L170P, L221P, I252T, L262P, I384T,

residues; thus, these domains are much closer to each other than the corresponding domains of UMOD, which are separated by a long linker (red). As a result, whereas the ZP-N homodimerization surface of UMOD is available to pair with another molecule (black/gray), the equivalent region of ENG interacts with ZP-C and maintains the protein in a monomeric state. See also Figures S1, S5, and S6 and Table S1.

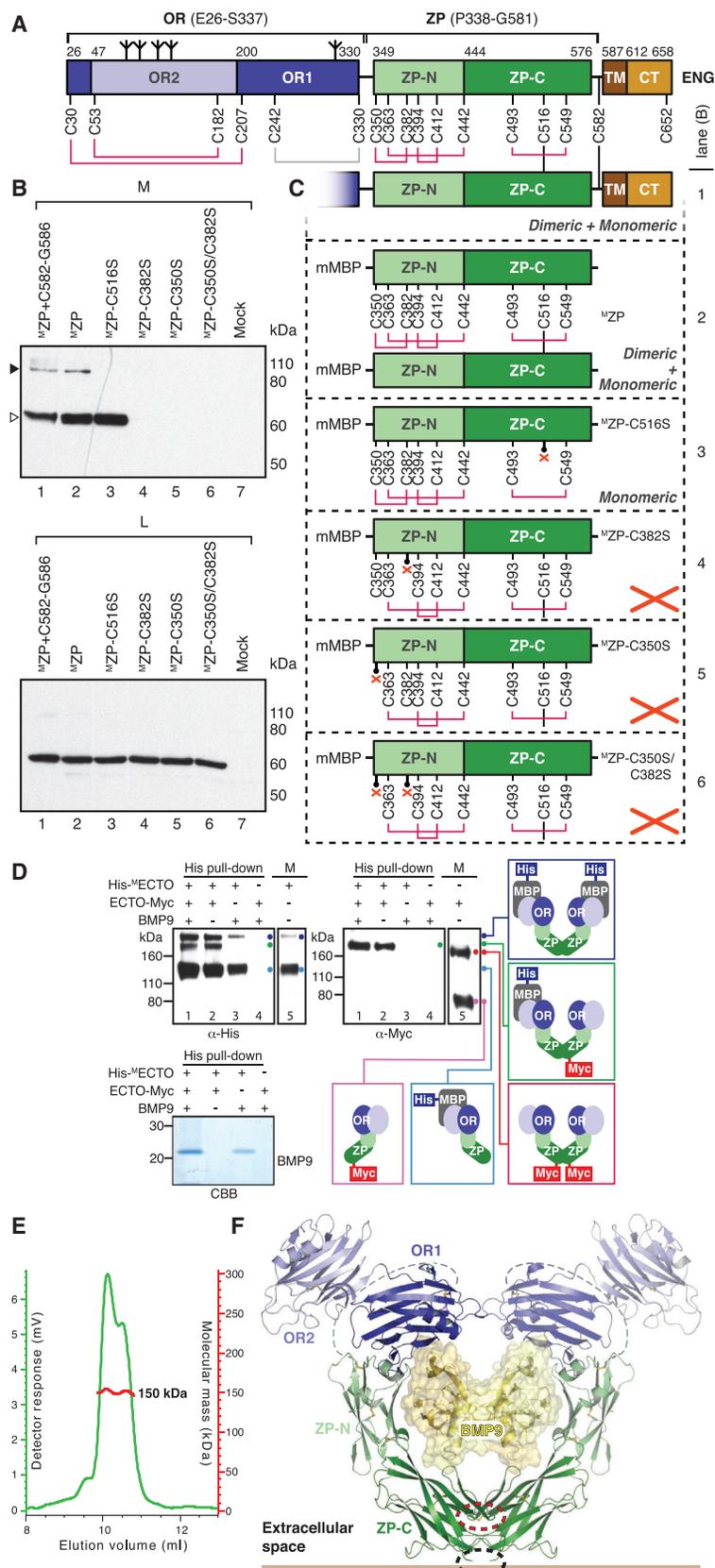


Figure 4. Intermolecular Disulfide Bonding of ENG ZP C516 Generates a Homodimeric Clamp for the BMP9 Ligand

(A) The disulfide bonds observed in the structures of OR (Figures 1A and 1B; Figure S2D) and ZP (Figure 3A; Figure S5D) are indicated in red. Because neither ENG E26-S337 (Castonguay et al., 2011) nor ^MOR (Figure 2B; Figure S1C) make intermolecular disulfide bonds, the pattern suggests that homodimerization of ENG is mediated by a C516-C516 disulfide in addition to C582-C582 (Guerrero-Esteo et al., 2002).

(B) A non-reducing anti-His immunoblot of conditioned medium (M) shows secretion of monomeric and dimeric ENG ZP constructs (white and black arrowheads, respectively). Immunoblot of cell lysates (L) indicates that all constructs are expressed in equal amounts. Consistent with the ENG architecture proposed in (A), mutation of C516 abolishes homodimerization of ^MZP. Moreover, the additional C350-C382 disulfide found in ENG ZP-N (Figure 3A; Figure S5D) is essential for ^MZP secretion.

(C) Disulfide connectivity and secretion state of the ^MZP constructs analyzed in (B). Lack of secretion is symbolized by a red cross.

(D) Immunoblot and Coomassie analyses of co-expression experiments under non-reducing conditions show that two ENG molecules, preferentially a disulfide-bonded dimer, bind homodimeric BMP9.

(E) SEC-MALS analysis of purified His-ECTO-BMP9 complex produced in HEK293S cells reveals two conformations whose common molecular mass (red line) is consistent with a 2:2 stoichiometry.

(F) Theoretical model of the complete extracellular region of homodimeric ENG in complex with BMP9. The relative orientation of the ZP-C domains cross-linked by the C516-C516 intermolecular disulfide (dashed red ellipse) is compatible with the C582-C582 disulfide (Guerrero-Esteo et al., 2002) at the C terminus of the ectodomain (dashed black ellipse). See also Figures S1 and S6.

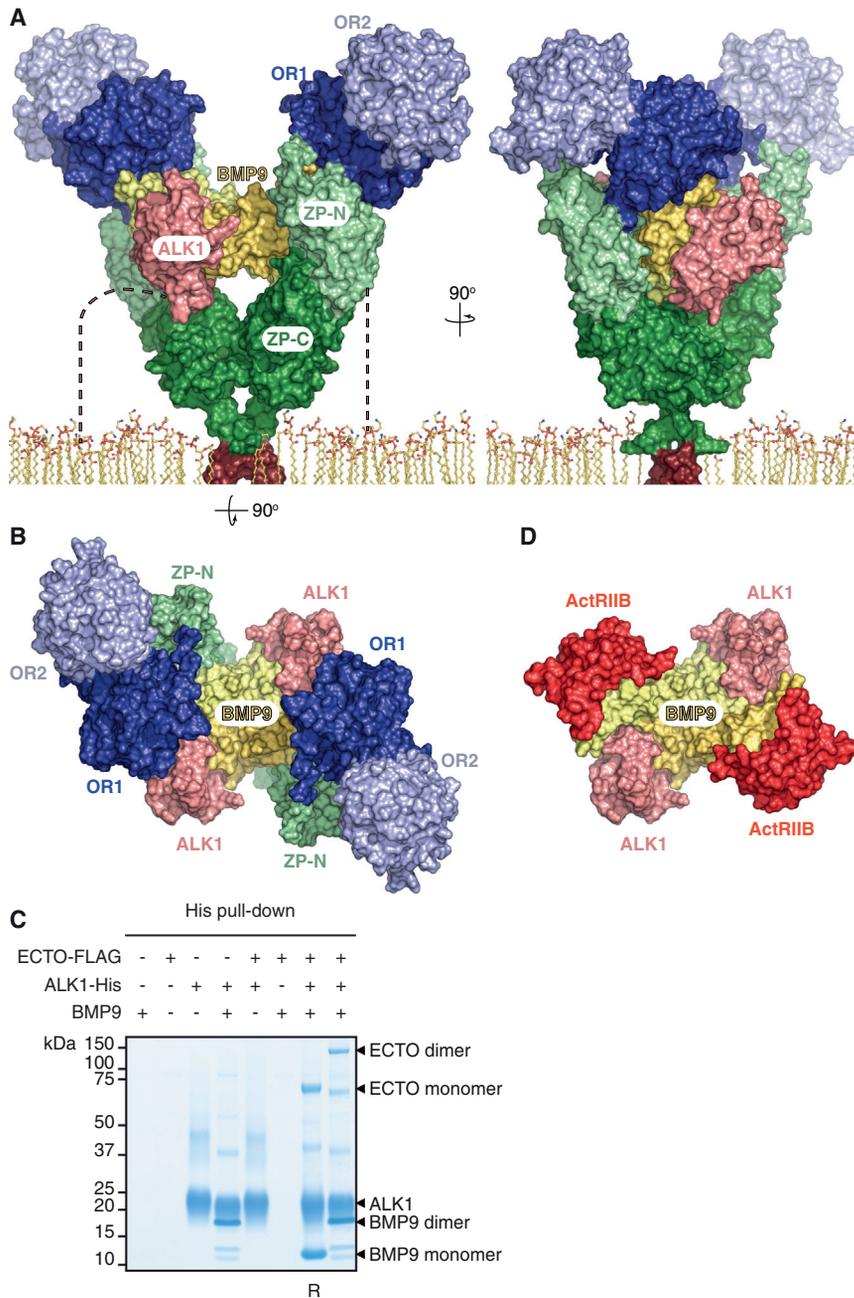


Figure 5. Theoretical Model of the ENG-BMP9-ALK1 Ternary Complex

(A) Combination of the homodimeric ENG-BMP9 model shown in Figure 4F with structural information on how the type I receptor ALK1 binds BMP9 (Townson et al., 2012) suggests a possible architecture for the co-receptor-BMP ligand-type I receptor ternary complex on the plasma membrane.

(B) Top view of the ENG-BMP9-ALK1 model.

(C) SDS-PAGE analysis of ENG ECTO-FLAG, ALK1-His, and BMP9 co-expression followed by His pull-down confirms that the proteins form a triple complex in solution. Consistent with the model depicted in (A) and (B), ALK1 is only able to pull down ECTO when BMP9 is present. All samples were analyzed under non-reducing conditions, except for that marked with R (reducing conditions).

(D) Top view of the crystal structure of the ALK1-BMP9-ActRIIB complex (Townson et al., 2012). Comparison with the ENG-BMP9-ALK1 model (B) suggests that ENG and ActRIIB compete for the same binding region on BMP9.

ficking of ENG to the cell surface (Ali et al., 2011; Pece-Barbara et al., 1999) by replacing a largely buried hydrophobic side chain (Figure 6A) with a reactive residue that could introduce a spurious disulfide bond. Finally, ENG secretion is also affected by mutations such as D391Y and R437W (Ali et al., 2011; Mallet et al., 2015), which substitute solvent-exposed residues (Figure 6B) with hydrophobic amino acids.

In summary, analysis of the ENG structure suggests that most HHT1 mutations impair its folding rather than directly hinder its ability to engage in protein-protein interactions. This leads to reduced amount of functional ENG at the cell surface, thereby causing HHT1 by haploinsufficiency (Paquet et al., 2001).

Conclusions

By suggesting how the bivalent architecture of ENG secures a pool of

and S407Q/N interfere with hydrophobic core packing, explaining the impaired expression and/or secretion of the corresponding mutants (Ali et al., 2011; Mallet et al., 2015; Paquet et al., 2001; Pece-Barbara et al., 1999). The same phenotype is also caused by mutations C53R, C207R, C363Y, C382W/G, C412Y/S, C493Y, and R529C, which—as observed in the case of the mutant C382S (Figure 4B)—disrupt folding by interfering with correct disulfide bond formation (Figures S2D and S5D; Ali et al., 2011; Mallet et al., 2015; Paquet et al., 2001; Pece-Barbara et al., 1999). The two effects are combined in the case of the mutation W149C, which severely impairs traf-

BMP9 molecules on the plasma membrane, where they can bind to type I receptor ALK1 and become transiently available for interaction with the type II receptors ActRIIB and BMPRII, our studies shed light on some of the mechanisms underlying BMP pathway regulation. At the same time, structural knowledge of human ENG and its interaction with BMP9 constitutes a valuable framework for understanding the molecular basis of HHT. This disease is caused primarily (>85%) by mutations in ENG (HHT1; Johnson et al., 1996; McAllister et al., 1994) and ALK1 (HHT2; Johnson et al., 1996; McAllister et al., 1994), but also by mutations in the common downstream signaling

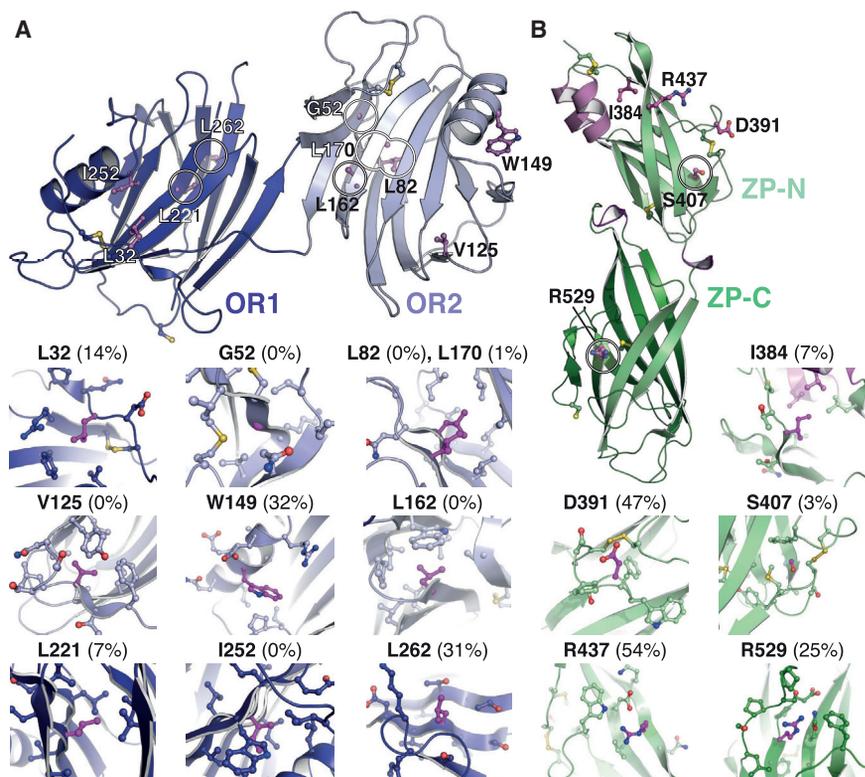


Figure 6. HHT1-Associated Mutations Disrupt the Folding of ENG

(A and B) Overview and close-ups of ENG OR (A) and ZP (B), with patient mutation sites highlighted in magenta. Disulfide bonds and C516 are shown in ball-and-stick representation; coloring is according to Figures 2 and 3. Relative percentages of solvent accessibilities of residue side chains are indicated in parenthesis. See also Figure S7.

Protein Expression

Large-scale production of ^MOR and the ^MOR-BMP9 complex for crystallographic studies was performed in HEK293S cells expressing Endoglycosidase H (Endo H)-sensitive, high mannose-type glycoproteins (Chang et al., 2007; Reeves et al., 2002). HEK293S cells were grown in Dulbecco's modified Eagle's medium (DMEM) with 4 mM L-glutamine (Life Technologies/Thermo Fisher Scientific) and 10% (v/v) fetal bovine serum (Saveen & Werner) at 37°C, 5% CO₂ using T flasks (150 cm², BD Biosciences) or cell factories (2,528 cm², Thermo Fisher Scientific). Transient transfections were essentially performed as described previously (Aricescu et al., 2006; Bokhove et al., 2016a). ^MOR-BMP9 and ECTO-BMP9 co-transfections were performed using a 3:7 DNA mass ratio of ^MOR versus BMP9. Because ^MZP does not contain N-glycosylation sites, it was expressed in HEK293T cells.

SMAD4 (Gallione et al., 2004) as well as in BMP9 itself in a very similar vascular anomaly syndrome (Wooderchak-Donahue et al., 2013). Our findings also contribute to clarifying the tumor-suppressing activity of monoclonal TRC105, which targets the ENG-BMP9 interface (Nolan-Stevaux et al., 2012). Future studies are needed to understand how ENG forms complexes with other ligands of the TGF- β superfamily, for which it has much lower affinities than for BMP9, and to elucidate how ENG contributes to type II receptor-mediated signaling.

EXPERIMENTAL PROCEDURES

DNA Constructs

For secreted expression of unfused proteins, cDNA fragments including tag-encoding sequences were cloned into mammalian expression vector pHLsec (Aricescu et al., 2006). Except in the case of ECTO and OR constructs, which were expressed using the native signal peptide of ENG (residues M1–A25), ENG inserts were cloned in-frame with the Cryp α signal peptide-encoding sequence of pHLsec. Constructs encoding non-mMBP-fused, C-terminally 8His-tagged OR1 with its native signal peptide (ENG M1-T46-(GGGS)₂-L203-S329-C242S) and extracellular C-terminally 6His-tagged ALK1 (residues D22–G114) with its native signal peptide (residues M1–G21) were synthesized (DNA2.0/ATUM) and cloned into pHLsec. mMBP fusion proteins were expressed using pHLmMBP vectors (Bokhove et al., 2016a), which are available at Addgene (https://www.addgene.org/Luca_Jovine/). C-terminally FLAG-tagged ECTO was expressed using a previously described pCMV5 vector (Gregory, 2011). The cDNA encoding residues M1–R429 of human BMP9 was cloned into pHLsec without any tags so that purification of mature BMP9 depended on its ability to form a complex with His-tagged ENG partners.

Mutations were generated by overlap extension PCR, and all constructs were verified by DNA sequencing (Eurofins Genomics).

For functional assays, complexes were obtained by co-expression of His-^MOR, Myc-^MOR, and BMP9 in HEK293T cells using a 2.5:2.5:5 DNA ratio. His-^MECTO, ECTO-Myc, and BMP9 complexes were prepared using a 2:3:5 DNA ratio.

Protein Purification

Three days post-transfection, conditioned medium containing secreted proteins was harvested, 0.22 μ m-filtered (Sarstedt), and adjusted to 20 mM sodium 2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonate (Na-HEPES) (pH 7.8), 150 mM NaCl, 5 mM imidazole (IMAC binding buffer). 10 mL nickel-nitrilotriacetic acid (Ni-NTA) agarose slurry (QIAGEN) was added per liter of medium and incubated overnight at 4°C on an overhead shaker (Heidolph). IMAC beads were then collected in a column, washed with binding buffer, and batch-eluted with a binding buffer containing 500 mM imidazole. The eluted material was concentrated using a spin filter (Millipore) with an appropriate molecular weight cutoff. Concentrated protein was subjected to SEC directly after IMAC (HEK293T-produced ^MZP) or after deglycosylation (HEK293S-produced ^MOR or ^MOR-BMP9 complex). Endo H deglycosylation of ^MOR and ^MOR-BMP9 was performed for 1 hr at 37°C in 120 mM sodium-potassium phosphate (pH 6.0) using a 1:10 mass ratio of enzyme to substrate. Concentrated proteins were applied to a Superdex 200 26/60 SEC column connected to an ÄKTAFPLC system (GE Healthcare). The composition of the SEC running buffer was 20 mM Na-HEPES (pH 7.8), 100 mM NaCl, and 1.5 mM D-maltose for ^MOR; 500 mM and 200 mM NaCl were used for the ^MOR-BMP9 complex and ^MZP, respectively. Protein-containing fractions were pooled, concentrated, and used for crystallization trials.

For functional assays, media were harvested from HEK293T cells transfected in 150 cm² T flasks (Sarstedt). Small-scale IMAC or IMAC/SEC pull-downs were performed using the same protocol as above, linearly scaled down. Following pull-down, IMAC beads with bound ENG or ^MZP constructs were incubated with medium of HEK293T cells expressing BMP9 (consecutive pull-downs). Media obtained from His-^MOR, Myc-^MOR, and BMP9 co-expression were purified using IMAC and SEC, whereas media obtained from

His-M^{ECTO}, ECTO-Myc, M^{ZP}, ZP, and BMP9 co-expression or ALK1-His, ECTO-FLAG, and BMP9 co-expression were purified by IMAC. In parallel, we established that mMBP does not pull down BMP9 (Figure S4E, lane 11).

Protein Analysis

Cell lysates were prepared by resuspending cells from a 10 cm² culture well in lysis buffer (50 mM Tris-HCl/50 mM 3-(N-Morpholino)propanesulfonic acid [MOPS] [pH 7.7], 0.1% SDS, and 1 mM EDTA) supplied with protease inhibitors (Roche) and benzonase (New England Biolabs), followed by 30 min incubation at 4°C and filtration with a 0.22 μm spin filter (Millipore). Conditioned medium was centrifuged for 5 min at 500 × *g*, and the cell pellet was discarded. The supernatant was filtered using a 0.22 μm syringe filter or centrifuged at 4°C at 18,000 × *g* for 15 min.

Proteins separated by SDS-PAGE were detected with SimplyBlue SafeStain (Invitrogen/Thermo Fisher Scientific) or transferred to nitrocellulose membranes (GE Healthcare) for immunoblotting with Penta-His mouse monoclonal antibody (1:1,000, QIAGEN) or 9E10 anti-Myc mouse monoclonal antibody (1:2,000, Sigma-Aldrich). Chemiluminescence detection was performed with Western Lightning ECL Plus (PerkinElmer).

Protein-Protein Interactions

Experimental details for BLI measurements and SEC-MALS analysis of the ENG-BMP9 interaction are reported in the [Supplemental Experimental Procedures](#).

Protein Crystallization

Crystallization experiments were carried out by hanging drop vapor diffusion at room temperature using a mosquito crystallization robot (TTP Labtech). Droplet-shaped crystals of C-terminally 6His-tagged M^{OR} (6 mg/mL) were obtained in 0.1 M Tris-HCl (pH 8.5) and 30% (w/v) PEG1000 (Figure S2A) and directly flash-cooled in liquid nitrogen for data collection. Because of the pH sensitivity of the ENG-BMP9 interaction, BMP9 crystals fortuitously appeared in drops set up with the M^{OR}-BMP9 complex (11 mg/mL) and a mother liquor consisting of 1.0 M LiCl, 4% (w/v) PEG6000, and 0.1 M Na-citrate (pH 3.5). For cryocooling, BMP9 crystals were adjusted to 1.0 M LiCl, 6% (w/v) PEG6000, 0.1 M Na-citrate (pH 4.0), and 30% (v/v) glycerol. Hexagonal bipyramidal crystals of N-terminally 6His-tagged M^{OR} in complex with BMP9 (M^{OR}-BMP9, 6 mg/mL) grew in 1.1 M ammonium tartrate (pH 7.0) and 25% saturated sucrose (Figure S4A). For data collection, crystals were directly flash-cooled in liquid nitrogen. Rod-shaped crystals of M^{ZP} (9 mg/mL) were obtained using 11.5% (v/v) 2-Methyl-2,4-pentandiol (MPD), 11.5% (w/v) PEG1000 and 11.5% (w/v) PEG3350, and 100 mM 2-(N-Morpholino)ethanesulfonic acid (MES)/imidazole (pH 6.5) (Figure S5B). Before flash-cooling in liquid nitrogen, crystals were stabilized using a 50% (v/v) precipitant mix containing 10% (v/v) glycerol.

X-Ray Diffraction Data Collection and Structure Determination

Experimental details for data collection, structure determination, as well as sequence and structure analysis are described in the [Supplemental Experimental Procedures](#). Data collection, refinement, and validation statistics are summarized in [Table S1](#).

ACCESSION NUMBERS

The accession numbers for the atomic coordinates and structure factors for human ENG OR, BMP9, the ENG OR-BMP9 complex, and ENG ZP reported in this paper are PDB: 5I04, 5I05, 5HZW, and 5HZV, respectively.

SUPPLEMENTAL INFORMATION

Supplemental Information includes Supplemental Experimental Procedures, eight figures, and one table and can be found with this article online at <http://dx.doi.org/10.1016/j.celrep.2017.05.011>.

AUTHOR CONTRIBUTIONS

T.S., M.B., R.C., S.Z.C., and L.H. expressed proteins and performed biological assays. T.S. and M.B. crystallized proteins. T.S., M.B., L.J., and D.d.S.

collected crystallographic data. M.B. and L.J. determined and analyzed structures. T.S., M.B., R.C., S.Z.C., M.L., and L.J. wrote the manuscript.

ACKNOWLEDGMENTS

We thank A. Gregory (University of Toronto) for initial work on this project, K. Nishimura (Karolinska Institutet) and M. Selmer (Uppsala University) for help with BMP9 expression and data collection, A. Stsiapanava (Karolinska Institutet) for comments, and the European Synchrotron Radiation Facility (ESRF, Grenoble) and Diamond Light Source (DLS, Oxford) for beamtime (ESRF, mx1639/mx1749; DLS, mx8492-34). We are also extremely grateful to Edwin de Grijjs and Tim Heiseler (PALL FortéBio) for help with BLI measurements and Christoph Albermann and Christian Ackerschott (Wyatt Technology Europe) for SEC-MALS analysis. This work was supported by the Karolinska Institutet, the Center for Innovative Medicine, Swedish Research Council grant 2012-5093, the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine, the Sven and Ebba-Christina Hagberg Foundation; an EMBO Young Investigator award, and the European Research Council (ERC) under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement 260759 (to L.J.). Crystallographic data collection was also supported by FP7/2007-2013 under BioStruct-X (grant agreement 283570).

Received: December 13, 2016

Revised: March 30, 2017

Accepted: May 2, 2017

Published: May 30, 2017

REFERENCES

- Abdalla, S.A., and Letarte, M. (2006). Hereditary haemorrhagic telangiectasia: current views on genetics and mechanisms of disease. *J. Med. Genet.* **43**, 97–110.
- Abdalla, S.A., Pece-Barbara, N., Vera, S., Tapia, E., Paez, E., Bernabeu, C., and Letarte, M. (2000). Analysis of ALK-1 and endoglin in newborns from families with hereditary hemorrhagic telangiectasia type 2. *Hum. Mol. Genet.* **9**, 1227–1237.
- Ali, B.R., Ben-Rebeh, I., John, A., Akawi, N.A., Milhem, R.M., Al-Shehhi, N.A., Al-Ameri, M.M., Al-Shamisi, S.A., and Al-Gazali, L. (2011). Endoplasmic reticulum quality control is involved in the mechanism of endoglin-mediated hereditary haemorrhagic telangiectasia. *PLoS ONE* **6**, e26206.
- Alt, A., Miguel-Romero, L., Donderis, J., Aristorena, M., Blanco, F.J., Round, A., Rubio, V., Bernabeu, C., and Marina, A. (2012). Structural and functional insights into endoglin ligand recognition and binding. *PLoS ONE* **7**, e29948.
- Aricescu, A.R., Lu, W., and Jones, E.Y. (2006). A time- and cost-efficient system for high-level protein production in mammalian cells. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1243–1250.
- Aykul, S., and Martinez-Hackert, E. (2016). Transforming Growth Factor-β family ligands can function as antagonists by competing for type II receptor binding. *J. Biol. Chem.* **291**, 10792–10804.
- Bellón, T., Corbí, A., Lastres, P., Calés, C., Cebrián, M., Vera, S., Cheifetz, S., Massague, J., Letarte, M., and Bernabéu, C. (1993). Identification and expression of two forms of the human transforming growth factor-β-binding protein endoglin with distinct cytoplasmic regions. *Eur. J. Immunol.* **23**, 2340–2345.
- Bernabeu, C., Lopez-Novoa, J.M., and Quintanilla, M. (2009). The emerging role of TGF-β superfamily coreceptors in cancer. *Biochim. Biophys. Acta* **1792**, 954–973.
- Bidart, M., Ricard, N., Levet, S., Samson, M., Mallet, C., David, L., Subileau, M., Tillet, E., Feige, J.-J., and Bailly, S. (2012). BMP9 is produced by hepatocytes and circulates mainly in an active mature form complexed to its prodomain. *Cell. Mol. Life Sci.* **69**, 313–324.
- Bokhove, M., Sadat Al Hosseini, H., Saito, T., Dioguardi, E., Gegenschatz-Schmid, K., Nishimura, K., Raj, I., de Sanctis, D., Han, L., and Jovine, L. (2016a). Easy mammalian expression and crystallography of maltose-binding protein-fused human proteins. *J. Struct. Biol.* **194**, 1–7.

- Bokhove, M., Nishimura, K., Brunati, M., Han, L., de Sanctis, D., Rampoldi, L., and Jovine, L. (2016b). A structured interdomain linker directs self-polymerization of human uromodulin. *Proc. Natl. Acad. Sci. USA* *113*, 1552–1557.
- Bourdeau, A., Dumont, D.J., and Letarte, M. (1999). A murine model of hereditary hemorrhagic telangiectasia. *J. Clin. Invest.* *104*, 1343–1351.
- Brown, M.A., Zhao, Q., Baker, K.A., Naik, C., Chen, C., Pukac, L., Singh, M., Tsareva, T., Parice, Y., Mahoney, A., et al. (2005). Crystal structure of BMP-9 and functional interactions with pro-region and receptors. *J. Biol. Chem.* *280*, 25111–25118.
- Castonguay, R., Werner, E.D., Matthews, R.G., Presman, E., Mulivor, A.W., Solban, N., Sako, D., Pearsall, R.S., Underwood, K.W., Seehra, J., et al. (2011). Soluble endoglin specifically binds bone morphogenetic proteins 9 and 10 via its orphan domain, inhibits blood vessel formation, and suppresses tumor growth. *J. Biol. Chem.* *286*, 30034–30046.
- Chang, V.T., Crispin, M., Aricescu, A.R., Harvey, D.J., Nettleship, J.E., Fennelly, J.A., Yu, C., Boles, K.S., Evans, E.J., Stuart, D.I., et al. (2007). Glycoprotein structural genomics: solving the glycosylation problem. *Structure* *15*, 267–273.
- Cheifetz, S., Bellón, T., Calés, C., Vera, S., Bernabeu, C., Massagué, J., and Letarte, M. (1992). Endoglin is a component of the transforming growth factor- β receptor system in human endothelial cells. *J. Biol. Chem.* *267*, 19027–19030.
- David, L., Mallet, C., Mazerbourg, S., Feige, J.-J., and Bailly, S. (2007). Identification of BMP9 and BMP10 as functional activators of the orphan activin receptor-like kinase 1 (ALK1) in endothelial cells. *Blood* *109*, 1953–1961.
- Diestel, U., Resch, M., Meinhardt, K., Weiler, S., Hellmann, T.V., Mueller, T.D., Nickel, J., Eichler, J., and Muller, Y.A. (2013). Identification of a Novel TGF- β -Binding Site in the Zona Pellucida C-terminal (ZP-C) Domain of TGF- β -Receptor-3 (TGF- β -R3). *PLoS ONE* *8*, e67214.
- Gallione, C.J., Repetto, G.M., Legius, E., Rustgi, A.K., Schelley, S.L., Tejpar, S., Mitchell, G., Drouin, E., Westermann, C.J.J., and Marchuk, D.A. (2004). A combined syndrome of juvenile polyposis and hereditary haemorrhagic telangiectasia associated with mutations in MADH4 (SMAD4). *Lancet* *363*, 852–859.
- Gougos, A., and Letarte, M. (1988a). Biochemical characterization of the 44G4 antigen from the HOON pre-B leukemic cell line. *J. Immunol.* *141*, 1934–1940.
- Gougos, A., and Letarte, M. (1988b). Identification of a human endothelial cell antigen with monoclonal antibody 44G4 produced against a pre-B leukemic cell line. *J. Immunol.* *141*, 1925–1933.
- Gougos, A., and Letarte, M. (1990). Primary structure of endoglin, an RGD-containing glycoprotein of human endothelial cells. *J. Biol. Chem.* *265*, 8361–8364.
- Gregory, A.L. (2011). Structural and Functional Characteristics of a Soluble Form of Endoglin in the Context of Preeclampsia. MSc (University of Toronto).
- Gregory, A.L., Xu, G., Sotov, V., and Letarte, M. (2014). Review: the enigmatic role of endoglin in the placenta. *Placenta* *35* (Suppl), S93–S99.
- Guerrero-Esteo, M., Sanchez-Elsner, T., Letamendia, A., and Bernabeu, C. (2002). Extracellular and cytoplasmic domains of endoglin interact with the transforming growth factor- β receptors I and II. *J. Biol. Chem.* *277*, 29197–29209.
- Han, L., Monné, M., Okumura, H., Schwend, T., Cherry, A.L., Flot, D., Matsuda, T., and Jovine, L. (2010). Insights into egg coat assembly and egg-sperm interaction from the X-ray structure of full-length ZP3. *Cell* *143*, 404–415.
- Hawinkels, L.J., Kuiper, P., Wiercinska, E., Verspaget, H.W., Liu, Z., Pardali, E., Sier, C.F., and ten Dijke, P. (2010). Matrix metalloproteinase-14 (MT1-MMP)-mediated endoglin shedding inhibits tumor angiogenesis. *Cancer Res.* *70*, 4141–4150.
- Hinck, A.P., Mueller, T.D., and Springer, T.A. (2016). Structural biology and evolution of the TGF- β family. *Cold Spring Harb. Perspect. Biol.* *8*, a022103.
- Horbelt, D., Denkis, A., and Knaus, P. (2012). A portrait of Transforming Growth Factor β superfamily signalling: Background matters. *Int. J. Biochem. Cell Biol.* *44*, 469–474.
- Johnson, D.W., Berg, J.N., Baldwin, M.A., Gallione, C.J., Marondel, I., Yoon, S.J., Stenzel, T.T., Speer, M., Pericak-Vance, M.A., Diamond, A., et al. (1996). Mutations in the activin receptor-like kinase 1 gene in hereditary haemorrhagic telangiectasia type 2. *Nat. Genet.* *13*, 189–195.
- Kienast, Y., Jucknischke, U., Scheiblich, S., Thier, M., de Wouters, M., Haas, A., Lehmann, C., Brand, V., Bernicke, D., Honold, K., and Lorenz, S. (2016). Rapid activation of bone morphogenic protein 9 by receptor-mediated displacement of pro-domains. *J. Biol. Chem.* *291*, 3395–3410.
- Lebrin, F., Goumans, M.-J., Jonker, L., Carvalho, R.L.C., Valdimarsdottir, G., Thorikay, M., Mummery, C., Arthur, H.M., and ten Dijke, P. (2004). Endoglin promotes endothelial cell proliferation and TGF- β /ALK1 signal transduction. *EMBO J.* *23*, 4018–4028.
- Li, D.Y., Sorensen, L.K., Brooke, B.S., Urness, L.D., Davis, E.C., Taylor, D.G., Boak, B.B., and Wendel, D.P. (1999). Defective angiogenesis in mice lacking endoglin. *Science* *284*, 1534–1537.
- Lin, S.J., Hu, Y., Zhu, J., Woodruff, T.K., and Jardetzky, T.S. (2011). Structure of betaglycan zona pellucida (ZP)-C domain provides insights into ZP-mediated protein polymerization and TGF- β binding. *Proc. Natl. Acad. Sci. USA* *108*, 5232–5236.
- Litscher, E.S., and Wassarman, P.M. (2015). A guide to zona pellucida domain proteins (Hoboken, New Jersey: John Wiley & Sons, Inc.).
- Llorca, O., Trujillo, A., Blanco, F.J., and Bernabeu, C. (2007). Structural model of human endoglin, a transmembrane receptor responsible for hereditary hemorrhagic telangiectasia. *J. Mol. Biol.* *365*, 694–705.
- Mahlawat, P., Ilangovan, U., Biswas, T., Sun, L.Z., and Hinck, A.P. (2012). Structure of the Alk1 extracellular domain and characterization of its bone morphogenetic protein (BMP) binding properties. *Biochemistry* *51*, 6328–6341.
- Mallet, C., Lamribet, K., Giraud, S., Dupuis-Girod, S., Feige, J.J., Bailly, S., and Tillet, E. (2015). Functional analysis of endoglin mutations from hereditary hemorrhagic telangiectasia type 1 patients reveals different mechanisms for endoglin loss of function. *Hum. Mol. Genet.* *24*, 1142–1154.
- McAllister, K.A., Grogg, K.M., Johnson, D.W., Gallione, C.J., Baldwin, M.A., Jackson, C.E., Helmbold, E.A., Markel, D.S., McKinnon, W.C., Murrell, J., et al. (1994). Endoglin, a TGF- β binding protein of endothelial cells, is the gene for hereditary haemorrhagic telangiectasia type 1. *Nat. Genet.* *8*, 345–351.
- Mi, L.Z., Brown, C.T., Gao, Y., Tian, Y., Le, V.Q., Walz, T., and Springer, T.A. (2015). Structure of bone morphogenetic protein 9 procomplex. *Proc. Natl. Acad. Sci. USA* *112*, 3710–3715.
- Miller, D.W., Graulich, W., Karges, B., Stahl, S., Ernst, M., Ramaswamy, A., Sedlacek, H.H., Müller, R., and Adamkiewicz, J. (1999). Elevated expression of endoglin, a component of the TGF- β -receptor complex, correlates with proliferation of tumor endothelial cells. *Int. J. Cancer* *81*, 568–572.
- Nithianantham, S., Le, S., Seto, E., Jia, W., Leary, J., Corbett, K.D., Moore, J.K., and Al-Bassam, J. (2015). Tubulin cofactors and Arl2 are cage-like chaperones that regulate the soluble $\alpha\beta$ -tubulin pool for microtubule dynamics. *eLife* *4*, e08811.
- Nolan-Stevaux, O., Zhong, W., Culp, S., Shaffer, K., Hoover, J., Wickramasinghe, D., and Ruefli-Brasse, A. (2012). Endoglin requirement for BMP9 signaling in endothelial cells reveals new mechanism of action for selective anti-endoglin antibodies. *PLoS ONE* *7*, e50920.
- Ohta, M., Greenberger, J.S., Anklesaria, P., Bassols, A., and Massagué, J. (1987). Two forms of transforming growth factor- β distinguished by multipotential haematopoietic progenitor cells. *Nature* *329*, 539–541.
- Paauwe, M., Heijkants, R.C., Oudt, C.H., van Pelt, G.W., Cui, C., Theuer, C.P., Hardwick, J.C.H., Sier, C.F.M., and Hawinkels, L.J.A.C. (2016). Endoglin targeting inhibits tumor angiogenesis and metastatic spread in breast cancer. *Oncogene* *35*, 4069–4079.
- Paquet, M.E., Pece-Barbara, N., Vera, S., Cymerman, U., Karabegovic, A., Shovlin, C., and Letarte, M. (2001). Analysis of several endoglin mutants reveals no endogenous mature or secreted protein capable of interfering with normal endoglin function. *Hum. Mol. Genet.* *10*, 1347–1357.

- Pece-Barbara, N., Cymerman, U., Vera, S., Marchuk, D.A., and Letarte, M. (1999). Expression analysis of four endoglin missense mutations suggests that haploinsufficiency is the predominant mechanism for hereditary hemorrhagic telangiectasia type 1. *Hum. Mol. Genet.* *8*, 2171–2181.
- Quackenbush, E.J., and Letarte, M. (1985). Identification of several cell surface proteins of non-T, non-B acute lymphoblastic leukemia by using monoclonal antibodies. *J. Immunol.* *134*, 1276–1285.
- Reeves, P.J., Callewaert, N., Contreras, R., and Khorana, H.G. (2002). Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. *Proc. Natl. Acad. Sci. USA* *99*, 13419–13424.
- Rossi, E., Lopez-Novoa, J.M., and Bernabeu, C. (2015). Endoglin involvement in integrin-mediated cell adhesion as a putative pathogenic mechanism in hereditary hemorrhagic telangiectasia type 1 (HHT1). *Front. Genet.* *5*, 457.
- Scharpfenecker, M., van Dinther, M., Liu, Z., van Bezooijen, R.L., Zhao, Q., Pukac, L., Löwik, C.W.G.M., and ten Dijke, P. (2007). BMP-9 signals via ALK1 and inhibits bFGF-induced endothelial cell proliferation and VEGF-stimulated angiogenesis. *J. Cell Sci.* *120*, 964–972.
- She, X.W., and Seon, B.K. (2001). Epitope mapping of endoglin, a TGF- β receptor, using recombinant fragments and twelve monoclonal antibodies. *Proc. Am. Assoc. Cancer Res.* *42*, 825.
- Stenson, P.D., Mort, M., Ball, E.V., Shaw, K., Phillips, A., and Cooper, D.N. (2014). The Human Gene Mutation Database: building a comprehensive mutation repository for clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum. Genet.* *133*, 1–9.
- Townson, S.A., Martinez-Hackert, E., Greppi, C., Lowden, P., Sako, D., Liu, J., Ucran, J.A., Liharska, K., Underwood, K.W., Sehra, J., et al. (2012). Specificity and structure of a high affinity activin receptor-like kinase 1 (ALK1) signaling complex. *J. Biol. Chem.* *287*, 27313–27325.
- Upton, P.D., Davies, R.J., Trembath, R.C., and Morrell, N.W. (2009). Bone morphogenetic protein (BMP) and activin type II receptors balance BMP9 signals mediated by activin receptor-like kinase-1 in human pulmonary artery endothelial cells. *J. Biol. Chem.* *284*, 15794–15804.
- Van Le, B., Franke, D., Svergun, D.I., Han, T., Hwang, H.Y., and Kim, K.K. (2009). Structural and functional characterization of soluble endoglin receptor. *Biochem. Biophys. Res. Commun.* *383*, 386–391.
- Venkatesha, S., Toporsian, M., Lam, C., Hanai, J., Mammoto, T., Kim, Y.M., Bdolah, Y., Lim, K.H., Yuan, H.T., Libermann, T.A., et al. (2006). Soluble endoglin contributes to the pathogenesis of preeclampsia. *Nat. Med.* *12*, 642–649.
- Wooderchak-Donahue, W.L., McDonald, J., O'Fallon, B., Upton, P.D., Li, W., Roman, B.L., Young, S., Plant, P., Fülöp, G.T., Langa, C., et al. (2013). BMP9 mutations cause a vascular-anomaly syndrome with phenotypic overlap with hereditary hemorrhagic telangiectasia. *Am. J. Hum. Genet.* *93*, 530–537.

Cell Reports, Volume 19

Supplemental Information

Structural Basis of the Human Endoglin-BMP9

Interaction: Insights into BMP Signaling and HHT1

Takako Saito, Marcel Bokhove, Romina Croci, Sara Zamora-Caballero, Ling Han, Michelle Letarte, Daniele de Sanctis, and Luca Jovine

This file includes:

SUPPLEMENTAL FIGURES

Figure S1. ENG Organization and Expression Constructs

Figure S2. Crystal Structure of mMBP-Fused ENG OR

Figure S3. Relationship between OR1 and OR2 and Possible Evolution of ENG OR

Figure S4. Structure of ENG OR in Complex with BMP9

Figure S5. Crystal Structure of ENG ZP

Figure S6. Sequence Alignment of Human ENG and BG

Figure S7. Mapping of HHT1 Missense Mutations onto the Structure of ENG

Figure S8. BLI Curve Fitting

SUPPLEMENTAL TABLE

Table S1. X-Ray Data Collection and Refinement Statistics

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

X-Ray Diffraction Data Collection

Structure Determination

Sequence and Structure Analysis

BLI Analysis of Protein-Protein Interaction

SEC-MALS Analysis of the ENG-BMP9 Complex

SUPPLEMENTAL REFERENCES

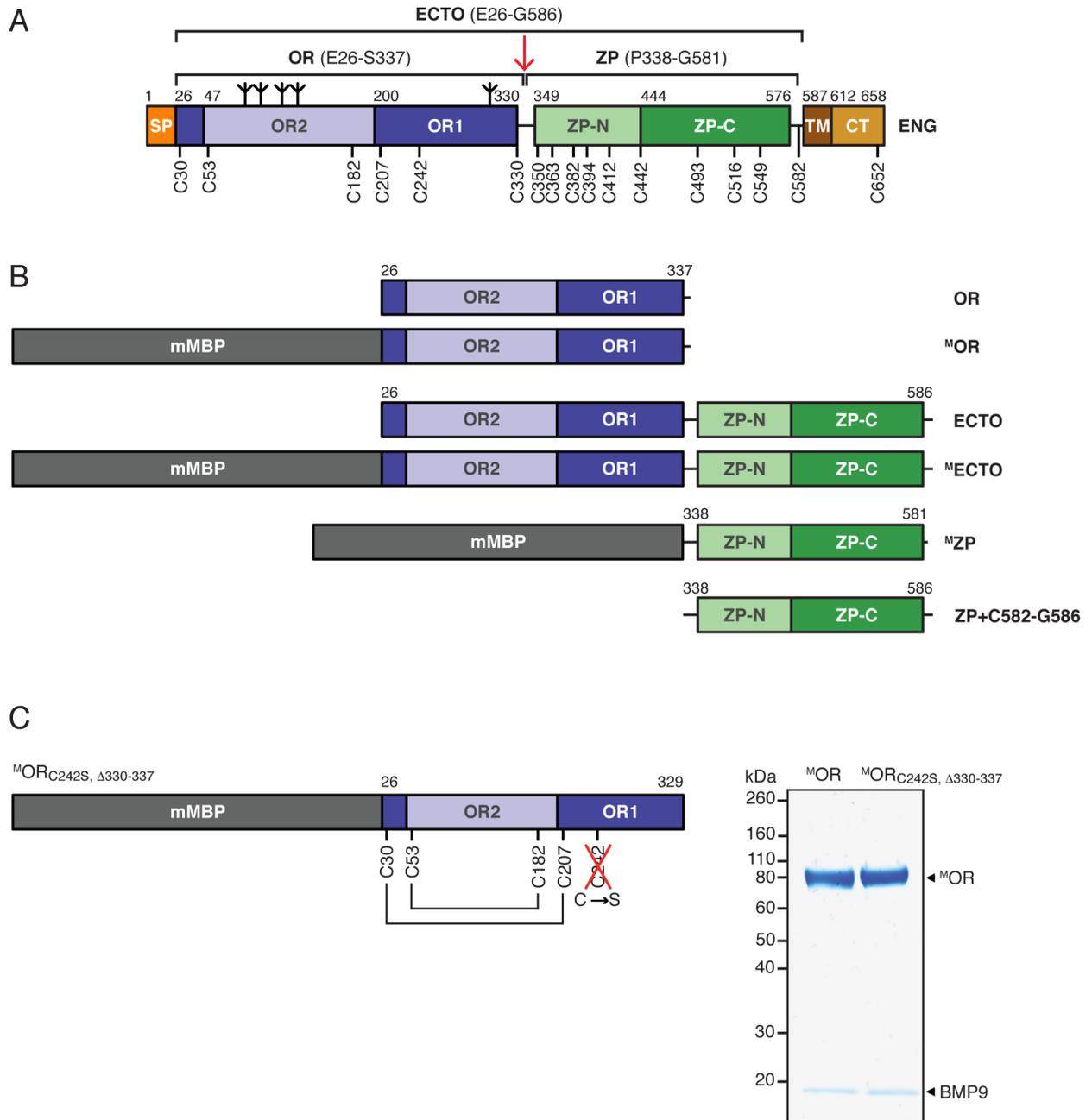


Figure S1. ENG Organization and Expression Constructs. Related to Figures 1, 3 and 4.

(A) Domain architecture of ENG. SP, signal peptide; TM, transmembrane domain; CT, cytoplasmic domain. N-glycosylation sites are indicated by inverted tripods and the OR/ZP interdomain linker is marked by a red arrow.

(B) ENG constructs used for crystallization and functional assays.

(C) Non-reducing SDS-PAGE after BMP9 co-expression followed by His pull-down indicates that both ^MOR and ^MOR_{C242S, Δ330-337} are secreted as monomers and bind BMP9.

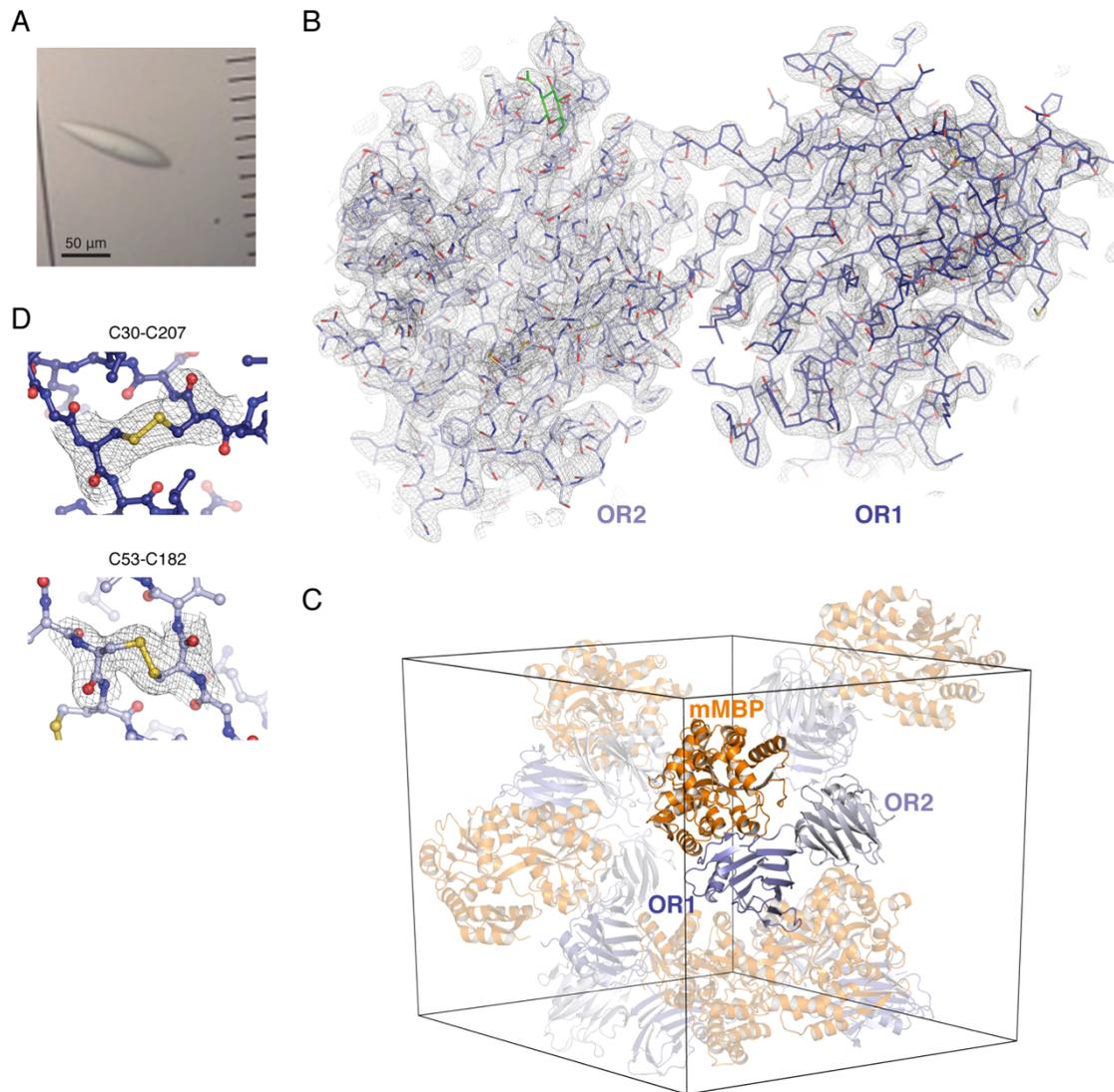


Figure S2. Crystal Structure of mMBP-Fused ENG OR. Related to Figure 1.

(A) Crystal of ^MOR.

(B) Stick representation of ENG OR1 and OR2 domains, colored dark blue and light blue, respectively. The $2mFo-DFc$ electron density map contoured at 1.0σ is shown as a black mesh.

(C) Crystal packing of mMBP-OR. mMBP is orange, OR1 and OR2 are colored as in panel B. The unit cell is shown, and symmetry-related molecules are semi-transparent.

(D) Detail of OR disulfide bonds and corresponding electron density map, using the same representation as in panel B.

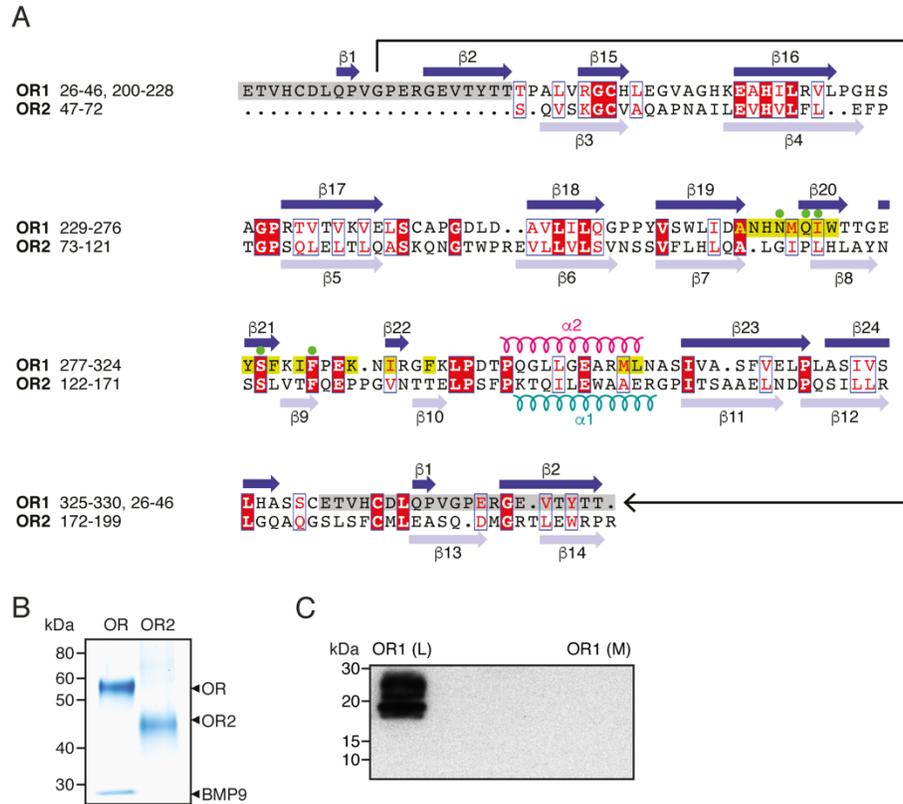


Figure S3. Relationship between OR1 and OR2 and Possible Evolution of ENG OR.

Related to Figure 1.

(A) Structure-based sequence alignment of OR1 and OR2, showing that the N-terminal part of OR1 (gray box) aligns with the C-terminal part of OR2. Secondary structure elements are shown; conserved residues are highlighted in red, residues within 4.5 Å of BMP9 are indicated in yellow and mutagenized sites are marked by green circles.

(B) Coomassie-stained gel of BMP9 pull-down with purified OR or OR2 (ENG T45-L203). BMP9 binds OR but not OR2, narrowing down the ligand binding site to OR1.

(C) Anti-His immunoblot shows that OR1 (ENG E26-T46-(GGGS)₂-L203-S329 C242S) is expressed but not secreted by mammalian cells. L, cell lysate; M, conditioned medium.

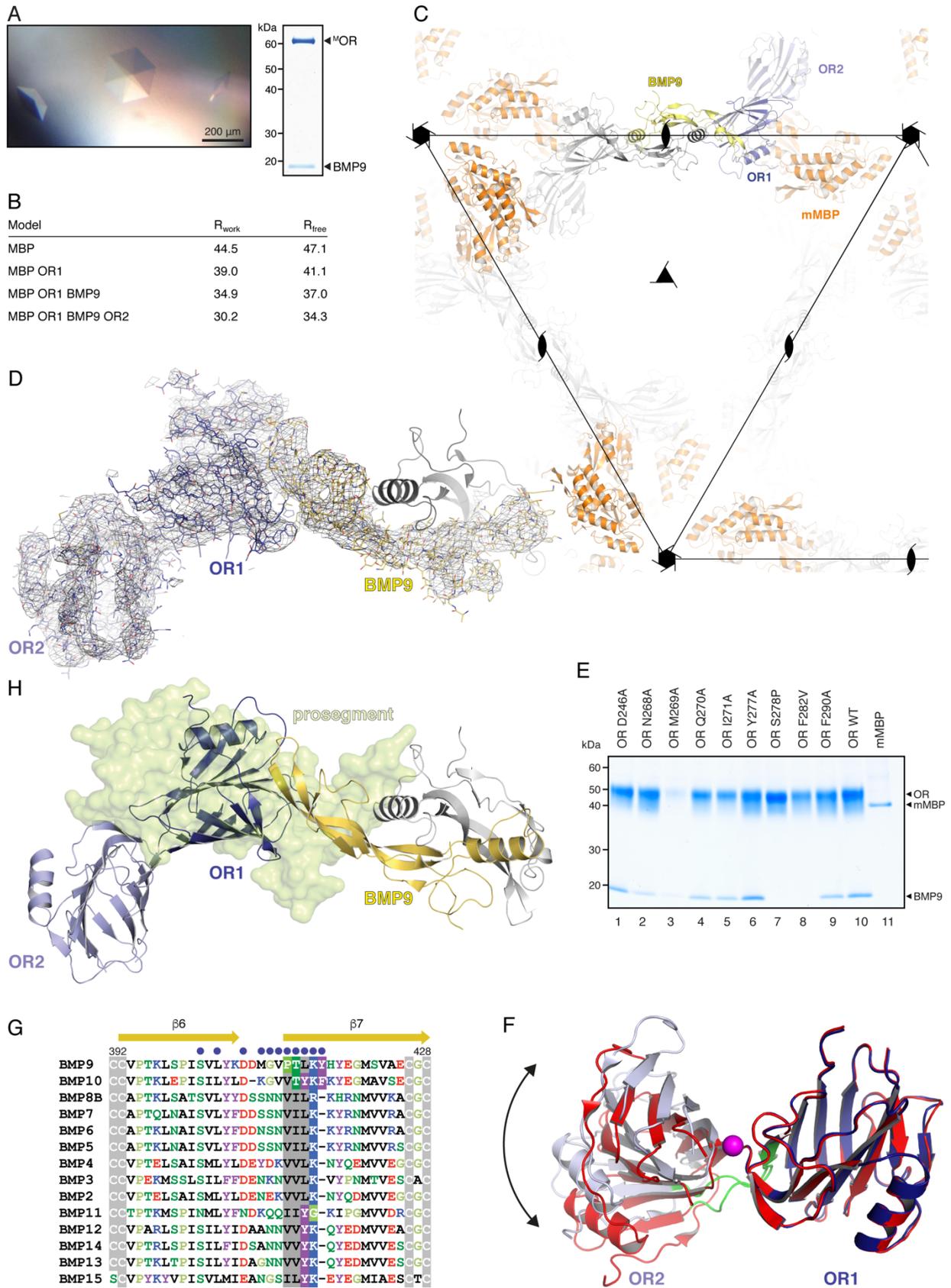


Figure S4. Structure of ENG OR in Complex with BMP9. Related to Figure 2.

(A) Hexagonal crystals of ^MOR-BMP9 (left). Despite their perfect shape, the complex crystals diffract very weakly and several hundred specimens were screened in order to obtain the 4.45 Å resolution data used for structure determination. Coomassie-stained SDS-PAGE analysis (right) indicates that the crystals contain both ^MOR and BMP9.

(B) Refinement statistics (phenix.refine R-factors) improve when additional model components are added.

(C) Correct placement of the ^MOR-BMP9 complex (OR1: dark blue; OR2: light blue; BMP9: yellow) in the unit cell, together with the indicated symmetry elements, generates the typical BMP9 homodimer (grey/yellow). ENG OR and BMP9 molecules in neighboring asymmetric units are shown as transparent light grey cartoons; mMBP moieties are orange. Note how the BMP9-bound mMBP-OR fusion protein packs completely differently from its unbound counterpart (Figure S2C).

(D) $2mFo-DFc$ electron density map shown for one asymmetric unit of the complex contoured at 1.0 σ . Proteins are shown in stick representation and colored as in panel C. The second BMP9 molecule of the dimer is grey.

(E) Coomassie-stained SDS-PAGE analysis of BMP9 pull-down by non-mMBP-fused, C-terminally 6His-tagged OR and mutants thereof. A single batch of conditioned medium from HEK293T cells transfected with BMP9 alone was used as input for all pull-down experiments; C-terminally 6His-tagged mMBP was used as a negative control.

(F) Comparison of free and BMP9-bound OR structures (dark blue/light blue and red, respectively) reveals a ~20-degree rotation of OR2 relative to OR1. Bending residues

(Y44-T46, W196-V204) are green and the rotation axis, which is perpendicular to the plane of the page, is magenta.

(G) Sequence alignment of human BMPs. Secondary structure is marked; residues within 4.5 Å of ENG are indicated with blue circles.

(H) Superposition of the BMP9 precursor ([Mi et al., 2015](#)) with BMP9 bound to ENG OR indicates that the prosegment of the precursor is not compatible with ENG binding. OR-BMP9 is oriented as in panel D, with OR1, OR2 and mature BMP9 colored as in panels C and D. The BMP9 prosegment is shown as a yellow surface.

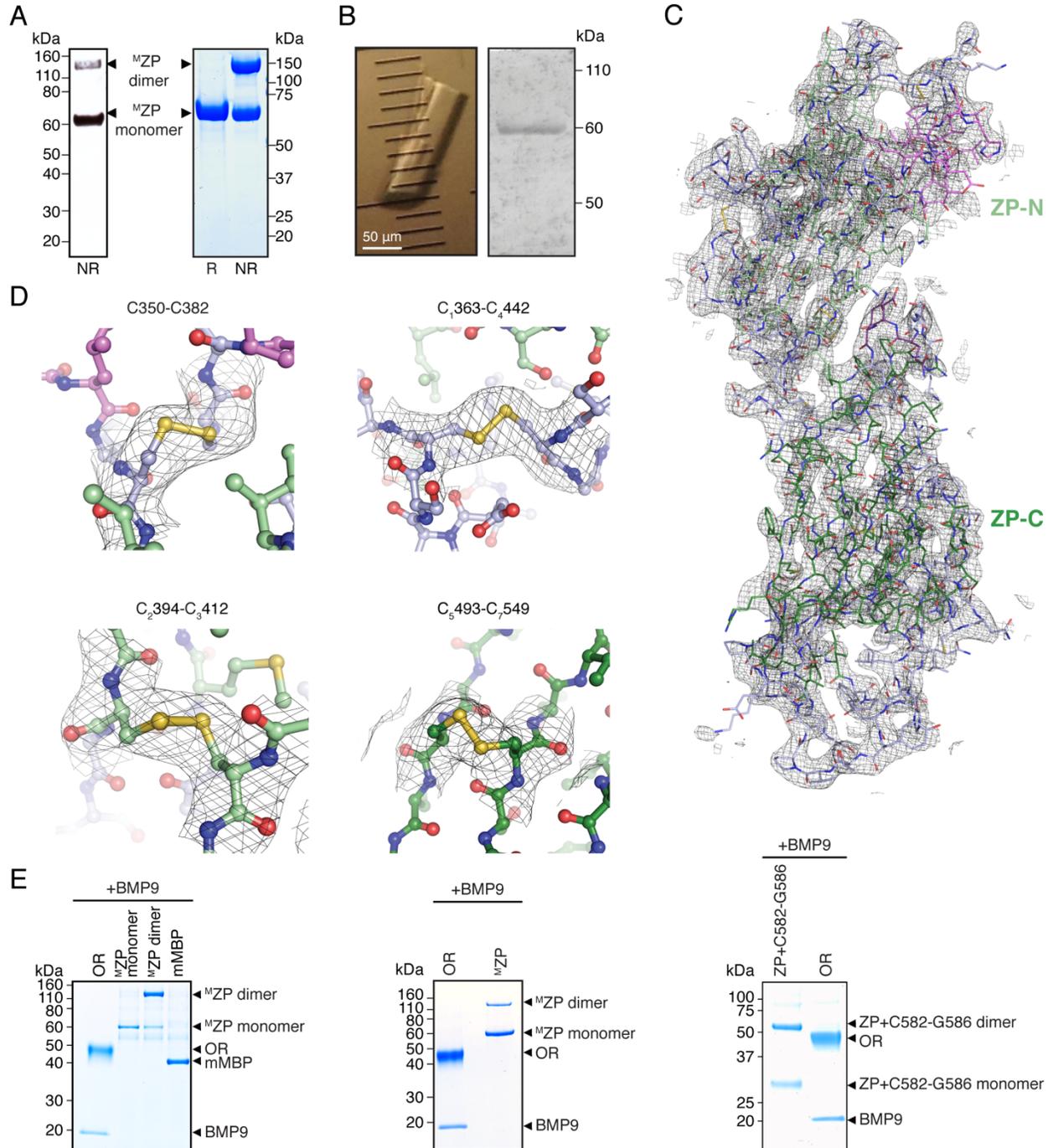


Figure S5. Crystal Structure of ENG ZP. Related to Figure 3.

(A) Anti-His immunoblot of ^MZP secreted in conditioned medium and Coomassie-stained SDS-PAGE of purified ^MZP. Although the protein is mostly secreted as a monomer (see also [Figure 4B](#), lane 2), after concentration it is present as a roughly equimolar mixture of monomer and intermolecularly disulfide-bonded homodimer.

(B) Crystal of ^MZP and non-reducing silver-stained SDS-PAGE analysis of washed ^MZP crystals, indicating that they only consist of monomeric protein.

(C) *2mFo-DFc* electron density map of ZP contoured at 1.0 σ . The model is shown as sticks. Loops are light blue; ZP-N strands and helices are pale green and pink, respectively; ZP-C strands are dark green.

(D) Details of ENG ZP disulfide bonds, represented as in [Figure S2D](#) and colored as in panel B. Both signature ZP module disulfides (C₁-C₄, C₂-C₃ and C₅-C₇) and the additional disulfide found in ZP-N (C350-C382) are shown.

(E) In contrast to ^MOR, ^MZP does not bind BMP9, as indicated by a pull-down with IMAC and SEC-purified monomeric and dimeric ^MZP (left) or co-expression followed by IMAC pull-down (middle). mMBP does not interfere with a possible interaction between ZP and BMP9, because co-expression of unfused ZP+C582-G586 and BMP9 followed by His pull-down also does not detect binding between the proteins (right).

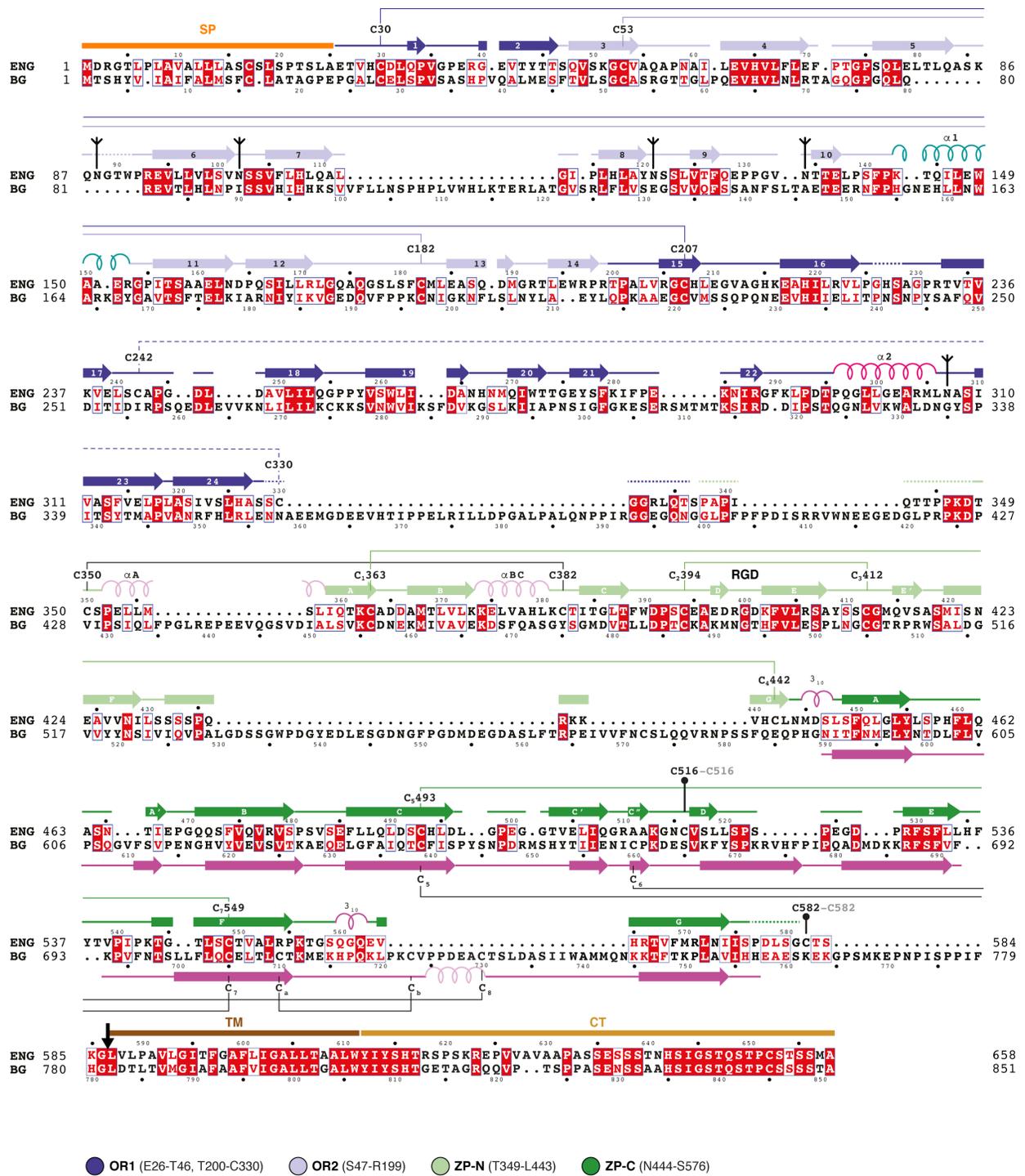
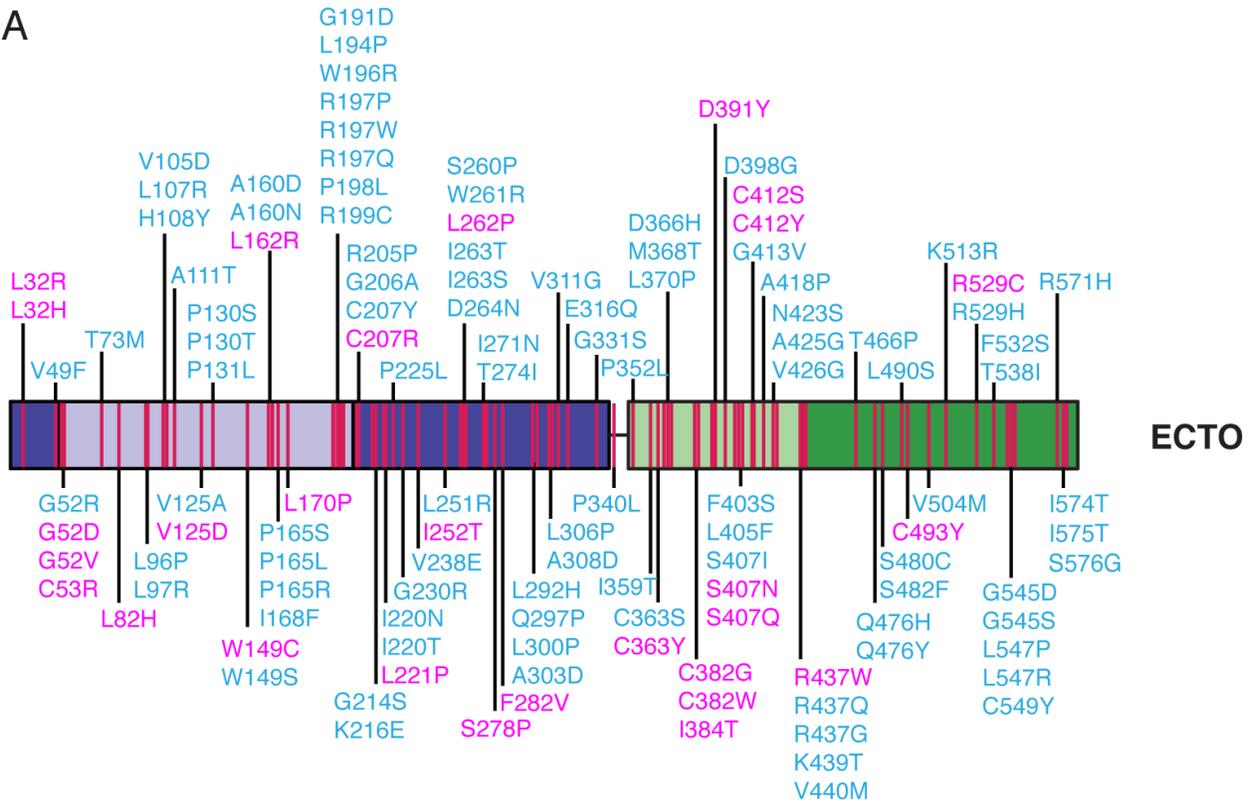


Figure S6. Sequence Alignment of Human ENG and BG. Related to Figures 1, 3 and 4.

Conserved cysteine residues and experimentally determined secondary structure elements are indicated (human ENG: this work; rat BG ZP-C domain: [Lin et al., 2011](#)). The ENG signal peptide and transmembrane domain are indicated by an orange line and a brown line, respectively; the MMP-14 cleavage site ([Hawinkels et al., 2010](#)) is indicated by a black arrow. Inverted tripods and lollipops represent ENG N-glycans and Cys residues involved in the C516-C516 and C582-C582 intermolecular disulfides of homodimeric ENG, respectively. Dashed lines within the secondary structure indicate residues that are included in the crystallized constructs but are not resolved in the corresponding electron density maps.

A



B

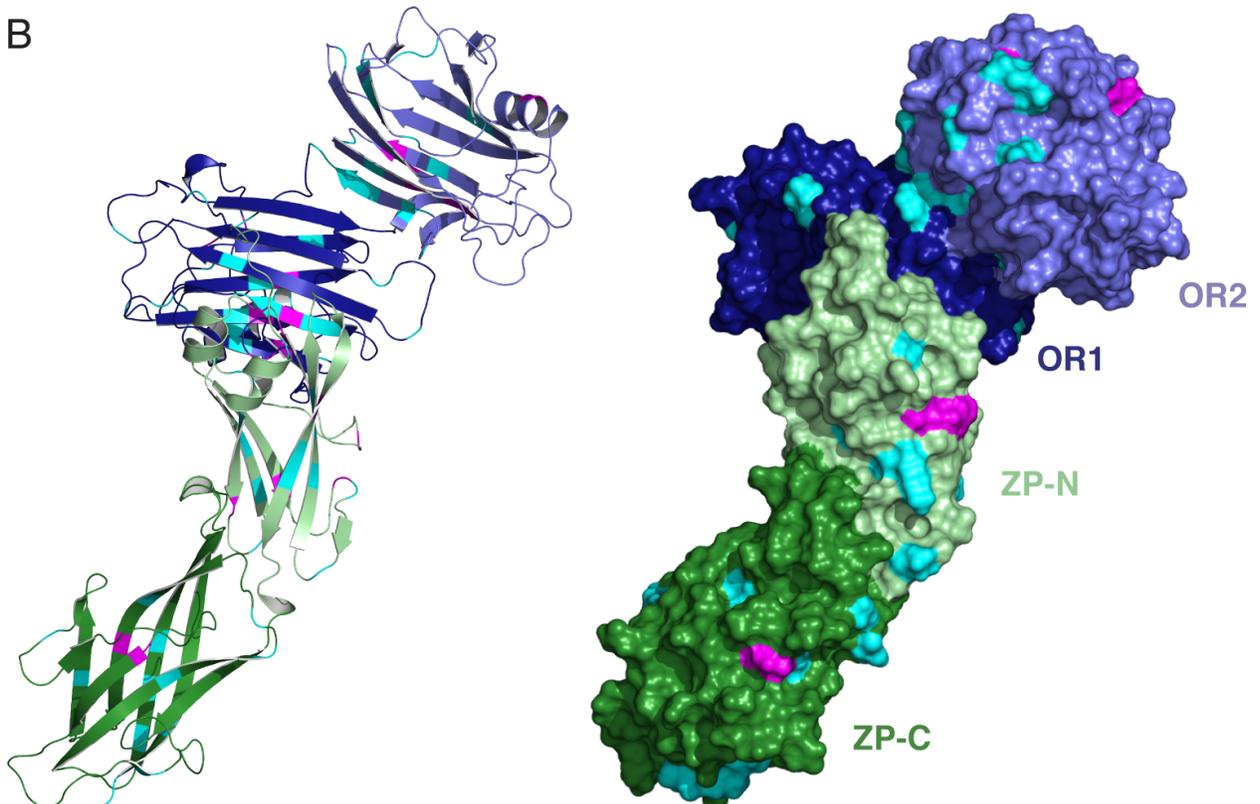


Figure S7. Mapping of HHT1 Missense Mutations onto the Structure of ENG.

Related to Figure 6.

(A) Domain architecture of ENG, showing the location of missense mutations associated with HHT1 (Ali et al., 2011; Cymerman et al., 2003; Mallet et al., 2014; Paquet et al., 2001; Pece-Barbara et al., 1999; University of Utah Department of Pathology HHT and ENG database). Mutations discussed in the main text and/or included in Figure 6 are highlighted in magenta.

(B) Mapping of the mutations listed in panel A onto a model of full-length ENG, with the same relative orientation of OR and ZP as in Figure 4F. The structure of ENG is shown both in cartoon (left) and surface (right) representations, with mutation sites colored as in panel A.

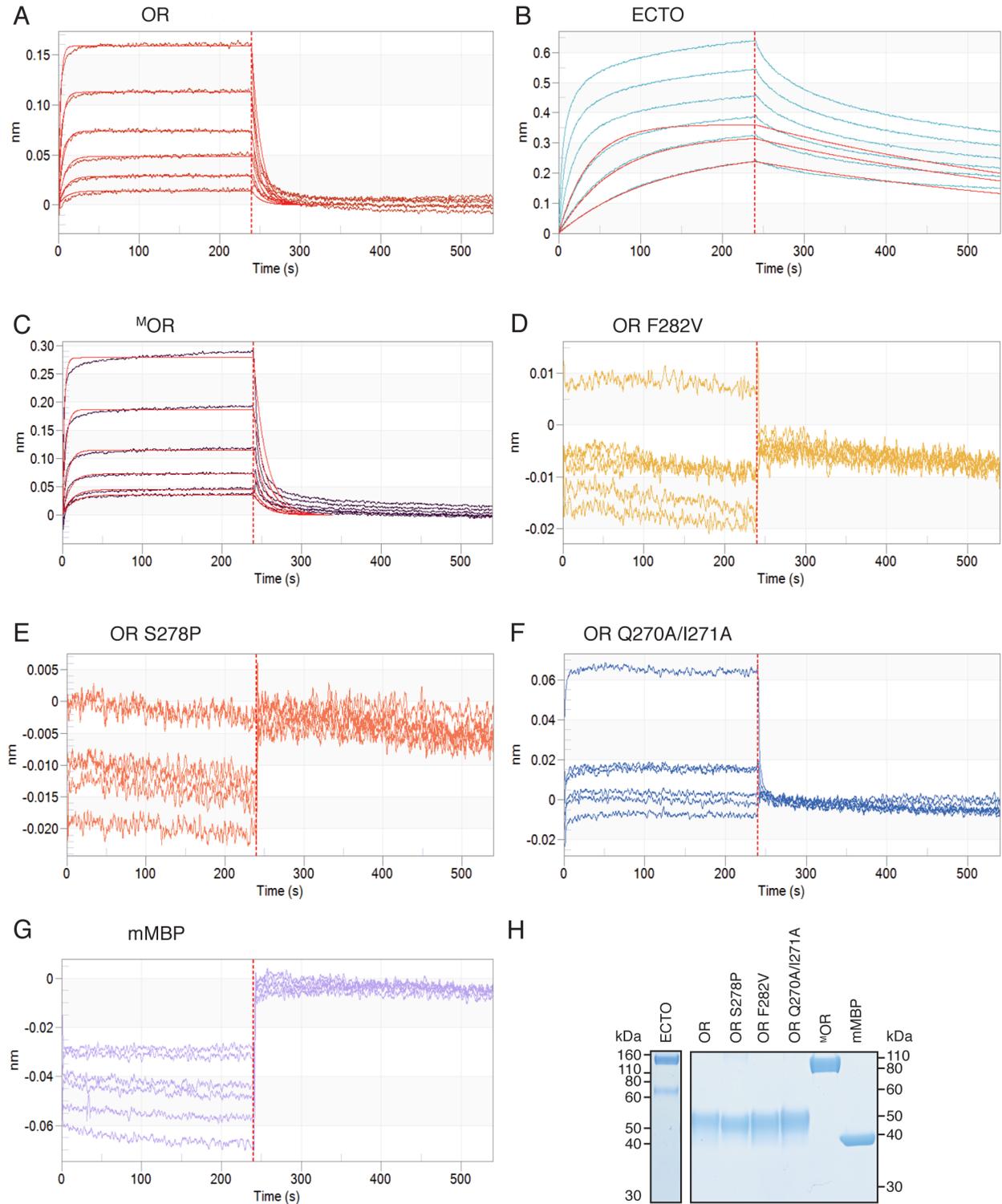


Figure S8. BLI Curve Fitting. Related to Table 1.

(A-G) Association and dissociation curves for the interaction of OR (A), ECTO (B), ^MOR (C), OR F282V (D), OR S278P (E) OR Q270A/I271A (F) and mMBP (G) with BMP9 obtained by BLI. Kinetic parameters for the interaction (K_D , k_{on} and k_{off}) were determined by global fitting (red curves) using 1:1 stoichiometry.

(H) Non-reducing SDS-PAGE analysis of purified proteins used for BLI measurements.

Table S1. X-Ray Data Collection and Refinement Statistics. Related to Figures 1-3.

Crystal PDB ID	ENG OR 5I04	BMP9 5I05	ENG OR-BMP9 Complex 5HZW	ENG ZP 5HZV
Experiment				
Beamline	ESRF ID29	DLS I02	DLS I02	ESRF ID23-1
Wavelength, Å	0.9762	0.9795	0.9794	1.0000
Number of crystals	1	1	1	1
Data Collection				
Space group	<i>P</i> 6 ₅ (170)	<i>I</i> 4 ₂₂ (98)	<i>P</i> 6 ₄ (172)	<i>P</i> 6 ₅ (170)
Unit cell dimensions				
<i>a</i> , <i>b</i> , <i>c</i> , Å	114.15, 114.15, 120.62	71.06 71.06 145.67	211.24, 211.24, 53.16	125.16, 125.16, 88.54
α , β , γ , °	90, 90, 120	90, 90, 90	90, 90, 120	90, 90, 120
Molecules/A.U.	1	1	1 + 1 (1 complex)	1
Solvent content, %	57.8	66.9	67.6	56.8
Wilson B factor, Å ²	58.26	32.58	219	64.09
Resolution range, Å	45.7-2.42 (2.51-2.42)	41.4-1.87 (1.93-1.87)	51.1-4.45 (4.61-4.45)	28.5-2.70 (2.81-2.70)
Total reflections	113614 (11586)	90525 (8280)	81763 (8123)	70991 (7028)
Unique reflections	33737 (3337)	15763 (1396)	8634 (838)	21504 (2328)
Multiplicity	3.4 (3.5)	5.7 (5.9)	9.5 (9.7)	3.3 (3.0)
Completeness, %	99.0 (98.4)	99.4 (99.7)	99.7 (99.4)	98.4 (96.3)
Mean <i>I</i> / σ (<i>I</i>)	11.8 (1.3)	16.0 (1.4)	3.8 (0.5)	7.7 (0.7)
<i>R</i> _{pim} , %	4.6 (84.2)	3.3 (84.3)	12.0 (134.1)	6.2 (68.9)
CC(1/2), %	99.9 (54.1)	99.9 (52.8)	99.7 (25.9)	99.7 (46.3)
CC*, %	100.0 (83.8)	100.0 (83.1)	99.9 (64.1)	99.9 (79.5)
Refinement				
Reflections	33699 (3333)	15756 (1396)	8611 (833)	21434 (2319)
Free reflections	1636 (153)	1576 (139)	847 (70)	1140 (110)
<i>R</i> _{work} , %	22.58 (39.39)	21.77 (36.51)	28.61 (36.69)	23.38 (40.81)
<i>R</i> _{free} , %	26.33 (43.90)	23.59 (36.94)	31.83 (41.89)	27.02 (42.03)
CC _{work} , %	95.7 (67.4)	95.0 (68.9)	96.5 (50.0)	96.8 (61.3)
CC _{free} , %	94.5 (77.8)	93.8 (59.2)	90.3 (25.5)	93.7 (54.8)
ML coordinate error, Å	0.47	0.23	0.66	0.51
ML phase error, °	35.15	26.53	39.52	36.22
rmsd				
Bond lengths, Å	0.004	0.006	0.005	0.004
Bond angles, °	0.97	1.08	1.01	0.96
Ramachandran plot				
Favored, %	98.6	99.1	96.6	99.3
Allowed %	1.4	0.9	3.4	0.7
Outlier, %	0	0	0	0
Rotamer outliers, %	0.2	0	0.9	0.4
Clashscore	2.9	0.6	5.3	1.3
No of atoms				
Total	5210	901	5985	4626
Protein	5117	837	5934	4582
Ligand	67	12	51	29
Water	26	52	0	15
Protein residues	668	107	773	598
Average B factor, Å ²				
Total	76.6	51.4	256.0	84.8
Protein	76.5	51.5	256.1	85.0
Ligand	94.2	61.2	251.9	64.3
Water	47.7	48.5	-	71.8

Parameters for the outermost shell are shown in parentheses. A.U., asymmetric unit; CC(1/2), percentage of correlation between intensities from random half-datasets; CC*, $\sqrt{(2CC(1/2)+CC(1/2))}$; CC_{work}, correlation of the experimental intensities of free reflections excluded from the refinement with the intensities calculated from the refined molecular model; CC_{free}, correlation of the experimental intensities with the intensities calculated from the refined molecular model; DLS, Diamond Light Source; ESRF, European Synchrotron Radiation Facility; *I*/ σ *I*, signal-to-noise ratio; ML, maximum likelihood; *R*_{pim}, $\sum_{hkl} \sqrt{(1/n-1)} \sum |I(hkl) - \langle I(hkl) \rangle| / \sum I(hkl)$, where *I*(*hkl*) is the intensity of a reflection and $\langle I(hkl) \rangle$ is the average intensity of all symmetry-related observations of a reflection; *R*_{free}, $\sum |F_o - k|F_c| / \sum |F_o|$; *R*_{work}, same as *R*_{free} calculated from free reflections excluded from refinement. Statistics were calculated using phenix.table_one, phenix.xtriage (solvent content) and phenix.refine (ML coordinate and phase errors).

SUPPLEMENTAL EXPERIMENTAL PROCEDURES

X-Ray Diffraction Data Collection

Datasets of ^MOR and ^MZP were collected at 100 K at beamlines ID29 ([de Sanctis et al., 2012](#)) and ID23-1 ([Nurizzo et al., 2006](#)) of the European Synchrotron Radiation Facility (ESRF, Grenoble), equipped with PILATUS 6M-F detectors (Dectris). BMP9 and ^MOR-BMP9 complex datasets were collected at 100 K at beamline I02 of Diamond Light Source (DLS, Didcott), also equipped with a PILATUS 6M detector. Data was integrated and scaled with XDS ([Kabsch, 2010](#)), using high resolution data cutoffs chosen based on statistical indicators $CC_{1/2}$ and CC^* ([Evans and Murshudov, 2013](#); [Karplus and Diederichs, 2015](#)). Processing statistics are summarized in [Table S1](#).

Structure Determination

The structure of ^MOR was solved by MR with Phaser ([McCoy et al., 2007](#)), using as search model an ensemble of MBP structures extracted from PDB entries 3SEX, 3SET ([Laganowsky et al., 2011](#)) and 4WRN ([Bokhove et al., 2016a, 2016b](#)). After autobuilding with PHENIX AutoBuild ([Terwilliger et al., 2008](#)) and manual building in Coot ([Emsley et al., 2010](#)), the model was refined using phenix.refine ([Afonine et al., 2012](#)). Protein and carbohydrate structure validation was performed with MolProbity ([Chen et al., 2010](#)) and Privateer ([Agirre et al., 2015](#)), respectively.

PDB entry 1ZKZ (Brown et al., 2005) was used to phase the data of BMP9 by MR. Rebuilding, refinement and validation were carried out as described for ^MOR.

The structure of the ^MOR-BMP9 complex was also determined by MR with Phaser. Despite the limited resolution of the dataset, correctness of the solution was indicated by a progressive decrease in refinement R-factors every time a new model component was added to the structure (Figure S4B). An independent indicator of the correct MR result was the placement of the BMP9 molecule in the unit cell which, together with 2-fold crystallographic symmetry, generated the known BMP9 homodimer (Brown et al., 2005) (Figures S4C and D). Due to the low resolution of the data, a conservative refinement strategy was followed, consisting of strictly constrained coordinate refinement together with group and translation, libration and screw (TLS) B-factor refinement (Winn et al., 2001), using one TLS group per domain.

MR of ^MZP was performed using PDB entry 3SEX (Laganowsky et al., 2011) as a search model. After obtaining initial MR phases, density modification was carried out with RESOLVE (Terwilliger, 2004) in the PHENIX package (Adams et al., 2010). Model building was performed with Buccaneer (Cowtan, 2006) and Coot. Refinement and validation were performed as described for ^MOR.

Refinement statistics are summarized in Table S1.

Sequence and Structure Analysis

Sequence alignments were calculated using T-Coffee (Notredame et al., 2000) and compiled with ESPript (Robert and Gouet, 2014) or SeaView4 (Gouy et al., 2010).

The circular permutation of ENG OR was first recognized using Chimera (Pettersen et al., 2004). Superposition of OR1 and OR2 was performed using PDBefold (Krissinel and Henrick, 2005), which was also used to search for structural homologues of the OR domains.

PDBePISA analysis (Krissinel and Henrick, 2007) indicates that the interface between OR and BMP9 plays a significant role in complex formation (complex formation significance score (CSS) = 0.648). However, in both the OR and the OR-BMP9 structures the CSS of the interface between mMBP and OR is 0.000, indicating that there is no significant interface between the two moieties of the fusion protein. Furthermore, the packing of mMBP and OR is different between the liganded and unliganded state, which suggests that mMBP does not interfere with the structure of OR.

Solvent accessible area calculations and domain motion analysis were carried out with Naccess (Hubbard and Thornton, 1993) and DynDom (Hayward and Lee, 2002), respectively.

Figures were created with PyMOL (Schrödinger, LLC). The ENG surface in Figure 2C was colored according to the normalized consensus hydrophobicity scale (Eisenberg et al., 1984).

BLI Analysis of Protein-Protein Interaction

An Octet Red96 system (PALL FortéBio) was used to measure binding constants of recombinant human BMP9 (carrier free; R&D Systems) to purified ENG constructs and an mMBP control expressed in HEK293T cells (Figure S8H).

Before immobilization to High Precision Streptavidin (SAX) Dip and Read Biosensors (PALL FortéBio) hydrated in sample buffer (20 mM Na-HEPES pH 7.4, 150 mM NaCl, 0.05% Tween), BMP9 was biotinylated on ice for 2 hours and diluted to 5 $\mu\text{g/ml}$. Subsequently, sensors were blocked with biocytin (10 $\mu\text{g/ml}$) for 1 minute to avoid nonspecific binding. In order to calculate K_D , binding in the presence of increasing concentrations of each ENG construct (0 nM, 37.5 nM, 75 nM, 125 nM, 250 nM, 500 nM and 1 μM) was monitored. Two 60-second and 120-second baseline measurements were performed prior to association measurements. Association and dissociation were monitored during 240 and 300 seconds respectively. Nonspecific binding was monitored using a sensor not coated with BMP9 and ENG constructs at 125 nM concentration. Contribution of the buffer to the signal was measured with a BMP9-coated sensor. Both controls were subtracted from the binding signal in order to calculate K_D values.

SEC-MALS Analysis of the ENG-BMP9 Complex

The chromatographic system consisted of an HPLC system and autoinjector (Agilent 1260 Infinity, Agilent Technologies) coupled online with a MALS detector (miniDawn Treos, Wyatt Technology; 658 nm wavelength), a dRI detector (Optilab rEX, Wyatt

Technology; 658 nm wavelength) and a UV detector (Agilent 1260 Infinity, Agilent Technologies; 280 nm wavelength). Separation was performed by a Superdex 200 Increase 10/300 GL column (GE Healthcare), using a flow rate of 0.75 ml/min and a mobile phase consisting of 50 mM Na-HEPES pH 7.5, 25 mM KCl, 20 mM MgCl₂, 200 ppm NaN₃. Data processing and weight-averaged molecular mass calculation were performed using the ASTRA 7.0.2 software (Wyatt Technology).

SUPPLEMENTAL REFERENCES

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W., et al. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66, 213–221.

Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H., and Adams, P.D. (2012). Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* 68, 352–367.

Agirre, J., Iglesias-Fernandez, J., Rovira, C., Davies, G.J., Wilson, K.S., and Cowtan, K.D. (2015). Privateer: software for the conformational validation of carbohydrate structures. *Nat. Struct. Mol. Biol.* 22, 833–834.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66, 12–21.

Cowtan, K. (2006). The Buccaneer software for automated model building. 1. Tracing protein chains. *Acta Crystallogr. D Biol. Crystallogr.* 62, 1002–1011.

Cymerman, U., Vera, S., Karabegovic, A., Abdalla, S., and Letarte, M. (2003). Characterization of 17 novel endoglin mutations associated with hereditary hemorrhagic telangiectasia. *Hum. Mutat.* *21*, 482–492.

Eisenberg, D., Schwarz, E., Komaromy, M., and Wall, R. (1984). Analysis of membrane and surface protein sequences with the hydrophobic moment plot. *J. Mol. Biol.* *179*, 125–142.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 486–501.

Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? *Acta Crystallogr. D Biol. Crystallogr.* *69*, 1204–1214.

Gouy, M., Guindon, S., and Gascuel, O. (2010). SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol Biol Evol* *27*(2), 221–224.

Hayward, S., and Lee, R.A. (2002). Improvements in the analysis of domain motions in proteins from conformational change: DynDom version 1.50. *J. Mol. Graph. Model.* *21*, 181–183.

Hubbard, S.J., and Thornton, J.M. (1993). Naccess. Computer Program, Department of Biochemistry and Molecular Biology, University College London.

Kabsch, W. (2010). XDS. *Acta Crystallogr. D Biol. Crystallogr.* *66*, 125–132.

Karplus, P.A., and Diederichs, K. (2015). Assessing and maximizing data quality in macromolecular crystallography. *Curr. Opin. Struct. Biol.* *34*, 60–68.

Krissinel, E., and Henrick, K. (2005). Multiple alignment of protein structures in three dimensions. In *Computational Life Sciences*, M. R. Berthold, R. C. Glen, K. Diederichs, O. Kohlbacher, and I. Fischer, eds. (Springer Berlin Heidelberg), pp. 67–78.

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

Laganowsky, A., Zhao, M., Soriaga, A.B., Sawaya, M.R., Cascio, D., and Yeates, T.O. (2011). An approach to crystallizing proteins by metal-mediated synthetic symmetrization. *Protein Sci.* *20*, 1876–1890.

McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. *J. Appl. Crystallogr.* *40*, 658–674.

Notredame, C., Higgins, D.G., and Heringa, J. (2000). T-Coffee: A novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.* *302*, 205–217.

Nurizzo, D., Mairs, T., Guijarro, M., Rey, V., Meyer, J., Fajardo, P., Chavanne, J., Biasci, J.C., McSweeney, S., and Mitchell, E. (2006). The ID23-1 structural biology beamline at the ESRF. *J. Synchrotron Radiat.* *13*, 227–238.

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

Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* *42*, W320–W324.

de Sanctis, D., Beteva, A., Caserotto, H., Dobias, F., Gabadinho, J., Giraud, T., Gobbo, A., Guijarro, M., Lentini, M., Lavault, B., et al. (2012). ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering. *J. Synchrotron Radiat.* *19*, 455–461.

Terwilliger, T. (2004). SOLVE and RESOLVE: automated structure solution, density modification and model building. *J. Synchrotron Radiat.* *11*, 49–52.

Terwilliger, T.C., Grosse-Kunstleve, R.W., Afonine, P.V., Moriarty, N.W., Zwart, P.H., Hung, L.W., Read, R.J., and Adams, P.D. (2008). Iterative model building, structure refinement and density modification with the PHENIX AutoBuild wizard. *Acta Crystallogr. D Biol. Crystallogr.* *64*, 61–69.

Winn, I., Isupov, M.N., and Murshudov, G.N. (2001). Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D Biol. Crystallogr.* *57*, 122–133.