Published in final edited form as:

Nat Struct Mol Biol. 2005 October ; 12(10): 910–914.

Structure of the calcium-rich signature domain of human thrombospondin-2

C. Britt Carlson^{1,2}, Douglas A. Bernstein², Douglas S. Annis¹, Tina M. Misenheimer¹, Blueleaf A. Hannah¹, Deane F. Mosher^{1,2}, and James L. Keck²

¹Department of Medicine, University of Wisconsin-Madison, 4285B Medical Sciences Center, 1300 University Avenue, Madison, WI 53706.

² Department of Biomolecular Chemistry, University of Wisconsin-Madison, 550 Medical Sciences Center, 1300 University Avenue, Madison, WI 53706

Abstract

Thrombospondins (TSPs) are secreted glycoproteins that play key roles in interactions between cells and the extracellular matrix. Here, we describe the 2.6 Å resolution crystal structure of the glycosylated signature domain of human TSP-2, which includes three epidermal growth factor-like (EGF-like) modules, 13 aspartate-rich repeats, and a lectin-like module. These elements interact extensively to form three striking structural regions termed the stalk, wire, and globe. The TSP-2 signature domain is stabilized by these interactions and by a network of 30 bound Ca²⁺ ions and 18 disulfide bonds. The structure suggests how genetic alterations of TSPs result in disease.

Keywords

calcium binding; extracellular matrix; lectin-like module; skeletal dysplasia; epidermal growth factor-like module; protein crystallography

Thrombospondins (TSPs) are a family of five secreted multimodular metallo/glycoproteins that have diverse roles involving interactions between cells and the extracellular matrix {Adams, 2004 #3568}. For TSP-2, these functions are critical for such processes as synaptogenesis{Christopherson, 2005 #3578}, megakaryocytopoiesis{Kyriakides, 2003 #3576}, and the foreign body reaction{Kyriakides, 2001 #3588}. All TSPs contain a highly conserved “signature domain” consisting of tandem epidermal growth factor-like (EGF-like) modules, aspartate-rich repeats, and a lectin-like module at their C-termini (Fig. 1a). These modules are conserved with remarkable fidelity as is evidenced by a 458-residue stretch of fly TSP and human TSP-2 in the signature domain that are 60% identical without an insertion or deletion{Adams, 2003 #3571;LaBell, 1993 #340}. The signature domain contains the sites of polymorphisms and mutations linked to familial coronary artery disease and two skeletal disorders, pseudoachondroplasia (PSACH) and multiple epiphyseal dysplasia (EDM1){Topol, 2001 #3537;Kennedy, 2005 #3585;Posey, 2004 #3579}.

Here, we describe the 2.6 Å resolution crystal structure of the complete glycosylated signature domain of human TSP-2, which includes three epidermal growth factor-like (EGF-like) modules, 13 aspartate-rich repeats, and a lectin-like module, along with 30 bound Ca²⁺ ions. The structure reveals the highly intertwined nature of the signature domain that includes major interactions among the different parts of the molecule. The structure furthermore establishes a molecular explanation of how known genetic alterations of TSPs result in disease.

Results

Structure of the human TSP-2 signature domain

The crystal structure of the signature domain of human TSP-2 was solved using a combination of anomalous scattering and molecular replacement approaches. Electron-density maps, particularly those derived from molecular replacement using the structure of a fragment of the signature domain from TSP-1 {Kvansakul, 2004 #3561}, were of high quality (Fig. 1b), and allowed modelling of the entire domain structure.

The overall structure of the TSP-2 signature domain can be visually divided into three major features: the stalk, wire, and globe (Fig. 1c and Supplementary Fig. 1 online). The stalk comprises the two N-terminal EGF-like modules (EGF1 and EGF2) and extends 60 Å from the base of the molecule. The wire, with a contour length of ~170 Å, is formed by the repetitive aspartate-rich sequences draped around 26 Ca²⁺ ions. The third EGF-like module (EGF3) serves as a clasp to close the tortuous circle formed by the wire. The globe consists of the lectin-like module suspended beneath the wire and EGF3. This arrangement creates a prominent opening with dimensions of 14 Å x 40 Å between the wire and the globe.

Stalk, wire, and globe elements of the signature domain

The EGF-like modules in the stalk are characterized by unusual cysteine spacing and by binding a Ca²⁺ ion at their interface. Each module has the Cys1-Cys3, Cys2-Cys4, Cys5-Cys6 disulfide connectivity typical of the cEGF subtype 1 grouping {Wouters, 2005 #3583}. However, two residues separate the 4th and 5th cysteines in both, whereas a single residue separates the cysteines in other EGF-like modules in this grouping (Supplementary Fig. 2 online). This unusual spacing extends the axial lengths of EGF1 and EGF2 to 29.0 Å and 32.5 Å, respectively, and could prove important for forming distinctive interaction sites on their surfaces (Fig. 1c). In addition, EGF2 forms a class-A Ca²⁺-binding EGF-like module {Boswell, 2004 #3575} and binds a single Ca²⁺ ion (Fig. 2a and Supplementary Table 1 online). The Ca²⁺ ion is bound at the interface between the EGF-like modules in the stalk and likely stabilizes their co-axial alignment as noted in other tandem Ca²⁺-binding EGF-like modules {Boswell, 2004 #3575}.

The wire element comprises 13 aspartate-rich repeats that form individual globular folds connected in a disulfide bond-stabilized “beads on a string” arrangement (Fig. 1c). Each repeat envelops a Ca²⁺ core and can be classified into one of two classes, N-type and C-type, based on its sequence {Kvansakul, 2004 #3561}. N-type repeats are 13- or 15- residues in length and their tertiary structures superimpose with a root mean square deviation (rmsd) of ~0.35 Å for all C_α atoms (Fig. 2b). C-type repeats are 23 residues long and superimpose with a rmsd of ~0.36 Å (Fig. 2c). Each repeat binds two Ca²⁺ ions except 5N and 12N, which bind only one (Supplementary Table 1 online). In addition, single Ca²⁺ ions are bound at the interfaces of 2N and 3C and of 10N and 11C (Fig. 2d). The individual Ca²⁺-binding repeats are linked via disulfide bonds to each other in an arrangement that resembles struts along the wire axis, likely enhancing the overall rigidity of the wire element. Interestingly, the circular dichroic spectrum of a TSP-2 construct comprising solely its aspartate-rich repeats displays a Ca²⁺-dependent ellipticity minimum at 203 nm that is characteristic of the circular dichroic spectra of all TSPs studied to date {Misenheimer, 2003 #3556}. The repeat-containing protein is denatured only slightly more readily than the full signature domain {Misenheimer, 2003 #3556}. Thus, the wire element forms a stable Ca²⁺-dependent structure that is independent of the rest of the signature domain. Independent folding by the wire, coupled with the observations that the lectin-like domains of TSP-1 {Kvansakul, 2004 #3561} and TSP-2 {Misenheimer, 2001 #3406} are not soluble when expressed in isolation, implies that the wire may act as a scaffold in formation of the overall signature domain fold.

The globe lectin-like module is composed largely of β -sheets in a jelly-roll packing arrangement that is nearly identical in structure to the analogous fold from TSP-1 {Kvansakul, 2004 #3561}. Differences between the TSP-1 and TSP-2 structures of this module include the carbohydrate attached at Asn1069 in TSP-2 and the presence of three bound Ca^{2+} ions in the TSP-2 globe compared to four in TSP-1 (Fig. 1b,c and Supplementary Fig. 3 online). The residues that coordinate a fourth Ca^{2+} ion in the TSP-1 structure are present in TSP-2 but are part of a meandering loop in contrast to the tight circular loop observed in TSP-1.

Extensive interactions occur among the components of the signature domain in the Ca^{2+} -replete signature domain structure (Fig. 1c). The third EGF-like module functions as a major organizer of the signature domain through its interactions with the wire element (repeats 1C, 12N, and 13C), the second EGF-like module, and the globe element. The lectin-like and second EGF-like modules make additional direct contacts with the wire element. The discovery of these interaction sites helps to explain the structural basis of disease-linked TSP mutations (see below).

Discussion

Role of Ca^{2+} in the TSP signature domain

The large amount of Ca^{2+} bound by the TSP-2 signature domain is consistent with the documented sensitivity of TSP structures to changes in Ca^{2+} concentration. Rotary shadowing electron microscopy experiments with TSP-1 and -2 demonstrate gross Ca^{2+} -dependent dynamics within the signature domain {Lawler, 1982 #350; Lawler, 1985 #245; Chen, 1996 #1457}. Comparing the present structure and the rotary shadowing images, it appears that upon removal of Ca^{2+} , EGF3 and part of the wire dissociate from the globe module, forming an extended structure along with more N-terminal components. Solution spectroscopic experiments further demonstrate that there are dramatic Ca^{2+} -dependent changes in the TSP signature domain structure {Misenheimer, 2003 #3556}. In the Ca^{2+} -replete signature domain structure presented here, both the $\alpha\text{v}\beta 3$ integrin- {Lawler, 1989 #329; Chen, 1994 #595} and CD47-binding sites {Gao, 1994 #927} are buried and inaccessible. However, removal of Ca^{2+} could alter TSP conformation to expose these sites, making them available for binding and thereby modulating the activity of TSP, as proposed earlier {Kvansakul, 2004 #3561}. The Ca^{2+} -dependent conformational changes seen in the TSP signature domain structure suggest that TSPs may act as buffers as well as sensors of Ca^{2+} concentration in solution.

Implications of structural differences among TSPs

Two deviations from the canonical 23-residue C-type Ca^{2+} -binding repeat are tolerated by TSPs. First, repeat 1C of human TSP-2 contains a 13-residue insert between positions 10 and 11 (Fig. 2d,e). Despite this insertion, the residues of the interrupted halves of repeat 1C are positioned identically to C-type repeats without the insert (Fig. 2c). The insert forms a structured loop that is stabilized by a disulfide bond and interacts extensively with EGF3. An N-glycosylation recognition sequence is present at the apex of the inserted loop, and the proximal three saccharides are resolved in the structure (Fig. 1c, 2d). A similar insertion is present in the TSP-1 sequence, whereas in TSP-3, -4, and -5, the inserts lack the glycosylation sequence and are only 10 residues long. The second tolerated variation is an insertion of four residues between positions 10 and 11 of repeat 11C in TSP-3 and -4. From the structure presented here, it seems unlikely that an insertion this short could be accommodated without considerable disturbance of the 11C structure. This insert could distort the tertiary structure of the wire, making the structure of TSP-3 and -4 different from TSP-2. The importance of these variations is not clear but they could impart specialized functions upon the different TSP family members.

Structural interpretation of signature domain disease-linked mutations

Although a partial structure of the signature domain of TSP-1 has previously been solved {Kvansakul, 2004 #3561}, it lacks the three EGF-like modules, the seven N-terminal repeats of the wire (and the coordinated Ca^{2+} ions) and the four glycosylation sites of the signature domain (Supplementary Fig. 3 online). Importantly, because the TSP-1 structure is not the complete structure of the signature domain, it lacks the extensive interaction sites found in the TSP-2 structure. The importance of these interactions is evidenced by the identification of disease-linked genetic alterations at these sites in TSP-1, -4 and -5 (Fig. 3 and Supplementary Table 2 online). A polymorphic residue in TSP-1 (N700S) associated with premature coronary artery disease{Topol, 2001 #3537} maps to wire repeat 1C near the interaction site between EGF3, the wire and the globe. The homologous residue in TSP-2, Asn702, coordinates Ca^{2+} through a water (Fig. 2c,e), a finding consistent with the decrement in Ca^{2+} -binding associated with the Ser700 allele{Hannah, 2004 #3574}. A polymorphism in TSP-4 (A374P), also associated with premature coronary artery disease{Topol, 2001 #3537}, is in an extra EGF-like module that is most similar to EGF2 of TSP-2. The extra EGF-like module is in a position that would require it to subserve the role of EGF2 in the signature domain and interact with the wire, which could be altered in the TSP-4 variant.

Over 100 signature-domain mutations in TSP-5 (also known as cartilage oligomeric matrix protein or COMP) have been associated with the bone and joint diseases PSACH and EDM1 (Supplementary Table 2 online){Kennedy, 2005 #3585;Posey, 2004 #3579}. The mechanism by which these mutations cause disease is through the accumulation of unstable or misfolded protein intracellularly in chondrocytes{Dinser, 2002 #3587}. Of the disease-linked TSP-5 mutations, 81 are missense mutations that change residues at 62 different positions in the signature domain. Of the 62 positions, 53 are wire residues that are conserved between TSP-2 and -5 and all but nine are involved in coordination of Ca^{2+} , disulfide formation, or globe-wire interactions in TSP-2 (Fig. 3 and Supplementary Table 2 online). Twenty-two additional disease-linked mutations result in insertions or deletions of residues in the wire and would disrupt its proper spacing. Thus, the disease-causing mutations in TSP-5 target features that are critical for the Ca^{2+} -sensitive structure of the wire or the interactions between the wire and the globe.

The structure presented here provides perspective for examining the functions of TSPs. Until now, activities have been assigned to various parts of TSPs using antibody mapping and other techniques{Adams, 1995 #2328}. However it is now clear that TSPs cannot be considered a mosaic of independent domains, but instead have a convoluted structure stabilized by multiple interactions. Because Ca^{2+} clearly plays a dominant role in the formation of the final fold of the signature domain, fluctuations in Ca^{2+} concentration likely modify the function of TSPs by regulating the nature of the interactions mediated by the domain. It will be important to learn how TSPs change in structure and function as Ca^{2+} is depleted and identify compartments in which relevant changes in Ca^{2+} concentration occur. In short, the present structure offers both a rationale for the structural consequences of the numerous genetic changes in TSPs that are associated with disease and a foundation for understanding how TSPs normally function.

Methods

Expression and Purification of the human TSP-2 signature domain

Recombinant human TSP-2 signature domain protein (residues 551-1172 with *ADP* and *ARGHHHHHH* N- and C-terminal sequence tags) was expressed in conditioned insect cell media{Mosher, 2002 #3508}. The expression protocol was modified for incorporation of selenomethionine into the protein{McWhirter, 1999 #3572}·{Bellizzi, 1999 #3573} as follows. High5 insect cells in SF-900 II SFM media were infected at a multiplicity of infection

of 5. Twenty-four hours after infection, the cells were centrifuged and resuspended in SF-900 II SFM methionine- and cysteine-free media supplemented with dialyzed fetal calf serum (FBS) to 10% (v/v) and 0.15 g L⁻¹ cysteine. After four hours of growth at 27°C, the cells were centrifuged and resuspended in SF-900 II SFM methionine- and cysteine-free media supplemented with 0.15 g L⁻¹ cysteine, 10% FBS and 1 g L⁻¹ selenomethionine, and grown for 48 hours before harvesting. Soluble TSP-2 signature domain was purified by nickel-chelate affinity chromatography and dialyzed into low-salt buffer (20 mM MOPS, 50 mM NaCl, 2 mM CaCl₂). The dialysate was further purified by MonoQ ion exchange column. Pure protein was concentrated and dialyzed into 5 mM MOPS, 2 mM CaCl₂, 100 mM sodium acetate, pH 7.5 for crystallization.

Crystallization of selenomethionine-incorporated TSP-2 signature domain

The TSP-2 signature domain was crystallized using the hanging drop vapor diffusion method at room temperature. 1 µL of protein was mixed with 1 µL of mother liquor (100 mM Tris-Cl pH 8.5-9.0, 200 mM sodium acetate, 30% polyethylene glycol 4000) and 100 mM glycine and equilibrated with mother liquor. Selenomethionine-incorporated protein crystals required streak seeding from native crystals. Prior to data collection, crystals were stabilized by transfer to a cryoprotectant solution of the above mother liquor conditions supplemented with 25% ethylene glycol and frozen directly in liquid nitrogen.

Data collection and structure solution

Data were collected at 100 K at the Advanced Photon Source (BioCARS beamline 14-IDB) using X-rays of energies 12662 eV, 12958 eV, and 12658 eV. The structure was solved at 2.6 Å resolution using a combination of multiple-wavelength anomalous dispersion (MAD) phasing and molecular replacement (Supplementary Table 3). MAD data were indexed and scaled using HKL2000{Otwinowski, 1997 #3593}. Selenium sites were found using SOLVE{Terwilliger, 1999 #3590} and refined using MLPHARE{, 1994 #3597}. Solvent flattening with DM{, 1994 #3597} resulted in poor quality experimental electron density. To aid in building, a model of TSP-1 that includes the lectin-like domain and several of its proximal aspartate repeats{Kvansakul, 2004 #3561} was manually placed into the electron density maps by fitting the sulfurs from the TSP-1 methionines to the selenium sites determined above. Refinement with REFMAC5{, 1994 #3597} produced interpretable 2F_o-F_c and F_o-F_c electron density maps. The model was improved by rounds of refinement with REFMAC5 and manual rebuilding to an R_{factor} of 21.8 % (R_{free} = 28.4 %, Supplementary Table 3). The unbiased MAD-phased electron density maps were used to prevent model bias throughout refinement. 76.7% of all residues fall in the most favored Ramachandran category, with 21.0% in the allowed category, 1.9% in the generously allowed category and 0.4% in the disallowed category.

Coordinates

Model coordinates and structure factors have been deposited in the Protein Data Bank (1YO8 accession code).

Acknowledgements

We thank Erhard Hohenester for communicating results prior to publication and the Advanced Photon Source beamline staff (BioCARS beamline 14-IDB) for assistance in data collection. This work was supported by an NIH grant to DFM (HC54462) and a grant from Shaw Foundation for Medical Research to JLK. CBC and DAB were supported by NIH training grants (HC07899 and GM08293).

References

1. Adams JC, Lawler J. The thrombospondins. *Int J Biochem Cell B* 2004;36:961–8.

2. Christopherson KS, et al. Thrombospondins are astrocyte-secreted proteins that promote CNS synaptogenesis. *Cell* 2005;120 :421–33. [PubMed: 15707899]
3. Kyriakides TR, et al. Megakaryocytes require thrombospondin-2 for normal platelet formation and function. *Blood* 2003;101 :3915–23. [PubMed: 12732502]
4. Kyriakides TR, Zhu YH, Yang ZT, Huynh G, Bornstein P. Altered extracellular matrix remodeling and angiogenesis in sponge granulomas of thrombospondin 2-null mice. *Am J Pathol* 2001;159:1255–1262. [PubMed: 11583953]
5. Adams JC, et al. Characterisation of Drosophila thrombospondin defines an early origin of pentameric thrombospondins. *J Mol Biol* 2003;328 :479–94. [PubMed: 12691755]
6. LaBell TL, Byers PH. Sequence and characterization of the complete human thrombospondin 2 cDNA: potential regulatory role for the 3' untranslated region. *Genomics* 1993;17 :225–229. [PubMed: 8406456]
7. Topol EJ, et al. Single nucleotide polymorphisms in multiple novel thrombospondin genes may be associated with familial premature myocardial infarction. *Circulation* 2001;104:2641–2644. [PubMed: 11723011]
8. Kennedy J, et al. Novel and recurrent mutations in the C-terminal domain of COMP cluster in two distinct regions and result in a spectrum of phenotypes within the pseudoachondroplasia - multiple epiphyseal dysplasia disease group. *Hum Mutat* 2005;25:593–4. [PubMed: 15880723]
9. Posey KL, Hayes E, Haynes R, Hecht JT. Role of TSP-5/COMP in pseudoachondroplasia. *Int J Biochem Cell B* 2004;36 :1005–12.
10. Kvensakul M, Adams JC, Hohenester E. Structure of a thrombospondin C-terminal fragment reveals a novel calcium core in the type 3 repeats. *EMBO J* 2004;23 :1223–1233. [PubMed: 15014436]
11. Wouters MA, et al. Evolution of distinct EGF domains with specific functions. *Protein Sci* 2005;14:1091–1103. [PubMed: 15772310]
12. Boswell, E.J., Kurniawan, N.D., Downing, A.K. In *Handbook of Metalloproteins Vol. 3.* (ed. Messerschmidt, A., Bode, W., Cygler, M.) 553–570 (Wiley, Chichester; 2004).
13. Misenheimer TM, Hannah BL, Annis DS, Mosher DF. Interactions among the three structural motifs of the C-terminal region of human thrombospondin-2. *Biochemistry* 2003;42 :5125–32. [PubMed: 12718556]
14. Misenheimer TM, Hahr AJ, Harms AC, Annis DS, Mosher DF. Disulfide connectivity of recombinant C-terminal region of human thrombospondin 2. *J Biol Chem* 2001;276 :45882–45887. [PubMed: 11590138]
15. Lawler J, Chao FC, Cohen CM. Evidence for calcium-sensitive structure in platelet thrombospondin: Isolation and partial characterization of thrombospondin in the presence of calcium. *J Biol Chem* 1982;257 :12257–12265. [PubMed: 7118943]
16. Lawler J, Derick LH, Connolly JE, Chen JH, Chao FC. The structure of human platelet thrombospondin. *J Biol Chem* 1985;260 :3762–3772. [PubMed: 2579080]
17. Chen H, Aeschlimann D, Nowlen J, Mosher D. Expression and initial characterization of recombinant mouse thrombospondin 1 and thrombospondin 3. *FEBS Lett* 1996;387 :36–41. [PubMed: 8654563]
18. Lawler J, Hynes RO. An integrin receptor on normal and thrombasthenic platelets that binds thrombospondin. *Blood* 1989;74 :2022–2027. [PubMed: 2478219]
19. Chen H, Sottile J, O'Rourke KM, Dixit VM, Mosher DF. Properties of recombinant mouse thrombospondin 2 expressed in Spodoptera cells. *J Biol Chem* 1994;269:32226–32232. [PubMed: 7798222]
20. Gao AG, Frazier WA. Identification of a receptor candidate for the carboxyl-terminal cell binding domain of thrombospondins. *J Biol Chem* 1994;269:29650–29657. [PubMed: 7525586]
21. Hannah BL, Misenheimer TM, Pranghofer MM, Mosher DF. A polymorphism in thrombospondin-1 associated with familial premature coronary artery disease alters Ca²⁺ binding. *J Biol Chem* 2004;279:51915–22. [PubMed: 15456750]
22. Dinser R, et al. Pseudoachondroplasia is caused through both intra-and extracellular pathogenic pathways. *J Clin Invest* 2002;110 :505–513. [PubMed: 12189245]
23. Adams, J.C., Tucker, R.P. & Lawler, J. In *The thrombospondin gene family.* (R.G. Landes Company, Austin; 1995).

24. Mosher DF, Huwiler KG, Misenheimer TM, Annis DS. Expression of recombinant matrix components using baculoviruses. *Method Cell Biol* 2002;69:69–81.
25. McWhirter SM, et al. Crystallographic analysis of CD40 recognition and signaling by human TRAF2. *P Natl Acad Sci USA* 1999;96:8408–13.
26. Bellizzi JJ, Widom J, Kemp CW, Clardy J. Producing selenomethionine-labeled proteins with a baculovirus expression vector system. *Struct Fold Des* 1999;7:R263–7.
27. Otwinowski, Z., Minor, W. in *Methods in Enzymology*. (ed. Carter, C.W., Sweet, R.M.) 307–326 (Academic Press, New York; 1997).
28. Terwilliger TC, Berendzen J. Automated MAD and MIR structure solution. *Acta Crystallogr D* 1999;55 (Pt 4):849–61. [PubMed: 10089316]
29. The CCP4 suite: programs for protein crystallography. . *Acta Crystallogr D* 1994;50:760–3. [PubMed: 15299374]

Figure 1.

Structure of human TSP-2. (a) Schematic diagram of TSP-2. Each TSP-2 monomer is composed of a N-terminal module (N), an oligomerization domain (O), a procollagen module (C), and type-1 or properdin modules (P123) followed by the signature domain common to all TSPs. In human TSP-2, this domain consists of tandem EGF-like modules (EGF1 (residues 551–589, yellow), EGF2 (590–647, orange), and EGF3 (648–692, wheat)), aspartate-rich repeats (693–956, blue), and a lectin-like module (957–1172, magenta). Glycosylation sites are shown as green ovals. The arrowhead above the EGF-like repeats indicates the site of extra EGF-like module in TSP-3, -4 and -5. (b) Stereo diagram of the glycosylated residue Asn1069. The 2F_o-F_c electron-density map for the region is shown contoured at 1.3 σ (c) Four views of the ribbon diagram (top) and corresponding surface representation (bottom) of the crystal structure of the signature domain of human TSP-2. Modules are colored as in Fig. 1a, with Ca²⁺ ions (red), carbohydrates (green), and disulfide bonds (stick representation) shown.

Figure 2.

Ca²⁺ coordination in the stalk and wire elements of TSP-2. (a) Ca²⁺ coordination at the interface between EGF1 and EGF2, colored as in Figure 1. Ca²⁺-coordinating residues are shown and labeled. Residue labelled with (mc) ligand via their main chain carbonyl groups. (b) Overlay of the five N-type Ca²⁺-binding motifs. Ca²⁺-binding residues are labeled. “Xxx” represents any residue. Ca²⁺ ions are shown as red spheres. Residues labeled “(mc)” coordinate Ca²⁺ via main chain carbonyl groups. (c) Overlay of the eight C-type Ca²⁺-binding motifs. The insert sequence from repeat 1C has been removed for the alignment. (d) The glycosylated insertion element in repeat 1C and coordination of an additional Ca²⁺ between repeats 2N and 3C are highlighted by showing 1C, 2N, 3C, and 4C in stereo. (e) Alignment of Ca²⁺-binding repeats in the wire. The Ca²⁺ coordination scheme for each residue is colored-coded as follows: side chain binding to two Ca²⁺ ions (red), side chain binding to one Ca²⁺ ion (purple), main chain carbonyl binding to one Ca²⁺ ion (green or underlined), and water-mediated binding (blue or italicized). Disulfide bonds are shown in red lines. The symbols + or - indicate N-type repeats that bind three or one Ca²⁺, respectively, rather than two. The arrowhead above repeat 1C indicates the site of the 13-residue insertion, which is shown. The arrowhead above repeat 11C indicates the location of the 4-residue insertion that is present in TSP-3 and -4. Residue numbers are listed to the left of the sequence and coordination numbering is shown above.

Figure 3.

Disease-associated mutations or polymorphisms of TSPs mapped onto the signature domain of human TSP-2. Residues homologous to positions in TSP-5 that are sites where missense mutations are linked to PSACH or EDM1, or to the polymorphism in TSP-1 that is linked to coronary artery disease are colored yellow and shown in stick form. The wire module is shown in cartoon form with the remainder the signature domain shown as a surface representation to highlight the 62 positions of the missense mutations. Fifty-two map to the wire region and 10 map to surfaces on the lectin-module. Twenty-two additional non-missense mutations are not highlighted but also map to the wire module (Supplementary Fig. 2 online).

Supplementary Figure 1.

Stereo ribbon diagram of the crystal structure of the signature domain of human TSP-2. Structure is color-coded to indicate the positions of the tandem EGF-like modules (EGF1 (residues 551–589, yellow), EGF2 (590–647, orange), and EGF3 (648–692, wheat)), aspartate-rich repeats (693–956, blue), and a lectin-like module (957–1171, magenta). Ca²⁺ ions (red), carbohydrates (green), and disulfide bonds (stick representation) are shown.

Supplementary Figure 2.

Sequence and secondary structure of the signature domain of human TSP-2. Helices (boxes) and β -strands (arrows) are numbered sequentially above the sequence and are color-coded as in Fig. 1. Cysteines connected by lines form disulfide bonds in the structure. Residues colored green are glycosylated. Underlined or bold-face residues are sites that are mutated or deleted, respectively, in either EDM-1 or PSACH. Two insertion mutations map to the position labeled with a "Y".

Supplementary Figure 3.

Comparison of the crystal structures of the TSP-2 signature domain (A) and the TSP-1 signature domain fragment¹ (B). Both structures are colored and shown in the same orientations as for the TSP-2 signature domain in Figure 1 (main text).

¹Kvansakul, M., Adams, J. C. & Hohenester, E. Structure of a thrombospondin C-terminal fragment reveals a novel calcium core in the type 3 repeats. *Embo Journal* **23**, 1223–1233 (2004).

Supplementary Table 1**TSP-2 coordination of the 30 bound Ca²⁺**

The coordination scheme for each of the 30 Ca²⁺ ions is colored-coded as follows: side chain binding to two Ca²⁺ ions (red box, labeled “bidentate”), side chain binding to one Ca²⁺ ion (purple, labeled “side chain”), main chain carbonyl binding to one Ca²⁺ ion (green, labeled “main chain”), water-mediated binding (blue, labeled “via water”), and water or unidentified liganding element (white, labeled “WATER” or “NO DENSITY”). The structural element within TSP-2 that is responsible for binding is indicated in the right-most column.

EGF-2							EGF-2
Ca1	D590 side chain	L591 main chain	E593 side chain	N612 side chain	T613 main chain	G616 main chain	Wire
Ca2	D694 side chain	D696 bidentate	D698 bidentate	W700 main chain	N702 via water	D718 bidentate	1C
Ca3	D696 bidentate	D698 bidentate	D718 bidentate	C720 main chain	L723 main chain	N725 side chain	Wire
Ca4	D730 side chain	D732 bidentate	D734 bidentate	I736 main chain	D738 via water	D741 bidentate	1C
Ca5	D732 bidentate	D734 bidentate	D741 bidentate	D743 main chain	NO DENSITY	N746 side chain	Wire
Ca6	D738 main chain	D741 main chain	D744 side chain	WATER	WATER	D751 side chain	2N
Ca7	D743 side chain	D745 bidentate	D747 bidentate	V749 main chain	NO DENSITY	D754 bidentate	2N and
Ca8	D745 bidentate	D747 bidentate	D754 bidentate	C756 main chain	L759 main chain	N761 side chain	3C
Ca9	D766 side chain	D768 bidentate	D770 bidentate	V772 main chain	D777 bidentate	N784 via water	Wire
Ca10	D768 bidentate	D770 bidentate	D777 bidentate	C779 main chain	V782 main chain	N784 side chain	3C
Ca11	D789 side chain	D791 side chain	N793 via water	E795 main chain	D797 via water	WATER	Wire
Ca12	D802 side chain	D804 bidentate	D806 bidentate	V808 main chain	N810 via water	D813 bidentate	4C
Ca13	D804 bidentate	D806 bidentate	D813 bidentate	C815 main chain	V818 main chain	N820 side chain	Wire
Ca14	D825 side chain	D827 bidentate	D829 bidentate	V831 main chain	D833 via water	D836 bidentate	6C
Ca15	D827 bidentate	D829 bidentate	D836 bidentate	C838 main chain	V841 main chain	N843 side chain	Wire
Ca16	D848 side chain	D850 bidentate	D852 bidentate	V854 main chain	D856 via water	D859 bidentate	7C
Ca17	D850 bidentate	D852 bidentate	D859 bidentate	N861 side chain	Q870 via water	NO DENSITY	Wire
Ca18	D863 side chain	D865 bidentate	D867 bidentate	H869 main chain	N871 via water	D874 bidentate	8N
Ca19	D865 bidentate	D867 bidentate	D874 bidentate	C876 main chain	I879 main chain	N881 side chain	9C
Ca20	D886 side chain	D888 bidentate	D890 bidentate	Q892 main chain	NO DENSITY	D897 bidentate	Wire
Ca21	D888 bidentate	D890 bidentate	D897 bidentate	D899 main chain	WATER	N902 side chain	10N
Ca22	D894 main chain	D897 main chain	D900 side chain	WATER	WATER	D907 side chain	Wire
Ca23	D899 side chain	D901 bidentate	D903 bidentate	V905 main chain	D907 via water	D910 bidentate	10N and
Ca24	D901 bidentate	D903 bidentate	D910 bidentate	C912 main chain	V915 main chain	N917 side chain	11C
Ca25	D922 side chain	D924 side chain	D926 side chain	R928 main chain	WATER	D930 side chain	Wire
Ca26	D935 side chain	D937 bidentate	D939 bidentate	I941 main chain	D943 via water	D946 bidentate	11C
Ca27	D937 bidentate	D939 bidentate	D946 bidentate	C948 main chain	N951 main chain	I954 main chain	Wire
Lectin							12N
Ca28	D1019 main chain	D1021 main chain	Q1044 side chain	D1106 side chain	WATER	WATER	Wire
							13C
							Lectin
							Single

Ca29	D1021 bidentate	D1022 side chain	D1023 bidentate	Q1047 side chain	WATER	WATER	Lectin Double Lectin Double
Ca30	N993 side chain	D1021 bidentate	D1023 bidentate	S1156 side chain	S1156 main chain	WATER	

Supplementary Table 2

Disease-linked TSP-family mutations

List of currently known mutations within TSP-5 signature domain that are linked to PSACH or EDM-1 and their apparent roles derived from the presented crystal structure of TSP-2. Mutations are divided into “Missense” and “Deletion or Insertion” mutations in the first and second groupings. Mutations are listed with TSP-5 numbering, followed by the associated disease, the analogous residue(s) in the TSP-2 signature domain, the roles played by the residue(s) in the TSP-2 signature domain structure, and the predicted role in TSP-5.

Missense Mutations TSP-5 Mutation	Disease	Equivalent TSP-2 residue	Deduced Role in TSP-2	Predicted Role in TSP-5
Wire				
D271H	PSACH	D696	Ca 2 and 3 coordination	Ca 2 and 3 coordination
L272P	PSACH	L697	interaction with lectin	interaction with lectin
P276R	EDM1	P701	?	
D290N	PSACH	D718	Ca 2 and 3 coordination	Ca 2 and 3 coordination
C292W	PSACH	C720	Ca 3 coordination and disulfide bond	Ca 3 coordination
S298L	PSACH	S726	interaction with lectin	interaction with lectin
G299R	PSACH	G727	interaction with lectin	interaction with lectin
D302V	EDM1	D730	Ca 4 coordination	Ca 4 coordination
G309R	PSACH	G737	?	
D310V	EDM1	D738	Ca 4 and 6 coordination	Ca 4 and 6 coordination
C328R	PSACH	C756	Ca 8 coordination and disulfide bond	Ca 8 coordination
D342Y	EDM1	D770	Ca 9 and 10 coordination	Ca 9 and 10 coordination
D349G/V	PSACH	D777	Ca 9 and 10 coordination	Ca 9 and 10 coordination
C351Y	PSACH	C779	Ca 10 coordination and disulfide bond	Ca 10 coordination
D361Y/A/V	EDM1	D789	Ca 11 coordination	Ca 11 coordination
G366R	EDM1	G794	?	
G368R	EDM1	G796	?	
C371S/F	EDM1	C799	disulfide bond	disulfide bond
D376V	PSACH	D804	Ca 12 and 13 coordination	Ca 12 and 13 coordination
D378Y	EDM1	D806	Ca 12 and 13 coordination	Ca 12 and 13 coordination
D385N	EDM1	D813	Ca 12 and 13 coordination	Ca 12 and 13 coordination
C387G	PSACH	C815	Ca 13 coordination and disulfide bond	Ca 13 coordination
D399N	EDM1	D827	Ca 14 and 15 coordination	Ca 14 and 15 coordination
D401N	EDM1	D829	Ca 14 and 15 coordination	Ca 14 and 15 coordination
C407F/S	EDM1	C835	disulfide bond	disulfide bond
D408Y	EDM1	D836	Ca 14 and 15 coordination	Ca 14 and 15 coordination
D420A	EDM1	D848	Ca 16 coordination	Ca 16 coordination
G427E	PSACH	G855	?	
D437G	PSACH	D865	Ca 18 and 19 coordination	Ca 18 and 19 coordination
D439N	EDM1	D867	Ca 18 and 19 coordination	Ca 18 and 19 coordination
G440R/E	PSACH	G868	?	
D446N	PSACH	D874	Ca 18 and 19 coordination	Ca 18 and 19 coordination
P449T	PSACH	P877	?	
N453S	EDM1	N881	Ca 19 coordination	Ca 19 coordination
S454R	EDM1	A883	interaction with lectin	interaction with lectin
G465R/S	PSACH	G893	?	
C468Y	PSACH	C896	disulfide bond	disulfide bond
D471G	PSACH	D899	Ca 21 and 23 coordination	Ca 21 and 23 coordination
D472Y	PSACH	D900	Ca 22 coordination	Ca 22 coordination
D473G/N/Y	EDM1/ PSACH	D901	Ca 23 and 24 coordination	Ca 23 and 24 coordination
D475N	PSACH	D903	Ca 23 and 24 coordination	Ca 23 and 24 coordination
D479H/Y	EDM1/ PSACH	D907	Ca 22 and 23 coordination	Ca 22 and 23 coordination
D482G/H	PSACH	D910	Ca 23 and 24 coordination	Ca 23 and 24 coordination
C484G	EDM1	C912	Ca 24 coordination and disulfide bond	Ca 24 coordination
N489K	EDM1	N917	Ca 24 coordination	Ca 24 coordination
Q492P	EDM1	Q920	?	
C504S	PSACH	C932	disulfide bond	disulfide bond
D507G	PSACH	D935	Ca 26 coordination	Ca 26 coordination
D509A/G/E	PSACH	D937	Ca 26 and 27 coordination	Ca 26 and 27 coordination
D511H	PSACH	D939	Ca 26 and 27 coordination	Ca 26 and 27 coordination
D518G/N/H	PSACH	D946	Ca 26 and 27 coordination	Ca 26 and 27 coordination
N523K	EDM1	N951	Ca 27 coordination	Ca 27 coordination
T527A	PSACH	S955	Possible Ca ²⁺ coordination	Possible Ca ²⁺ coordination
Lectin				
T5291	PSACH	T957	interaction with wire	interaction with wire

Missense Mutations TSP-5 Mutation	Disease	Equivalent TSP-2 residue	Deduced Role in TSP-2	Predicted Role in TSP-5
N555K	EDM1	N993	Ca 30 coordination	Ca 30 coordination
E583K	PSACH	S1011	interaction with wire	interaction with wire
T585R/M	PSACH/ EDM1	T1013	interaction with wire	interaction with wire
H587R	PSACH	Y1015	interaction with wire	interaction with wire
D605N	EDM1	S1033	interaction with wire	interaction with wire
S681C	EDM1	A1109	interaction with wire	interaction with wire
R718P/W	EDM1	A1146	interaction with wire	interaction with wire
G719D/S	PSACH	G1147	interaction with wire	interaction with wire
Deletions or insertions				
DP277-283del	PSACH	NA702-708	?	spacing
ED341-342del	PSACH	KD769-770	Ca 9 and 10 coordination	Ca 9 and 10 coordination
RG367-368del	EDM1	EG795-796	Ca 11 coordination	Ca 11 coordination
D372del	PSACH	S800	?	spacing
D373del	PSACH	V801	?	spacing
D374del	PSACH	D802	Ca 12 coordination	Ca 12 coordination
N386del	EDM1	N814	?	
PNSD391-394C	PSACH	YNTD819-822	Ca 13 coordination	Ca 13 coordination
QKD395-397del	PSACH	QRD823-825	Ca 14 coordination	Ca 14 coordination
E457del	PSACH	A885	hairpin turn/structural	hairpin turn /structural
S459del	PSACH	H887	hairpin turn/structural	hairpin turn /structural
D469insertD	EDM1	D897	Ca 20 and 21 coordination	Ca 20 and 21 coordination
D469del	PSACH	D897	Ca 20 and 21 coordination	Ca 20 and 21 coordination
D469insertDD	PSACH	D897	Ca 20 and 21 coordination	Ca 20 and 21 coordination
DD469-70delDD	PSACH	DP897-8	Ca 20 and 21 coordination	Ca 20 and 21 coordination
D470del	PSACH	P898	?	spacing
D471del	PSACH	D899	Ca 21 and 23 coordination	Ca 21 and 23 coordination
D472del	PSACH	D900	Ca 22 coordination	Ca 22 coordination
D473del	PSACH	D901	Ca 23 and 24 coordination	Ca 23 and 24 coordination
GDV501-3del	EDM1	GDI929-31	Ca 25 coordination	Ca 25 coordination
ADKV510-513del	PSACH	NDNI938-941	Ca 26 and 27 coordination	Ca 26 and 27 coordination
VVDK513-516del	PSACH	IPDI941-944	Ca 26 coordination	Ca 26 coordination
N742stop	EDM1	R1170	interaction with E3	interaction with E3

Supplementary Table 3

Data collection, phasing and refinement statistics

Data collection			
Space group		I222	
Cell dimensions			
<i>a</i> , <i>b</i> , <i>c</i> (Å)		93.44,	
		121.59,	
		155.33	
α β γ (°)		90, 90, 90	
	<i>Peak</i>	<i>Inflection</i>	<i>Remote</i>
Wavelength (Å)	0.9792	0.9795	0.9568
Resolution (Å)	50-2.6	50-2.6	50-2.6
<i>R</i> _{sym} or <i>R</i> _{merge} (%)	8.3 (27.5)	8.2 (31.9)	8.9 (29.6)
<i>I</i> / σ <i>I</i>	18.6 (2.8)	17.6 (2.5)	23.2 (3.3)
Completeness (%)	93.8 (60.9)	89.7 (35.3)	96.2 (79.6)
Redundancy	4.9 (3.2)	4.8 (2.6)	9.3 (6.2)
Refinement			
Resolution (Å)		20-2.6	
No. reflections		23,531	
<i>R</i> _{work} / <i>R</i> _{free} (%)		21.8/28.4	
No. atoms		4939	
Protein/glycosylation			
Ca ⁺⁺ ions		30	
Water		116	
B-factors		68	
Protein/glycosylation			
Ca ⁺⁺ ions		62	
Water		54	
R.m.s deviations			
Bond lengths (Å)		0.012	
Bond angles (°)		1.15	

* Highest resolution shell (2.69-2.6 Å) is shown in parenthesis.