JCB Comment

The face of TSR revealed: an extracellular signaling domain is exposed

Roy L. Silverstein

Division of Hematology and Medical Oncology, Department of Medicine, Weill Medical College of Cornell University, New York, NY 10021

In this issue, Tan et al. (2002) report the first high resolution (1.9 Å) structural data for thrombospondin (TSP)-1, a large multifunctional protein that regulates cell adhesion, angiogenesis, cell proliferation and survival, TGFβ activation, and protease function (for review see Chen et al., 2000). Because TSP-1 has multiple binding partners and many functions, precise structural information is crucial to understanding its biology. The structure now reported, derived from crystals of the second and third type I repeats of TSP-1 is of particular interest because of the specific functions attributed to these repeats and because domains homologous to the repeats appear in many other proteins in nature. The novel layered fold motif described brings great insight into how the complicated functions of TSP-1 and related molecules are affected.

TSP-1 is the prototypical member of a group of secreted, extracellular matrix (ECM)* proteins referred to as matricellular proteins (Bornstein, 1995). This name signifies their major role as matrix-based cellular signaling molecules rather than as determinants of matrix structure. These proteins are deposited in the matrix in a highly regulated manner, e.g., during embryonic development or postnatal remodeling events such as wound healing, and function via specific interactions with growth factors, proteases, and receptors on migrating cells. TSP-1, a 450-Kd homotrimer was first identified as a protein secreted in large amounts from the α-granules of activated platelets. Platelet TSP is easily purified and has been very well studied. In vitro experiments have identified numerous binding partners, including fibrinogen, fibronectin, plasminogen, thrombin, elastase, matrix metalloproteases, heparin, sulfated glycolipids, and cellular surfaces (Chen et al., 2000). Despite seemingly critical homeostatic roles, the phenotype of the murine TSP-1 knock-out strain is subtle (Lawler et al., 1998), perhaps the result of redundancies among the matricellular proteins as a group.

Address correspondence to Roy Silverstein, Division of Hematology and Medical Oncology, Department of Medicine. Weill Medical College of Cornell University, 1300 York Ave., New York, NY 10021. Tel.: (212) 746-2060. Fax: (212) 746-8866. E-mail: rlsilve@med.cornell.edu

*Abbreviations used in this paper: ECM, extracellular matrix; GAG, glycosaminoglycans; HB-GAM, heparin binding growth-associated molecule.

The myriad effects of TSP-1 on cellular functions may seem confusing and inconsistent. For example, TSP-1 has been reported in some studies to promote cell adhesion, cell proliferation, angiogenesis, and tumor progression, whereas other equally compelling studies show disruption of cell adhesion, induction of apoptosis, inhibition of angiogenesis, and inhibition of tumor growth. These inconsistencies, however, are readily explainable by its ability to interact specifically with a range of cellular receptors expressed differentially by different cell types. These include at least three integrins (α3β1, αvβ3, and αIIbβ3), CD47 (integrin-associated protein), CD36, LDL receptor–related protein, and cellular glycosaminoglycans (GAGs). Also, TSP-1 has been shown to regulate protease and growth factor functions, adding a further level of complexity to its effects on cells.

Understanding of the complex biological functions of TSP-1 followed to a large extent from analysis of its structure. Initial studies using limited proteolysis and SDS-PAGE suggested a divalent cation-dependent modular structure, similar to many other proteins of the ECM. This was confirmed with the generation and mapping of a series of monoclonal antibodies. Rotary shadowing electron microscopy revealed that the monomers were tethered at their NH_2 termini forming a globular domain, from which extended three long flexible regions, each ending in a COOH-terminal smaller globular domain (Galvin et al., 1985). The NH_2-terminal globular region was found to be a heparin-binding domain and the COOH-terminal globular regions a cell binding domain. Monoclonal antibodies and peptides directed to the intervening regions also revealed a myriad of biological actions for these domains.

These early studies correlating structure and function were followed by the pioneering work of Lawler and Hynes (1986), solving the primary structure by cDNA cloning. From this, two major insights followed. First, platelet TSP was part of a small gene family consisting of four other members, all of which encoded secreted matrix proteins (for review see Adams and Tucker, 2000). Second, a unique feature of two members of the family, TSP-1 and -2, was the presence of three copies of the type I repeat, each encoded by a separate exon. Subsequent sequence analyses revealed that the type I repeat was highly homologous to the so-called properdin repeat sequence that had been identified in many

components of the complement system and in the circumsporozoite protein encoded by the genome of the unicellular parasite *falciparum* malaria. Comparative genomics now reveals that the primordial exon encoding the type I repeat is ancient and has been extensively duplicated and shuffled during evolution; >40 human genes contain one or more copies of the repeat, along with >10 in the fly, 20 in the worm, and 2 in *falciparum* malaria. Among these are the ADAM-TS family of metalloproteinases, F- and M-spondins, semaphorins, Unc5, heparin binding growth-associated molecule (HB-GAM), and brain angiogenesis inhibitor (BAI)-1. Of those whose functions have been probed, most seem to be involved in cell matrix interactions and in the control of migration, axonal guidance, and/or matrix remodeling during development (Adams and Tucker, 2000). It is the type I repeat in many of these proteins that seems to be a critical mediator of cell–matrix interactions.

The type I/properdin repeat has now been renamed TSR (thrombospondin structural homology repeat). It is ~60 amino acids in length, ≥12 of which are highly conserved, including 2–3 tryptophans, 5–6 cysteines, and 2 arginines. Solution spectroscopic studies of a TSR from distant TSP "cousins" HB-GAM and midkine were reported in 2000 (Kilpelainen et al., 2000) and suggested an elongated structure with a large surface area. Tan et al. have now successfully crystallized the second and third TSR from TSP-1 and solved the structure at high resolution. This work sheds considerable light on the often confusing and paradoxical biology of TSP-1.

Their studies show that TSR is a highly structured domain organized as a long right-handed spiraling ribbon containing three antiparallel strands (two are β sheets and the third is a less structured "rippled" strand). The most novel aspect of the structure is that the three antiparallel strands are held together by a series of interlocking stacks of amino acid side chains made up of six alternating layers of tryptophans (W layer) and arginines (R layer) sandwiched between cystine disulfides on the top and bottom. This pattern of stacked planar cationic guanidium groups and aromatic groups is stabilized by multiple cation-π interactions and forms a continuously positive charged face containing a groove-like structure that could easily be imagined forming a recognition site for ligands. Several testable hypotheses thus emerge from this work. The authors speculate that an anionic disaccharide unit from a right-handed spiraling heparin molecule could "fit" well into the cationic grove formed in the 20-Å distance between the first W layer and the third R layer. It is tempting to speculate that the degree of positive charge and shape of the groove within the charged face of the domain along with the number of sequential repeats may regulate the specificity and affinity for GAGs. ADAM-TS4 (aggrecanase) and F-spondin have additional arginines in the C strand of some of their TSRs and thus a more positive charged face than TSP-1. F-spondin and Sema5 have a larger number (seven) of sequential TSRs. Whether these features relate to the ability of ADAM-TSY to bind and specifically cleave aggrecan or of F-spondin to guide neurite outgrowth on the developing floor plate, remains to be seen.

Another role for the TSR in TSP-1 and -2 is as a ligand for the cell surface receptor CD36 (Asch et al., 1987). This interaction, although initially controversial, is of considerable interest in that it has now been shown convincingly to mediate the antiangiogenic, proapoptotic effects of TSP on microvascular endothelial cells (Dawson et al., 1997). Several other TSR proteins are also antiangiogenic, although it is not yet known whether all operate via CD36, nor are the structural factors known that determine whether a TSR protein binds CD36. Tan et al. have divided the TSR proteins into two large groups based on putative structural differences, including the position of the first cystine layer. Although data are not available for all, the known antiangiogenic members are all in one of the groups. Studies using synthetic peptides have suggested that the conserved CSVTCG sequence in TSRs is the likely binding site for CD36 (Asch et al., 1992); CSVTCG peptides inhibit angiogenesis and block TSP-1 binding to CD36-expressing cells. More difficult to explain are studies showing that peptides containing sequences flanking, but not including, CSVTCG are also antiantiangiogenic (Tolsma et al., 1993). In fact a D-isoleucine enantiomer of the highly conserved GVITRIR sequence is nearly as potent as TSP-1 itself and is being developed as an antiangiogenesis-based anticancer drug. The crystal structure of TSR reveals that the cationic GVITRIR sequence makes up part of the positively charged face of the domain, whereas the CSVTCG sequence is not readily available with its two cysteines involved in separate disulfide bonds. The hypothesis that the cationic face of the TSR serves as a CD36 ligand is consistent with work from our laboratory and others showing that the binding site on CD36 for TSP is between amino acids 90 and 110 (Pearce et al., 1995) and contains a negatively charged sequence that is conserved in murine CD36. Furthermore, this sequence is conserved in CD36 homologous domains found in other TSP binding proteins, including histidine-rich glycoprotein, LIMP2, and HIV gp120 (Crombie and Silverstein, 1998). Cocrystallization of the TSR with its binding domain from CD36 or CD36 homologues is the next challenge and may allow us to establish parameters to predict which TSR proteins are likely to be antiangiogenic.

Submitted: 30 September 2002
Accepted: 7 October 2002

References

Adams, J.C., and R.P. Tucker. 2000. The thrombospondin type 1 repeat (TSR) superfamily: Diverse proteins with related roles in neuronal development. *Dev. Dyn.* 218:280–299.

Asch, A.S., S. Silbeger, E. Heimer, and R.L. Nachman. 1992. Thrombospondin sequence motif (CSVTCG) is responsible for CD36-binding. *Biochem. Biophys. Res. Commun.* 182:1208–1217.

Asch, A.S., J. Barnwell, R.L. Silverstein, and R.L. Nachman. 1987. Isolation of the thrombospondin membrane receptor. *J. Clin. Invest.* 79:1054–1061.

Bornstein, P. 1995. Diversity of function is inherent in matricellular proteins: An appraisal of thrombospondin-1. *J. Cell Biol.* 130:503–506.

Chen, H., M.E. Herndon, and J. Lawler. 2000. The cell biology of thrombospondin-1. *Matrix Biol.* 19:597–614.

Crombie, A.R., and R.L. Silverstein. 1998. Lysosomal integral membrane protein II binds thrombospondin1: Evidence of functional homology with the cell adhesion molecule CD36. 1998. *J. Biol. Chem.* 273:4855–4864.

Dawson, D.W., S.F.A. Pearce, R. Zhong, R.L. Silverstein, W.A. Frazier, and N.P. Bouck. 1997. CD36 mediates the inhibitory effects of thrombospondin-1 on endothelial cells. *J. Cell Biol.* 138:707–717.

Galvin, N.J., V.M. Dixit, K.M. O'Rourke, S.A. Santoro, G.A. Grant, and W.A.

Frazier. 1985. Mapping of epitopes for monoclonal antibodies against human platelet thrombospondin with electron microscopy and high sensitivity amino acid sequencing. *J. Cell Biol.* 101:1434–1441.

Kilpelainen, I., M. Kaksonen, T. Kinnunen, H. Avikainen, M. Fath, R.J. Linhardt, E. Raulo, and H. Rauvala. 2000. Heparin-binding growth-associated molecule contains two heparin-binding β-shhet domains that are homologous to the thrombospondin type 1 repeat. *J. Biol. Chem.* 275:13564–13570.

Lawler, J., and R.O. Hynes. 1986. The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins. *J. Cell Biol.* 103:1635–1648.

Lawler, J., M. Sunday, V. Thibert, M. Duquette, E.L. George, H. Rayburn, and R.O. Hynes. 1998. Thrombospondin-1 is required for normal murine pulmonary homeostasis and its absence causes pneumonia. *J. Clin. Invest.* 101: 982–992.

Pearce, S.F.A., J. Wu, and R.L. Silverstein. 1995. Recombinant fusion proteins define a thrombospondin binding domain: Evidence for a single calcium-dependent binding site on CD36. *J. Biol. Chem.* 270:2981–2986.

Tan, T., M. Duquette, J. Liu, Y. Dong, R. Zhang, A. Joachimiak, J. Lawler, and J. Wang. 2002. Crystal structure of the TSP-1 type 1 repeats: a novel layered fold and its biological implication. *J. Cell Biol.* 159:373–382.

Tolsma, S.S., O.V. Volpert, D.J. Good, W.A. Frazier, P.J. Polverini, and N. Bouck. 1993. Peptides derived from two separate domains of the matrix protein thrombospondin-1 have anti-angiogenic activity. *J. Cell Biol.* 122: 497–511.