

# Crystal structure of the TSP-1 type 1 repeats: a novel layered fold and its biological implication

Kemin Tan,<sup>1,3</sup> Mark Duquette,<sup>2,4</sup> Jin-huan Liu,<sup>1,3</sup> Yicheng Dong,<sup>1,3</sup> Rongguang Zhang,<sup>7</sup> Andrzej Joachimiak,<sup>7</sup> Jack Lawler,<sup>2,4</sup> and Jia-huai Wang<sup>1,5,6</sup>

<sup>1</sup>Dana-Farber Cancer Institute, Boston, MA 02115

<sup>2</sup>Beth Israel Deaconess Medical Center, Boston, MA 02215

<sup>3</sup>Departments of Medicine, <sup>4</sup>Pathology, and <sup>5</sup>Pediatrics and <sup>6</sup>Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115

<sup>7</sup>Biosciences Division, Argonne National Laboratories, Argonne, IL 60439

**T**hrombospondin-1 (TSP-1) contains three type 1 repeats (TSRs), which mediate cell attachment, glycosaminoglycan binding, inhibition of angiogenesis, activation of TGF $\beta$ , and inhibition of matrix metalloproteinases. The crystal structure of the TSRs reported in this article reveals a novel, antiparallel, three-stranded fold that consists of alternating stacked layers of tryptophan and arginine residues from respective strands, capped by disulfide bonds on each

end. The front face of the TSR contains a right-handed spiral, positively charged groove that might be the “recognition” face, mediating interactions with various ligands. This is the first high-resolution crystal structure of a TSR domain that provides a prototypic architecture for structural and functional exploration of the diverse members of the TSR superfamily.

## Introduction

The thrombospondins (TSPs)\* are a family of five matricellular glycoproteins that regulate extracellular matrix structure and cellular phenotype (for reviews see Bornstein, 1995; Chen et al., 2000). These modular proteins act by bringing together cytokines, growth factors, other matrix components, membrane receptors, and extracellular proteases. TSP-1, for instance, regulates cell proliferation, migration, and apoptosis in a variety of physiological and pathological settings, such as wound healing, inflammation, angiogenesis, and neoplasia (Chen et al., 2000). The inhibition of angiogenesis involves the inhibition of endothelial cell migration and the induction of endothelial cell apoptosis (Tolsma et al., 1993; Jimenez et al., 2000). In addition, TSP-1 inhibits the mobilization of

VEGF from the extracellular matrix by matrix metalloproteinase 9 (Rodriguez-Manzanique et al., 2001). TSP-1 can also suppress the growth of TGF $\beta$ -responsive tumor cells through activation of this cytokine (Miao et al., 2001). The 420-kD TSP-1 molecule is trimeric and each protomer is composed of multiple domains: NH<sub>2</sub>- and COOH-terminal globular domains; a procollagen-like domain; and three types of repeated sequence motifs, designated type 1, 2, and 3 repeats (Lawler and Hynes, 1986). The type 1 repeats (designated TSP-1 domain in the pfam database or thrombospondin repeat [TSR] elsewhere) were initially identified in human endothelial cell TSP-1 (Lawler and Hynes, 1986). Database searches at that time revealed that a similar structural motif was present in the complement factors C8 and C9 and in the circumsporozoite protein of the *Plasmodium falciparum* parasite. Since then, the TSR domain has been identified in multiple different protein families (for review see Adams and Tucker, 2000). The human genome has 41 proteins that contain TSRs and the *Caenorhabditis elegans* and *Drosophila* genomes have 27 and 14 proteins with TSRs, respectively (Venter et al., 2001). The copy number of the TSR in these proteins varies from 1 to 18. All of the TSRs are in secreted proteins or in the extracellular portion of transmembrane proteins. They appear to be involved in cellular migration, communication, and tissue remodeling in complex tissues. The TSR domain consists of ~60 amino

Address correspondence to Jia-huai Wang, Dana-Farber Cancer Institute, Harvard Medical School, 44 Binney St., Room SM-1036, Boston, MA 02115. Tel.: (617) 632-3983. Fax: (617) 632-4393. E-mail: jwang@red.dfci.harvard.edu; or Jack Lawler, Beth Israel Deaconess Medical Center, Harvard Medical School, 99 Brookline Ave., Boston, MA 02215. Tel.: (617) 667-1694. Fax: (617) 667-3591. E-mail: lawler@mbcrr.harvard.edu

\*Abbreviations used in this paper: GAG, glycosaminoglycan; MAD, multiple-wavelength anomalous diffraction; TRAP, thrombospondin-related anonymous protein; TSP, thrombospondin; TSR, type 1 repeat.

Key words: thrombospondin; TSR domain; X-ray structure; angiogenesis; GAG binding

Table I. Data collection, structure determination, and refinement

	Data set			
	Hg peak	Hg inflection	Hg remote	Native
Data collection				
Space group	C2			C2
Unit cell (Å, °)	a = 67.94, b = 83.97, c = 37.17, β = 108.37			a = 67.63, b = 84.41, c = 37.17, β = 108.67
X-ray source	APS			APS
Wavelength (Å)	1.00803	1.00808	1.05969	1.1000
Resolution (Å)	20 – 2.50	20 – 2.50	20 – 2.50	20 – 1.90
Observations	65741	65699	66576	162978
(unique)	(6903) <sup>a</sup>	(6872) <sup>a</sup>	(6996) <sup>a</sup>	(15485)
I/σ overall	12.5 (3.5) <sup>b</sup>	11.7 (4.1) <sup>b</sup>	12.8 (2.6) <sup>b</sup>	14.6 (3.8) <sup>b</sup>
Completeness (%)	99.9 (99.9) <sup>b</sup>	99.9 (99.9) <sup>b</sup>	99.6 (97.1) <sup>b</sup>	99.8 (99.7) <sup>b</sup>
R <sub>merge</sub> (%)	10.5 (31.0) <sup>b</sup>	11.4 (24.0) <sup>b</sup>	10.5 (39.5) <sup>b</sup>	7.5 (46.9) <sup>b</sup>
Structure determination				
Resolution (Å)	10 – 3.0			
Figure of merit	0.535			
Phasing power	(1.73/1.09) <sup>c</sup>	(2.11/1.56) <sup>c</sup>		
R <sub>Cullis</sub> (dispersive)	0.95	0.92		
R <sub>Cullis</sub> (isomorphous)	(0.71/0.78) <sup>c</sup>	0.64/0.65) <sup>c</sup>		
Structure refinement				
Resolution (Å)				20 – 1.90
Number of work/test reflections				13590/1533
Nonhydrogen protein/carbohydrate/solvent atoms				844/20/154
R <sub>work</sub> /R <sub>free</sub> (%)				23.8/28.2
Bond length (Å)/angle (°) rms deviation from ideal geometry				0.008/1.631
Ramachandran statistics (%)				
(Favorable/additional/generous/forbidden)				83.9/16.1/0/0
Protein atoms average B value (Å <sup>2</sup> ), mainchain/sidechain				39.70/41.42

<sup>a</sup>Including Bijvoet pairs.<sup>b</sup>(Last resolution bin).<sup>c</sup>(Acentric/centric).

acids, of which 12 are highly conserved. Most notably, the NH<sub>2</sub>-terminal portion of the TSR contains two or three tryptophan residues separated by two to four amino acids each (Klar et al., 1992; Adams and Tucker, 2000; Kilpelainen et al., 2000). The majority of TSRs have six cysteine residues; however, those found in some complement factors and in the malaria proteins contain fewer cysteines. In addition, two arginines and two glycines are also highly conserved.

The functions of the three TSRs that are present in TSP-1 have been extensively studied. They reportedly function as (a) attachment sites for many cell types, (b) inhibitors of angiogenesis, (c) protein binding sites, and (d) glycosaminoglycan (GAG) binding sites (for review see Chen et al., 2000). Synthetic peptides have been used to identify amino acid residues within the TSRs of TSP-1 that are important for the anti-angiogenic activity. Tolsma et al. (1993) found that peptides from the second and third TSRs that included the CSVTCG sequence inhibit angiogenesis. This sequence is reportedly involved in the binding of TSP-1 to CD36, and CD36 on endothelial cells reportedly mediates the anti-angiogenic effect of TSP-1 (Dawson et al., 1997). However, two groups have shown that synthetic peptides that flank either side of the CSVTCG sequence, but do not include it, also inhibit angiogenesis (Dawson et al., 1999; Iruela-Arispe et al., 1999). One of these includes the W\*\*W\*\*W (the asterisk

indicates positions that are occupied by various amino acids) sequence that has been proposed to be a GAG binding site (Guo et al., 1992). The other sequence, GVITRIR, contains the highly conserved arginine residues (Dawson et al., 1999).

A unique function has been mapped to the sequence between the first and second TSRs of TSP-1. This region binds and activates TGFβ (for review see Murphy-Ullrich and Poczatek, 2000). This activity is mediated by the tripeptide RFK of TSP-1 and is conserved in all TSP-1 sequences determined to date. By contrast, this sequence is not found in the TSRs of most other proteins, including the closely related TSP-2 protein. An RFK sequence is found in the sixth TSR of F-spondin but it falls in a different location within the sequence and, to date, it has not been shown to activate TGFβ (Klar et al., 1992).

Taken together, the data indicate that the TSR-containing proteins participate in development, angiogenesis, tumor progression, axon guidance, activation of TGFβ, and wound healing. Some of the amino acid sequences that mediate TSR functions have been identified. However, several nonoverlapping synthetic peptides have been reported to have similar functions. To better understand the biological function of the TSRs, we have determined the three-dimensional structure of the TSRs of human TSP-1 by X-ray crystallography.

## Results

### Crystal structure determination of the h3TSR

A soluble recombinant protein containing all three TSRs of human TSP-1 (designated h3TSR) was expressed in *Drosophila* S2 cells. We have previously shown that this protein, h3TSR, is active in that it activates TGF $\beta$  and inhibits endothelial cell migration, angiogenesis, and tumor growth (Miao et al., 2001). The h3TSR contains 186 amino acid residues, including Asp358–Ile530 of human TSP-1, six vector-derived amino acids (RSPWPG) at the NH<sub>2</sub> terminus, and two vector-derived amino acids (TG) followed by a His tag at the COOH terminus. Crystals were obtained that diffracted to 1.9 Å resolution, and the structure was solved by the multiple-wavelength anomalous diffraction (MAD) method (Hendrickson, 1991) using a Hg-derivatized crystal (Table I). Only residues from TSR2 and TSR3 are seen in the electron density maps. SDS-PAGE run on crystal samples revealed a single band with a molecular mass of around 17 kD, corresponding to a two TSR domain protein (unpublished data). Although the three TSRs in the intact TSP-1 protein are resistant to protease digestion, apparently the high concentration of h3TSR in the crystallization droplet exposes the highly basic sequence of KRFK between TSR1 and TSR2 to a trypsin-like protease for degradation. The final model hence just includes the second and third TSRs arranged in tandem with a total of 117 residues. For the following discussion, we will refer to this model as hTSR[2,3].

### The overall hTSR[2,3] structure

Fig. 1 A is a ribbon diagram of the molecular structure of hTSR[2,3]. Each TSR folds into a long, thin, spiraling, anti-parallel, three-stranded domain with dimensions of  $\sim 15 \times 20 \times 55$  Å. The core of each domain is a novel eight-layered structure that will be described in detail below. The tilt and twist angles between TSR2 and TSR3 are 40° and 180°, respectively. The 180° twist places the TSR2 and TSR3 domains facing in opposite directions. DALI search with the TSR domain (using the website <http://www2.ebi.ac.uk/dali/>) failed to identify any homologous structures in the database.

The TSR2 and TSR3 domains are the products of exons 9 and 10 of the human TSP-1 gene, respectively. The exception is Ile473, which is encoded by exon 9 but does not belong to TSR2. Rather, Ile473 forms two main chain hydrogen bonds to Gln507, joining the top portions of the TSR3 domain. The three residues at the COOH terminus of exon 10 are not visible in the electron density map of the current crystal structure. Pro472 appears to be a linker between the TSR2 and TSR3 domains. Immediately before Pro472 is the disulfide bridge of Cys471–Cys433, which stabilizes the conformation of the bottom portion of the TSR2 domain. The residue Pro472 (the linker between TSR2 and TSR3) and its spatially neighboring residues Ala470, Cys471, Ile473, Phe508, and Gly509 create a hydrophobic interface between the two domains. These data suggest that there is a relatively rigid linkage between the TSR2 and TSR3 domains. Based on the amino acid sequence alignment shown in Fig. 1 B, the TSR1 domain should have a very similar fold as that of the TSR2 and TSR3 domains. A distinct feature of the linkage between the TSR1 and TSR2 domains is that there is a five-

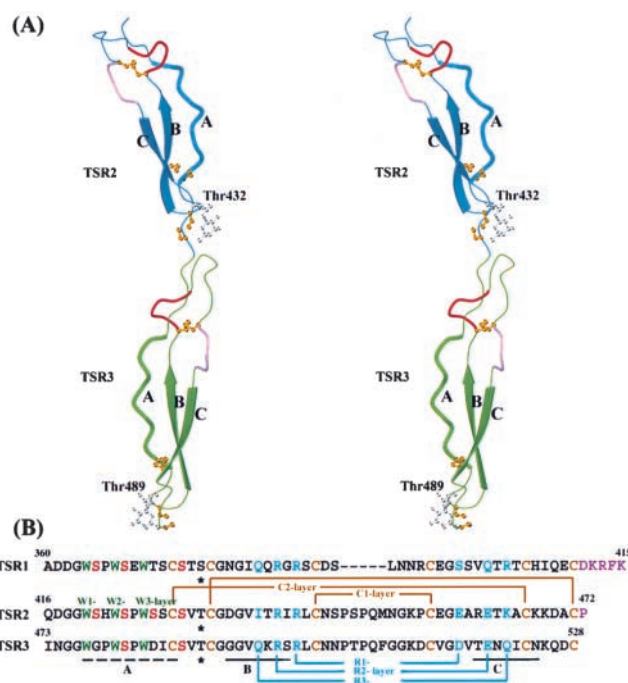
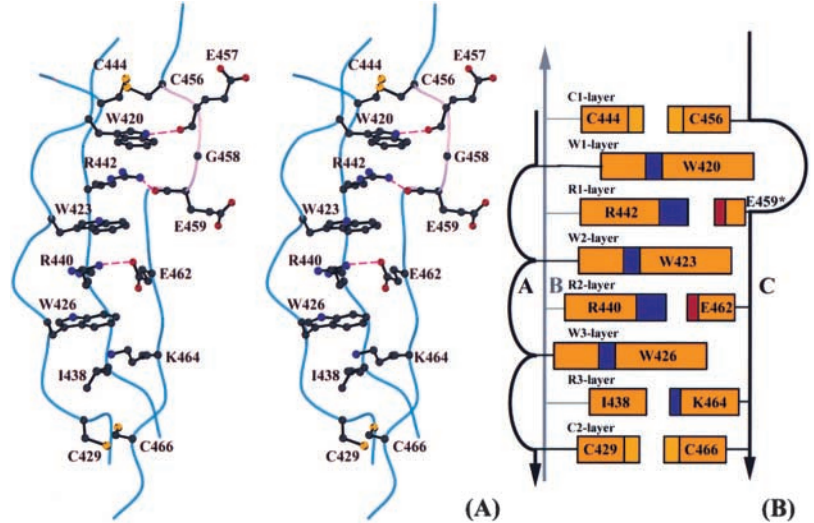


Figure 1. **The structure of hTSR[2,3].** (A) A stereo view of the ribbon diagram of hTSR[2,3]. The disulfide bonds (in yellow) and O-linked fucose moieties (in gray) at Thr432 and Thr489 are drawn in ball and stick representation. The small and large jar handle motifs in each TSR domain are highlighted in dark red and purple, respectively. (B) The sequence alignment of hTSR domains. The  $\beta$  strands are underlined for strands B and C, whereas the bulged strand A is marked by a dashed line. The disulfide bonds are marked in dark yellow lines connecting paired cysteines. The tryptophans that form the W layers are shown in green and the amino acids that form the R layers are shown in blue. The pairing between amino acids from the B and C strands that form the R layers is marked by blue lines. The serine residues of the A strand that contribute to hydrogen bonds to the backbone of the B strand (see Fig. 4) are in red. The linker residues between TSR domains are in purple. Fig. 1 A was prepared using the program MOLSCRIPT (Kraulis, 1991).

residue linker (D411KRFK), as opposed to just the one residue that is between TSR2 and TSR3. Presumably, the linkage between TSR1 and TSR2 is very flexible. The structures of TSR2 and TSR3 are so similar that the rmsd value for the superposition of TSR2 and TSR3 is only 0.62 Å (two loops are excluded). Therefore, unless otherwise indicated, only TSR2 is used for the discussion below.

Two types of glycosylation have been identified biochemically in the TSRs (Hofsteenge et al., 2001). Hofsteenge et al. (2001) have demonstrated that a fucose is O-linked to the threonine in the CSVTCG sequence motif on each TSR. Our electron density maps have clearly shown that one fucose moiety each can be built onto the TSR2 and TSR3 domains and no electron density beyond this sugar unit is visible. The modeled glycan is located on the AB loop at the bottom of each domain (Fig. 1 A). Another potential glycosylation is the C-mannosylation of the first tryptophan of the W\*\*W motif (Hofsteenge et al., 1999, 2001). In platelet TSP-1, for instance, Trp367, Trp420, Trp423, and Trp480 are C-mannosylated (Hofsteenge et al., 2001). The S2 expression system used in this work, however, does not allow

**Figure 2. CWR-layered core structure of the TSR domain.** (A) A stereoview of C, W, and R layers in TSR2 of TSP-1. Displayed residues that are directly involved in forming the layered structure are drawn in ball and stick representation with salt bridges, and hydrogen bonds drawn as dashed lines. The big jar handle motif, which is associated with the first W layer is highlighted in pink. (B) A schematic drawing of the CWR-layered structure with each layer and layer-forming residue(s) labeled. The residue Glu459 that is marked with an asterisk forms a hydrogen bond between its main chain carbonyl group and the side chain of Arg442 in the R1 layer. The three antiparallel strands are drawn in lines schematically with arrowheads indicating their polarities. The three bulges associated with the rippled strand A and the big jar handle are also shown. Fig. 2 A was prepared using the program MOLSCRIPT (Kraulis, 1991).



for this posttranslational glycosylation (Hofsteenge et al., 2001). There is indeed no electron density to be seen for any C-manosylation in our calculated maps. However, the side chains of all three tryptophans on each TSR are so oriented that the Cδ1 atoms are exposed, ready for accommodation of mannoses (Hartmann and Hofsteenge, 2000) (Fig. 2). The function of this glycosylation is not well studied.

**The CWR-layered core structure of the TSR fold**

Small domains like the TSR (<100 amino acid residues) very often have special folds that lack regular secondary structure and are maintained by the rich presence of disulfide bonds or metal ions (Richardson, 1981). The exceptions are some recently identified small modules, such as SH3 (Pawson, 1995) and WW (Macias et al., 1996) domains, that are present in a number of signal transduction proteins and that do have reg-

ular secondary structures as basic folding elements. The TSR domain described in this article, however, represents a novel fold, for which the interdigitating side chain stacking of cysteine, tryptophan, and arginine in layers (the CWR layer) from three strands comprises the core of the structure (Fig. 2). Only the B and C strands form limited regular β structures (Fig. 1). By contrast, the A strand assumes a unique rippled conformation (Fig. 1 A; Fig. 2) and has a conserved sequence motif of W\*\*W\*\*W (Fig. 1 B). As shown in Fig. 2 A, the side chains of the three tryptophans (Trp420, Trp423, and Trp426 of TSR2) make up three tryptophan layers (the W layers) that play a central role in the fold. Sandwiched between each of the two adjacent W layers is one guanidinium group of an arginine from strand B. A comparison of the amino acid sequences of the TSRs from various proteins reveals that the sequence R\*R\*R is commonly found in strand

**Figure 3. Multiple sequence alignment of TSRs.**

The alignment, in particular, compares the amino acid sequence of layer-forming residues in two major groups of TSRs as discussed in the text. The prototypical disulfide bond patterns of groups 1 and 2 are drawn schematically in the left margin with disulfide bonds as yellow lines. Group 1 is in the top panel, and group 2 is in the middle panel. For group 1, two hydrogen bonds between the small jar handle and NH<sub>2</sub> terminus are drawn as red lines. In the sequences, cysteines that form disulfide bond are also yellow, whereas those that are unpaired or involved in "extra" disulfide bonds are magenta. Tryptophans or tryptophan substituents in W layers are in green, and arginines or arginine substituents as well as their paired residues in R layers in cyan. In the bottom panel, the sequences of TSR-like domains from midkine and HB-GAM are aligned.

		A	B	C
TSP-1	1	ADDDNSPHEVTS-RT	SGM-GI	QDQSGD
RAI-1	1	QDQNSPHEVTS-SV	TGD-GV	QDQNSP
ADAMTS-1	1	INGGGRSPDPI-SV	TGG-GI	QDQNSP
ADAMTS-4	1	ATDQKLSLWGE-TR	DGG-GL	QDQNSP
Properdin	1	VHGANDSPHSL-SV	TGR-GP	QDQNSP
UNC-5	1	VDDGNSPHEVTS-SV	TGD-GV	QDQNSP
C6	1	NDVDDGNSPHEVTS-SV	TGD-GV	QDQNSP
C7	1	GVDDGNSPHEVTS-SV	TGD-GV	QDQNSP
C8β	1	AKADGNSPHEVTS-SV	TGD-GV	QDQNSP
C8β	1	FIDKNSPHEVTS-SV	TGD-GV	QDQNSP
C9	1	VVDGNSPHEVTS-SV	TGD-GV	QDQNSP
F-spondin	1	HTC-IYVNSPHEVTS-SV	TGD-GV	QDQNSP
N-spondin	1	LDVDDGNSPHEVTS-SV	TGD-GV	QDQNSP
Midkine	1	LDVDDGNSPHEVTS-SV	TGD-GV	QDQNSP
HB-GAM	1	LDVDDGNSPHEVTS-SV	TGD-GV	QDQNSP

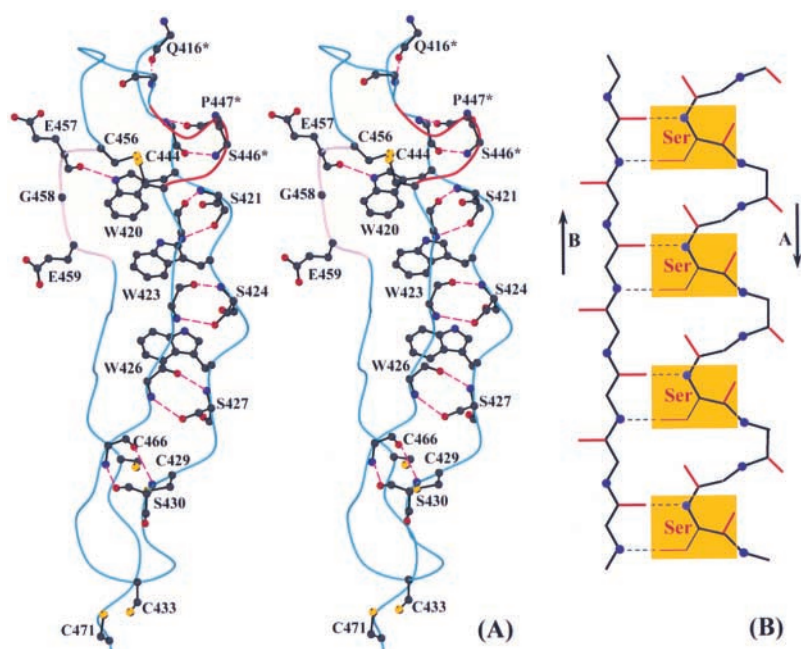


Figure 4. **Rippled A strand and jar handle motifs.**

(A) A stereoview of the bulged strand A, the jar handle motifs, and their associated hydrogen bond networks. Displayed side chains of staggering tryptophans, serines, and other relevant residues are drawn in ball and stick representation. For the residues marked with asterisk, only main chain atoms that are involved in hydrogen bonds are shown. The big and small jar handles are highlighted in pink and dark red, respectively. (B) A schematic drawing of the hydrogen bond network between the antiparallel bulged strand A and the regular  $\beta$  strand B. The serine residues that are involved in the formation of side chain to main chain hydrogen bonds between the two strands are shaded in yellow. Fig. 4 A was prepared using the program MOLSCRIPT (Kraulis, 1991).

B (Fig. 3). These arginines comprise the R layers. The alternate stacking of the planar cationic guanidinium groups of the arginines and the aromatic side chains of the tryptophans forms multiple cation- $\pi$  interactions (Flocco and Mowbray, 1994; Gallivan and Dougherty, 1999), which may provide vital stabilization energy for TSR folding. The spacings between the R and W layers are observed to be around 4.6 Å throughout TSR2 and TSR3. A similar structural element is found in the class I cytokine receptor family. The so-called “WSXWS box” is observed in this family, but its function is controversial (Somers et al., 1994). In the human growth hormone receptor, for instance, the stacked tryptophans and arginines come from four different strands on the same face of a membrane proximate Ig-like domain. This face is not involved in hormone binding, nor receptor dimerization (de Vos et al., 1992). The erythropoietin receptor has the WSXWS sequence motif in a  $\beta$  bulge, and the pyrrole rings of two Trps sandwich a guanidinium group of an Arg from a neighboring  $\beta$  strand (Livnah et al., 1996). The uniqueness of the TSR domain is that, to our knowledge, it is the first example in which the alternately stacked W and R layers capped by C layers constitute the major structural feature of the domain. In the TSRs of other proteins, a glutamine or lysine sometimes replaces the first arginine in the sequence (the R3 layer in Fig. 3). In TSP-1, TSR1 and TSR3 have a glutamine in the R3 layer, whereas TSR2 has an isoleucine. The amino group from a lysine and the amide group from a glutamine can also participate in cation- $\pi$  interaction with a tryptophan residue (Flocco and Mowbray, 1994; Gallivan and Dougherty, 1999).

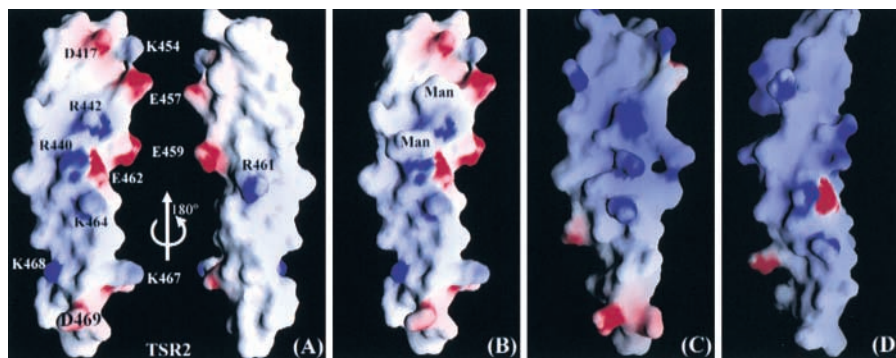
The arginines in the R layers also always interact with a residue from the C strand laterally, either in a salt bridge like the Arg440–Glu462 pair, or in a hydrogen bond to the carbonyl group like the Arg442–Glu459 pair. In TSR2, Ile438 in the R3 layer makes hydrophobic contacts with Lys464 of the C strand between the aliphatic portions of their side chains. The sequence alignment shown in Fig. 3 demonstrates that

the residues in the R layers from the B strand and their interacting partners in the C strand frequently have large side chains. This suggests that the contacts between hydrophobic portions of the paired side chains may also play a role in stabilizing the TSR domain. In TSR3 of TSP-1, Gln495 in the R3 layer and Gln521 from strand C form a typical hydrogen bond between the NE2 atom and OE1 atom of their respective amide groups. Again, we see a spectrum of variation in the R3 layer among different TSR domains.

The six alternated W and R layers are capped by disulfide bonds at the top and bottom, hence the two C layers. These are Cys444–Cys456 and Cys429–Cys466 in TSR2 (Fig. 2), Cys501–Cys513 and Cys486–Cys523 in TSR3, and their conserved counterparts in other TSRs (Fig. 3). Fig. 2 B depicts the eight CWR-layered fold of the TSR2 domain in a schematic manner. Within the eight layers (in the order of C1, W1, R1, W2, R2, W3, R3, and C2 from top to bottom), all of the side chains that form the CWR-layered structure are in a planar and parallel conformation. Even the side chain of Ile438 in the R3 layer of TSR2 assumes the best possible orientation for stacking. Due to the inherent strand twist, the whole CWR-layered structure is in a right-handed spiral conformation, reminiscent of a DNA structure.

The rippled A strand with its multiple bulges is another interesting feature of the eight-layered TSR domain. As mentioned previously, the A strand is not a regular  $\beta$  strand. It nevertheless has a unique hydrogen bonding network to the B strand. Fig. 4 shows how a serine after each of the tryptophans in the W layers and after the COOH-terminal Cys429 on the A strand forms two hydrogen bonds to the backbone of the B strand: one via its hydroxyl side chain and the other with its amide group. The formation of side chain to main chain hydrogen bonds between the two antiparallel strands creates even larger bulges in comparison to a well-defined wide bulge (Richardson, 1981). A threonine or a glycine sometimes replaces the serine in these positions in other TSRs (Fig. 3). The former can act like a serine, and

Figure 5. **Electrostatic potential surface representation of TSRs.** (A) The front and back faces of TSP-1 TSR2 with charged residues labeled. The two orientations are related by a rotation of 180°. (B) The front face of a modeled C-mannosylated TSP-1 TSR2 with two mannoses attached to W420 and W423 as labeled. (C) The front face of a homologous model of ADAMTS-4 TSR. (D) The front face of a homologous model of F-spondin TSR6. This figure was prepared using the program GRASP (Nicholls et al., 1991).



the latter allows a sharp turn. They all favor the bulged conformation for the A strand. This hydrogen bonding network between the A and B strands, the regular  $\beta$  structure of the B and C strands, the disulfide bonds, and the side chain stacking of the CWR-layered structure ensure a stable folding unit for this small domain.

### The “jar handle” structures

The CWR-layered structure is the core of TSR domain. At the bottom of both the TSR2 and TSR3 domains, a cysteine on the AB loop forms the third disulfide bond with a cysteine in the COOH terminus of strand C (Fig. 1 A; Fig. 4 A). This disulfide bond consolidates the bottom of the domains. At the top of the domains, there is a relatively long BC loop that tends to be enriched in proline, glycine, or charged amino acids in the TSRs of various proteins. In the TSR structure of TSP-1, the loop features a regular  $\beta$  turn in the middle and two “jar handle”-like structures at the beginning (the dark red-colored smaller one in Fig. 1 A and Fig. 4 A) and the end (the purple-colored larger one in Fig. 1 A and Fig. 4 A). The small jar handle is at the beginning of the BC loop and above the C1 layer (the Cys444–Cys456 pair in TSR2; Fig. 4 A). It appears somewhat like a conventional  $\beta$  turn composed of four residues (Asn445–Ser448), but the characteristic hydrogen bond within the turn is missing. Instead, the amide group of the second residue (Ser446) and the carbonyl group of the third residue (Pro447) turn to form hydrogen bonds to the main chain of the NH<sub>2</sub> terminus of the A strand (Fig. 4 A). This unique local structure brings the beginning of the A strand closer to the BC loop, while the nearby disulfide bond in the C1 layer cross-links the B and C strands. Together, the interactions stabilize the top of the domain. The third residue is a cis-proline in both TSR2 and TSR3, which may favor the formation of a jar handle. In TSR1 and the TSRs of some other proteins, this proline is absent (Fig. 1 B; Fig. 3). It is difficult to convincingly predict if this small jar handle exists in these proteins. The unusual handle structure acts just like the serine hydrogen bonding network to integrate the A strand into the domain (Fig. 4 A). As discussed later, a disulfide bond may serve this function in the TSRs of some other proteins.

The large handle (purple in Fig. 1 A and Fig. 4 A) is at the end of the BC loop, and immediately below the C1 layer. It consists of five residues (Cys456–Ala460). The interesting features of this handle are that the third position is a conserved Gly458, and the carbonyl group of the second resi-

due, Glu457, forms a hydrogen bond to the NH1 atom of the tryptophan in the W1 layer (Fig. 4 A). This allows this tryptophan’s bulky side chain to be accommodated within the handle. By contrast, the tryptophans in the W2 and W3 layers have their polar NH1 atoms exposed and available for potential ligand binding.

### A positively charged groove on the “front” face of the TSR: the proposed “recognition” face

The above-mentioned exposed tryptophans from the W layers along with the exposed arginines from the R layers define the “front” face of the domain. This is best seen in Fig. 2 A, where the three strands run along the “back,” providing a framework to support a continuous positively charged front face containing side chains from the W and R layers. Other residues that are located at the edge of the A and C strands also help to create a groove-like structure within this positively charged region. Fig. 5 A is a molecular surface representation of TSR2 with a few residues labeled on the surface. This figure clearly depicts this DNA-like right-handed spiral groove. We propose that this is likely to be the “recognition face” of the domain. The molecular orientation in Fig. 1 A sets the TSR2 domain’s recognition face away from the reader, whereas the TSR3 domain’s recognition face is toward the reader.

Having the structure of TSP-1 TSRs at hand enables us to interpret some functional data that are largely based on peptide studies. In general, a short linear peptide cannot easily mimic its counterpart in a folded TSR domain. Nevertheless, some peptide investigations are still suggestive, as discussed below.

## Discussion

### GAG binding properties of TSRs

GAGs and proteoglycans constitute one major component of extracellular matrix. GAGs are long, unbranched polysaccharide molecules composed of repeating disaccharide units. One of the disaccharides is always a uronic acid, whereas the second one is an *N*-acetylglucosamine or *N*-acetylgalactosamine, which are usually sulfated, rendering the molecule very negatively charged. The ability to bind GAGs is important to the function of many extracellular and membrane proteins. TSRs have been reported to mediate GAG binding in several proteins. Crystal structures of heparin–protein complexes indicate that a cluster of positively charged resi-

dues on a protein's surface can interact with the negatively charged carboxylate and sulfate groups of heparin (Pellegrini et al., 2000; Lietha et al., 2001). Heparin molecules are in a right-handed helical conformation, and the disaccharide repeat of heparin spans  $\sim 9$  Å (Mulloy et al., 1993). The distance between the W1 layer and R3 layer in the TSR is close to 20 Å. It is conceivable that two spiral disaccharide units may fit well into the right-handed groove in the recognition face of the TSR. The C-mannosylation on some tryptophans (Hofsteenge et al., 2001) of TSP-1 TSRs does not significantly change their surface electrostatic potential, but creates protrusions along the edge of the grooves (Fig. 5 B). The mannose groups may help to define the groove in the recognition face. Because the separation of two arginines in the groove is also  $\sim 9$  Å, the speculation might be that the negatively charged groups of GAGs can form electrostatic interactions with these arginines. TSRs of TSP-1 have a relatively low affinity for heparin that may not be detectable at physiological salt concentrations (Panetti et al., 1999), suggesting that residues other than the tryptophans and arginines, in particular residues from the C strand, may contribute significantly to GAG binding.

The protein ADAMTS-4 or aggrecanase-1 is a newly identified cartilage-degrading enzyme. It contains one TSR, which is highly homologous to the TSRs of TSP-1. The strong binding of the TSR of ADAMTS-4 to the GAG side chains of aggrecan (the major sulfate proteoglycan in cartilage) appears to be necessary for the enzymatic cleavage of aggrecan that occurs during disease-related cartilage destruction (Tortorella et al., 2000). It has been reported that synthetic peptides that include sequences found in the A and C strands of the TSR of ADAMTS-4 have weak affinities for aggrecan. The C strand of the ADAMTS-4 TSR contains more arginine residues than the TSRs of TSP-1 (Fig. 3). Fig. 5 C shows the surface representation of the ADAMTS-4 TSR domain, which has been modeled based on the TSR structure of TSP-1 reported here. This figure demonstrates that positive charges cover nearly the entire front face of this TSR domain due to the extra positively charged residues from the C strand. This model may explain the high binding affinity of ADAMTS-4 for the GAG chains of aggrecan.

F-spondin is a matrix-associated protein that is expressed in the floor plate and involved in patterning the axonal trajectory of commissural and motor neurons. F-spondin has multiple TSRs in its COOH terminus as well as reelin and mindin homologous regions in the NH<sub>2</sub> terminus (Klar et al., 1992; Tzarfaty-Majar et al., 2001). The interaction of the F-spondin TSRs with extracellular matrix and neuronal growth cones has been reported to involve binding to GAGs (Klar et al., 1992; Tzarfaty-Majar et al., 2001). This binding is mediated by the fifth and sixth TSRs, whereas the first four TSRs are not involved. Comparing the sequence of the TSRs of F-spondin with those of TSP-1 reveals that the fifth and sixth TSRs of F-spondin have more basic residues on the C strand and BC loop (Fig. 3). These basic residues contribute positive charges to the front face of these TSRs as they do in the ADAMTS-4 TSR. For example, the residues of Glu457 and Glu459 (Fig. 5 A) in TSR2 of TSP-1 are substituted by Lys794 and Lys796 in TSR6 of F-spondin,

respectively (Fig. 3). The only two acidic residues (an aspartic and a glutamic acid) on the C strand of TSR6 in F-spondin point away from the front surface. The front face of TSR6 is therefore dominated by a cluster of positively charged residues, as shown in Fig. 5 D.

The interaction of TSRs with GAGs is also involved in cellular invasion by apicomplexa. This process is central to the life cycle of *P. falciparum* and to the etiology of malaria. Host cell invasion involves the binding of circumsporozoite protein and thrombospondin-related anonymous protein (TRAP) on the parasite surface to GAGs on the surface of salivary glands in the mosquito and liver hepatocytes in the vertebrate host (Matuschewski et al., 2002). Mutations in the TSR sequence of TRAP result in a decrease in the ability of the parasite to invade salivary glands and liver cells (Wengelnik et al., 1999; Matuschewski et al., 2002). Mutation of the tryptophan in the W3 layer to alanine or substitution of the positively charged amino acids in and around the R layers with alanine residues decreased sporozoite infectivity of the salivary gland (Matuschewski et al., 2002). Because these residues are integral components of the CWR-layered structure, our data suggest that these mutations would disrupt the overall folding of the TSR and the formation of the recognition face.

Taken together, these data suggest that positive charges across the front face of a TSR are essential for high-affinity binding to GAGs in many important biological processes. More basic residues in the C strand and elsewhere around the front face probably contribute to stronger binding.

### The binding of TSP-1 to CD36

TSP-1 is one of several naturally occurring inhibitors of angiogenesis. The inhibition of angiogenesis by TSP-1 is reportedly due in part to CD36-mediated inhibition of endothelial cell migration and induction of apoptosis (Dawson et al., 1997; Jimenez et al., 2000). Peptides having sequences from strand A and the sequence of GVITRIR from strand B of the second TSR have been shown to have the inhibitory activities (Tolsma et al., 1993; Dawson et al., 1999; Iruela-Arispe et al., 1999; Jimenez et al., 2000). The major portions of these peptides can now be mapped on the front face of the TSR structure presented here. On the other hand, the region between amino acids 90 and 120 of CD36 has been implicated in TSP-1 binding (Leung et al., 1992; Pearce et al., 1995). Secondary structure predictions (using the website <http://www.embl-heidelberg.de/predict-protein/predictprotein.html>) suggest that the sequence T(103)QDAEDN of CD36 falls into a negatively charged loop. This sequence is conserved in human and mouse CD36. We hypothesize that this negatively charged loop of CD36 may interact with the positively charged groove of the TSR's front face. The CSVTCG sequence has also been implicated in the interaction of TSP-1 with CD36. Upon careful examination of the folded TSR, one sees that the two cysteines in the CSVTCG sequence are involved in two separate disulfide bonds and that the threonine is O-glycosylated (Hofsteenge et al., 2001). Thus, the activities of synthetic peptides may not correlate with the function of the CSVTCG sequence in the context of the three-dimensional structure of the TSR.

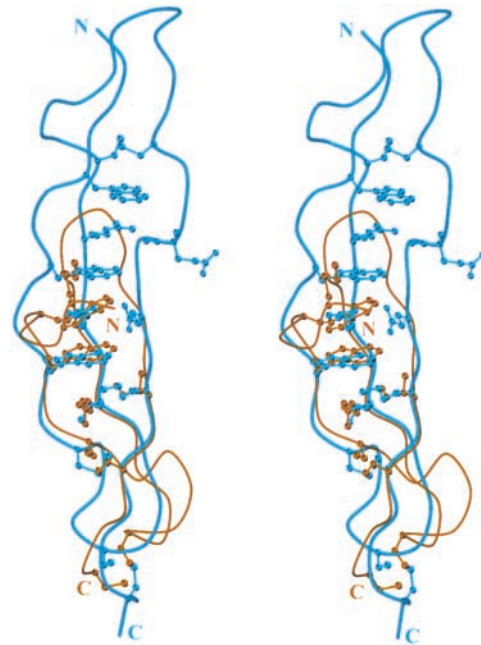
### Disulfide bond pattern and TSR grouping

A close scrutiny of the TSRs in the database reveals considerable variability (Fig. 3). Based on the disulfide bond patterns and multiple sequence alignment algorithms, we tentatively divide TSRs into two major groups represented by the TSRs of TSP-1 (group 1) and F-spondin (group 2). The major difference is the disulfide bond at the top of TSR's layered structure (the C1 layer), as illustrated at the left margin of Fig. 3.

The TSRs of TSP-1 are the prototype of group 1. Within this group (Fig. 3, top), the C1 layer at the top of the domain is formed by cysteines from the end of the B strand and the beginning of the C strand. In some of these TSRs, such as TSRs found in TSP-2, BAI-1, -2, and -3, ADAMTS family members, properdin, and others, there is a third disulfide bond at the bottom of the domain that helps to consolidate the bottom of the domain as described before. There are also TSRs found in the COOH terminus of the complement factors C6, C7, C8 $\alpha$ , C8 $\beta$ , and C9, in which this third disulfide bond is absent. Instead, there appears an unpaired cysteine residue between the first and second tryptophans on the A strand (Fig. 3).

The structure of group 2 can be predicted for the TSRs of F-spondin. In these TSRs, the cysteines that form the C1 layer are predicted to be from the A and C strands, rather than from the B and C strands (Fig. 3, bottom). This disulfide bond pattern is found in midkine (Iwasaki et al., 1997). Thus, the beginning of the A strand is linked to the top portion of the domain through this disulfide bond. Note that in group 1, it is a jar handle-like element that integrates the A strand into the top of the domain (represented by two hydrogen bonds in red in Fig. 3, top). Like those TSRs in TSP-1, the TSRs of F-spondin also have a third disulfide bond at the bottom of the domain. TSRs of M-spondin, mindin-1 and -2, the NH<sub>2</sub>-terminal of complement factors, and TRAP/CTRP (circumsporozoite protein from *P. falciparum*) may fall into the same category. In the TRAP TSR and CTRP TSR1, 2, 3, 5, 6, and 7, the first tryptophan on the A strand is absent, presumably resulting in a seven-layered TSR core structure. An extra disulfide bond may exist in a few TSRs, like TSR2 of properdin, in which the NH<sub>2</sub> terminus of the A strand is probably linked to the beginning of the C strand (see two cysteines in red for TSR2 of properdin in Fig. 3).

Midkine and HB-GAM have two TSR-like domains (Iwasaki et al., 1997; Kilpelainen et al., 2000). Sequence alignment demonstrates their large deviation from TSRs of other proteins (Fig. 3, bottom). The NMR structure of midkine's TSR-like domains (Iwasaki et al., 1997) showed that although their overall topology is reminiscent of that of TSP-1's TSRs reported here, they depart substantially in the CWR-layered structure of a TSR domain. Not only are the layers significantly fewer, but they can hardly be structurally compared with that found in TSP-1's TSRs. Fig. 6 is the superposition of the NH<sub>2</sub>-terminal TSR-like domain of midkine from the NMR structure onto TSR2 of the TSP-1. The layered positioning of tryptophan and arginine residues to form multiple cation- $\pi$  interactions is a key feature of the TSP-1 TSRs. This does not seem to exist in midkine where there is only one potential cation- $\pi$  interaction. A DALI



**Figure 6. A stereo view of a structure superposition of TSP-1 TSR2 (in light blue) and the NH<sub>2</sub>-terminal domain of midkine (in brown).** Residues mostly from the  $\beta$  strand regions of the two structures, including W426–S430, D435–R440, and R461–K468 from TSP-1 TSR2 and W20–T24, V32–G37, and T43–V50 from the NH<sub>2</sub> terminus of midkine (Protein Data Bank code, 1mkn), were used for the least squares structure alignment. Residues with displayed side chains in ball and stick representation are those highlighted in the TSP-1 TSR2 and NH<sub>2</sub>-terminal midkine sequences aligned in Fig. 3. The NH<sub>2</sub> and COOH termini of the two structures are labeled. This figure was prepared using the program MOLSCRIPT (Kraulis, 1991).

search with TSP-1's TSR domain did not detect similarity to the structure of midkine, indicating that the structures are distinct.

The disulfide bond-based grouping might provide a framework to aid functional studies. For example, some proteins in groups 1 and 2 seem to differ in their effects on angiogenesis. The TSRs that are present in TSP-1, TSP2, BAI-1, and ADAMTS-1 in group 1 have been shown to inhibit angiogenesis (Nishiomori et al., 1997; Vázquez et al., 1999). By contrast, some of the proteins in group 2, including two members of the CCN family and HB-GAM, reportedly stimulate angiogenesis (for review see Lau and Lam, 1999; Papadimitriou et al., 2001).

## Materials and methods

### Protein production, crystallization, and X-ray data collection

A recombinant version of all three TSRs of human TSP-1, designated h3TSR, was prepared by PCR using the full-length cDNA of human TSP-1, as described previously (Miao et al., 2001). For crystallization, the purified protein was dialyzed against 20 mM Na<sub>3</sub>PO<sub>4</sub> (pH 7.8) containing 500 mM NaCl, and was reappplied to a column of ProBond resin. The protein eluted with 500 mM imidazole was dialyzed against 20 mM Na<sub>3</sub>PO<sub>4</sub> (pH 7.0) and 500 mM NaCl, and 1% sucrose was added before storage at  $-80^{\circ}\text{C}$ .

Single crystals of h3TSR were grown from a crystallization buffer containing 0.5 M sodium potassium tartrate and 0.1 M sodium acetate at pH 5.5 using the vapor diffusion hanging drop method. For the data collection at cryogenic temperature, the crystals were treated with a cryoprotectant solution (25% glycerol, 0.5 M sodium potassium tartrate, 0.1 M sodium acetate, pH 5.5), and then frozen and stored in liquid nitrogen. Mercury



derivatives were prepared by soaking the crystals in the same cryoprotectant solution containing 1 mM ethylmercury chloride for 4 h. X-ray diffraction data were collected from pre-frozen crystals at APS SBC 19ID at a temperature of 100 K. A native crystal diffracted to a resolution of 1.9 Å, with one molecule in one asymmetric unit. A MAD dataset of the mercury derivative was obtained to a resolution of 2.1 Å. All the raw data were indexed and reduced with HKL2000 (Table I).

### Structure determination and refinement

Using programs in the CCP4 suite (CCP4, 1994), we located one major Hg binding site in one asymmetric unit in both isomorphous and anomalous difference Patterson maps. Considering possible nonisomorphism between the native and Hg derivative crystal, the MAD data alone from Hg derivative were actually used for heavy atom parameter refinement and phasing with the high-energy dataset as reference. After the refinement and initial phasing at 3 Å resolution with the program MLPHARE in the CCP4 suite, an additional minor Hg binding site was identified and two sites were refined together. Phase extension was performed subsequently using the native data to 2.5 Å by solvent flattening and histogram matching with DM (Cowtan, 1994). An improved Fourier map was obtained and used for initial model building with the program O (Jones et al., 1991). After each step of model building, the combined phases from experimental and calculated from the partial model were used to perform density modification. The model building cycle was repeated until ~90% of the residues from TSR2 and TSR3 were built into the model.

The native dataset was eventually used for the completion of model building and refinement. The final model was refined at 1.9 Å resolution to an Rfree factor of 28.2% and Rwork of 23.8% (Table I) using the Xplor (Brunger, 1992). At a 1.5 sigma contour level in 2Fo-Fc map, there was continuous density for the main chain backbone except for part of the glycine-rich AB loop and the COOH terminus of TSR3. The final model contains 113 residues (from Gln416 to Cys528) with two O-linked fucoses associated with residues Thr432 and Thr489. There was no indication on the electron density maps for the existence of the first domain as mentioned before. The current model also includes a total of 154 water molecules. The coordinates have been deposited in the Protein Data Bank under the accession code 1LSL.

We thank Richard Hynes (Massachusetts Institute of Technology, Cambridge, MA) for his critical reading. We are grateful to Alexis Bywater (Beth Israel Deaconess Medical Center, Boston MA) for preparation of the manuscript.

This work was supported by the grants HL68003, HL49081, and CA92644 from the National Institutes of Health and also in part by GM56008 to J.-h. Wang.

Submitted: 13 June 2002

Revised: 4 September 2002

Accepted: 6 September 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.
- Bornstein, P. 1995. Diversity of function is inherent in matricellular proteins: an appraisal of thrombospondin 1. *J. Cell Biol.* 130:503–506.
- Brunger, A. 1992. X-PLOR version 3.1: A system for crystallography and NMR. Yale University Press, New Haven, CT. 382 pp.
- CCP4. 1994. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr.* D50:760–763.
- Chen, H., M.E. Herndon, and J. Lawler. 2000. The cell biology of thrombospondin-1. *Matrix Biol.* 19:597–614.
- Cowtan, K. 1994. DM: an automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography.* 31:34–38.
- Dawson, D.W., S.F.A. Pearce, R. Zhong, R.L. Silverstein, W.A. Frazier, and N.P. Bouck. 1997. CD36 mediates the in vitro inhibitory effects of thrombospondin-1 on endothelial cells. *J. Cell Biol.* 138:707–717.
- Dawson, D.W., O.V. Volpert, S.F.A. Pearce, A.J. Schneider, R.L. Silverstein, J. Henkin, and N.P. Bouck. 1999. Three distinct D-amino acid substitutions confer potent antiangiogenic activity on an inactive peptide derived from a thrombospondin-1 type 1 repeat. *Mol. Pharmacol.* 55:332–338.
- de Vos, A.M., M. Ultsch, and A.A. Kossiakoff. 1992. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science.* 255:306–312.
- Flocco, M.M., and S.L. Mowbray. 1994. Planar stacking interactions of arginine and aromatic side-chains in proteins. *J. Mol. Biol.* 235:709–717.
- Gallivan, J.P., and D.A. Dougherty. 1999. Cation-pi interactions in structural biology. *Proc. Natl. Acad. Sci. USA.* 96:9459–9464.
- Guo, N.H., H.C. Krutzsch, E. Negre, V.S. Zabrenetzky, and D.D. Roberts. 1992. Heparin-binding peptides from the type I repeats of thrombospondin. Structural requirements for heparin binding and promotion of melanoma cell adhesion and chemotaxis. *J. Biol. Chem.* 267:19349–19355.
- Hartmann, S., and J. Hofsteenge. 2000. Properdin, the positive regulator of complement, is highly C-mannosylated. *J. Biol. Chem.* 275:28569–28574.
- Hendrickson, W.A. 1991. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. *Science.* 254:51–58.
- Hofsteenge, J., M. Blommers, D. Hess, A. Furmanek, and O. Miroschnichenko. 1999. The four terminal components of the complement system are C-mannosylated on multiple tryptophan residues. *J. Biol. Chem.* 274:32786–32794.
- Hofsteenge, J., K.G. Huwiler, B. Macek, D. Hess, J. Lawler, D.F. Mosher, and J. Peter-Katalinic. 2001. C-mannosylation and O-fucosylation of the thrombospondin type 1 module. *J. Biol. Chem.* 276:6485–6498.
- Iruela-Arispe, M.L., M. Lombardo, H.C. Krutzsch, J. Lawler, and D.D. Roberts. 1999. Inhibition of angiogenesis by thrombospondin-1 is mediated by two independent regions within the type 1 repeats. *Circulation.* 100:1423–1431.
- Iwasaki, W., K. Nagata, H. Hatanaka, T. Inui, T. Kimura, T. Muramatsu, K. Yoshida, M. Tasumi, and F. Inagaki. 1997. Solution structure of midkine, a new heparin-binding growth factor. *EMBO J.* 16:6936–6946.
- Jimenez, B., O.V. Volpert, S.E. Crawford, M. Febbraio, R.L. Silverstein, and N. Bouck. 2000. Signals leading to apoptosis-dependent inhibition of neovascularization by thrombospondin-1. *Nat. Med.* 6:41–48.
- Jones, T.A., J.-Y. Zou, S.W. Cowan, and M. Kjeldgaard. 1991. Improved methods for building protein models in electron density maps and location of errors in these models. *Acta Crystallogr. A.* 47(Pt. 2):110–119.
- Kilpelainen, I., M. Kaksonen, H. Avikainen, M. Fath, R.J. Linhardt, E. Raulo, and H. Rauvala. 2000. Heparin-binding growth-associated molecule contains two heparin-binding  $\beta$ -sheet domains that are homologous to the thrombospondin type I repeat. *J. Biol. Chem.* 275:13564–13570.
- Klar, A., M. Baldassare, and T.M. Jessell. 1992. F-spondin: a gene expressed at high levels in the floor plate encodes a secreted protein that promotes neural cell adhesion and neurite extension. *Cell.* 69:95–110.
- Kraulis, P.J. 1991. Molscript: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallography.* 24:946–950.
- Lau, L.F., and S.C. Lam. 1999. The CCN family of angiogenic regulators: the integrin connection. *Exp. Cell Res.* 248:44–57.
- 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.
- Leung, L.L.K., W.-X. Li, J.L. McGregor, G. Albrecht, and R.J. Howard. 1992. CD36 peptides enhance or inhibit CD36-thrombospondin binding. A two-step process of ligand-receptor interaction. *J. Biol. Chem.* 267:18244–18250.
- Lietha, D., D.Y. Chirgadze, B. Mulloy, T.L. Blundell, and E. Gherardi. 2001. Crystal structures of NK1-heparin complexes reveal the basis for NK1 activity and enable engineering of potent agonists of the MET receptor. *EMBO J.* 20:5543–5555.
- Livnah, O., E.A. Stura, D.L. Johnson, S.A. Middleton, L.S. Mulcahy, N.C. Wrighton, W.J. Dower, L.K. Jolliffe, and I.A. Wilson. 1996. Functional mimicry of a protein hormone by a peptide agonist: the EPO receptor complex at 2.8 Å. *Science.* 273:464–471.
- Macias, M.J., M. Hyvonen, E. Baraldi, J. Schultz, M. Sudol, M. Saraste, and H. Oschkinat. 1996. Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature.* 382:646–649.
- Matuschewski, K., A.C. Nunes, V. Nussenzeig, and R. Menard. 2002. *Plasmodium* sporozoite invasion into insect and mammalian cells is directed by the same dual binding system. *EMBO J.* 21:1597–1606.
- Miao, W., W.L. Seng, M. Duquette, C. Laus, M. Detmar, and J. Lawler. 2001. Thrombospondin-1 type 1 repeat recombinant proteins inhibit tumor growth. *Blood.* 97:1689–1696.
- Mulloy, B., M.J. Forster, C. Jones, and D.B. Davies. 1993. N.m.r. and molecular-modelling studies of the solution conformation of heparin. *Biochem. J.* 293:849–858.
- Murphy-Ullrich, J.E., and M. Poczatek. 2000. Activation of latent TGF- $\beta$  by

- thrombospondin-1: mechanisms and physiology. *Cytokine Growth Factor Rev.* 11:59–69.
- Nicholls, A., K.A. Sharp, and B. Honig. 1991. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. *Proteins.* 11:281–296.
- Nishiomori, H., T. Shiratsuchi, T. Urano, Y. Kimura, K. Kiyono, K. Tatsumi, S. Yoshida, M. Ono, M. Kuwano, Y. Nakamura, and T. Tokino. 1997. A novel brain-specific p53-target gene, BAI1, containing thrombospondin type 1 repeats inhibits experimental angiogenesis. *Oncogene.* 15:2145–2150.
- Panetti, T.S., B.J. Kudryk, and D.F. Mosher. 1999. Interaction of recombinant procollagen and properdin modules of thrombospondin-1 with heparin and fibrinogen/fibrin. *J. Biol. Chem.* 274:430–437.
- Papadimitriou, E., A. Polykratis, J. Courty, P. Koolwijk, M. Heroult, and P. Katsoris. 2001. HARP induces angiogenesis in vivo and in vitro: implication of N or C terminal peptides. *Biochem. Biophys. Res. Commun.* 282:306–313.
- Pawson, T. 1995. Protein modules and signalling networks. *Nature.* 373:573–580.
- Pearce, S.F., J. Wu, and R.L. Silverstein. 1995. Recombinant GST/CD36 fusion proteins define a thrombospondin binding domain. *J. Biol. Chem.* 270:2981–2986.
- Pellegrini, L., D.F. Burke, F. von Delft, B. Mulloy, and T.L. Blundell. 2000. Crystal structure of fibroblast growth factor receptor ectodomain bound to ligand and heparin. *Nature.* 407:1029–1034.
- Richardson, J.S. 1981. The anatomy and taxonomy of protein structure. *Adv. Protein Chem.* 34:167–339.
- Rodriguez-Manzanique, J.C., T.F. Lane, M.A. Ortega, R.O. Hynes, J. Lawler, and M.L. Iruela-Arispe. 2001. Thrombospondin-1 suppresses tumor growth by a novel mechanism that includes blockade of matrix metalloproteinase-9 activation. *Proc. Natl. Acad. Sci. USA.* 98:12485–12490.
- Somers, W., M. Ultsch, A.M. De Vos, and A.A. Kossiakoff. 1994. The X-ray structure of a growth hormone-prolactin receptor complex. *Nature.* 372:478–481.
- 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.
- Tortorella, M., M. Pratta, R.Q. Liu, I. Abbaszade, H. Ross, T. Burn, and E. Arner. 2000. The thrombospondin motif of aggrecanase-1 (ADAMTS-4) is critical for aggrecan substrate recognition and cleavage. *J. Biol. Chem.* 275:25791–25797.
- Tzarfaty-Majar, V., R. Lopez-Aleman, Y. Feinstein, L. Gombau, O. Goldshmidt, E. Soriano, P. Munoz-Canoves, and A. Klar. 2001. Plasmin-mediated release of the guidance molecule F-spondin from the extracellular matrix. *J. Biol. Chem.* 276:28233–28241.
- Vázquez, F., G. Hastings, M. Ortega, T.F. Lane, S. Oikemus, M. Lombardo, and M.L. Iruela-Arispe. 1999. METH-1, a human ortholog of ADAMTS-1, and METH-2 are members of a new family of proteins with angi-inhibitory activity. *J. Biol. Chem.* 274:23349–23357.
- Venter, J.C., M.D. Adams, E.W. Myers, P.W. Li, R.J. Mural, G.G. Sutton, H.O. Smith, M. Yandell, C.A. Evans, R.A. Holt, et al. 2001. The sequence of the human genome. *Science.* 291:1304–1351.
- Wengelnik, K., R. Spaccapelo, S. Naitza, K.J. Robson, C.J. Janse, F. Bistoni, A.P. Waters, and A. Crisanti. 1999. The A-domain and the thrombospondin-related motif of *Plasmodium falciparum* TRAP are implicated in the invasion process of mosquito salivary glands. *EMBO J.* 18:5195–5204.