Structure of a thrombospondin C-terminal fragment reveals a novel calcium core in the type 3 repeats

Marc Kvansakul¹, Josephine C Adams²
and Erhard Hohenester^{1,*}

¹Department of Biological Sciences, Imperial College London, South Kensington Campus, London, UK and ²Department of Cell Biology, Lerner Research Institute, Cleveland Clinic Foundation, Cleveland, OH, USA

Thrombospondins (TSPs) are extracellular regulators of cell–matrix interactions and cell phenotype. The most highly conserved region of all TSPs are the calcium-binding type 3 (T3) repeats and the C-terminal globular domain (CTD). The crystal structure of a cell-binding TSP-1 fragment, spanning three T3 repeats and the CTD, reveals a compact assembly. The T3 repeats lack secondary structure and are organised around a core of calcium ions; two DxDxDGxxDxxD motifs per repeat each encapsulate two calcium ions in a novel arrangement. The CTD forms a lectin-like β -sandwich and contains four strictly conserved calcium-binding sites. Disruption of the hairpin structure of T3 repeats 6 and 7 decreases protein secretion and stability. The availability for cell attachment of an RGD motif in T3 repeat 7 is modulated by calcium loading. The central architectural role of calcium explains how it is critical for the functions of the TSP C-terminal region. Mutations in the T3 repeats of TSP-5/COMP, which cause two human skeletal disorders, are predicted to disrupt the tertiary structure of the T3–CTD assembly.

The EMBO Journal (2004) **23**, 1223–1233. doi:10.1038/sj.emboj.7600166; Published online 11 March 2004
Subject Categories: structural biology; cell & tissue architecture

Keywords: calcium binding; cell adhesion; extracellular matrix; L-type lectin; protein folding

Introduction

The thrombospondins (TSPs) are large, secreted, multimeric glycoproteins that modulate extracellular matrix (ECM) structure and cell behaviour. There are five TSPs in vertebrates, each the product of a different gene. TSP-1 and TSP-2 are homotrimers, whereas TSP-3, TSP-4 and TSP-5 (usually referred to as cartilage oligomeric matrix protein, COMP) are homopentamers (for reviews, see Lawler, 2000; Adams, 2001; Armstrong and Bornstein, 2002). There exists a single

pentameric TSP in *Drosophila* (Adams *et al.*, 2003). TSP-1 has been the most intensively studied family member and the prototype for functional analysis of other TSPs. Each subunit contains binding sites for cellular receptors, cytokines and ECM molecules. Its multifunctionality allows TSP-1 to assemble multiprotein complexes at the cell surface and thereby modulate cell behaviour. TSP-1 is involved in platelet aggregation, cell adhesion and migration, the regulation of proliferation, angiogenesis, wound healing and tumour growth (for reviews, see Lawler, 2000; Adams, 2001; Armstrong and Bornstein, 2002).

Each protomer of TSP-1 and TSP-2 consists of an N-terminal domain, a coiled-coil oligomerisation domain, a von Willebrand factor-type C or procollagen homology domain, three TSP type 1 or properdin domains of known structure (Tan *et al.*, 2002), three TSP type 2 (T2) or EGF-like domains, seven calcium-binding type 3 (T3) repeats and a C-terminal domain (CTD) of ≈ 230 residues that shows no sequence similarity to any other protein domain (Figure 1A). The pentameric TSPs have a different N-terminal domain organisation and lack the type 1 repeats, but have in common with TSP-1 and TSP-2 the T2–T3–CTD arrangement. The T3 repeats and CTD are the most highly conserved region between all TSPs (Adams and Lawler, 1993a; Adams *et al.*, 2003). The structure of this region is thus of high interest for understanding the molecular interactions and functions of the TSP family.

The most striking property of the C-terminal region of TSPs is the profound dependence of protein conformation and function on calcium binding. The T3 repeats are a tandem of aspartic acid-rich motifs, which resemble EF hands in the spacing of acidic side chains (Lawler and Hynes, 1986). Unlike EF hands, however, the calcium-binding loops of the T3 repeats are not flanked by secondary structure elements. TSPs have been shown to bind 10–12 calcium ions per subunit with high cooperativity and moderate affinity (average $K_D \approx 0.1$ mM) (Lawler and Simons, 1983; Misenheimer and Mosher, 1995; Chen *et al.*, 2000). Chelation of calcium by EDTA leads to a large shape change, visualised by rotary shadowing of intact molecules as an elongation of the C-terminal ‘arms’ and shrinking of the C-terminal globule (Lawler *et al.*, 1985). This has been attributed to collapse of the T3 repeats and indeed the entire T2–T3–CTD assembly (Misenheimer *et al.*, 2003). Other functions of the C-terminal region include cell adhesion and migration (for review, see Adams, 2001). Crucially, cell attachment activity depends on the calcium loading of TSP-1 (Lawler *et al.*, 1988). The cytoskeletal rearrangements associated with cell spreading are only elicited when C-terminal fragments of TSP-1 or TSP-2 are presented in trimeric form (Anilkumar *et al.*, 2002). Candidate receptors for this region include integrins that bind an RGD motif in T3₇ of TSP-1 (Lawler *et al.*, 1988), CD47 (Gao *et al.*, 1996) and syndecan-1 (Adams *et al.*, 2001).

*Corresponding author. Biophysics Group, Blackett Laboratory, Imperial College London, South Kensington Campus, London SW7 2AZ, UK.
Tel.: +44 20 7594 7701; Fax: +44 20 7589 0191;
E-mail: e.hohenester@imperial.ac.uk

Received: 6 October 2003; accepted: 19 February 2004; published online: 11 March 2004

Figure 1 Cell binding to T3_{5.7}-CTD. (A) Domain organisation of TSP-1 and its T3_{5.7}-CTD fragment. NTD, N-terminal domain; cc, coiled-coil domain; PC, procollagen homology domain; T1, type 1 repeats; T2, type 2 repeats; T3, type 3 repeats; CTD, C-terminal domain. (B) Binding of C2C12 myoblasts (left) and rat vascular smooth muscle cells (right) to immobilised TSP-1, a recombinant T2₃-T3_{1.7}-CTD fragment (Anilkumar *et al*, 2002) and the T3_{5.7}-CTD C974S fragment.

The C-terminal regions of COMP and TSP-4 bind ECM molecules including collagens I, II and IX, and are thus implicated in ECM assembly (Rosenberg *et al*, 1998; Narouz-Ott *et al*, 2000; Adams, 2001; Holden *et al*, 2001).

The C-terminal regions of TSPs are also involved in human disease. Two related autosomal dominant skeletal disorders, pseudoachondroplasia (PSACH) and multiple epiphyseal dysplasia (MED), are linked to mutations in COMP (Briggs *et al*, 1995; Hecht *et al*, 1995). Notably, the great majority of mutations affect calcium-binding residues within the T3 repeats (for review, see Briggs and Chapman, 2002). Mutant COMP aggregates in the endoplasmic reticulum of chondrocytes, leading to impaired COMP secretion, reduced cell viability and altered pericellular matrix (Delot *et al*, 1998; Holden *et al*, 2001; Dinser *et al*, 2002; Hashimoto *et al*, 2003). The T2-T3-CTD region appears to be extraordinarily sensitive to mutation, with single missense or deletion mutations typically resulting in the loss of about half the exchangeable calcium ions (Chen *et al*, 2000; Maddox *et al*, 2000; Thur *et al*, 2001; Kleerekoper *et al*, 2002). Recently, polymorphism of TSPs (N700S at the N-terminus of the T3 repeats of TSP-1 and A387P in the last T2 repeat of TSP-4) has been associated with familial premature coronary heart disease (Topol *et al*, 2001). Like the COMP mutations, the TSP-1 polymorphism reduces calcium binding (Hannah *et al*, 2003), but how the structural change predisposes to coronary heart disease remains to be established.

We have determined the crystal structure of a recombinant C-terminal fragment of TSP-1, spanning the last three T3 repeats and the CTD. The T3 repeats are devoid of secondary structure and are held together by a core of 12 calcium ions. The three T3 repeats and the lectin-like β -sandwich of the CTD are assembled into a compact unit. These results demonstrate how calcium binding so dramatically affects the properties of TSPs and indicate how mutations of calcium-binding residues in COMP can cause misfolding and disease.

Results

Construct design and crystallisation

To search for a crystallisable C-terminal fragment of TSP-1, we systematically extended from the N-terminus of the CTD and expressed these constructs in human embryonic kidney cells. The smallest construct that was both highly expressed and soluble spans the last three T3 repeats and the CTD (T3_{5.7}-CTD; amino-acid residues 816-1152). To prevent disulphide isomerisation, which is known to occur in TSP-1 (Sun *et al*, 1992) and would frustrate crystallisation attempts, we mutated the free cysteine at position 974 to serine, the amino acid at this position in most other TSPs. Far-UV circular dichroism, fluorescence spectroscopy and limited trypsin digestion indicated that T3_{5.7}-CTD C974S bound calcium and was saturated at ≈ 2 mM calcium. Calcium binding was associated with a considerable compaction of the protein, as evidenced by a 0.9 ml shift in elution volume from a 24 ml gel filtration column (not shown). To show that the T3_{5.7}-CTD C974S mutant retained the biological functions of its parent molecule, we examined its cell-binding properties. C2C12 skeletal myoblasts and rat vascular smooth muscle cells attached to immobilised calcium-replete T3_{5.7}-CTD C974S as strongly as to a previously studied recombinant T2₃-T3_{1.7}-CTD fragment (Figure 1B). Being monomeric, T3_{5.7}-CTD C974S only supported cell attachment but not spreading (Anilkumar *et al*, 2002).

Because we failed to obtain crystals of T3_{5.7}-CTD C974S, we removed the N-linked glycan by mutagenesis of the putative acceptor site at Asn1049. The N1049K mutation abolished glycosylation, as indicated by a shift to lower molecular mass on SDS-PAGE gels, but did not affect protein solubility or cell binding (not shown). The T3_{5.7}-CTD C974S/N1049K double mutant could be crystallised in the presence of 5 mM calcium and its structure was determined to a resolution of 1.9 Å (Table I).

Overall structure

We observed clear electron density for almost the entire T3_{5.7}-CTD fragment. Only the N-terminal tag and residues 962-966 in the β 2- β 3 hairpin of the CTD are not defined in our structure. The T3_{5.7}-CTD fragment adopts a compact structure of approximate dimensions $70 \times 50 \times 35$ Å (Figure 2). At its heart is the globular CTD, which folds into a curved 15-stranded β -sandwich. T3₅ makes multiple contacts with loop regions at the top of the concave β -sandwich face (view of Figure 2A), while T3₆ and T3₇ form a tight hairpin that projects from the body of the structure. The interface between T3₅ and the CTD buries 940 Å² of solvent-accessible surface area, and a further 1240 Å² is buried in the hairpin interface. The glycosylation site at residue 1049, which was mutated to lysine to obtain crystals, is located below and behind T3₅. An N-linked glycan could be accommodated at this position without causing steric clashes. There are four disulphide bridges in the T3_{5.7}-CTD structure. Three of them link consecutive T3 repeats (see below for a definition of the repeats) and a fourth bridge links T3₇ to the extreme C-terminus of the CTD. The same sequential disulphide linkage was determined chemically for the T2₃-T3_{1.7}-CTD region of TSP-2 (Misenheimer *et al*, 2001). Residue 974, which is a free cysteine in native TSP-1 and serine in our construct (and in most other TSPs), is located

Table 1 Crystallographic statistics

Crystal	Native 2 ^a	Native 1	K ₂ PtCl ₄	Sm(CH ₃ COO) ₃
<i>Data collection and phasing</i>				
Resolution range (Å)	20–1.9	20–2.4	20–2.8	20–2.8
Unique reflections	32 099	16 840	10 725	10 637
Multiplicity	5.4 (4.0)	14.9	10.6	7.2
Completeness (%)	95.3 (92.9)	99.8	99.7	99.7
R_{merge}^b	0.086 (0.37)	0.104	0.079	0.081
R_{deriv}^c			0.196	0.141
R_{Cullis}^d (centric/acentric) ^d			0.82/0.88	0.76/0.84
Phasing power (centric/acentric) ^e			0.72/0.81	0.82/0.98
<i>Refinement</i>				
Resolution range (Å)	20–1.9 Å			
Reflections (working set/test set)	28 902/3197			
Protein atoms	2667			
Solvent atoms	262 H ₂ O, 16 Ca ²⁺			
$R_{\text{cryst}}/R_{\text{free}}^f$	0.199/0.227			
r.m.s.d. bonds (Å)	0.005			
r.m.s.d. angles (deg)	1.4			
r.m.s.d. B-factors (Å ²)	1.6			
Ramachandran plot (%) ^g	80.4/19.3/0/0.4			

^aNumbers in parentheses refer to the highest resolution shell (2.00–1.90 Å).

^b $R_{\text{merge}} = \sum_h \sum_i |I_i(h) - \langle I(h) \rangle| / \sum_h \sum_i I_i(h)$, where $I_i(h)$ is the i th measurement of reflection h and $\langle I(h) \rangle$ is the weighted mean of all measurements of h .

^c $R_{\text{deriv}} = \sum_h ||F_{\text{PH}}| - |F_P|| / \sum_h |F_P|$, where F_P and F_{PH} are the native and derivative structure factors, respectively. The native 1 data were used for phasing.

^d $R_{\text{Cullis}} = \sum_h ||F_{\text{PH}}| - |F_P| - |F_H|| / \sum_h ||F_{\text{PH}}| - |F_P||$, where F_H is the calculated heavy atom structure factor.

^eThe phasing power is defined as (r.m.s. F_H /r.m.s. lack-of-closure).

^f $R = \sum_h |F_{\text{obs}} - F_{\text{calc}}| / \sum_h F_{\text{obs}}$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively. R_{cryst} and R_{free} were calculated using the working and test set, respectively.

^gResidues in most favoured, additionally allowed, generously allowed and disallowed regions (Laskowski *et al*, 1993).

between strands $\beta 3$ and $\beta 4$, and is buried by the calcium-binding $\beta 1$ – $\beta 2$ loop. A total of 16 metal ions were identified in the T3_{5–7}–CTD structure and assigned as calcium. There are four calcium ions in T3₅, five ions in T3₆ and three ions in T3₇. Four additional ions are bound to the edge of the CTD opposite the T3_{6–7} hairpin. The calcium concentration in our crystals (5 mM) suggests that all observed metal sites are physiologically relevant.

Structure of the C-terminal domain

The globular CTD is a β -sandwich of two curved antiparallel β -sheets. The fold is an elaboration of the jelly roll topology, with strands $\beta 3$ – $\beta 7$, $\beta 11$ and $\beta 14$ – $\beta 15$ forming the eight-stranded jelly roll motif. The smaller seven-stranded sheet presents a convex face to the solvent, whereas the more radically curved eight-stranded sheet creates a deep concave cleft, which is completely filled by a short α -helix and residues contributed by four loops. The long $\beta 7$ – $\beta 8$ loop is particularly prominent, closing off the pocket and all but burying the α -helix. The CTD contains two large hydrophobic cores, one between the two β -sheets and one within the concave cleft. Despite its unique sequence, the CTD was found to belong to a larger family of lectin-like folds. The DALI server (Holm and Sander, 1997) retrieved two calcium-binding lectin domains, which match the CTD of TSP-1 in its entirety. These structural homologues are p58/ERGIC-53 (Velloso *et al*, 2002), an animal lectin functioning in the secretory pathway (r.m.s. deviation 2.7 Å for 164 C α atoms, 7% sequence identity) and a plant lectin from *Griffonia simplicifolia* (Delbaere *et al*, 1993) (r.m.s. deviation 2.6 Å for 162 C α atoms, 12% sequence identity), both members of the L-type lectin family (Dodd and Drickamer, 2001). The

CTD of TSP-1 and the lectins are most similar in the β -sheets, but differ markedly in many loop regions, in particular those shaping the concave cleft.

There are four calcium ions bound to the lower rim of the CTD sandwich. The first ion is coordinated by three main-chain carbonyl groups in the $\beta 1$ – $\beta 2$ loop and the side chains of Asp956 and Asp975. The second and third ions are bound in close proximity to each other by the Asp1001–Asp1002–Asp1003 sequence in the $\beta 5$ – $\beta 6$ loop (Figure 3). Each ion receives five ligands from the CTD. In the crystal lattice, the side chain of Asp825 from a neighbouring molecule completes the metal coordination by bridging the two calcium ions. Finally, a fourth ion is bound loosely by two main-chain carbonyl groups in $\beta 5$ – $\beta 6$ loop and the side chains of Gln1024 and Asp1086. The location of the double-calcium site (Ca2 and Ca3) is intriguing as it broadly coincides with the calcium-dependent carbohydrate-binding site of the distantly related lectins. The long $\beta 7$ – $\beta 8$ loop positions the solvent-exposed side chain of Trp1030 just above the double-calcium site, highlighting this region as a prime candidate for an interaction site. Crucially, residues involved in calcium binding to the CTD are strictly conserved across the human TSP family, indicating that they are required for the structure or function of all TSPs. There are no other striking surface features of the CTD offering clues about the location of receptor-binding sites.

Calcium coordination by the type 3 repeats

The most remarkable feature of the T3_{5–7}–CTD structure is the intricate folding of the T3 repeats. The 120 residues preceding the CTD in our construct contain $\approx 30\%$ aspartic acid and almost no aliphatic or aromatic residues, yet they are fully ordered in the crystal. The polypeptide chain, which

Figure 2 Overall structure of T3₅₋₇-CTD. (A) Cartoon drawing of the T3₅₋₇-CTD structure. T3₅, T3₆ and T3₇ are in red, yellow and green, respectively. The CTD is in cyan and its β -strands are sequentially labelled 1–15. Disulphide bridges and residue 974 are shown in ball-and-stick representation and are in yellow. Calcium ions are represented by pink spheres. (B) Sequence alignment of human TSPs. The sequence numbering and secondary structure elements of TSP-1 are included above the alignment. Strictly conserved residues are in bold. Cysteines and calcium-binding residues are highlighted in yellow and pink, respectively. An RGD sequence in T3₇ (Lawler and Hynes, 1986) and two putative cell-adhesive sequences in the CTD of TSP-1 (Kosfeld and Frazier, 1993) are in green. TSP-5/COMP residues affected by missense mutations (see text) are indicated by red asterisks.

Figure 3 Double-calcium site in the CTD. Selected loop regions are shown as C α traces. Calcium-binding residues and Trp1030 are shown in ball-and-stick representation and are labelled. Calcium ions are represented as pink spheres. Calcium-ligand bonds are shown as thin black lines. Water molecules have been omitted for clarity.

apart from β -turns and two turns of 3_{10} helix is devoid of secondary structure, wraps around a series of calcium ions, creating a highly unusual protein structure organised around a metal core. The lack of secondary structure in the T3 repeat region is in reasonable agreement with circular dichroism spectra of TSP-1 and COMP T3 constructs (Thur *et al*, 2001; Misenheimer *et al*, 2003).

To describe the calcium coordination by the T3 repeats, we must first define the repeating unit. There are 12 D₁x D₃x D₅GxxD₉xxD₁₂ motifs in every TSP sequence, which commonly have been grouped into seven repeats (Lawler and Hynes, 1986). In this grouping, T3₂ and T3₄ are short and contain only one motif, whereas the other five repeats contain two motifs, which we will refer to as N-type and C-type motifs. In our structure-based definition of the T3 repeats, disulphide bridges link adjacent repeats, and the single motifs in T3₂ and T3₄ are identified as C-type motifs (Figure 4).

The structurally most conserved part of the T3 repeats is the C-type motif, and T3₅, T3₆ and T3₇ superimpose almost perfectly in this region. Each C-type motif encapsulates two calcium ions, which are likely to be bound with

Figure 4 Calcium coordination by the type 3 repeats. **(A)** Sequence alignment of T3 repeats in TSP-1. The calcium-binding aspartates of the D₁x D₃x D₅GxxD₉xxD₁₂ motifs are highlighted in pink. A putative calcium-binding motif in the linker between T2₃ and T3₁ is also marked. Cysteines are highlighted in yellow and the disulphide linkages are indicated by yellow lines. **(B)** Structural superposition of T3₅ (red), T3₆ (yellow) and T3₇ (green). The repeats were superimposed by fitting their C-type motifs. Calcium ions are represented as small spheres. **(C)** Stereoviews showing the calcium coordination in the C-type motif (upper panel) and N-type motif (lower panel) of T3₆. Calcium-binding residues are shown in ball-and-stick representation and are labelled. Calcium ions are represented as pink spheres. Calcium-ligand bonds are shown as thin black lines. Water molecules have been omitted for clarity. **(D)** Schematic drawing of the calcium coordination in the C-type motif (see text). Metal ligands are numbered according to their position in the consensus sequence for comparison with (A). Some ligands are also labelled with the corresponding sequence numbers in T3₆ to facilitate comparison with (C).

high affinity. The first (upper) calcium ion is coordinated by the side chains at positions 1, 3, 5 and 12 of the D₁x D₃x D₅GxxD₉xxD₁₂ motif (Figure 4). The six-fold coordination is completed by the main-chain carbonyl group at

position 7 and a water molecule bridging to the aspartate at position 9. The residue at position 6 adopts a main-chain conformation that is favourable only for glycine. The second (lower) calcium ion is seven-fold coordinated by the side

chains at positions 3 and 5 (monodentate and bridging to the first calcium), the aspartate at position 12 (bidentate) and three carbonyl groups. The first and second carbonyl group always is donated by main-chain peptide bonds. In T3₇ the last ligand is also a peptide carbonyl group, but in T3₅ and T3₆ this position is taken over by the side-chain oxygen atom of a conserved asparagine (position 19 of the C-type motif). The polypeptide chain wraps around the double-calcium core, with a turn just before the crucial aspartate at position 12 providing the connection between the upper and lower loops. This arrangement is further stabilised by main-chain interactions and conserved residues not engaged in metal binding. The upper loop is stabilised by the conserved aspartate at position 9, and the lower loop is stabilised by a glutamine or glutamic acid at position 22 of the C-type motif.

The N-type motif resembles the C-type motif, but seems to allow for more variation. A first calcium ion is coordinated as in the C-type motif. In T3₇, this is the only ion bound to the N-type motif, because a lysine replaces the crucial aspartic acid at position 12. In T3₅ and T3₆, a second calcium ion is bound by the aspartates at positions 3, 5 and 12, with the latter interaction being bidentate as in the C-type motif. However, unlike in the C-type motif, the polypeptide chain does not fully encircle the second calcium ion, leaving the metal coordination sphere incomplete and probably resulting in a lower affinity for calcium. In T3₅ and T3₆, residues at positions 14 and 15, respectively, provide bonds to the second calcium ion. As in the C-type motif, position 6 of the N-type motif usually is a glycine. Additionally, there is an invariant glycine at position 8, which allows the close approach of the preceding T3 repeat to form the inter-repeat disulphide bridge. T3₆ is distinct from T3₅ and T3₇ in that a third calcium ion is bound to the outside of the N-type motif by the side chains of Asp880 (bidentate) and Asp887 (monodentate) and two main-chain carbonyl groups.

The D₁x D₃x D₅GxxD₉xxD₁₂ motifs of TSP T3 repeats resemble the classic EF-hand calcium-binding motif D₁x D₃x D₅GxxxxxE₁₂ (Lawler and Hynes, 1986; Nakayama and Kretsinger, 1994). Structural superposition reveals that the calcium coordination is indeed similar up to position 7. However, there are profound differences in the remainder of the metal coordination, the most obvious distinction being the different number of calcium ions bound (two by the TSP motifs, one by EF hands). EF hands also usually occur as interlocked pairs embedded in an α -helical fold, which sharply contrasts with the dearth of secondary structure in the T3 repeats. The superficial similarity between the T3 and EF-hand motifs is probably due to the constraints of creating a calcium-binding pocket from a contiguous polypeptide segment, which seem to favour regularly spaced aspartic acids.

Tertiary structure of the type 3 repeat region

The relative orientation of N- and C-type motifs within repeats is highly variable and seems to depend upon the tertiary structure context. We have retained the original definition of a total of seven repeats in TSPs (Lawler and Hynes, 1986), but note that the T3 repeat portion of T3₅₋₇-CTD could equally well be described as a chain of alternating high-affinity (C-type) and low-affinity (N-type) calcium-binding motifs.

The disulphide bridges between consecutive repeats are clearly important for the tertiary structure of the T3 repeat

portion of T3₅₋₇-CTD, but they appear to be insufficient to dictate the unique arrangement of repeats. T3₅ and T3₆ are joined in a relatively extended conformation, whereas T3₆ and T3₇ form a hairpin structure (Figure 5A). Remarkably, this hairpin is stabilised by a well-formed hydrophobic core consisting of residues (Leu894 and Val895 from T3₆ and Phe916 from T3₇) that are strictly conserved in all human TSPs and likely to be important for the structure. There is no interaction between T3₅ and T3₆, apart from the Cys856–Cys876 disulphide bridge. Thus, the position of T3₅ in the T3₅₋₇-CTD structure appears to depend largely on its contacts with the CTD. The majority of these contacts involve polar residues that are not conserved in all TSPs, and the disposition of T3₅ relative to the CTD could be different in other TSPs.

To determine the contribution of the T3₆₋₇ hairpin to TSP-1 folding and stability, we produced an N-terminally truncated construct lacking all of T3₅ and the N-type motif of T3₆ (T3₇-CTD C974S). The shorter construct was produced less efficiently by the 293 cells than T3₅₋₇-CTD C974S, indicating impaired folding and/or secretion. Gel filtration chromatography of T3₇-CTD C974S revealed extensive protein aggregation, in sharp contrast to T3₅₋₇-CTD C974S, which is 100% monomeric (Figure 5B). We conclude that disruption of the native T3₅₋₇ tertiary structure exposes hydrophobic residues, leading to poor protein solubility. In agreement with this interpretation, an NMR study of a COMP T3₅₋₇ construct has shown this region to be a stable subdomain with calcium-dependent tertiary structure (Kleerekoper *et al*, 2002).

Location of the RGD motif in type 3 repeat 7

The single RGD motif in T3₇ of human TSP-1 has been identified as a binding site for α v β 3 and α IIb β 3 integrins (Lawler *et al*, 1988; Lawler and Hynes, 1989); however, not all cell types undergo RGD-dependent attachment to TSP-1 (e.g., Long and Dixit, 1990; Adams and Lawler, 1993b). In the structure, the RGD sequence is located near the tip of the T3₆₋₇ hairpin and embedded in the N-type calcium-binding site of T3₇ (Figure 5A). The main-chain carbonyl group of Arg908 coordinates the calcium ion, while its guanidinium group interacts with Asp923 in the C-type motif, thus stabilising the extended conformation of T3₇. Gly909 is largely buried and Asp910 makes two interactions with Ser903. Both Arg908 and Gly909 adopt an extended main-chain conformation. Because integrin binding requires the RGD motif to be presented in a β -hairpin context (Xiong *et al*, 2002), native calcium-saturated TSP-1 is unlikely to be cell adhesive directly via its RGD motif. A related KGD motif in T3₆, which is conserved in TSP-1 and TSP-2 species orthologues, is also located in a calcium-binding loop and equally unlikely to be involved in cell attachment.

The possible roles of calcium ions and disulphide bridges in RGD-dependent cell attachment to T3₅₋₇-CTD were tested by comparing cell attachment under different calcium-loading conditions. For intact TSP-1 in 2 mM CaCl₂ buffers, C2C12 cell attachment is not RGD inhibitable (Adams and Lawler, 1994), whereas attachment of human intestinal smooth muscle cells (HISM) is RGD- and α v β 3 integrin-dependent (Adams and Lawler, 1993b). C2C12 cell attachment to T3₅₋₇-CTD was maintained under different calcium ion concentrations, whereas HISM cell attachment increased at lower calcium ion concentrations but was decreased on EDTA-

Figure 5 T3₆₋₇ hairpin and RGD site. (A) Cartoon drawing. T3₆ and T3₇ are in yellow and green, respectively. Selected residues are shown in ball-and-stick representation and are labelled. The RGD motif in T3₇ is in grey. Calcium ions are represented as pink spheres. (B) Gel filtration chromatograms of T3₅₋₇-CTD C974S (blue) and T3₇-CTD C974S (red), showing extensive aggregation of the shorter construct. Both proteins were injected at a concentration of 1 mg/ml. The running buffer was 25 mM Na-HEPES pH 7.5, 140 mM NaCl and 1 mM CaCl₂. The elution volumes of globular molecular mass standards are indicated by labelled arrows. (C) Cell attachment of C2C12 or HISM cells at different calcium concentrations. Solid lines indicate the percentage of cell attachment relative to attachment at 2 mM calcium (UT, untreated). Dotted lines indicate the percentage of cell attachment remaining in the presence of 1 mM GRGDSP peptide for each calcium ion concentration and cell type (non-RGD). GRGESp peptide at 1 mM was non-inhibitory. Each point is the mean from three experiments, bars indicate s.e.m. (D) Effect of calcium ion conditions and reduction on C2C12 cell morphology. (i-iv) Confocal XY images of C2C12 cells plated for 1 h on T3₅₋₇-CTD C974S/N1049K coated under the indicated conditions, after fixation and processing for phalloidin staining of F-actin. Inset panels in (ii) show the unchanged morphology of cells in the presence of 1 mM GRGESp peptide and the loss of spreading by cells attached in the presence of 1 mM GRGDSP peptide. Bar = 10 μm .

treated T3₅₋₇-CTD (Figure 5C). However, protrusive spreading of C2C12 cells was increased at 0.1 mM CaCl₂ (compare Figure 5D, panels i and ii). Whereas attachment of HISM cells was strongly RGD-inhibitable under all conditions, the RGD-inhibitable fraction of C2C12 cell attachment increased at lower calcium ion concentrations (Figure 5C). Moreover, C2C12 cells that attached to T3₅₋₇-CTD prepared in 0.1 mM CaCl₂ were rounded in the presence of 1 mM GRGDSP pep-

tide; control GRGESp peptide did not have this effect (Figure 5C, panel ii insets). We also compared the effects of disulphide reduction on C2C12 cell attachment for two calcium-loading conditions of T3₅₋₇-CTD. At 2 mM CaCl₂, cell spreading was marginally increased (Figure 5D, panel iii). At 0.1 mM CaCl₂, reduction of T3₅₋₇-CTD resulted in enhanced protrusive spreading of cells and elongated morphologies (Figure 5D, panel iv). This spreading was

inhibited by GRGDSP peptide (not shown). Thus, calcium loading and disulphide bond status both contribute to RGD-dependent cell adhesion on a cell-type dependent basis.

Mapping of TSP-5/COMP mutations onto the structure

Mutations in TSP-5/COMP that cause PSACH or MED are scattered along the entire T3 repeat region, but are significantly concentrated in the T3₅₋₇ portion; six mutations within the CTD have also been described. Most mutations are missense mutations or in-frame deletions of single residues (Briggs and Chapman, 2002; Mabuchi *et al*, 2003). In view of the high sequence conservation of the T3-CTD region of TSPs (Figure 2B), COMP missense mutations were mapped onto the T3₅₋₇-CTD structure (Figure 6). A total of 24 mutations map to the T3 repeats, 16 of which affect calcium-binding residues. Notably, many more mutations affect the C-type motifs than the N-type motifs (13 versus 3). This strong bias is consistent with our interpretation of the C-type motifs as the high-affinity calcium-binding sites critical for TSP folding. Mutations within a sequence of five aspartates in T3₆ of COMP produce particularly severe phenotypes (Mabuchi *et al*, 2003). The corresponding region in TSP-1 (Asp877–Asp881) lies at the junction of the N- and C-type motifs and binds a unique calcium ion on the outside of the canonical calcium-binding loops. From sequence comparison, it seems likely that this arrangement is conserved in COMP and helps organise the T3₆₋₇ hairpin. Curiously, the adjacent N-type motif of T3₇, which contains the RGD motif in TSP-1, is spared of disease mutations in COMP.

Figure 6 Mapping of COMP mutations. Cartoon drawing of the T3₅₋₇-CTD structure with sites of COMP missense mutations (see text) shown as large C α atom spheres. The mutations are labelled using COMP sequence numbering (compare Figure 2B). Red spheres and labels indicate mutations of calcium-binding residues; yellow spheres and labels indicate all other mutations. Calcium ions have been omitted for clarity.

The effects of the six COMP missense mutations in the CTD are more difficult to rationalise. Two mutations, R718W and G719D, map to the β 13– β 14 loop. This loop is part of the interface between the CTD and T3₅, raising the possibility that these mutations destabilise the T2–T3-CTD unit. The four remaining missense mutations affect solvent-exposed residues on the β 5 strand, well away from T3₅₋₇. A convincing structural explanation for the deleterious effect of these mutations will require a structure determination of the complete T3-CTD region.

Discussion

It has been appreciated since the first studies of TSP-1 that calcium binding evokes major conformational changes (Lawler *et al*, 1982). Subsequent cloning revealed that all TSPs contain a tandem of putative calcium-binding motifs, termed T3 repeats. Many studies have demonstrated the importance of calcium binding to all levels of function of TSPs, from binding interactions, to cell attachment, to human genetic disease in the case of COMP (for reviews, see Adams, 2001; Briggs and Chapman, 2002). However, the mechanisms by which alterations to calcium binding result in profound structural changes of the C-terminal region have remained obscure. Here we report the structure of a calcium-replete C-terminal TSP-1 fragment (T3₅₋₇-CTD) and provide a structural interpretation for the tight coupling of calcium binding and folding of TSPs.

The T3 repeats represent a novel type of protein structure, in which the polypeptide chain wraps around a core of metal ions and secondary structure is notable only by its absence. Crucially, the T3 portion of the T3₅₋₇-CTD fragment has many features that characterise the tertiary structures of conventional globular proteins, namely long-range interactions and the interdependence of local and global structure. Thus, we can begin to see how calcium binding underlies the acquisition of a complex folded structure in TSPs.

Equilibrium dialysis experiments have shown that 10–12 calcium ions are bound by each TSP chain (Misenheimer and Mosher, 1995; Chen *et al*, 2000). The total calcium content determined by atomic absorption spectroscopy, however, is more than twice as high (Misenheimer *et al*, 2003). Our structure affords a plausible explanation for this observation. The cage-like structure of the T3 repeat C-type motifs suggests a high affinity for calcium, and the majority of the C-type calcium ions may exchange very slowly, in agreement with observations on a COMP T3₅₋₇ construct (Kleerekoper *et al*, 2002). The moderate ($K_D \approx 0.1$ mM) binding seen by equilibrium dialysis would then correspond mainly to ions bound to the N-type motifs and the CTD. From the structure and from sequence analysis of the missing T3 repeats, we predict a total of 26 calcium ions per full-length TSP chain (14 C-type and eight N-type ions, plus four ions in the CTD), which is in excellent agreement with the latest experimental data (Misenheimer *et al*, 2003). Electron micrographs show the calcium-replete T3-CTD region of TSPs to be a single globule (Lawler *et al*, 1985). Furthermore, indirect evidence suggests the presence of multiple calcium-mediated tertiary contacts within the T2–T3-CTD unit (Misenheimer *et al*, 2003). We are clearly unable to predict the location of the T2 and T3 repeats not included in our construct, but speculate that the double-calcium site in the CTD might be

involved in such a tertiary contact, which could explain its strict conservation in the TSP family.

Calcium binding to TSP-1 modulates the reactivity of the free Cys974, which has been suggested to be important for the reduction of multimeric von Willebrand factor (Pimanda *et al*, 2002). In calcium-depleted TSP-1, Cys974 can exchange with the disulphide bridges in the T3 repeats, which unmasks an integrin-binding RGD sequence in T3₇ (Sun *et al*, 1992). Reductions of TSP-1 and TSP-2 are also known to increase RGD-dependent attachment of endothelial cells (Sun *et al*, 1992; Chen *et al*, 1994). Whether integrin binding to the T3 repeats of TSPs is regulated by disulphide isomerisation *in vivo* is not known. The structure shows that Cys974 is buried and at >30 Å distance from the disulphide bridges of T3₅₋₇. However, it is easy to envisage how removal of calcium from the β1-β2 loop, which is shielding Cys974 from bulk solvent, would lead to a loosening of the CTD structure. Calcium-depleted T3 repeats are known to be random coils (Kleerekoper *et al*, 2002) and would be highly susceptible to attack by the exposed reactive thiol of Cys974. In the other vertebrate TSPs, Cys974 is replaced by a serine. TSP-2 has an even number of cysteines, whereas the pentameric TSPs have an unpaired cysteine near the C-terminus of the CTD. The corresponding TSP-1 residue at this position, Val1134, is deeply buried in the hydrophobic core of the CTD, suggesting that the pentameric TSPs do not undergo physiological disulphide reshuffling.

Our structure also provides a new perspective on putative cell-binding sites within the C-terminal region of TSP-1. Cell attachment to this region appears to involve multiple receptors and to depend on cell type (e.g., Long and Dixit, 1990; Adams and Lawler, 1993b; Chen *et al*, 1994; Adams, 2001). The putative integrin-binding RGD sequence in T3₇ (Lawler *et al*, 1988; Lawler and Hynes, 1989) is part of a calcium-binding motif and evidently of restricted availability for integrin binding in calcium-saturated TSP-1. Indeed, we found that the RGD-inhibitable attachment of smooth muscle cells to T3₅₋₇-CTD was quantitatively increased under low calcium ion conditions. C2C12 cells, for which attachment to intact TSP-1 or T2₃-T3-CTD is not RGD inhibitable, gained an RGD-dependent component of cell attachment under conditions of low calcium ions and reduction. For this cell type under these conditions, the joint contributions of RGD- and putative CTD-dependent attachment mechanisms resulted in partial protrusive cell spreading on the monomeric T3₅₋₇-CTD. It will be of future interest to determine if there are physiological mechanisms that regulate RGD accessibility. We note that there are cases of allosteric RGD peptide inhibition (e.g., fibrinogen) and it will be important to determine if the TSP-1 RGD motif directly mediates integrin binding.

Studies with synthetic peptides have implicated two CTD sequence regions in cell adhesion and binding to the receptor CD47/IAP (Kosfeld and Frazier, 1993; Gao *et al*, 1996). One sequence (RFYVVMWK) maps to the central strand β7 that is completely buried in the CTD structure; the other sequence (IRVVM) maps to strand β12 and is also largely buried. Neither sequence appears available for direct CD47 binding. One explanation for this discrepancy may be that recombinant CTD of bacterial origin contains non-native conformations (McDonald *et al*, 2003), which may display spurious activities. It will be important to conduct further studies into

the mechanisms of cell adhesion to the C-terminal region of TSP-1 in the light of our structure.

Finally, our finding that the T3 repeat region assumes a complex tertiary structure explains why single missense or deletion mutations in this region of TSP-5/COMP have global consequences on calcium binding and protein function (Chen *et al*, 2000; Maddox *et al*, 2000; Thur *et al*, 2001; Kleerekoper *et al*, 2002). COMP mutations have been found in all T3 repeats. However, the majority of mutations affect the T3 repeats contained in our structure (Briggs and Chapman, 2002; Mabuchi *et al*, 2003), suggesting that the T3₆₋₇ hairpin structure is important for COMP folding, stability or function. Characteristic lamellar aggregates of misfolded mutant COMP are observed within the endoplasmic reticulum of chondrocytes from PSACH and MED patients (Maynard *et al*, 1972; Briggs and Chapman, 2002). Interestingly, however, 293 cells readily secrete mutant COMP (Chen *et al*, 2000; Maddox *et al*, 2000; Thur *et al*, 2001), indicating that intracellular COMP aggregation is a chondrocyte-specific pathology (Maddox *et al*, 1997).

In conclusion, the high-resolution structure of the T3₅₋₇-CTD portion of TSP-1 has revealed in detail the unique and critical role of calcium in assembling the C-terminal region of TSPs and has established a structural homology of the CTD with members of a family of animal and plant lectins. These structural insights provide a new framework for functional studies of this highly conserved region in all TSPs.

Materials and methods

Protein production

The complete cDNA of human TSP-1 was used to amplify by polymerase chain reaction (PCR) the regions coding for residues 816–1152 (T3₅₋₇-CTD) and 877–1152 (T3₇-CTD) of mature TSP-1. The forward primers introduced *EcoRI* and *NheI* restriction sites, and the reverse primers introduced a stop codon followed by *XhoI* and *BamHI* sites (primer sequences available upon request). The PCR products were cloned into the pUC18 vector using *EcoRI* and *BamHI*. Mutations were introduced in T3₅₋₇-CTD using the QuikChange (Stratagene) kit, and transferred into T3₇-CTD using *Clal* and *XhoI*. The sequence-verified cDNAs were cloned into a modified pCEP-Pu expression vector (Kohfeldt *et al*, 1997) using *NheI* and *XhoI*. The pCEP-Pu vector used codes for a fusion protein consisting of the BM-40 signal sequence, followed by a His₆ tag and the respective insert sequence. After secretion and cleavage of the signal sequence, an APLVHHHHHHALA sequence remains fused to the N-terminus of the recombinant protein.

Human embryonic kidney (293-EBNA) cells, maintained in Dulbecco's modified Eagle's medium (DMEM) plus 10% fetal bovine serum, were transfected with the pCEP-Pu vector using Fugene reagent (Roche). Transfected cells were selected with 1 µg/ml puromycin. Proteins were isolated from conditioned serum-free medium supplemented with 1 mM calcium acetate (CaAc₂) using TALON affinity resin (Clontech) and dialysed against 25 mM Na-HEPES pH 7.5, 140 mM NaCl and 1 mM CaAc₂. T3₅₋₇-CTD C974S/N1049K was subjected to preparative gel filtration chromatography on an S200 column (Amersham) and eluted as a single peak corresponding to monomeric protein. The yield was 5 mg of pure protein/litre of culture medium.

Cell attachment assay

Cell attachment to T3₅₋₇-CTD C974S/N1049K was determined as described for C2C12 mouse skeletal myoblasts, rat vascular smooth muscle and human intestinal smooth muscle cells (Adams and Lawler, 1993b, 1994; Anilkumar *et al*, 2002). Recombinant monomeric fragments were coated at a concentration of 1 µM and trimeric TSP-1 was coated at 0.1 µM. As described in Anilkumar *et al* (2002), monomeric proteins were non-adhesive below 0.5 µM. To prepare T3₅₋₇-CTD C974S/N1049K with different calcium loadings,

protein samples were diluted 1:200 from a 9.8 mg/ml stock into TBS containing either 2, 0.5 or 0.1 mM CaCl₂, or 1 mM EDTA, each in the absence or presence of 1 mM dithiothreitol. Blocking and washing buffers contained the same additions. Cells were added for the 1 h assay in standard serum-free DMEM. GRGDSP or GRGESF synthetic peptides (>98% purity; AnaSpec, San Jose, CA) were added to final concentrations of 1 mM, based on previous titration experiments (Adams and Lawler, 1993b, 1994). In some experiments, cells were fixed and processed for phalloidin staining of F-actin as described (Anilkumar *et al*, 2002). Digital images were captured in XY from a Leica TCS-SP/SP-AOSB laser scanning confocal microscope under a $\times 63$ objective using Leica confocal software v2.5 and arranged for figures in Adobe Photoshop.

Crystallisation and structure determination

Crystals of T3₅₋₇-CTD C974S/N1049K were obtained by the hanging drop method at room temperature using a protein concentration of 6–10 mg/ml and 60% MPD, 100 mM HEPES pH 7.5 and 5 mM CaAc₂ as precipitant. The crystals belong to space group R32 with $a = b = 162.34$ Å, $c = 85.14$ Å (hexagonal setting); the asymmetric unit contains one protein molecule and 54% solvent. Diffraction data were collected from frozen crystals at 100 K using a MAR image plate detector mounted on a Rigaku RU-H3R rotating anode X-ray generator equipped with OSMIC focusing mirrors (CuK α radiation; $\lambda = 1.54$ Å). A high-resolution native data set was collected using an ADSC Quantum-4 CCD detector on beamline 9.6 at the SRS Daresbury ($\lambda = 0.87$ Å). The diffraction data were processed with MOSFLM (Leslie, 1994) and programs of the CCP4 suite (CCP4,

1994). Heavy atom derivatives for phasing by multiple isomorphous replacement were obtained by soaking the crystals in artificial mother liquor supplemented with either 2 mM samarium acetate (for 3 days) or 3 mM K₂PtCl₄ (for 18 h). Heavy atom sites were found by standard Patterson and Fourier techniques and refined with MLPHARE (CCP4, 1994) including the anomalous data (2.8 Å resolution; mean figure-of-merit 0.45). Although the phasing information was strong only to ≈ 5 Å, the electron density map after density modification with DM (CCP4, 1994) was of good quality. The model was built with O (Jones *et al*, 1991) and refined with CNS (Brünger *et al*, 1998). Data collection, phasing and refinement statistics are summarised in Table I. The coordinates have been deposited in the Protein Data Bank (accession code 1ux6). BOBSCRIPT (Esnouf, 1997) and RASTER3D (Merritt and Bacon, 1997) were used to prepare the figures.

Acknowledgements

We thank Alexander Becker, Frank Zaucke and Mats Paulsson (University of Cologne, Germany) for help with the calcium-binding experiments; Naomi Clout and the staff at beamline 9.6 at the SRS Daresbury for help with X-ray data collection; Amit Vasanji, CCF Imaging core, for assistance with confocal microscopy; and Peter Brick, Birgit Leitinger, Patrik Maurer and members of the lab for helpful comments on the manuscript. This work was supported by the Wellcome Trust. MK acknowledges receipt of a Wellcome Prize Studentship. EH is a Wellcome Senior Research Fellow.

References

- Adams JC (2001) Thrombospondins: multifunctional regulators of cell interactions. *Annu Rev Cell Dev Biol* **17**: 25–51
- Adams JC, Kureishy N, Taylor AL (2001) A role for syndecan-1 in coupling fascin spike formation by thrombospondin-1. *J Cell Biol* **152**: 1169–1182
- Adams JC, Lawler J (1993a) The thrombospondin family. *Curr Biol* **3**: 188–190
- Adams JC, Lawler J (1993b) Diverse mechanisms for cell attachment to platelet thrombospondin. *J Cell Sci* **104**: 1061–1071
- Adams JC, Lawler J (1994) Cell-type specific adhesive interactions of skeletal myoblasts with thrombospondin-1. *Mol Biol Cell* **5**: 423–437
- Adams JC, Monk R, Taylor A, Ozbek S, Fascetti N, Baumgartner S, Engel J (2003) Characterisation of *Drosophila* thrombospondin defines an early origin of pentameric thrombospondins. *J Mol Biol* **328**: 479–494
- Anilkumar N, Annis DS, Mosher DF, Adams JC (2002) Trimeric assembly of the C-terminal region of thrombospondin-1 or thrombospondin-2 is necessary for cell spreading and fascin spike organisation. *J Cell Sci* **115**: 2357–2366
- Armstrong LC, Bornstein P (2002) Thrombospondin 1 and 2 function as inhibitors of angiogenesis. *Matrix Biol* **22**: 63–71
- Briggs MD, Chapman KL (2002) Pseudoachondroplasia and multiple epiphyseal dysplasia: mutation review, molecular interactions, and genotype to phenotype correlations. *Hum Mut* **19**: 465–478
- Briggs MD, Hoffman SMG, King LM, Olsen AS, Mohrenweiser H, Leroy JG, Mortier GR, Rimoin DL, Lachman RS, Gaine ES, Czekleniak JA, Knowlton RG, Cohn DH (1995) Pseudoachondroplasia and multiple epiphyseal dysplasia due to mutations in the cartilage oligomeric matrix protein gene. *Nat Genet* **10**: 330–336
- Brünger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, Jiang JS, Kuszewski J, Nilges M, Pannu NS, Read RJ, Rice LM, Simonson T, Warren GL (1998) Crystallography, and NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr D* **54**: 905–921
- Chen H, Deere M, Hecht JT, Lawler J (2000) Cartilage oligomeric matrix protein is a calcium-binding protein, and a mutation in its type 3 repeats causes conformational changes. *J Biol Chem* **275**: 26538–26544
- Chen H, Sottile J, O'Rourke KM, Dixit VM, Mosher DF (1994) Properties of recombinant mouse thrombospondin-2 expressed in *Spodoptera* cells. *J Biol Chem* **269**: 32226–32232
- CCP4-Collaborative Computing Project No 4 (1994) The CCP4 suite: programs for protein crystallography. *Acta Crystallogr D* **50**: 760–763
- Delbaere LT, Vandonselaar M, Prasad L, Quail JW, Wilson KS, Dauter Z (1993) Structures of the lectin IV of *Griffonia simplicifolia* and its complex with the Lewis b human blood group determinant at 2.0 Å resolution. *J Mol Biol* **230**: 950–965
- Delot E, Brodie SG, King LM, Wilcox WR, Cohn DH (1998) Physiological and pathological secretion of cartilage oligomeric matrix protein by cells in culture. *J Biol Chem* **273**: 26692–26697
- Dinser R, Zaucke F, Kreppel F, Hultenby K, Kochanek S, Paulsson M, Maurer P (2002) Pseudoachondroplasia is caused through both intra- and extracellular pathogenic pathways. *J Clin Invest* **110**: 505–513
- Dodd RB, Drickamer K (2001) Lectin-like proteins in model organisms: implications for evolution of carbohydrate-binding activity. *Glycobiology* **11**: 71R–79R
- Esnouf RM (1997) An extensively modified version of MOLSCRIPT which includes greatly enhanced colouring facilities. *J Mol Graph* **15**: 132–134
- Gao AG, Lindberg FP, Finn MB, Blystone SD, Brown EJ, Frazier WA (1996) Integrin-associated protein is a receptor for the C-terminal domain of thrombospondin. *J Biol Chem* **271**: 21–24
- Hannah BA, Misenheimer TM, Annis DS, Mosher DF (2003) A polymorphism in thrombospondin-1 associated with familial premature coronary heart disease causes a local change in conformation of the Ca²⁺-binding repeats. *J Biol Chem* **278**: 8929–8934
- Hashimoto Y, Tomiyama T, Yamano Y, Mori H (2003) Mutation (D472Y) in the type 3 repeat domain of cartilage oligomeric matrix protein affects its early vesicle trafficking in endoplasmic reticulum and induces apoptosis. *Am J Pathol* **163**: 101–110
- Hecht JT, Nelson LD, Crowder E, Wang Y, Elder FFB, Harrison WR, Francomano CA, Prange CK, Lennon GG, Deere M, Lawler J (1995) Mutation in exon 17B of cartilage oligomeric matrix protein (COMP) cause pseudoachondroplasia. *Nat Genet* **10**: 325–329
- Holden P, Meadows RS, Chapman KL, Grant ME, Kadler KE, Briggs MD (2001) Cartilage oligomeric matrix protein interacts with type IX collagen, and disruption to these interactions identify a pathogenic mechanism in a bone dysplasia family. *J Biol Chem* **276**: 6046–6055
- Holm L, Sander C (1997) DALI/FSSP classification of three-dimensional protein folds. *Nucleic Acids Res* **25**: 231–234

- Jones TA, Zhou JY, Cowan SW, Kjeldgaard M (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr A* **47**: 110–119
- Kleerekoper Q, Hecht JT, Putkey JA (2002) Disease-causing mutations in cartilage oligomeric matrix protein cause an unstructured Ca²⁺-binding domain. *J Biol Chem* **277**: 10581–10589
- Kohfeldt E, Maurer P, Vannahme C, Timpl R (1997) Properties of the extracellular calcium binding module of the proteoglycan testican. *FEBS Lett* **414**: 557–561
- Kosfeld MD, Frazier WA (1993) Identification of a new cell adhesion motif in two homologous peptides from the COOH-terminal cell binding domain of human thrombospondin. *J Biol Chem* **268**: 8806–8814
- Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. *J Appl Crystallogr* **26**: 283–291
- Lawler J (2000) The functions of thrombospondin-1 and -2. *Curr Opin Cell Biol* **12**: 634–640
- Lawler J, Chao FC, Cohen CM (1982) Evidence for a calcium-sensitive structure in platelet thrombospondin. *J Biol Chem* **257**: 12257–12265
- Lawler J, Derick LH, Connolly JE, Chen JH, Chao FC (1985) The structure of human platelet thrombospondin. *J Biol Chem* **260**: 3762–3772
- Lawler J, Hynes RO (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, Hynes RO (1989) An integrin receptor on normal and thrombasthenic platelets that binds thrombospondin. *Blood* **74**: 2022–2027
- Lawler F, Simons ER (1983) Cooperative binding of calcium to thrombospondin. *J Biol Chem* **258**: 12098–12101
- Lawler J, Weinstein R, Hynes RO (1988) Cell attachment to thrombospondin: the role of Arg-Gly-Asp, calcium, and integrin receptors. *J Cell Biol* **107**: 2351–2361
- Leslie AGW (1994) *MOSFLM Users Guide*. Cambridge, UK: MRC-LMB
- Long MW, Dixit VM (1990) Thrombospondin functions as a cytoadhesion molecule for human hematopoietic progenitor cells. *Blood* **75**: 2311–2318
- Mabuchi A, Manabe N, Hag N, Kitoh H, Ikeda T, Kawaji H, Tamai K, Hamada J, Nakamura S, Brunetti-Pierri N, Kimizuka M, Takatori Y, Nakamura K, Nishimura G, Ohashi H, Ikegawa S (2003) Novel types of *COMP* mutations and genotype-phenotype association in pseudoachondroplasia and multiple epiphyseal dysplasia. *Hum Genet* **112**: 84–90
- Maddox BK, Keene DR, Sakai LY, Charbonneau NL, Morris NP, Ridgway CC, Boswell BA, Sussmann MD, Horton WA, Bächinger HP (1997) The fate of cartilage oligomeric matrix protein is determined by the cell type in the case of a novel mutation in pseudoachondroplasia. *J Biol Chem* **272**: 30993–30997
- Maddox BK, Mokashi A, Keene DR, Bächinger HP (2000) A cartilage oligomeric matrix protein mutation associated with pseudoachondroplasia changes the structural and functional properties of the type 3 domain. *J Biol Chem* **275**: 11412–11417
- Maynard JA, Cooper RR, Ponseti IV (1972) A unique rough surfaced endoplasmic reticulum inclusion in pseudoachondroplasia. *Lab Invest* **26**: 40–44
- McDonald JF, Dimitry JM, Frazier WA (2003) An amyloid-like C-terminal domain of thrombospondin-1 displays CD47 agonist activity requiring both VVM motifs. *Biochemistry* **42**: 10001–10011
- Merritt EA, Bacon DJ (1997) Raster3D: photorealistic molecular graphics. *Methods Enzymol* **277**: 505–524
- Misenheimer TM, Hahr AJ, Harms AC, Annis DA, Mosher DF (2001) Disulfide connectivity of recombinant C-terminal region of human thrombospondin-2. *J Biol Chem* **276**: 45882–45887
- Misenheimer TM, Hannah BA, Annis DS, Mosher DF (2003) Interactions among the three structural motifs of the C-terminal region of human thrombospondin-2. *Biochemistry* **42**: 5125–5132
- Misenheimer TM, Mosher DF (1995) Calcium ion binding to thrombospondin-1. *J Biol Chem* **270**: 1729–1733
- Nakayama S, Kretsinger RH (1994) Evolution of the EF-hand family of proteins. *Annu Rev Biophys Biomol Struct* **23**: 473–507
- Narouz-Ott L, Maurer P, Nitsche PD, Smyth N, Paulsson M (2000) Thrombospondin-4 binds specifically to both collagenous and non-collagenous extracellular matrix proteins via its C-terminal domains. *J Biol Chem* **275**: 37110–37117
- Pimanda JE, Annis DS, Raftery M, Mosher DF, Chesterman CN, Hogg PJ (2002) The von Willebrand factor-reducing activity of thrombospondin-1 is located in the calcium-binding/C-terminal sequence and requires a free thiol at position 974. *Blood* **100**: 2832–2838
- Rosenberg K, Olsson H, Mörgelin M, Heinegård D (1998) Cartilage oligomeric matrix protein shows high affinity zinc-dependent interaction with triple helical collagen. *J Biol Chem* **273**: 20397–20403
- Sun X, Skorstengaard K, Mosher DF (1992) Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin. *J Cell Biol* **118**: 693–701
- Tan K, Duquette M, Lin J, Dong Y, Zhang R, Joachimiak A, Lawler J, Wang J (2002) Crystal structure of the TSP-1 type 1 repeats: a novel layered fold and its biological implication. *J Cell Biol* **159**: 373–382
- Thur J, Rosenberg K, Nitsche DP, Pihlajamaa T, Ala-Kokko L, Heinegård D, Paulsson M, Maurer P (2001) Mutations in cartilage oligomeric matrix protein causing pseudoachondroplasia and multiple epiphyseal dysplasia affect binding of calcium and collagen I, II, and IX. *J Biol Chem* **276**: 6083–6092
- Topol EJ, McCarthy J, Gabriel S, Moliterno DJ, Rogers WJ, Newby LK, Freedman M, Metivier J, Cannata R, O'Donnell CJ, Kottke-Marchant K, Murugesan G, Plow EF, Stenina O, Daley GQ (2001) Single nucleotide polymorphisms in multiple novel thrombospondin genes may be associated with familial premature myocardial infarction. *Circulation* **104**: 2641–2644
- Velloso LM, Svensson K, Schneider G, Pettersson RF, Lindqvist Y (2002) Crystal structure of the carbohydrate recognition domain of p58/ERGIC-53, a protein involved in glycoprotein export from the endoplasmic reticulum. *J Biol Chem* **277**: 15979–15984
- Xiong JP, Stehle T, Zhang R, Joachimiak A, Frech M, Goodman SL, Arnaut MA (2002) Crystal structure of the extracellular segment of integrin $\alpha V\beta 3$ in complex with an Arg-Gly-Asp ligand. *Science* **296**: 151–155