

Identification and Characterization of Thrombospondin-4, a New Member of the Thrombospondin Gene Family

Jack Lawler,* Mark Duquette,* Charles A. Whittaker,‡ Josephine C. Adams,* Katherine McHenry,*
and Douglas W. DeSimone‡

*The Vascular Research Division, Department of Pathology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts 02115; and ‡Department of Anatomy and Cell Biology and the Molecular Biology Institute, University of Virginia Health Sciences Center, Charlottesville, Virginia 22908

Abstract. A new member of the thrombospondin gene family, designated thrombospondin-4, has been identified in the *Xenopus laevis* genome. The predicted amino acid sequence indicates that the protein is similar to the other members of this gene family in the structure of the type 3 repeats and the COOH-terminal domain. Thrombospondin-4 contains four type 2 repeats and lacks the type 1 repeats that are found in thrombospondin-1 and 2. The amino-terminal domain of thrombospondin-4 has no significant homology with the other members of the thrombospondin gene family or with other proteins in the database. RNase protection analysis establishes that the initial expression of

Xenopus thrombospondin-4 is observed during neurulation. Levels of mRNA expression increase twofold during tailbud stages but decrease by the feeding tadpole stage. The size of the thrombospondin-4 message is 3.3 Kb and 3.4 Kb in the frog and human, respectively. Northern blot analysis of human tissues reveals high levels of thrombospondin-4 expression in heart and skeletal muscle, low levels in brain, lung and pancreas and undetectable levels in the placenta, liver and kidney. These data establish the existence of a new member of the thrombospondin gene family that may participate in the genesis and function of cardiac and skeletal muscle.

PLATELET thrombospondin is a 420,000-D glycoprotein that is structurally and functionally similar to the adhesive glycoproteins (Lawler and Hynes, 1986). Multiple structural domains can be discerned on an electron microscopic level (Galvin et al., 1985; Lawler et al., 1985). Globular domains at the NH₂ and COOH-terminals are involved in the interaction of thrombospondin with the cell surface (Gartner et al., 1984; Dixit et al., 1985; Clezardin et al., 1988; Sun et al., 1989; Kosfeld et al., 1991). The central portion of the molecule is composed of multiple copies of structural motifs found in other proteins (Lawler and Hynes, 1986). Amino acid sequences that have been shown to mediate cellular attachment are also present in the central portion of the molecule (Rich et al., 1990; Asch et al., 1992; Prater et al., 1991; Tuszynski et al., 1992). In addition, thrombospondin contains a region that is rich in calcium binding sites and that contains the arg-gly-asp (RGD) sequence that promotes the adhesion of some cell types (Lawler and Simons, 1983; Lawler and Hynes, 1986; Lawler et al., 1988).

Thrombospondin has been shown to modulate the attachment of a variety of cell types in vitro (Varani et al., 1986, 1988; Roberts et al., 1987; Santoro, 1987; Tuszynski et al., 1987; Lawler et al., 1988; Neugebauer et al., 1991; O'Shea et al., 1991; Stomski et al., 1992). For some cells, the interaction with thrombospondin involves the concurrent partici-

pation of several domains within the molecule and multiple receptors on the cell surface (Asch et al., 1991; Stomski et al., 1992). The NH₂-terminal heparin-binding domain binds to proteoglycans including syndecan and to cell surface sulfatides (Roberts et al., 1985; Sun et al., 1989). Thrombospondin also interacts with CD36 or platelet glycoprotein IV (Legrand et al., 1991; Asch et al., 1991; Stomski et al., 1992). Several integrin receptors, including $\alpha_{\text{IIb}}\beta_3$, $\alpha_v\beta_3$, and $\alpha_2\beta_1$ have been reported to bind thrombospondin (Lawler et al., 1988; Lawler and Hynes, 1989; Tuszynski et al., 1989, 1991; Karczewski et al., 1989; Stomski et al., 1992). In addition, $\alpha_v\beta_{100k}$ and a complex of β_1 with a novel α subunit have been reported to be involved in neurite outgrowth (Neugebauer et al., 1991). Through these, and yet to be identified interactions, thrombospondin can modulate cell migration, angiogenesis and neurite outgrowth.

The 185,000-D thrombospondin polypeptides are encoded by two distinct genes, designated *THBS1* and *THBS2* (Bornstein et al., 1990, 1991a, 1991b; Wolf et al., 1990). To date, partial or complete cDNA sequences are available for human, mouse, and frog thrombospondin-1, and human, mouse and chicken thrombospondin-2 (Lawler and Hynes, 1986; Bornstein et al., 1990, 1991a, b; Lawler et al., 1991a, b; Urry et al., 1991). The overall molecular architecture of thrombospondin-1 and 2 are equivalent. The predicted amino acid sequences are very similar in the type 1, 2 and

3 repeats and the COOH-terminal domain, but the NH₂-terminal domains of mouse thrombospondin-1 and 2 are only 32% identical (Bornstein et al., 1991a, b). Phylogenetic analysis based on progressive sequence alignment and comparison based on parsimony indicates that the gene duplication that produced the thrombospondin-1 and 2 genes occurred 583 million years ago (J. Lawler, M. Duquette, L. Urry, K. McHenry, and T. Smith, submitted for publication).

In this paper, we describe the cloning and sequencing of a new member of the thrombospondin gene family. This protein appears to have been produced by a gene duplication event that occurred 925 million years ago (J. Lawler et al., 1993). Whereas the type 3 repeats and the COOH-terminal domain are similar to those of thrombospondin-1 and 2, this protein contains four type 2 repeats, instead of three, and lacks the type 1 repeats and the region of homology with procollagen. We also show that this new member of the thrombospondin gene family is expressed at high levels in adult heart and skeletal muscle. This distribution distinguishes this protein from the recently identified protein, designated thrombospondin-3 (Vos et al. 1992). Based on the number of differences between a partial sequence of mouse thrombospondin-3 and the protein that we have cloned, and on differences in the tissue distribution, we propose that the protein that is described here should be designated thrombospondin-4. In addition, we report that thrombospondin-4 mRNA is first expressed in neurula stage *Xenopus laevis* embryos consistent with a possible role for this protein in heart and skeletal muscle formation.

Materials and Methods

Degenerate Oligonucleotides

The mouse thrombospondin-1, chicken thrombospondin-2 and frog thrombospondin-1 clones were isolated by screening libraries with existing probes for other species at low stringency (Lawler et al., 1991a, b; Urry et al., 1991). We then aligned these sequences with the human thrombospondin-1 sequence and identified two regions that are highly conserved in terms of the nucleotide sequence. Based on these sequences, degenerate oligonucleotides were synthesized and used as primers for the PCR. These sequences fall in the type 3 repeats and the COOH-terminal region of the molecule (Table I).

Polymerase Chain Reaction

Aliquots (1, 5 and 25 μ l) of a *Xenopus laevis* stage 45 cDNA library (DeSimone, unpublished observations) were brought to a final volume of 71.5 μ l with H₂O. The samples were heated to 70°C for 5 min then cooled on ice. To each sample, 10 μ l of 10 \times reaction buffer (Cetus Corp., Emeryville, CA), 6 μ l of 25 mM MgCl₂, 16 μ l of dNTPs and 300 pmol of primers were added. The reaction mixture was heated to 95°C for 5 min and then equilibrated to the annealing temperature (37–48°C). TAQ polymerase (2.5 U) was added and the sample was heated to 72°C for 3 min. The amplification cycles were (a) incubate at 94°C for 1 min and 20 s, (b) incubate at 48°C for 2 min, (c) ramp to 72°C over 2 min, and (d) incubate at 72°C for 3 min. This cycle was repeated 30–40 times; finally the sample was incubated at 72°C for 7 min. The PCR products were separated by agarose gel electrophoresis and the appropriately sized products, based on the thrombospondin-1 and 2 sequences, were subcloned into pBluescript KS or SK (Stratagene Corp., LaJolla, CA).

Cloning and Sequencing

To establish the validity of the thrombospondin-4 clone and to complete the sequence, we screened the *Xenopus laevis* stage 45 library with the PCR

Table I. Oligonucleotides for Degenerate PCR

Forward primer (106T)											
	T	T	T			A	T	T	T	A	T
GACTGAATTCC	AA	GC	AACCAGGCTGA	CA	GA	AA	GA	GG			
	C	C	C			C	C	C	C	G	C
Reverse primer (107T)											
				T	A	A			T	A	
CTAGGAATTCTCG	CCTGG	GTGTTTCC	GT	TGCCA							
	G	G	G						G	G	

product as the probe. The probe was labeled with digoxigenin-dUTP, and hybridization performed using the Genius Kit following the supplier's protocols (Boehringer Mannheim Corp., Indianapolis, IN). Positive plaques were taken through successive rounds of screening with the same probe at progressively lower plaque densities. The purified plaques were amplified to yield high titre plate stocks. Because the library was constructed in the λ ZAPII vector pBluescript II SK, the inserts were excised with helper phage and grown up directly following the supplier's protocols (Stratagene Corp.). *Bam*HI and *Eco*R1 fragments were subcloned into pBluescript II SK and KS. All sequencing was done by the chain termination method of Sanger et al. (1977) with Sequenase reagents (US Biochemical Corp., Cleveland, OH). The ends of all clones and subclones were sequenced with the remainder of the sequence being determined using synthetic oligonucleotides as primers. The sequence of thrombospondin-4 was obtained on both strands.

The largest clone that we obtained from screening the stage 45 library was 2.8 kb. To complete the sequence, we have used a 580 bp *Pst*I-*Stu*I fragment of the XS9 clone to screen a *Xenopus laevis* stage 22 library (the generous gift of Dr. Douglas Melton, Department of Biology and Molecular Biology, Harvard University, Cambridge, MA). Eight clones were isolated and phage DNA was prepared as described previously (Lawler and Hynes, 1986). Of these clones, one clone (designated XM15) contains additional sequence at the 5' end of the TSP-4B family (Fig. 1).

Preparation of Probe for Human Thrombospondin-3

The genomic clone GPEM-2 was kindly provided by Dr. Sandra Gendler (Imperial Cancer Research Fund, London; Lancaster et al., 1990). *Bam*HI fragments of GPEM-2 were subcloned into pBluescript KS and the ends of each clone were sequenced. One of these clones contained sequences that were homologous to the 3' end of the thrombospondin-1, 2 and 4. Based on this homology, the position of the 5' end of the last exon was determined. The 3' end of this exon was taken to be the polyadenylation signal. Oligonucleotides that primed at the 5' and 3' ends of the last exon were used to amplify and clone a 293 bp DNA segment that corresponds to the last exon of the human thrombospondin-3.

Northern Blot Analysis

The *Xenopus* thrombospondin-4 clone XF3 was digested with *Eco*R1 and *Xho*I and the insert purified. The human thrombospondin-1 probe was the human full-length cDNA (Lawler et al., 1992). The PCR product for the last exon of thrombospondin-3 and the β -actin probe (Clontech, Palo Alto, CA) were radiolabeled directly. All probes were radiolabeled with the Multiprime DNA Labeling System (Amersham Corp., Arlington Heights, IL).

A Northern blot that was prepared with Poly A+ RNA from adult human heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas was obtained from Clontech. The blot was prehybridized and hybridized as described previously (Lawler and Hynes, 1986).

Construction of Probes for RNase Protection

Templates for the synthesis of antisense thrombospondin-4 transcripts were prepared by linearizing a *Bam*HI subclone of clone XF3 with *Eco*R1. A standard in vitro transcription reaction was carried out using 0.5 μ g of template and 20 U of T7 RNA polymerase (Promega Corp., Madison, WI) as described in DeSimone et al. (1992). Transcripts prepared using this template were ~355 nucleotides in length and protected a fragment of ~300 nucleotides. EF1 α template was prepared by digesting a *Pst*I-*Sac*I fragment of *Xenopus* EF1 α cloned into pSP65 (Krieg et al., 1989) with *Alu*I. Antisense

transcripts ~90 nucleotides in length were synthesized using Sp6 RNA polymerase (Promega Corp.). This probe protects a fragment of ~75 nucleotides. Probes were labeled by the incorporation of [α - 32 P]rUTP during the synthesis reaction (800 Ci/mmol; ICN Biomedicals Inc., Costa Mesa, CA) and purified on a 5% polyacrylamide denaturing gel.

RNAse Protection Analysis

Total RNA was isolated from embryos exactly as described in DeSimone et al. (1992). Embryo equivalents of total RNA were then hybridized to 50,000 cpm of purified probe. Hybridization was carried out in the presence of 80% formamide, 0.4 M NaCl, 40 mM Pipes (piperazine-N,N'-bis-2-ethanesulfonic acid), and 1 mM EDTA at 45°C for at least 12 h. Digestion of unhybridized probe and electrophoresis was carried out exactly as described in DeSimone et al. (1992), except that 47 U of RNase A and 344 U RNase T₁ (Boehringer Mannheim Corp.) were used per 350 μ l RNase reaction.

Results

Identification of Thrombospondin-4

Degenerate PCR using the *Xenopus laevis* stage 45 library has produced four distinct sequences that are related to the thrombospondins. Two of the four sequences correspond to the two copies of the thrombospondin-1 gene that are present in the *Xenopus* genome (Urry et al., 1991). The *Xenopus* genome duplicated ~30 million years ago (Bisbee et al., 1977). In some cases, both copies of the gene are expressed (e.g., DeSimone and Hynes, 1988). To date, the thrombospondin-1 sequences represent the majority of the products that we have obtained. However, two PCR products comprise sequences that are related to, but clearly distinct from thrombospondin-1. The sequences of these two PCR products are very similar to each other suggesting that they represent the two copies of a newly identified gene in the *Xenopus* genome.

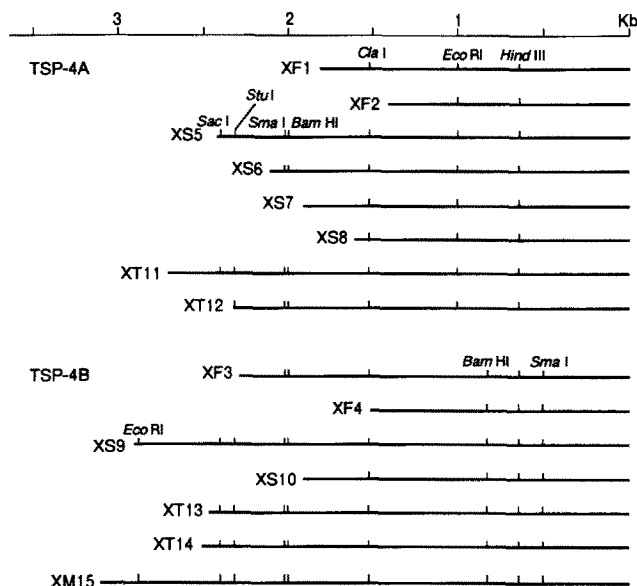


Figure 1. Alignment of *Xenopus* thrombospondin-4 clones. Restriction endonuclease sites are indicated for the two families (TSP-4A and TSP-4B). The clones that have been isolated in the first (XF1-XF4), second (XS5-XS10), and third (XT11-XT14) rounds of screening have been grouped into their appropriate family by restriction endonuclease mapping and nucleotide sequencing.

To establish that these sequences are in fact derived from the *Xenopus* library and to obtain more nucleotide sequence, we have prepared a probe from the PCR product and screened the library. A screen of 120,000 plaques produced four positive clones that range in size from 1.7 to 2.3 kb (Fig. 1, XF1-XF4). As shown in Fig. 1, the restriction maps of the clones indicate that two distinct gene products can be identified. The longest clone for each gene (XF1 and XF3) has been sequenced on both strands. The sequence of the PCR products is included in the sequences of these clones. These data confirm that the PCR product is derived from the *Xenopus* library and not from another contaminating source. When clone XF3 is used to probe a Northern blot of *Xenopus* stage 17 RNA, a 3.3 kb band is observed (Fig. 2). Because the message size is greater than the length of clone XF3 and the reading frame is open at the 5' end of the predicted amino acid sequence for clone XF3, the library was rescreened with the *Eco*R1 fragment of clone XF1. This screen produced six additional clones (XS5-XS10, Fig. 1). Clone XS9 has been sequenced on both strands. Clone XS9 is ~469 nucleotides smaller than the message and the reading frame is open at the 5' end of the predicted amino acid sequence. The library has been rescreened with the *Eco*R1 to *Bam*HI fragment of XS9. Four additional clones have been isolated (XT11-XT14); however, they do not contain additional nucleotide sequence.

To obtain additional 5' sequence, a *Xenopus laevis* stage 22 library was screened. Restriction endonuclease mapping indicated that one of the clones (XM15) that was isolated contained additional 5' sequence for the TSP-4B family (Fig. 1).

Determination of Nucleotide Sequence

The clones XF3, XS9, and XM15 have been used to determine the nucleotide sequence of *Xenopus* thrombospondin-4 on both strands. The sequence is 3,120 nucleotides in length. A single open reading frame exists between nucleotides 103 and 2,970 (Fig. 3). There is a short (140 bp) 3' untranslated region that ends with a continuous series of adenosines. An AATAAA consensus polyadenylation signal is observed upstream of the poly A sequence. The first 120 bp of the sequence contains stop codons in all three reading frames. The sequence is 180 bp smaller than the message size that is

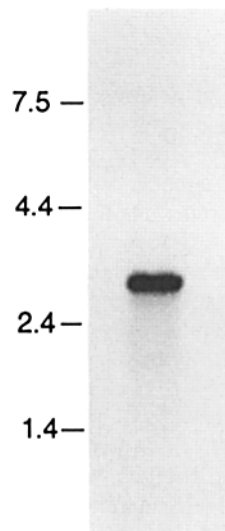


Figure 2. Northern blot of *Xenopus* stage 17 RNA. Two micrograms of total stage 17 mRNA was electrophoresed and blotted. The blot was probed with the XF3 clone. The positions and sizes of the markers are indicated on the left.

1: AGTCTTCTCTGCTGCCATCCACTAAGCCTGCTCACTGCTGAGGAGCTGGGATATTTACA
61: TCTGCTGAGCTGCTAAAGGCTAATCCATTGCAGAAGTCAAGATGCCTAGACGAAAAGGC
-24: M P R R K G
121: TTGTGCTATTCTACAGATGTTACTTCTTTCATCTTTATGGAGTGTGCAAGCCAGCCA
-18: L C L F L Q M L L L H L Y G V C Q A Δ P
181: AATTATCAAGTCTTCGATCTTCTATCCGTGTCGGTCAAAGGCAGGTCACTCTTTTCCTT
3: N Y Q V F D L L S V S V Q R Q V T S F L
241: CAAAGCCACTGAGCAATCCCTCCATGAATGAGGTGTATGCTTGTCAACTTCAAACTG
23: Q S A L S (N) P L S M N E V Y V L S T F K L
301: CAGCCCAAGTCCACAGTTCAGCTCTTTGGACTTTATCCACCAAGTGAACAACAGCAGGTT
43: Q P K S T V T L F G L Y S T S D N S R F
361: TTTGAATTCACAGTTCATGGTTCGTTAAACAAAGCCTCTTACAGATACCTCCGGAGTGAT
53: F E F T V H L R G T L N K A S Q L R Y K L F I D
421: GGAAGTTACACTCAGTCTTCTTTAATAAGCTTGACATAGCTGATGGGAAGCAGCAGCGG
83: G K L H S V F F N K L D I A D G K Q H A
481: CTTCTGTGCACTGAGCGGCTTACACGGGGCGCAACGTTTGAAGAGCTTACATAGAC
103: L L L H L S P L G R L N K A S Q L R Y K L F I D
541: TGTAAATCCGACAGTGTGTGTAAGACTACCCCGGCGTATCAGGGATAAGGCTCAAC
123: C (N) P T G V V E D L P R P L S G I R L N
601: ACAGGCTCTGCACTTAAGAACAACACAGAAAAGGACAGGATCCATGGATGAATTA
143: T S E V H L R G T L N K A S Q L R Y K L F I D
661: AAATGGTAATGGAGGCACTCTGTCAGGAGTGAAGCAACAAGAATGTTTATGACAG
163: K L V M G G T L S E V G A I Q E C F M Q
721: AAAAGTGAAGCCGACAGCAGGTCAGCTCAGCAGCAGTGTATTGGCCAGATAAC
183: K S E A G L R G V R C I (D) T E G G F Q C G
781: CAAATGAATCAGATGCTGGGAGAGCTCCGAGATGTCATGAGACAGCAGGTGAAGAGACC
203: Q M N Q M L G E L R D V M R Q Q V K E T
841: ATGTTCTTGAAGAACCACTTACAGATGCCAGGCTGTGGCTTAGTCTGACTTCCCA
223: M F L R N T I A E A C Q C G L G S F P
901: TTGCCAACCAAGTTCGCCAGCGCTAGCCACCACCTACACCTCCAAAGCCTCGATGTGAT
243: L P T K V P Q R L A T T T P P K P R C D
961: GCAACTTCATGTTTCAGAGAGTGGGTCATTGATACAGAGGGCGCTTCCAATGTGGG
263: A T S C F R G V R C I (D) T E G G F Q C G
1021: CCGTCTCTGAAGCTATACAGCCAAAGGTCATTTGTACTGATGTGGATGAGTGTCCG
283: P C P E G Y T G N G V I C T D V D E C R
1081: TTGAATCCATGTTTCTTGTGTACCTGTCATAAACACTTCTCCGGTTCAAATGTGAG
303: L N P C F L R G V R C I (N) T E G G F Q C G
1141: AGCTGCCCTCCCGGTACACTGGATCCCAAAATCAAGGGATGGCATTAACTTTGCCAAG
323: S C P P G Y T G S T I Q G I G I N F A K
1201: CAAAATAAGCAGGTTGCACAGATCAACATGAATGTAAAATGGAAGAAATGGAGGGTGT
343: Q N K Q V T C T D T N E C T E N G R N G G C

1261: ACATCCAATCTCTTTCATCAATACGATGGGATCTTCCCGTGTGGGGGCTGCAAACT
363: T S N S L C I (N) T M G S F R C G G C K P
1321: GGTATGTGCGGGATCAAAATAAAGGCTGCAAACTGAAAAAGCTCCGTCATGGACAG
383: G Y V G D Q I K G C K P E K S C R H G Q
1381: AATCCGTGTCATGCAAGTGTCTAGTGTTCAGAGGAAAAGGACGGTGAACCTGCACCT
403: N P C H A S A Q C S E E K D G D V T C T Y
1441: TGTTCAGTGGTGGGCGGCAATGGCTACCTCTGTGGCAAGATACTGATATTGATGGC
423: C S V G W A G N G Y L C G K D T D I D G
1501: TACCCGGATGAAGCCCTGCCATGTCCAGATAAGAAGCTGAAAAAGGCAACTGTGTATAT
443: Y P D E A L P C P D K N C K D N C V Y
1561: GTTCTAACTCGGTCAGGAAGCACTGATAAAGATAACATGGAGATGCTTGTGATGAA
463: V P N S G Q E D T D K D N I G D A C D E
1621: GATCGGATGGAGATGGTATCTAAATGAGCAGGACAACTGTGTGCTGGTCCCAACATC
483: D S D G D G I L R P D K N C K D N C V Y
1681: GATCAGAAAAACAGTGACCAAGATATATTGGGAGCGCTGTGCAACTGCCGCTTAACC
503: D Q K N S D Q D I F G D A C D N C R L T
1741: CTCACCAATGACCAAGGGACACAGACAATGACGGAAAAGGAGATGCTTGTGACGATGAC
523: L N N D Q R D T D N C D K D N C V Y
1801: ATGGATGGAGATGGCATCAAGAATATCTGGATAACTGCCAGAGATGCCAATGTGGAC
543: M D G D G I K N I L D N C Q R V P N V D
1861: CAGAAAGCAAAAGATGGAGATGGTGTGATATATGTGACAGCTCTCTGACATCATI
563: Q S D G D G D G V G D I N C D P L A I N I
1921: AATCCAAACAGTCAGACATGACAATGACCTTGTGGAGATCTCTGTGATACTAACCAA
583: N P (N) Q S D I D N D L V G D S C D T N Q
1981: GACAGCGATGGTGTGTCACAGGACAGCAGCAACTGCCCAAGTGTATAAACAGC
603: D S D G D G H Q D S C D N C D P L A I N I
2041: AACCAGCTCGACACAGCAAGGACCGCATCGGAGATGAATGTGACGATGATGATAAC
623: N Q L D T D K D G I G D E C D D D D D N
2101: GATGAACTCCCGGATCTTCTCCCGGACCTGATAACTGAACTGGTTCGCCAACCCA
643: D G I P D T V P P P G P D N C K L V P N Y
2161: GGGCAGGAGGATGACAAATGATGGATCGGAGACGTCGTGAGGCGGATTTTGACCAG
663: G Q E D D N N D G V G D V C E A D F D Q
2221: GACACGGTCATTGACCGAATGACGTTTGCCTGAAAATGCAGAGATCACCTGCAGAT
683: D T V I D R I D V C P E N A E I T L T D
2281: TTCAGAGCTTATCAAATGATGTTCTGGATCCGAGGAGATGCCAAATGATCCAAAC
703: F R A Y Q T V V L D P E G D A Q I D P N
2341: TGGATGTTTGAACCCAGGAAATGGAGATGTGTCAGACGATGAACAGTGCACCTGGACTG
723: W I V L N Q G M E I V G P T M S S P F Y
2401: GCAGTTGGTTACACAGCATTTAATGGAGTTGATTCGAGGGCACATTCACCGTGAACCC
743: A V G Y T A F N G V D F E G T F H V N T
2461: ATGACGGATGATGATACCGTGGTTCATCTTGGTTATCAGGACAGTCAAGCTTTAT
763: M T D D D Y A G F I V G Q D S S S F Y

Figure 3. The nucleotide sequence and corresponding amino acid sequence of *Xenopus* thrombospondin-4. The predicted site for signal sequence cleavage is indicated (open triangle). The termination codon TAA and the polyadenylation signal AATAAA are underlined. The sites for potential N-linked glycosylation are circled and the site for potential β -hydroxylation is boxed. Note that asparagine(314) could be either glycosylated or β -hydroxylated. These sequence data are available from EMBL/GenBank/DBJ under accession number Z19091.

predicted by Northern blotting. Thus, the sequence that is presented here appears to lack ~180 bp of 5' untranslated sequence.

Amino Acid Sequence

An open reading frame that is 931 amino acids long is predicted from the nucleotide sequence of *Xenopus* thrombospondin-4 (Fig. 3). The first 24 amino acids appear to comprise a signal sequence. The predictive algorithm of von Heijne (1986) indicates that the signal sequence would be cleaved after alanine(-1) and that glutamine(1) is the NH₂-terminal of the mature thrombospondin-4 peptide. The mature thrombospondin-4 peptide has a molecular weight of 102,143 and an isoelectric point of 4.2, before post-translational modifications. As with the other members of the thrombospondin gene family that have been sequenced, the majority of cysteine residues are located in the center one third of the molecule. Two types of repeating sequences, designated homology type 2 and 3 occur within the amino acid sequence (see below). There are five potential sites for N-linked glycosylation, although two of them include a proline residue and are less likely to be used and one of them

also fits the consensus for β -hydroxylation (Fig. 3; Hubbard and Ivatt, 1981; Stenflo et al., 1988).

The first 282 amino acids of the thrombospondin-4 sequence have very little homology with human thrombospondin-1 (Lawler and Hynes, 1986). Whereas amino acid sequences GKQH and LYDC are present in both sequences and the positions of the cysteine residues are similar, we have not detected regions of homology with the heparin-binding motifs of thrombospondin-1 (Lawler et al., 1992).

Four adjacent type 2 repeats can be identified on the basis of the positions of the cysteine residues (Fig. 4). The overall homology with other thrombospondins is low in this region and the introduction of several gaps is necessary to optimize the alignments. However, the second type 2 repeat of thrombospondin-4 is similar to those of thrombospondin-1 and 2, in that thirteen residues are inserted between the last two cysteine residues. In addition, a consensus sequence for β -hydroxylation can be identified in the second type 2 repeat of thrombospondin-1, 2 and 4 (see Discussion). The first and third type 2 repeats of thrombospondin-4 also contain consensus sequences for β -hydroxylation (Fig. 3).

The type 3 repeats of *Xenopus* thrombospondin-4 are

2521: GTGGTGTGTGGAAGCAGACTGAGCAGACTTACTGGCAGGCAACCCCTTCAGAGCAGTT
 783: V V M W K Q T E Q T Y W Q A T P F R A V

2581: GCAGAGCCTGGAATCCAACCTGAAGGCTGTGAAATCCAAGTCAGGACCCGGGAACATCTG
 803: A E P G I Q L K A V K S K S G P G E H L

2641: AGGAACGCTCTGTGGCAGCAGGAGACACCAATGATCAAGTGAGGCTGCTCTGGAAGAC
 823: R N A L W H T G D T N D Q V R L L W K D

2701: CCCAGGAATGTCGGCTGGAAAGACAAAGTCTCCTACCGCTGGTTCTTACAGCACAGGCCA
 843: P R N V G W K D K V S Y R W F L Q H R P

2761: CAAGTCGGCTACATCAGAGCCAGATTTTATGAAGGCCACCGAGCTGGTGGCTGACTCTGGA
 863: Q V G Y I R A R F Y E G T E L V A D S G

2821: GTCACGTGGACACCACCATGCGAGGAGGAAGACTGGGAGTATTCTGCTTTTCACAGGAA
 883: V T V D T T M R G G R L G V F C F S Q E

2881: AACATAATTTGGTCCAATCTGAAATACCGGTGTAATGATACAATCCAGAGGATTTCCAG
 903: N I I W S N L K Y R C (N) D T I P E D F Q

2941: GCATTTCAAGCACAAACAGTTTCCAGTTAAACAGAACCCACACAATATCCGGTGATTTTT
 923: A F Q A Q Q F S S

3001: TTTTGTGATTTTTTTTTTGTAGTAATATGAGAAAACGTTATTTTCATGCAGCCTTGT
 3061: CTACCAACTGTACAATAATGTCTGTAATAAATGGATACAAAATGAGAAAAA

Figure 3.

61.4% identical to the type 3 repeats of human thrombospondin-1 (Fig. 4). The consensus sequence and overall organization of the type 3 repeats of thrombospondin-4 are equivalent to thrombospondin-1, with the second and fourth type 3 repeats being truncated after the second cysteine. Thrombospondin-4 contains four amino acids (PPGP) at the end of the sixth type 3 repeat that do not align with sequences in thrombospondin-1 or 2.

Alignment of the COOH-terminal of the *Xenopus* thrombospondin-4 sequence with the last 227 amino acids of human thrombospondin-1 reveals that 60.8% of the amino acids are identical and no insertions or deletions are required. The thrombospondin-4 sequence includes fifteen amino acids that extend beyond the stop codon for human thrombospondin-1. The cysteine residue in position 898 aligns with a valine residue in thrombospondin-1 and 2. Because thrombospondin-4 has an odd number of cysteine residues in the predicted mature peptide, it is possible that cysteine(898) has a free sulfhydryl group.

A. Type 2 Repeats of Thrombospondin-4:

PRCDATS---CFRGRVRCIDTEGGFQ-CGFCPEGYTG-----NGVICTDV
 DECRLL--NP-CFLGVRCINTSPGFK-CESCFPGYTGSTIQQGIGINFARQNKQVCTDT
 NECENGRNGGCTSNLSLINTMGSFR-CGGCKPGYVG-----DQIKGCKPE
 KSCRHGQNP-CHASAQCSEKVGVDVCT-CSVGWAG-----NGYLCKG

B. Type 3 Repeats of Thrombospondin-4:

DNCVVVPSNGQEDTDKDNIGDACDE--DADGDGILNEQ
 DNCVLAANIDQKNSDQDIFGDAC
 DNCRLTLNNDQRDTDNDGKGDACDD--DMDGDGIKNIL
 DNCQRVFNVDQKDKDGDVGDIC
 DSCFDIINPNQSDIDNDLVGDSCTDNQDSGDGHDQDST
 DNCPTVINSNQLDTRDKDGIQDCCDD--DDDNDGIPDTPVPPGP
 DNCKLVPNPGQEDDNDGVDVCEA--DFDQDVTIDRI

D. C...N..Q.D.D.D..GD.C...D.D.D..... Consensus
 DNC...N..Q.D.D.D..GD.C...D.D.D..... TSP-1 and 2 Consensus

Figure 4. The amino acid sequence of the type 2 (A) and 3 (B) repeats of thrombospondin-4. The amino acid sequences for the four type 2 repeats P(217)-K(394) and seven type 3 repeats (D[416]-I[647]) have been aligned. The positions of gaps are marked by dashes. The consensus sequence for the type 3 repeats of *Xenopus* thrombospondin-4 is compared to that for human and mouse thrombospondin-1 and chicken thrombospondin-2 at the bottom of the figure. The underline indicates that an N occupies one of the positions that is occupied by a D.

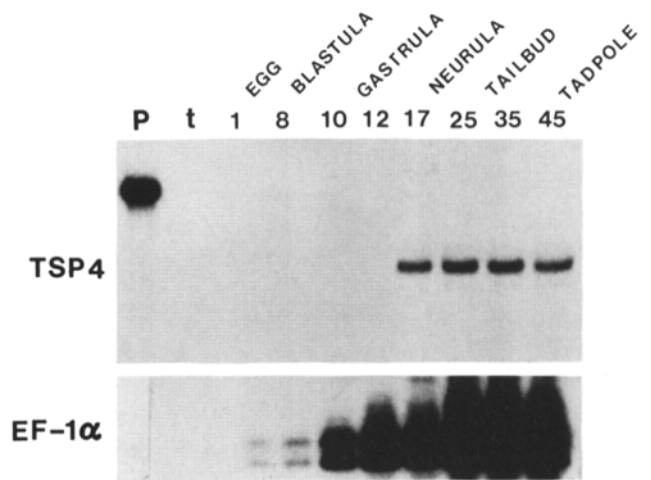


Figure 5. Developmental expression of thrombospondin-4 mRNAs during *Xenopus* development as determined by RNase protection analyses. Ten embryo equivalents total mRNA per lane were hybridized with a ³²P-labeled thrombospondin-4 antisense transcript synthesized in vitro. EF1-α mRNA analysis is included as a control for RNA loading. The relative stage-specific levels of EF1-α shown here are equivalent to those described by Krieg et al. (1989). Embryos were staged according to Nieuwkoop and Faber (1967). Stage 8 corresponds to the mid-blastula, which is marked by the onset of zygotic transcription in *Xenopus* (Newport and Kirschner, 1982). (P), probe alone, unhybridized transcript, no RNase; (t) tRNA control lane, no protected fragments after RNase digestion. Stages include: (1) fertilized egg, (8) mid-blastula, (10-12) early-late gastrula, (17) neurula, (25-35) tailbud stages, (45) feeding tadpole. Thrombospondin-4 protected fragment is ~300 nucleotides. EF1-α protected fragment is ~75 nucleotides. Full length EF1-α probe (90 nucleotides) is not shown.

Expression of Thrombospondin-4 during *Xenopus* Development

RNase protection analysis has been used to establish the temporal pattern of thrombospondin-4 expression during *Xenopus* development (Fig. 5). 10 embryo equivalents of fertilized eggs and representative developmental stages have been hybridized with ³²P-labeled antisense thrombospondin-4 and EF1-α transcripts followed by digestion with RNase as described in Materials and Methods. EF1-α has been included as a control for RNA loading (Krieg et al., 1989). No thrombospondin-4 mRNA is detected in egg, cleavage stage, blastula or gastrula stage embryos, even after prolonged exposure. These data indicate that thrombospondin-4 is not maternally stored in *Xenopus*. Furthermore, thrombospondin-4 transcripts are not among those mRNAs synthesized at the onset of zygotic transcription, which commences at the mid-blastula stage in *Xenopus* (Newport and Kirschner, 1982). Initial expression of thrombospondin-4 is observed during neurulation (Fig. 4, stage 17). Levels of mRNA expression increase ~twofold during tailbud stages but decrease by the feeding tadpole stage (Fig. 5, stages 25, 35, and 45).

Expression of Thrombospondin-4 in Adult Human Tissues

A Northern blot of poly A+ selected RNA from eight adult human tissues is shown in Fig. 6. The size of the human thrombospondin-4 message is 3.4 kb. Thrombospondin-4

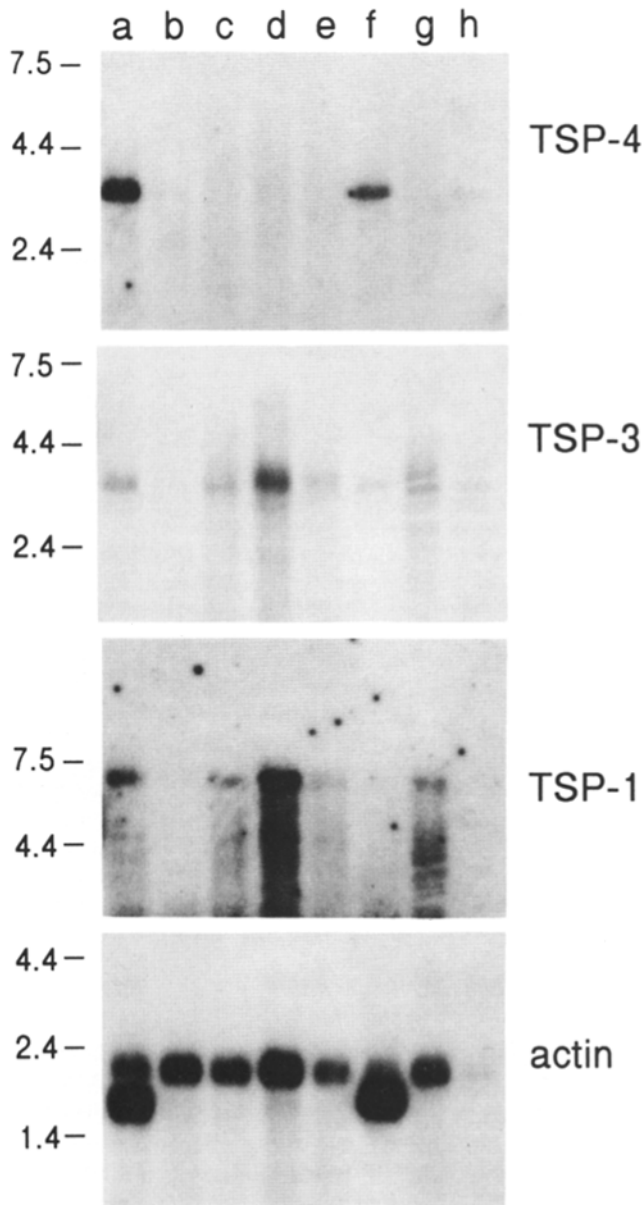


Figure 6. The expression of thrombospondin-4 in adult human tissue. A northern blot of poly A + RNA from adult human heart (a), brain (b), placenta (c), lung (d), liver (e), skeletal muscle (f), kidney (g), and pancreas (h). The blot was probed with a 2.2-kb fragment of *Xenopus* thrombospondin-4. The positions and sizes (kb) of the markers are indicated on the left.

shows a restricted pattern of expression. High levels of expression are observed in the heart and skeletal muscle (Fig. 6). On longer exposures, a faint band is detectable in the tissue from the brain, lung and pancreas. No expression has been detected in the placenta, liver or kidney. Comparable levels of the 2.0 kb form of β -actin are observed in all of the lanes except the pancreas (Fig. 6). Because a considerable fraction of the total mRNA in the pancreas encodes preproinsulin and α -amylase, other mRNAs give a lower hybridization signal. Thus, although the thrombospondin-4 signal is weak in the pancreas, the relative level of expression may be significant.

When the same blot is probed for thrombospondin-3, the

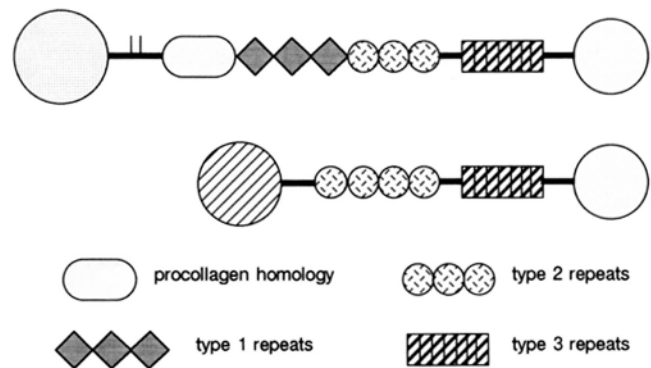


Figure 7. Schematic model of thrombospondin-4. The molecular architecture of thrombospondin-4 is shown below that of thrombospondin-1 or 2. The NH₂ terminal is to the left and the COOH terminal is to the right. Whereas the NH₂-terminal domains are similar in size, they share very little sequence homology. Two interchain disulfides of thrombospondin-1 and 2 are depicted as vertical lines between the NH₂-terminal domain and the region of homology with procollagen. Whereas two cysteine residues are present in a similar location in thrombospondin-4, the structure of the protein that contains the thrombospondin-4 polypeptide remains to be determined.

strongest signal is observed in the lung (Fig. 6). The size of the thrombospondin-3 message is also 3.4 kb. Lower levels of thrombospondin-3 expression are observed in most of the lanes with the brain displaying the weakest hybridization signal (Fig. 6). The adult lung tissue also produces the strongest signal when the blot is probed with a human thrombospondin-1 probe (Fig. 6). Varying levels of thrombospondin-1 are observed in all of the tissues on the blot (Fig. 6). In this case, the principal message is 6.0 kb with faint bands at 4.5 and 3.6 kb.

Discussion

In this paper, we have described a new member of the thrombospondin gene family, designated thrombospondin-4. This member is similar to thrombospondin-1 and 2 in that the type 3 (calcium-binding) repeats and the COOH-terminal domain are conserved. By contrast, the NH₂-terminal domain is different from those of thrombospondin-1 and 2, and the regions of homology with procollagen and the type 1 repeats are absent (Fig. 7). Thrombospondin-4 appears to have a more restricted pattern of tissue expression than the other members of the thrombospondin gene family. The highest levels of thrombospondin-4 message are in adult human cardiac and skeletal muscle. These data indicate that specific tissues are capable of synthesizing specific forms of thrombospondin and that biodiversity in the thrombospondin gene family is generated at a transcriptional level with multiple genes participating.

The difference in tissue distribution between thrombospondin-3 and 4 was the first indication that they represented different gene products (Vos et al., 1992). Several other observations support this conclusion. The amino acid identity between the coding sequences of mouse thrombospondin-3 and *Xenopus* thrombospondin-4 is lower than would be predicted based on the high level of sequence identity between mouse and *Xenopus* thrombospondin-1 in an equiva-

lent region (Vos et al., 1992; Lawler et al., 1991b; Urry et al., 1991). Thrombospondin-3 and 4 are also distinct in the size and sequence of the 3' untranslated regions. However, because our clones have been isolated from a *Xenopus* library and the initial characterizations of thrombospondin-3 have been performed in the mouse and human, differences in sequence or gene structure might be due to species variations (Vos et al., 1992). Using oligonucleotide primers that are based on the *Xenopus* sequence, we have isolated and sequenced a human thrombospondin-4 clone (J. Lawler, M. Duquette and K. McHenry, unpublished data). This establishes that all four forms of thrombospondin are present in the human genome. We are currently mapping the thrombospondin-4 gene in the human and mouse genomes.

The overall molecular architecture of thrombospondin-4 is similar to that of thrombospondin-1 and 2 in that the NH₂ and COOH-terminal portions do not contain multiple copies of sequence motifs and are not homologous to other proteins in the database. When the database is searched with the COOH-terminal domain of thrombospondin-4, only significant homology with thrombospondin-1 and 2 is detected. However, a search with the first 282 amino acids of the sequence reported here fails to identify thrombospondin-1 and 2. This gradient of homology seems to represent a general feature of the thrombospondin gene family: the levels of identity between the thrombospondin-1 sequences in the human and mouse, or the thrombospondin-1 and 2 sequences in the mouse, are significantly greater at the 3' end of the genes.

The molecular architecture of the central portion of thrombospondin-4 is distinct from that of thrombospondin-1 or 2 (Fig. 7). Whereas the type 2 and type 3 repeats are present in all of the forms of thrombospondin that have been sequenced to date, thrombospondin-4 lacks the region of homology with procollagen and the type 1 repeats. This suggests that the divergence of thrombospondin-4 from the branch that gave rise to thrombospondin-1 and 2 occurred before the exon shuffling events that led to the insertion of the region of homology with procollagen and the type 1 repeats. We have proposed that this gene duplication occurred 925 million years ago (J. Lawler et al., 1993). The thrombospondin-1 and 2 branches appear to have diverged 583 million years ago (J. Lawler et al., 1993). By this time, the region of homology with procollagen and the type 1 repeats were apparently present. This proposal is consistent with the existence of the type 1 repeats in the invertebrate, *C. elegans* (Leung-Hagesteijn et al., 1992). In vitro mutagenesis experiments and expression of variant forms of the thrombospondin-1 polypeptide indicate that the region of homology with procollagen is involved in trimer assembly (Lawler et al., 1992). The absence of this region from thrombospondin-4 may indicate that this molecule is not trimeric.

Whereas thrombospondin-1 and 2 have three type 2, EGF-like repeats, thrombospondin-4 has four type 2 repeats. A proposed consensus sequence for β -hydroxylation is similar to a sequence that is found in the second type 2 repeat of all of the members of the thrombospondin gene family (Stenflo et al., 1988). This sequence is found between the cysteine residues at positions 312 and 324 of *Xenopus* thrombospondin-4 (Fig. 3). The asparagine at position 314 would be predicted to be β -hydroxylated. In thrombospondin-4, the

consensus sequence for β -hydroxylation is also found in the first and third type 2 repeats (Fig. 3). The only variation from the consensus sequence is the insertion of an additional residue between the fourth and fifth cysteine residues in the thrombospondin type 2 repeats. Human thrombospondin-1 has been reported to contain erythro- β -hydroxyasparagine (Przywiecki et al., 1987). In factor IX and protein C, β -hydroxylation of an aspartic acid residue of the EGF-like domain is involved in calcium binding to this domain (Ohlin et al., 1988; Hanford et al., 1991). Mutation of the aspartic acid to asparagine in the EGF-like repeat of factor IX markedly reduces calcium binding. However, if a glutamic acid residue is present in a specific position that is upstream of the asparagine, this repeat is able to support calcium binding (Fig. 3). This has been designated a type II consensus sequence by Hanford and co-workers (1991). All of the members of the thrombospondin gene family that have been sequenced to date contain this motif (Fig. 2). These data suggest that the second EGF-like repeat of thrombospondin contains a calcium-binding site that has been conserved during evolution.

Thrombospondin-4 mRNA is expressed early in embryogenesis, suggesting a role for this protein in development. The highest levels of thrombospondin-4 expression that are observed coincide with the appearance of the major organ systems and tissue types in the embryo. These events include the formation of the heart, vascular system and skeletal muscle. Interestingly, after reaching maximal levels of expression during tailbud stages a reproducible drop in thrombospondin-4 mRNA levels is noted by the feeding tadpole stage. Further analysis will be required to determine the spatial expression of both thrombospondin-4 mRNA and protein during these stages using in situ hybridization and immunohistochemical methods. However, given the timing of expression in the embryo and the adult tissue distribution of thrombospondin-4 mRNA (Figs. 5 and 6), it is reasonable to propose that thrombospondin-4 may have an important function in heart and muscle morphogenesis. Immunocytochemistry with monoclonal and polyclonal antibodies that have been prepared against human platelet thrombospondin indicates that thrombospondin(s) are widely expressed in the developing mouse embryo (O'Shea and Dixit, 1988; Corless et al., 1992).

The data presented here identify a fourth member of the thrombospondin gene family. Comparisons of the sequences of *Xenopus* thrombospondin-4 with mouse thrombospondin-3 suggest that they were produced by a gene duplication that occurred before the gene duplication that separated their branch from the thrombospondin-1 and 2 branch (J. Lawler et al., 1993). The sequences suggest that the type-2 and 3 repeats and the COOH-terminal domain are conserved and therefore, serve a function that is common to all of the members of the thrombospondin gene family. By contrast, the NH₂-terminal domain, the regions of homology with procollagen and the type 1 repeats may have tissue-specific functions. These data indicate that thrombospondin-4 has activities in common with thrombospondin-1 and 2, that are mediated by the type 2 and type 3 repeats or the COOH-terminus, and also has unique functions that are mediated by the divergent NH₂-terminal domain. Assuming that thrombospondin-4 is a secreted protein, as the presence of an apparent signal sequence indicates, it is likely that its interac-

tions with the extracellular matrix and with the cell surface will be different from those of thrombospondin-1 and 2. The ability of thrombospondin-1 to interact with other extracellular matrix and with the cell surface components has been located to the 70,000-D chymotryptic fragment that corresponds to the procollagen homology region, and the type 1 and type 2 repeats (Lawler et al., 1986). The type 1 repeats contain consensus heparin-binding sequence motifs and indeed bind heparin (Guo et al., 1992).

For thrombospondin-1, it is known that cell-binding activity involves the amino-terminal heparin-binding domain; the CSVTCG sequence that is located within the type 1 repeats; the RGD sequence that is located within the last type 3 repeat, and the COOH-terminal domain (Asch et al., 1991; Kosfeld et al., 1991; Stomski et al., 1992). These sites are also present in thrombospondin-2. Because *Xenopus* thrombospondin-4 does not contain the CSVTCG or RGD sequences, and possibly does not bind heparin, it is clear that its interactions with cells will be different from those of thrombospondin-1 and 2. Efforts are currently under way to determine if thrombospondin-4 is indeed a cell adhesion molecule. Such studies will yield information regarding the functions of thrombospondin-4 and will also lead to fresh insights concerning the functions of the other members of the thrombospondin gene family.

We wish to thank Dr. Sandra Gendler for providing the GP2M-2 clone that contains the human thrombospondin-3 gene and Dr. Douglas Melton for providing the *Xenopus laevis* stage 22 library. The manuscript was typed by Marybeth Muckian and Pamela Caffey, and edited by Sami Lawler.

J. C. Adams was supported by long-term fellowships from the European Molecular Biology Organization and latterly from the Human Frontiers of Science Program Organization. D. W. DeSimone is supported by a Pew Scholars Award in the Biomedical Science and Junior Faculty Research Award from the American Cancer Society. This work was supported by the National Institutes of Health Research grants HL28749 and HL42443 from the National Heart, Lung and Blood Institute and by HD26402 from the National Institute of Child Health and Human Development.

Received for publication 21 July 1992 and in revised form 13 November 1992.

References

- Asch, A. S., J. Tepler, S. Silbiger, and R. L. Nachman. 1991. Cellular attachment to thrombospondin: cooperative interactions between receptor systems. *J. Biol. Chem.* 266:1740-1745.
- Asch, A. S., S. Silbiger, E. Heimer, and R. L. Nachman. 1992. Thrombospondin sequence motif (CSVTCG) is responsible for CD36 binding. *Biochem. Biophys. Res. Commun.* 182:1208-1217.
- Bisbee, C. A., M. A. Baker, and A. C. Wilson. 1977. Albumin phylogeny for clawed frogs (*Xenopus*). *Science (Wash. DC)*. 195:785-787.
- Bornstein, P., D. Alfi, S. Devarayalu, P. Framson, and P. Li. 1990. Characterization of the mouse thrombospondin gene and evaluation of the role of the first intron in human gene expression. *J. Biol. Chem.* 265:16691-16698.
- Bornstein, P., S. Devarayalu, P. Li, C. M. Distech, and P. Framson. 1991a. A second thrombospondin gene in the mouse is similar in organization to thrombospondin-1 but does not respond to serum. *Proc. Natl. Acad. Sci. USA.* 88:8636-8640.
- Bornstein, P., K. O'Rourke, K. Wikstrom, F. W. Wolf, R. Katz, P. Li, and V. M. Dixit. 1991b. A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome. *J. Biol. Chem.* 266:12821-12824.
- Clezardin, P., M-C. Bourdillon, N. R. Hunter, and J. L. McGregor. 1988. Cell attachment and fibrinogen binding properties of platelet and endothelial cell thrombospondin are not affected by structural differences in the 70 and 18 kDa protease-resistant domain. *FEBS (Fed. Eur. Biochem. Soc.) Lett.* 228:215-218.
- Corless, C. L., A. Mendoza, T. Collins, and J. Lawler. 1992. Colocalization of thrombospondin and syndecan during murine development. *Dev. Dynamics.* 193:346-358.
- DeSimone, D. W., and R. O. Hynes. 1988. *Xenopus laevis* integrins: structural conservation and evolutionary divergence of integrin β subunits. *J. Biol. Chem.* 263:5333-5340.
- DeSimone, D. W., P. A. Norton, and R. O. Hynes. 1992. Identification and characterization of alternatively spliced fibronectin mRNAs expressed in early *Xenopus* embryos. *Dev. Biol.* 149:357-369.
- Dixit, V. M., D. M. Haverstick, K. M. O'Rourke, S. W. Hennessy, G. A. Grant, S. A. Santoro, and W. A. Frazier. 1985. A monoclonal antibody against human thrombospondin inhibits platelet aggregation. *Proc. Natl. Acad. Sci. USA.* 82:3472-3476.
- Galvin, N. J., V. M. Dixit, K. M. O'Rourke, S. A. Santoro, G. A. Grant, and W. A. Frazier. 1985. Mapping of epitopes for monoclonal antibodies against human platelet thrombospondin with electron microscopy and high sensitivity amino acid sequence. *J. Cell Biol.* 101:1434-1441.
- Gartner, T. K., D. A. Walz, M. Aiken, L. Starr-Spires, and M. L. Ogilvie. 1984. Antibodies against a 23 kd heparin-binding fragment of thrombospondin inhibit platelet aggregation. *Biochem. Biophys. Res. Commun.* 124:290-295.
- Guo, N.-H., H. C. Krutzsch, E. Negre, T. Vogel, D. A. Bleke, and D. D. Roberts. 1992. Heparin- and sulfatide-binding peptides from the type 1 repeats of human thrombospondin promote melanoma cell adhesion. *Proc. Natl. Acad. Sci. USA.* 89:3040-3044.
- Hanford, P. A., M. Mayhew, M. Baron, P. R. Winship, I. D. Campbell, and G. G. Brownlee. 1991. Key residues involved in calcium-binding motifs and EGF-like domains. *Nature (Lond.)*. 351:164-167.
- Hubbard, S. C., and R. J. Ivatt. 1981. Synthesis and processing of asparagine-linked oligosaccharides. *Annu. Rev. Biochemistry.* 50:555-583.
- Karczewski, J., K. A. Knudsen, L. Smith, A. Murphy, V. L. Rothman, and G. P. Tuszyński. 1989. The interaction of thrombospondin with platelet glycoprotein GPIIb-IIIa. *J. Biol. Chem.* 264:21322-21326.
- Kosfeld, M. D., T. V. Pavlopoulos, and W. A. Frazier. 1991. Cell attachment activity of the carboxyl-terminal domain of human thrombospondin expressed in *Escherichia coli*. *J. Biol. Chem.* 266:24257-24259.
- Krieg, P. A., S. M. Varnum, W. M. Wormington, and D. A. Melton. 1989. The mRNA encoding elongation factor-1 alpha (EF-1 alpha) is a major transcript at the midblastula transition in *Xenopus*. *Dev. Biol.* 133:93-100.
- Lancaster, C. A., N. Peat, T. Duhig, D. Wilson, J. Taylor-Papadimitriou, and S. J. Gendler. 1990. Structure and expression of the human polymorphic epithelial mucin gene: an expressed VNTR unit. *Biochem. Biophys. Res. Commun.* 173:1019-1029.
- Lawler, J., J. E. Connolly, P. Ferro, and L. H. Derick. 1986. Thrombin and chymotrypsin interactions with thrombospondin. *Ann. NY Acad. Sci.* 485:273-287.
- Lawler, J., L. H. Derick, J. E. Connolly, J. H. Chen, and F. C. Chao. 1985. The structure of human platelet thrombospondin. *J. Biol. Chem.* 260:3762-3774.
- Lawler, J., M. Duquette, and P. Ferro. 1991a. Cloning and sequencing of chicken thrombospondin. *J. Biol. Chem.* 266:8039-8043.
- Lawler, J., M. Duquette, P. Ferro, N. G. Copeland, D. J. Gilbert, and N. A. Jenkins. 1991b. Characterization of the murine thrombospondin gene. *Genomics.* 11:587-600.
- Lawler, J., M. Duquette, L. Urry, K. McHenry, and T. F. Smith. 1993. The evolution of the thrombospondin gene family. *J. Mol. Evol.* In press.
- Lawler, J., P. Ferro, and M. Duquette. 1992. Expression and mutagenesis of thrombospondin. *Biochemistry.* 31:1173-1180.
- Lawler, J., and R. O. Hynes. 1989. An integrin receptor on normal and thrombasthenic platelets that binds thrombospondin. *Blood.* 74:2022-2027.
- Lawler, J., and R. O. Hynes. 1986. The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins. *J. Cell Biol.* 103:1635-1648.
- Lawler, J., and E. Simons. 1983. Cooperative binding of calcium to thrombospondin: The effect of calcium on the circular dichroism and limited tryptic digestion of thrombospondin. *J. Biol. Chem.* 258:12098-12101.
- Lawler, J., R. Weinstein, and R. O. Hynes. 1988. Cell attachment to thrombospondin: The role of arg-gly-asp, calcium, and integrin receptors. *J. Cell Biol.* 107:2351-2361.
- Legrand, C., D. Pidard, P. Beiso, D. Tenza, and L. Edelman. 1991. Interaction of a monoclonal antibody to glycoprotein IV (CD36) with human platelets and its effect on platelet function. *Platelets.* 2:99-105.
- Leung-Hagstjijn, C., A. M. Spence, B. D. Stern, Y. Zhou, M.-W. Su, E. M. Hedgecock, and J. G. Culotti. 1992. Unc-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 domains, guides cell and pioneer axon migrations in *C. elegans*. *Cell.* 71:289-299.
- Neugebauer, K. M., C. J. Emmett, K. A. Venstrom, and L. F. Reichardt. 1991. Vitronectin and thrombospondin promote retinal neurite outgrowth: Developmental regulation and role of integrins. *Neuron.* 6:345-358.
- Newport, J., and M. Kirschner. 1982. A major developmental transition in early *Xenopus* embryos. 1. Characteristics and timing of cellular changes at the med-blastula stage. *Cell.* 30:675-686.
- Nieuwkoop, P., and J. Faber. 1967. Normal Table of *Xenopus laevis* (Daudin). North Holland, Amsterdam. 125-152.
- Ohlin, A. K., G. Landes, P. Bourdon, C. Oppenheimer, R. Wydro, and J. Stenflo. 1988. β -hydroxyaspartic acid in the first epidermal growth factor-like domain of protein C. *J. Biol. Chem.* 263:19240-19248.
- O'Shea, K. S., and V. M. Dixit. 1988. Unique distribution of the extracellular matrix component thrombospondin in the developing mouse embryo. *J. Cell*

- Biol.* 107:2737-2748.
- O'Shea, K. S., L.-H. J. Liu, and V. M. Dixit. 1991. Thrombospondin and a 140 kd fragment promote adhesion and neurite outgrowth from embryonic central peripheral neurons and from PC12 cells. *Neuron*. 7:231-237.
- Prater, C. A., J. Plotkin, D. Jaye, and W. A. Frazier. 1991. The properdin-like type 1 repeats of human thrombospondin contain a cell attachment site. *J. Cell Biol.* 112:1031-1040.
- Przywiecki, C. T., J. E. Stagers, H. G. Ramjit, D. G. Musson, A. M. Stern, C. D. Bennett, and P. A. Friedman. 1987. Occurrence of β -hydroxylated asparagine residues in non-vitamin K-dependent proteins containing epidermal growth factor-like domains. *Proc. Natl. Acad. Sci. USA*. 84:7856-7860.
- Rich, K. A., F. W. George IV, J. L. Law, and W. J. Martin. 1990. Cell-adhesive motif in region II of malarial circumsporozoite protein. *Science (Wash. DC)*. 249:1574-1577.
- Roberts, D. D., D. M. Haverstick, V. M. Dixit, W. A. Frazier, S. A. Santoro, and V. Ginsburg. 1985. The platelet glycoprotein thrombospondin binds specifically to sulfated glycolipids. *J. Biol. Chem.* 260:9405-9411.
- Roberts, D. D., J. A. Sherwood, and V. Ginsburg. 1987. Platelet thrombospondin mediates attachment and spreading of human melanoma cells. *J. Cell Biol.* 104:131-139.
- Sanger, F., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA*. 74:5463-5467.
- Santoro, S. A. 1987. Thrombospondin and the adhesive behavior of platelets. *Semin. Thromb. Hemostasis*. 13:290-297.
- Stenflo, J., A. K. Ohlin, W. G. Owen, and W. J. Schneider. 1988. β -hydroxyaspartic acid or β -hydroxyasparagine in bovine low density lipoprotein receptor and in bovine thrombomodulin. *J. Biol. Chem.* 263:21-24.
- Stomski, F. C., J. Gani, R. C. Bates, and G. F. Burns. 1992. Adhesion to thrombospondin by human embryonic fibroblasts is mediated by multiple receptors and includes a role for glycoprotein 88 (CD36). *Exp. Cell Res.* 198:85-92.
- Sun, X., D. F. Mosher, and A. Rapraeger. 1989. Heparan sulfate-mediated binding of epithelial cell surface proteoglycan to thrombospondin. *J. Biol. Chem.* 264:2885-2889.
- Tuszynski, G. P., J. Karczewski, L. Smith, A. Murphy, V. L. Rothman, and K. A. Knudsen. 1989. The GPIIb-IIIa-like complex may function as a human melanoma cell adhesion receptor for thrombospondin. *Exp. Cell Res.* 182:473-481.
- Tuszynski, G. P., and M. A. Kowalska. 1991. Thrombospondin-induced adhesion of human platelets. *J. Clin. Invest.* 87:1387-1394.
- Tuszynski, G. P., V. L. Rothman, A. H. Deutch, B. K. Hamilton, and J. Eyal. 1992. Biological activities of peptides and peptide analogues derived from common sequences present in thrombospondin, properdin and malarial proteins. *J. Cell Biol.* 116:209-217.
- Tuszynski, G. P., V. Rothman, A. Murphy, K. Siegler, L. Smith, S. Smith, J. Karczewski, and K. A. Knudsen. 1987. Thrombospondin promotes cell-substratum adhesion. *Science (Wash. DC)*. 104:131-139.
- Urry, L. A., J. Ramos, M. Duquette, D. W. DeSimone, and J. Lawler. 1991. Cloning, characterization and expression of thrombospondin in *Xenopus laevis* embryos. *J. Cell Biol.* 115:295a. (Abstr.)
- Varani, J., V. M. Dixit, S. E. G. Fligel, P. E. McKeever, and T. E. Carey. 1986. Thrombospondin-induced attachment and spreading of human squamous carcinoma cells. *Exp. Cell Res.* 167:376-390.
- Varani, J., B. J. Nickoloff, B. L. Riser, R. S. Mitra, K. O'Rourke, and V. M. Dixit. 1988. Thrombospondin-induced adhesion of human keratinocytes. *J. Clin. Invest.* 81:1537-1544.
- Vos, H. L., S. Devarayalu, Y. de Vries, and P. Bornstein. 1992. Thrombospondin 3 (*Thbs3*), a new member of the thrombospondin gene family. *J. Biol. Chem.* 267:12192-12196.
- von Heijne, G. 1986. A new method for predicting signal-sequence cleavage sites. *Nucleic Acid Res.* 14:4683-4690.
- Wolf, F. W., R. L. Eddy, T. B. Shows, and V. M. Dixit. 1990. Structure and chromosomal localization of the human thrombospondin gene. *Genomics*. 6:685-691.