Biophysical Characterization, Including Disulfide Bond Assignments, of the Anti-Angiogenic Type 1 Domains of Human Thrombospondin-1†

Kristin G. Huwiler,‡ Martha M. Vestling,§ Douglas S. Annis,‡ and Deane F. Mosher*,‡

Departments of Medicine and Chemistry, University of Wisconsin, 1300 University Avenue, MSC 4459, Madison, Wisconsin 53706

Received July 16, 2002; Revised Manuscript Received September 13, 2002

ABSTRACT: Thrombospondin-1 (TSP1), a modular secreted glycoprotein, possesses anti-angiogenic activity both in vitro and in vivo. This activity has been localized to the thrombospondin type 1 repeats/domains (TSR). A TSP1 monomer contains three TSRs, each with a hydrophobic cluster with three conserved tryptophans (**W**xx**W**xx**W**), a basic cluster with two conserved arginines (**R**x**R**), and six conserved cysteines. Using the baculovirus system, we expressed TSRs of human TSP1 as either the three domains in tandem (P123) or the third domain alone (P3) and demonstrated that both P123 and P3 at nanomolar concentrations inhibit either basic fibroblast-growth-factor or sphingosine-1-phosphate induced endothelial cell migration. Far-UV circular dichroism (CD) indicated that P123 and P3 have a common global fold that is very similar to properdin, a protein with six TSRs. Near-UV CD and fluorescence quenching studies indicated the conserved tryptophans are in a structured, partially solvent-accessible, positively charged environment. N-terminal sequence and mass spectrometry analysis of trypsin-digested TSRs indicated that the RFK linker sequence between P1 and P2 is readily proteolyzed and the conserved arginines are solvent accessible. By a combination of proteolysis and mass spectrometry, the recombinant TSRs were determined to be fully disulfide bonded with a connectivity of 1−5, 2−6, and 3−4 (cysteines are numbered sequentially from N− to C-terminus). TSRs are found in numerous extracellular proteins. These TSRs share the hydrophobic and basic clusters of the TSP TSRs but some have quite different placement of cysteine residues. We propose a sorting of TSRs into six groups that reconciles our results with information about other TSRs.

Thrombospondin-1 (TSP1)[1] belongs to a family of large oligomeric proteins consisting of multiple modules (*1*). Analysis of the number and type of modules indicates there are two sub-families: TSP1 and TSP2 form the first branch; TSP3, TSP4, and TSP5/COMP (cartilage oligomeric matrix protein) form the second branch (*1*). TSP1 and TSP2 contain one procollagen and three type 1 domains (TSR) that are missing from the other TSPs. Numerous studies have shown that TSP1 and TSP2 can serve as anti-angiogenic agents by their ability to specifically inhibit endothelial cell proliferation, migration, and tubulogenesis in vitro and to inhibit new blood vessel growth in vivo (*2*). In addition, overexpression of TSP1 in several carcinomas decreases the in vivo growth of these tumors and decreases the vascularization of these tumors (*3*−*6*). As a result, TSP1 is recognized as a potent and relevant angiogenesis inhibitor.

Attempts to localize the anti-angiogenic region(s) of TSP1 using in vitro as well as in vivo assays have implicated the TSRs of TSP1 (*2*). Specifically, with regard to an ability to inhibit stimulant-induced endothelial cell (EC) migration, synthetic peptides based on TSP1 TSRs and containing the three conserved tryptophans (Figure 1A,B, shaded) or the conserved basic residues (Figure 1A,B, outlined) were found to be active (*7*−*10*). How well the peptides mimic the activity of native TSP-1, however, is open to question. With the exception of peptides bearing D-amino acids, the concentrations of these synthetic peptides required for anti-migratory activity were between 6 and 30 μM. TSP1 from human platelets and the three TSRs in tandem as well as the second TSR of TSP1 when expressed in S2 insect cells inhibit EC migration in the low nanomolar range (*8, 9, 11*−*13*).

TSP TSRs are found alone and in multiple arrays within numerous extracellular receptors and secreted proteins in diverse species, from flies and round worms to humans (*14*). TSRs within proteins other than TSP1, such as brain-

† K.G.H. was supported by the U.S. Army Medical Research and Materiel Command under DAMD-17-96-6151. This work was supported by funds from the National Institute of Health Grant HL54462 and utilized the University of Wisconsin−Madison Biophysics Instrumentation Facility, which is supported by National Science Foundation Grant BIR-9512577.

* To whom correspondence should be addressed. Tel.: 608 262 1576. Fax: 608 263 4969. E-mail: dfmosher@facstaff.wisc.edu

‡ Department of Medicine, University of Wisconsin.

§ Department of Chemistry, University of Wisconsin.

[1] Abbreviations: TSP1, thrombospondin-1; bFGF, basic fibroblast growth factor; BME, β-mercaptoethanol; **C**^{1}, **C**^{2}, **C**^{3}, **C**^{4}, **C**^{5}, and **C**^{6}, six conserved cysteines of TSRs numbered as in Experimental Procedures; CD, circular dichroism; DTT, dithiothreitol; EC, endothelial cell; ESI-MS, electrospray ionization mass spectrometry; Fuc, fucose; Glc, glucose; GuHCl, guanidine hydrochloride; HB-GAM, heparin-binding growth associated molecule; LC, liquid chromatography; MALDI-TOF-MS, matrix-assisted laser desorption ionization time-of-fight mass spectrometry; P123, the baculovirally expressed first, second, and third type 1 domain of human TSP1; P3, the baculovirally expressed third type 1 domain of human TSP1; S1P, shingosine-1-phosphate; TSR, thrombospondin type 1 repeat/domain (historically this domain has been referred to as a TSR).

10.1021/bi026463u CCC: $22.00
Published on Web 11/05/2002

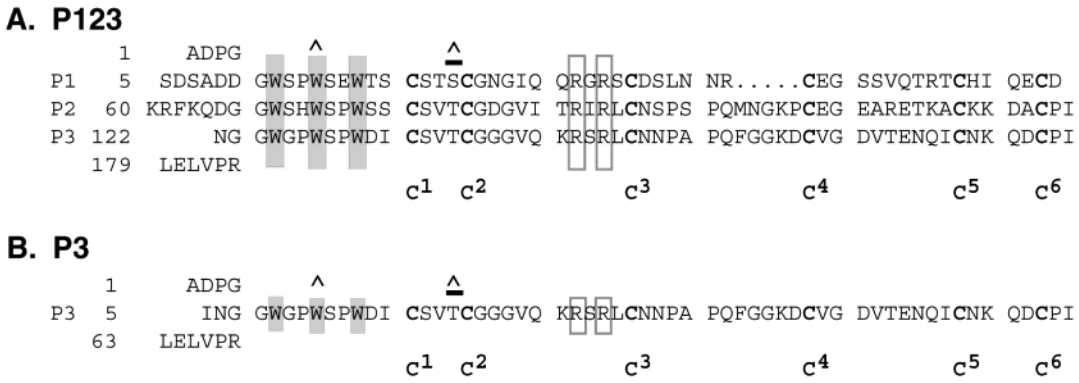


FIGURE 1: Recombinant human TSP1 TSRs, P123, and P3: the amino acid sequences of the recombinant TSRs P123 (A) and P3 (B) are shown. The sequences of the three TSRs of P123 are aligned according to the conserved cysteines (bold) **C**^{1}, **C**^{2}, **C**^{3}, **C**^{4}, **C**^{5}, **C**^{6}. The N- and C-terminal sequences (ADPG and LELVPR, respectively) are shown for the secreted proteins following thrombin cleavage to remove the C-terminal His tag. The three conserved Trps in each TSR are shaded, while the two conserved basic residues of TSP1 TSR are outlined. Sites of C-mannosylation (∧) and O-linked glycosylation (≙) are marked (*20*).

angiogenic inhibitor and ADAMTS-1 (METH-1), have been reported to possess anti-angiogenic activity (*15*−*16*). However, no single function is attributable to all TSRs.

A decade ago, Smith and colleagues aligned 31 TSRs (*17*). On the basis of the conservation of the six cysteines, the authors concluded that the disulfide bond connectivity of a TSR is **C**^{1}−**C**^{3}, **C**^{2}−**C**^{6}, and **C**^{4}−**C**^{5} (*17*), where cysteines are numbered consecutively from N- to C-terminus based on the group 1 TSRs (see Figure 1A,B). Subsequently, an alternative disulfide bond connectivity of **C**^{1}−**C**^{5}, **C**^{2}−**C**^{6}, and **C**^{3}−**C**^{4} was proposed (*18*). Recently, 2 Cys → Gly mutations within TSRs of the ADAMTS-13 zinc metalloproteinase, in addition to mutations in other domains of ADAMTS-13, were linked to the familial form of the life-threatening blood disorder thrombotic thrombocytopenic purpura (*19*).

TSRs are targets of unique carbohydrate modifications (*20*). Each of the TSRs of TSP1 contains three conserved tryptophans in the form of **W**xx**W**xx**W** (Figure 1A,B, shaded). The **W**xx**W** is a recognition sequence for C-linked glycosylation (*21*) in which the first tryptophan in the motif becomes mannosylated in an enzyme-catalyzed reaction involving dolichyl-phosphate-mannose (*22*). The central tryptophan in the **W**xx**W**xx**W** sequence is mannosylated in the TSRs of platelet TSP1 and baculovirally expressed TSP1 TSRs (Figure 1A,B) (*20*). C-mannosylation of TSRs has also been observed in the complement components C6, C7, C8α, C8*β*, and C9 and properdin (*20*, *23*, *24*). The sequence CSV-(S/T)CG of the TSRs of platelet TSP1 carries the O-linked disaccharide Glc-Fuc-O-Ser/Thr (*20*).

The purpose of the present study was to characterize highly conserved essential features of TSRs, namely, the conserved tryptophans, arginines, and cysteines. First, to ensure that our recombinant TSRs were correctly folded and functionally active, we examined their far-UV CD characteristics and their ability to inhibit stimulant induced endothelial cell migration. The environment of the conserved tryptophans was probed by near-UV CD and fluorescence spectroscopy. The accessibility of a pair of arginines that are conserved among TSRs was assessed by susceptibility to trypsinization. Finally, the state of the cysteines and disulfide connectivity were investigated by mass spectrometry. The biophysical characterization and disulfide bond assignments facilitate the study of the role of TSRs in angiogenesis and enhance the understanding of the structure and function of TSRs in general.

EXPERIMENTAL PROCEDURES

Reagents. The following were purchased from Sigma (St. Louis, MO): *β*-mercaptoethanol (BME), $CaCl_2$, CsCl ultragrade, dithiothreitol (DTT), imidazole, KI, K_2PO_4, *N*-acetyl-tryptophan-amide, phenyl-methylsulfonyl fluoride, sodium dodecyl sulfate (SDS), NaCl, L-1-tosylamido-2-phenylethyl chloromethyl ketone treated trypsin, and trizma. Pefabloc SC was obtained from Roche. Ultrapure acrylamide and guanidine hydrochloride (GuHCl) were purchased from ICN (Costa Mesa, CA). TSP1 was isolated from fresh human platelets by heparin agarose affinity chromatography followed by gel filtration, as described previously (*25*).

Baculovirus Transfer Vector and Generation of Recombinant Baculoviruses. The construction of the pCOCO baculovirus transfer vector has been described previously (*26*, *27*). The DNA sequences encoding the first, second, and third type 1 domains in tandem (P123) or only the single third type 1 domain (P3) were amplified by the polymerase chain reaction (PCR) from human TSP1 cDNA. The primers used to amplify the P3 sequence are 5′-TAT**CCCGGG**ATCAATGGAGGCTGGGGTCCTTGG-3′ (P3 forward) and 5′-GGG**TCTAGA**ATTGGACAGTCCTGCTTGTTGC-3′ (P123 reverse). The primers used to amplify the P123 sequence are 5′-TCC**CCCGGG**AGCGACTCTGCGGACGATGG-3′ (P123 forward) and P123 reverse. The boldface sequences represent XmaI and XbaI restriction sites. Amplified DNA for P3 encoded for amino acids Ile 491 through Ile 548 of mature TSP1, while the DNA for P123 encoded for amino acids Ser 375 through Ile 548. The PCR products were cloned into the XmaI and XbaI sites in the multiple cloning site of pAcGP67.COCO to give pAcGP67P3.COCO and pAcGP-67P123.COCO. The DNA sequences were verified by sequencing prior to the generation of high titer ($>10^8$ pfu/mL) recombinant baculoviruses, as described previously (*26*).

Expression and Purification of the Type 1 Domains (P3 and P123). The expression of recombinant proteins using the pCOCO system has been described elsewhere (*26*, *27*). In brief, High Five (*Trichoplusia ni* Hi5) or Sf9 cells (Invitrogen) were grown in suspension at 27 °C with SF900

II (Gibco BRL) serum-free medium and infected at a density of 1×10^6 cells/mL. A multiplicity of infection (MOI) of 2−5 was routinely used, and the infection was allowed to proceed for ∼60−70 h. The processing of the conditioned media containing P3.COCO or P123.COCO, prior to purification using NiNTA resin (Qiagen), differ due to the level of expression of these two proteins (1−2 mg/L and >10 mg/L, respectively); these methods have been described elsewhere (*27*). Similarly the steps involving NiNTA resin incubation, washing, and elution are published (*27*). Eluate fractions were analyzed by SDS-PAGE, the appropriate fractions dialyzed, and the Histidine-tag (His-tag) removed with biotinylated-thrombin, as described previously (*26*, *27*). The following nomenclature for the expressed recombinant proteins is invoked. The removal of the suffix COCO indicates that the recombinant proteins have had the C-terminal His-tag removed by thrombin cleavage. The inclusion of Hi5 or Sf9 following P123 or P3 indicates the insect cells that were used to express the recombinant protein. Amino acid residues are numbered beginning with the first residue of the mature secreted recombinant protein that lacks the signal sequence.

Extinction Coefficient Determination. Calculated extinction coefficients (ϵ_{calc}) were determined for P3 and P123 by the method of Mach and co-workers (*28*). The values of ϵ_{calc} for P123 and P3 are 5.1×10^4 and 1.7×10^4 $M^{-1}cm^{-1}$, respectively. Experimentally determined values of the extinction coefficient for the P3 Hi5 and P3 Sf9 proteins were also determined by quantitative amino acid analysis. UV wavelength scans (240−340 nm) were obtained prior to hydrolysis and amino acid analysis (Medical College of Wisconsin's Shared Protein and Nucleic Acid Facility). The average extinction coefficient was determined using Beer's law, and the average concentration of each sample was determined by content analysis of 14 different amino acids. The experimentally determined values of the extinction coefficient (ϵ_{exp}) for P3 expressed in Hi5 and Sf9 cells are $1.7 \pm 0.2 \times 10^4$ and $1.8 \pm 0.1 \times 10^4$ $M^{-1}cm^{-1}$, respectively. The results indicated that P3 ϵ_{calc} agrees with the experimentally determined values. Thus, ϵ_{calc} values for P123 and P3 were used to determine the concentration based on the absorbance at 280 nm.

Molecular Mass Determination and Disulfide Bond Content. Molecular masses of P123 and P3 were analyzed by either electrospray-ionization mass spectrometry (ESI-MS) or matrix-assisted laser desorption ionization time-of-fight mass spectrometry (MALDI-TOF MS). For MALDI-TOF MS analysis, the matrix was α-cyano-4-hydroxycinnamic acid, and internal calibrants of insulin and ubiquitin were used to determine the masses. These data were collected on a Bruker ReflexII (Billerica, MA) MALDI-TOF MS equipped with delayed extraction, reflectron, and a 337 nm laser (University of Wisconsin−Madison, Dept of Chemistry) in positive linear mode. The determination of disulfide bond content of P3 and P123 were accomplished following a published procedure (*29*) and used iodoacetic acid (IA) to label the cysteines. Data were collected either by MALDI-TOF MS data, as described above, or by ESI-MS on an Applied Biosystems (Foster, CA) MDS Sciex API 365 LC/MS/MS triple quadrupole (University of Wisconsin-Madison, Biotechnology Center).

Fluorescence Quantum Yields and Quenching. Quantum yields for P123 Hi5 and P3 Hi5 were determined at 295 nm using *N*-acetyl-tryptophan-amide prepared in TBS as the reference compound, as described previously (*30*). Increasing concentrations of either acrylamide (0−140 mM), KI (0−200 mM), or CsCl (0−200 mM) were used to quench the fluorescence resulting from excitation at 295 nm at 25 °C. Emission scans (300−450 nm) were collected and analyzed using the Stern−Volmer equation, as described previously (*26*). Each Stern−Volmer constant (K_{SV}) was determined from two to five experiments.

Circular Dichroism (CD). An Aviv 62 ADS circular dichroism spectrophotometer (University of Wisconsin−Madison, Biophysics Instrumentation Facility) was used. For far-UV CD, P123 Hi5 and P3 Hi5 were dialyzed against 10 mM potassium phosphate, 100 mM NaCl, pH 7.5, and data collected in 0.1 cm quartz cuvette, while samples for near-UV CD were dialyzed against 10 mM Tris-Cl, 150 mM NaCl, pH 7.5, and used 1 cm cuvettes. Time constants of 5 or 10 s were used. Three scans (at 25 °C) were averaged for each experiment and baseline corrected, using spectra of the last dialysis buffer. The far-UV CD data in millidegrees was converted to molar ellipticity mean residue weight, $[\Theta]_{MRW}$. For near-UV CD, the data in millidegrees were first converted to molar ellipticity [Θ] with units of degrees·cm^2·$dmol^{-1}$ (*31*) and subsequently to units of [Θ] per TSR, by normalizing [Θ] for the number of TSRs present in P123 or P3. The following conversion was used, 1 mol TSR is equivalent to 0.33 mol P123 and 1 mol P3. The spectra were smoothed using Igor Pro software from Wavemetrics.

Endothelial Cell Culture and Migration Assays. Fetal bovine heart endothelial (FBHE) cells and bovine pulomonary artery endothelial (BPAE) cells were obtained from American Type Culture Collection and VEC Technologies (Rensselaer, NY), respectively. The FBHEs were maintained at 37 °C, 8% CO_2 in Dulbecco's modified Eagle's (DME) media supplemented with 10% fetal bovine serum, 1X penicillin-streptomycin (PS), and 25 ng/mL basic fibroblast growth factor (bFGF). The BPAEs were maintained at 37 °C, 8% CO_2 in DME media supplemented with 20% fetal bovine serum, PS. For the wounding assay, FBHEs were seeded near confluence in 24-well tissue culture plates using DMEM, 0.5% calf serum, PS. The next day, the cells were wounded using a sterile plastic pipet tip. Following wounding, the media was replaced with DME media, 0.25% calf serum, plus 5 ng/mL bFGF without (positive control) or with 10 nM TSP1, 30 nM P3 Hi5, or 30 nM P123 Hi5. Approximately 24 h after wounding, the cells were washed, fixed, and then stained with Diff Quik (Dade AG), according to manufacturer's instructions. The number of cells that migrated into the wound was determined by summing five random fields per wound. Endothelial cell migration was also determined using a 48-well micro-chemotaxis chamber (Neuroprobe) with 5.0 μm pore size polycarbonate nucleopore filters (Corning), as described previously by Tolsma et al. (*8*). Differences included the following: filters were coated with 20 ug/mL collagen I, the cells used were BPAEs, and the stimulant used was 1 μM sphingosine-1-phosphate (S1P). The concentration ranges tested were from 0.4 to 10 nM for TSP1 and 1 to 750 nM for both P3 and P123. For both the wounding and chemotaxis chamber assays, all samples were tested in triplicate within each experiment and

each experiment was repeated 2−4 times. For all migration experiments, to average the results from interday experiments, the data were normalized so that 100% represents the number of cells that migrated in response to the stimulant (bFGF or S1P) minus the number cells that migrated in the absence of stimulant (bFGF or S1P), while 0% represents the number cells that migrated in the absence of stimulant (bFGF or S1P) (*8*).

Analysis of Trypsin Digested P123. P123 Hi5 (20 μM) was digested with increasing amounts of TPCK-treated trypsin in 25 mM Tris-Cl, 150 mM NaCl, 2.5 mM $CaCl_2$, pH 7.8 at 37 °C for 1 h. The trypsin:P123 ratios used were between 0% and 4% (w/w). The digests were analyzed using SDS-PAGE, N-terminal sequencing, and mass spectrometry. For N-terminal sequencing, 1% (w/w) TPCK-trypsin digest of P123 was separated under nonreducing conditions on 14% SDS-PAGE, transferred to PVDF, and stained with 0.1% Amido Black. The samples underwent >10 cycles of N-terminal sequencing in the lab of Johan Stenflo, Lund University, Sweden. For tandem LC-ESI-MS analysis, a 2% (w/w) TPCK-trypsin digest of P123 was used.

Analysis of Thermolysin Digested P3. P3 Sf9 (0.5 mg/mL) was digested with 2% (w/w) thermolysin in 25 mM MOPS, 150 mM NaCl, 2.5 mM $CaCl_2$, pH 6.5 at 45 °C for 2 h. The reaction was stopped by addition of EDTA to a final concentration of 25 mM. The digest was immediately analyzed on 20% SDS-PAGE and separated by HPLC using a Vydac C_{18} SN910401 column (4.6 × 250 mm). The A and B solvents were 0.1% trifluoroacetic acid (TFA) and 0.088% TFA, 90% acetonitrile, respectively. The following linear gradient with flow of 1 mL/min was used: 0−5 min, 0% B; 5−75 min, 0−70% B; 75−95 min, 70−100% B. Detection was at 215 and 280 nm. HPLC fractions were analyzed by MALDI-TOF-MS in the presence or absence of 10 mM DTT using positive linear and reflectron modes. Several matrixes were utilized, including α-cyano-4-hydroxycinnamic acid, 2-(4-hydroxyphenylazo)-benzoic acid, and 2,5-dihydroxybenzoic acid (Sigma). External calibrants used included human adrenocorticotropic hormone fragment 18−39 and human angiotensin II (Sigma). Certain HPLC fractions resulting from separation of P3-thermolysin digest were subjected to trypsinization in 25 mM MOPS, 150 mM NaCl, 2.5 mM $CaCl_2$, pH 6.7 at 37°C for 3 h. The digests were analyzed by MALDI-TOF-MS in the presence or absence of 10 mM DTT, as described above.

RESULTS

Expression and General Characterization of P3 and P123. P123 and P3 (Figure 1A,B) were expressed and accumulated in the insect cell conditioned media at yields of approximately 10 and 1 mg/L, respectively. N-terminal sequencing verified that the GP67 signal sequence is cleaved at the expected location, such that each construct began with the sequence ADPG (Figure 1A,B). The expected average masses of P123 and P3 taking into account the presence of disulfides bonds are 19945 and 7250 Da, respectively. P123 and P3 were shown to be pure by SDS-PAGE (data not shown). MALDI-TOF MS revealed five masses for P3; the smallest mass observed was 7250 Da (data not shown). The larger masses are due to differential C-mannosylation and O-linked glycosylation (*20*) (Figure 1B, sites marked with ∧ and ⩓). ESI-MS of P123 revealed a series of masses. The smallest was equal to the expected mass of 19945 Da (data not shown), while the larger masses result from differential C- and O-linked glycosylation of each TSR of P123 (*20*) (Figure 1A, sites marked with ∧ and ⩓).

Table 1: Determination of Disulfide Bond Content of P3 and P123[a]

protein	treatment	mass of first peak (Da)	Δmass (Da)	Δmass/ mass CM
P3	GuHCl	7252	1	<0.1
	DTT	7597	346	6.0
	DTT + GuHCl	7599	347	6.0
P123	GuHCl	19947	1	<0.1
	DTT	20988	1042	18.0
	DTT + GuHCl	20991	1045	18.0

[a] P3 or P123 were labeled for 30 min with iodoacetic acid (IA) in the presence of guanidine hydrochloride (GuHCl), dithiotreitol (DTT), or DTT plus GuHCl and subsequently were analyzed by either MALDI-TOF-MS or ESI-MS, as described in the Experimental Procedures. P3 and P123 are heterogeneous glycoproteins (*20*). The masses listed in this table are for the nonglycosylated form determined by MS for each of the samples. ΔMass refers to the change in mass of each of these samples relative to the theoretical mass of either P3 or P123. On the basis of the presence of three disulfide bonds in P3 and 9 disulfide bonds in P123, the theoretical average masses $[M+H]^+_{AV}$ for P3 and P123 are 7251 and 19946 Da, respectively. Mass CM is the mass of a carboxymethyl (CM) group (58 Da).

The observed masses for P3 and P123 are consistent with complete disulfide bonding. Additional MS analyses of P123 and P3 were performed to verify the number of disulfide bonds and the absence of free thiols. No significant mass shifts were observed for P123 or P3 after labeling with iodoacetic acid under the denaturing condition of GuHCl (Table 1), indicating the absence of free thiols. Increases in the masses of P123 and P3 were observed after labeling under the reducing conditions of DTT, or DTT plus GuHCl (Table 1). Both of these reducing conditions resulted in the labeling of 18 cysteines for P123 and six cysteines for P3. Taken together, the data indicates that P123 and P3 lack free thiols and are fully disulfide bonded.

We examined the functional ability of our recombinant TSRs to inhibit migration of endothelial cells utilizing a wounding assay and micro-chemotaxis chamber assay. By the wounding assay, endothelial cells (EC) treated with bFGF plus 10 nM TSP1 had an 11 ± 13% migratory ability relative to bFGF alone (100%). ECs treated with bFGF plus 30 nM P3 demonstrated 25 ± 14% migration relative to bFGF alone, while 30 nM P123 plus bFGF had 28 ± 8% migration relative to bFGF alone. Sphingosine-1-phosphate (S1P) has been shown to be a potent inducer of EC migration (*32*). Using the micro-chemotaxis chamber assay, we determined TSP1 to inhibit in a dose-dependent manner S1P-induced EC migration. In the presence of S1P plus 0.4−10 nM TSP1, migration was found to be approximately 88% to 36% of that induced by S1P alone. Both P3 and P123 inhibited S1P-induced EC migration at all concentrations tested (1−750 nM). In the presence of S1P plus 1−750 nM P3, migration was found to be approximately 71% to 43% of that induced by S1P alone, while for S1P plus 1−750 nM P123, migration was approximately 64% to 50% of that induced by S1P alone. Taken together, these data indicate that the recombinant TSP1 TSRs expressed in the baculovirus system inhibit bFGF-induced and S1P-induced EC migration.

Table 2: Absorbance and Fluorescence Properties of P123, P3, and the Tryptophan Analogue *N*-Acetyl-tryptophan-amide

protein	ϵ_{280} x 10^{-4} (M^{-1} cm^{-1})	λmax (nm)	quantum yield	K_{SV} (M^{-1}) acrylamide	KI	CsCl
P123	5.1[b]	333	0.028 ± 0.001	3.2 ± 0.2	1.3 ± 0.1	0.1 ± 0.1
P3	1.7 ± 0.2	333	0.018 ± 0.002	2.3 ± 0.1	ND	ND
N-acetyl-tryptophan-amide		355	0.13[c]	24.5 ± 0.8	11.9 ± 0.2	2.1 ± 0.1

[a] Extinction coefficients (ϵ_{280}) were determined as described in the Experimental Procedures. Using an excitation wavelength of 295 nm, steady-state fluorescence emission spectra were collected and used to determine emission maximas (λmax), quantum yields, and Stern−Volmer constants. [b] Calculated extinction coefficient using the equation of Mach et al. (*28*) [c] The quantum yield of *N*-acetyl-tryptophan-amide in TBS was obtained from Isaacs et al. (*30*).

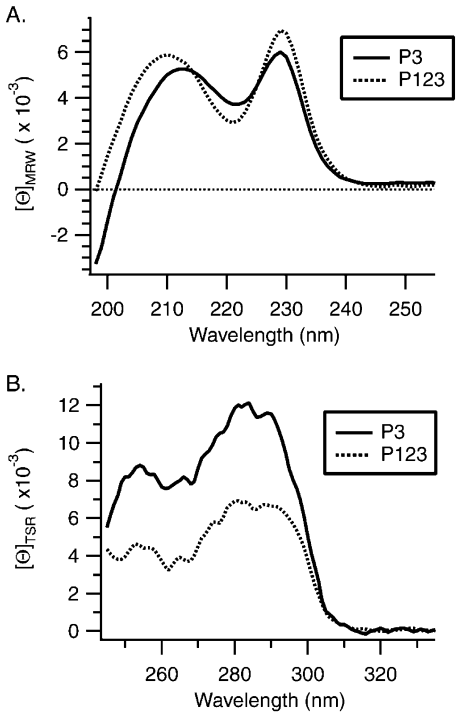


FIGURE 2: Circular dichroism of P123 and P3 at 25 °C: far-UV (A) and near-UV (B) CD spectra of P123 and P3. Each spectrum represents the average obtained from two independent experiments.

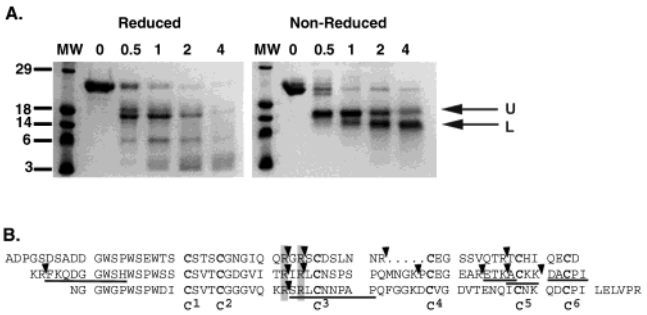


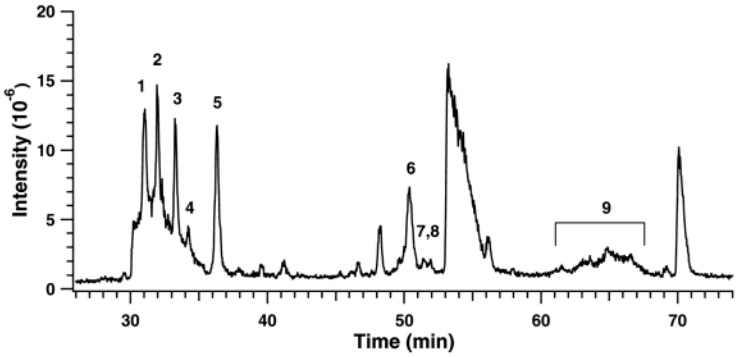
FIGURE 3: Trypsin sensitivity of P123: (A) P123 containing either 0%, 0.5%, 1%, 2%, or 4% trypsin (w/w) was digested at 37 °C as described in the Experimental Procedures and subsequently separated on 14% SDS-PAGE under reducing (+BME) or nonreducing (−BME) conditions. The gels were stained with Gel-Code Blue (Pierce, Rockford, IL). MW is the molecular weight markers. The arrows indicate the upper (U) and (L) bands from a 1% trypsin digest that were subjected to N-terminal sequencing. (B) The amino acid sequence of recombinant P123. The N-termini identified from sequencing of the upper (U) and lower (L) bands are underlined. Arrowheads mark the sites of trypsin cleavage identified by LC-ESI-MS analysis of trypsin digested P123, described in Figure 4.

The secondary structures of the recombinant TSRs were compared by far-UV CD. The spectra are dominated by positive ellipticity (Figure 2A). A sharp peak at 229 nm occurs for P3 and P123; a broader peak is maximal at 212−213 nm for P3 and 210 nm for P123. The spectra are similar to those published for properdin (maximums at 231 and 212 nm), which was purified from human serum (*33*) and thus indicate the TSRs are folded correctly. Since positive far-UV elliptical signals have been noted for proteins with predominantly β-sheet and β-turn secondary structure that contain aromatic clusters (*34*), our data are compatible with the presence of these elements in the TSRs of TSP1.

Environment of Conserved Tryptophans. The only tryptophans in each TSR are those clustered together in the sequence **W**xx**W**xx**W** (Figure 1). The environment of these tryptophans was examined by near-UV CD. Since P3 and P123 lack tyrosines, any ellipticity above 276 nm is due to the conserved tryptophans. The near-UV CD spectra of P3 and P123 at 25 °C are characterized by two very broad positive signals, beginning at 310 nm and continuing through 240 nm (Figure 2B). Integration of the near-UV CD spectra (240−340 nm) in the form of molar ellipticity [Θ] per TSR revealed that the intensity of P123 relative to P3 is 61% ± 7%. In addition, [Θ] per TSR for P123 at 292 nm is 61% of the value of [Θ] per TSR for P3 at 292 nm. These results indicate that the tryptophans are in a structured environment that varies somewhat among TSRs.

Fluorescence spectroscopy was also performed to examine the local environment of the conserved tryptophans. The wavelength of maximum emission for P123 and P3 was 333 nm (Table 2) and indicates neither an exclusively nonpolar nor polar environment. The quantum yield for P3 was ∼64% of the value obtained for P123 (Table 2), indicating that the three tryptophans in P3 are on the average slightly more quenched than the nine tryptophans in P123. Fluorescence quenching experiments using neutral, positively charged, and negatively charged quenchers were performed. Acrylamide and iodide were both effective quenchers of P123 fluorescence, with acrylamide being the most effective with a quenching constant of 3.2 M^{-1} (Table 2). Cesium had a quenching constant of 0.1 M^{-1}, which was less than that obtained with the control salt, NaCl (data not shown). The data indicate that the local environment of the conserved tryptophans within the TSRs of P123 possess a positive charge density. The quenching constant for acrylamide for P3 was 2.3 M^{-1}; this value is ∼72% of the value obtained for P123 (Table 2). The data indicate that the three tryptophans in P3 are on the average more densely packed than the nine tryptophans in P123.

Trypsin-Sensitive Sites. Increasing concentrations of trypsin at 37°C cleaved P123 progressively into several fragments that migrated as well-resolved bands on reducing SDS-PAGE (Figure 3A). Trypsinized P123 at each step in the digestion



HPLC Peak	Peptides		TSR Module (Cys #)	Predicted $[M]_{AV}$ Reduced	Predicted $[M]_{AV}$ with Disulfides (# S-S bonds)	Observed $[M]_{AV}$ Non-Reduced
1	$S_{35} \rightarrow R_{42}$ $C_{43} \rightarrow R_{51}$	*	P1 (C^3) P1 (C^4)	908.0 966.0	1872.0 (1)	1872.1
2	$S_{35} \rightarrow R_{51}$	*	P1 (C^3, C^4)	1856.0	1854.0 (1)	1854.4
3	$L_{91} \rightarrow R_{109}$	*	P2 (C^3, C^4)	2018.3	2016.3 (1)	2015.9
4	$I_{89} \rightarrow K_{102}$ $P_{103} \rightarrow R_{109}$	*	P2 (C^3) P2 (C^4)	1544.8 760.8	2303.7 (1)	2303.9
5	$I_{89} \rightarrow R_{109}$	*	P2 (C^3, C^4)	2287.6	2285.6 (1)	2285.9
6	$A_1 \rightarrow R_{32}$ $T_{52} \rightarrow R_{61}$		P1 (C^1, C^2) P1 (C^5, C^6)	3388.5 1232.4	4616.9 (2)	4617.2
7, 8	$A_1 \rightarrow R_{32}$ $T_{52} \rightarrow R_{61}$	Δ	P1 (C^1, C^2) P1 (C^5, C^6)	3550.5 1232.4	4800.9 (2) [M+Δ+Na]	4801.0 [M+Δ+Na]
9	$F_{62} \rightarrow R_{88}$ $A_{113} \rightarrow K_{116}$ $D_{117} \rightarrow R_{145}$ $S_{146} \rightarrow R_{184}$		P2 (C^1, C^2) P2 (C^5) P2 (C^6), P3 (C^1, C^2) P3 (C^3, C^4, C^5, C^6)	2984.3 448.6 3047.5 4272.9	10743.2 (5)	10743.8
	$F_{62} \rightarrow R_{88}$ $A_{113} \rightarrow K_{116}$ $D_{117} \rightarrow R_{145}$ $S_{146} \rightarrow R_{184}$	Δ	P2 (C^1, C^2) P2 (C^5) P2 (C^6), P3 (C^1, C^2) P3 (C^3, C^4, C^5, C^6)	2984.3 448.6 3047.5 4272.9	10927.2 (5) [M+Δ+Na]	10926.5 [M+Δ+Na]
	$F_{62} \rightarrow R_{88}$ $A_{113} \rightarrow K_{116}$ $D_{117} \rightarrow R_{145}$ $S_{146} \rightarrow R_{184}$	ΔΔ	P2 (C^1, C^2) P2 (C^5) P2 (C^6), P3 (C^1, C^2) P3 (C^3, C^4, C^5, C^6)	2984.3 448.6 3047.5 4272.9	11111.2 (5) [M+ Δ+Δ +2Na]	11111.1 [M+ Δ+Δ+2Na]

FIGURE 4: Disulfide-bonded peptides resulting from Trypsinization of P123: the digest was analyzed by LC-ESI-MS, and the nine peaks within the total ion current (TIC) trace that correspond to disulfide-bonded peptide fragments of P123 are labeled. All experimentally determined masses are within 0.03% of the predicted masses and the peptides identified cover A_1-R_{32}, $S_{35}-R_{109}$, $A_{113}-R_{184}$. Abbreviations: TSR, the first (P1), second (P2), or third (P3) type 1 domain of P123 from which the peptide derives; Cys #, the 6 conserved cysteines within a given TSR (**C^1**, **C^2**, **C^3**, **C^4**, **C^5**, **C^6**) that are found within the peptide; # S−S bonds, one-half the number of cysteines present in the peptide; $[M]_{AV}$, average isotopic mass; *, disulfide-bonded peptides from the first TSR (P1) containing **C^3-C^4**; , disulfide-bonded peptides from the second TSR (P2) containing **C^3-C^4** disulfide bond; Δ, +162 Da modification (consistent with Mannosyl-tryptophan).

migrated as two dominate bands, upper (U) and lower (L), under nonreducing conditions (Figure 3A). By Edman degradation, the upper band revealed a single sequence of F_{62}KQDG (Figure 3B), indicating cleavage at R61 in the KRFK linker sequence between P1 and P2. N-terminal sequencing of the lower band revealed a minimum of five N-termini that are underlined in the sequence shown in Figure 3B. On the basis of our LC-ESI-MS (Figure 4, discussed below) and Edman degradation of trypsin digested P123, we identified 12 protease-sensitive sites (Figure 3B, marked with ▼). These data suggest the most sensitive site was R61 within the P1−P2 interdomain linker (KRFK), followed by selected sites within each TSR.

Determination of Disulfide Bond Connectivity. To examine the disulfide bond connectivity of the six cysteines within a single TSR, P3 was digested with thermolysin at a pH value that minimizes the possibility for disulfide scrambling. A temperature study revealed increased susceptibility of P3 to thermolysin cleavage at temperatures of ≥45 °C (data not shown). A large-scale thermolysin digest of P3 at 45 °C was

Table 3: Peptides Identified by MALDI-TOF-MS Following C_{18} HPLC Separation of Thermolysin Digested P3[a]

peptides		Cys #	HPLC peak	predicted $[M+H]^+_{mono}$ non-reduced	observed $[M+H]^+_{mono}$ non-reduced	predicted $[M+H]^+_{mono}$ reduced	observed $[M+H]^+_{mono}$ reduced
I_5NGGWGPWSPWD$_{16}$		−	13	1371.6	1371.2	1371.6	1371.3
I_5NGGWGP**W**SPWD$_{16}$		−	12	1533.6	1533.3	1533.6	1533.3
L_{32}CNNPAPQFGGKDCVGDVTENQ$_{53}$	*	**C³, C⁴**	9−11	2304.0	2303.7	2306.0	2306.1
L_{32}CNNPAPQ$_{39}$ F_{40}GGKDCVGDVTENQ$_{53}$	*	**C³** **C⁴**	4	2322.0	2322.1	856.4 1468.6	856.0 1468.6
L_{32}CNNPAPQ$_{39}$ F_{40}GGKDCVGD$_{48}$	*	**C³** **C⁴**	3	1750.8	1750.7	856.4 897.4	856.2 897.2
L_{32}CNNPAPQ$_{39}$ F_{40}GGKDC$_{45}$	*	**C³** **C⁴**	2	1479.6	1479.4	856.4 626.2	856.1 626.0
I_{17}CSVTCGGGVQKR$_{29}$ I_{54}CNKQDCPILE$_{64}$		**C¹, C²** **C⁵, C⁶**	11	2578.3	2578.1	1307.7 1275.6	1307.5 1275.4
I_{17}CSV$\underline{T}$CGGGVQKR$_{29}$ I_{54}CNK$\underline{Q}$DCPILE$_{64}$		**C¹, C²** **C⁵, C⁶**	9−10	2724.3	2723.7	1453.7 1275.6	1453.5 1275.4
I_{17}CSV$\underline{T}$CGGGVQKR$_{29}$ I_{54}CNK$\underline{\underline{Q}}$DCPILE$_{64}$		**C¹, C²** **C⁵, C⁶**	8−9	2886.2	2886.2	1615.6 1275.6	1615.7 1275.5
I_{17}CS$_{19}$ V_{20}TCGGGVQKR$_{29}$ I_{54}CNKQDCPILE$_{64}$		**C¹** **C²** **C⁵, C⁶**	7	2596.3	2596.2	322.1 1004.5 1275.6	n/o 1004.4 1275.5
I_{17}CS$_{19}$ $V_{20}\underline{T}$CGGGVQKR$_{29}$ I_{54}CNKQDCPILE$_{64}$		**C¹** **C²** **C⁵, C⁶**	6	2742.3	2742.0	322.1 1150.5 1275.6	n/o 1150.4 1275.6
I_{17}CS$_{19}$ $V_{20}\underline{T}$CGGGVQKR$_{29}$ $I_{54}\underline{\underline{C}}$NKQDCPILE$_{64}$		**C¹** **C²** **C⁵, C⁶**	5	2904.2	2904.1	322.1 1312.5 1275.6	n/o 1312.7 1275.6
L_{65}VPR$_{68}$		−	1	484.3	484.2	484.3	484.1

[a] Coverage: $I_5-R_{29}L_{32}-R_{68}$. All experimentally determined masses are within 0.05% of the predicted masses. Abbreviations: Cys #, the six conserved cysteines within a given TSR (**C¹**, **C²**, **C³**, **C⁴**, **C⁵**, **C⁶**) that are found within the peptide, as numbered in Figure 1; **W**, mannosyltryptophan; $\underline{T}$, Fuc-O-Thr; $\underline{\underline{T}}$, Glc-Fuc-O-Thr; n/o, not observed. *, disulfide-bonded peptides that contain the **C³**−**C⁴** disulfide bond.

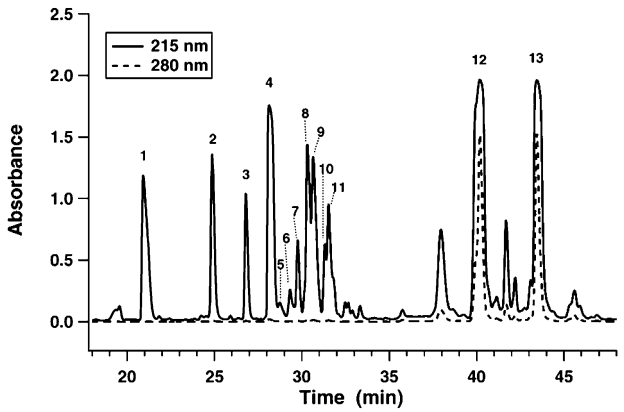


FIGURE 5: C_{18} HPLC chromatogram of thermolysin digested P3: P3 was digested with 2% (w/w) thermolysin and separated on a C_{18} column, as described in the Experimental Procedures. All HPLC fractions were analyzed by MALDI-TOF-MS, and the mass identities of the labeled peaks are presented in Table 3.

separated by reverse-phase HPLC to yield the chromatograph shown in Figure 5. All fractions were analyzed by MALDI-TOF-MS without and with reduction with DTT, which allowed identification of peptides held together by disulfide bonds. Peptides that span 62 out of the 68 amino acids of P3 were found (Table 3). A linkage between **C³** (Cys_{33}) and **C⁴** (Cys_{45}) of P3 was observed within four separate disulfide-bonded peptides: $L_{32} \rightarrow Q_{53}$; $L_{32} \rightarrow Q_{39} + F_{40} \rightarrow Q_{53}$; $L_{32} \rightarrow Q_{39} + F_{40} \rightarrow D_{48}$; and $L_{32} \rightarrow Q_{39} + F_{40} \rightarrow C_{45}$ (Table 3, marked with *).

To determine if disulfide bonds between **C³** and **C⁴** also occur within the first (P1) and second (P2) TSRs, P123 was subjected to trypsinization followed by LC-ESI-MS, as described in the Experimental Procedures. Tryptic peptides accounting for 179 of the 184 amino acids of P123 were found. These peptides span all but two short segments, one two and the other three amino acids long. Disulfide bonds between **C³** and **C⁴** within P1 and P2 were deduced on the basis of the observation of masses consistent with five separate disulfide-bonded peptides: $S_{35} \rightarrow R_{51}$; $S_{35} \rightarrow R_{42} + C_{43} \rightarrow R_{51}$; $I_{89} \rightarrow R_{109}$; $L_{91} \rightarrow R_{109}$; and $I_{89} \rightarrow K_{102} + P_{103} \rightarrow R_{109}$ (Figure 4, marked with $\underline{*}$ and $\underline{\underline{*}}$). Peptides containing only the **C³**−**C⁴** disulfide bond of P3 were not found in the tryptic digest of P123. This result is expected because Arg or Lys is not present in the P3 sequence between **C⁴** and **C⁵** (Figure 1B). Taken together, MS analyses of the P3-thermolysin cleavage products and P123-tryptic cleavage products provide strong evidence for the **C³**−**C⁴** disulfide linkage in the first, second, and third TSRs of TSP1.

Evidence for disulfide linkages of peptides containing **C¹** and **C²** with peptides bearing **C⁵** and **C⁶** was obtained by MALDI-TOF-MS analysis of thermolysin digested P3. Masses consistent with disulfide linkage of noncontiguous peptides ($I_{17} \rightarrow R_{29} + I_{54} \rightarrow E_{64}$) via **C¹** ($Cys_{18}$) and **C²** ($Cys_{22}$) in one peptide and **C⁵** (Cys_{55}) and **C⁶** (Cys_{60}) in another peptide were found in HPLC peaks 8−11 (Figure 5). HPLC peak 11 contained the unglycosylated disulfide-bonded peptide ($I_{17} \rightarrow R_{29} + I_{54} \rightarrow E_{64}$), while HPLC peaks

Table 4: Disulfide-Bonded Peptides Resulting from Trypsinization of HPLC Peaks 5, 7, and 8 of Thermolysin-Cleaved P3[a]

HPLC peak	disulfide-bonded peptides resulting from trypsin digestion		Cys #	$[M+H]^+_{mono}$ predicted non-reduced	$[M+H]^+_{mono}$ observed non-reduced	$[M+H]^+_{mono}$ predicted reduced	$[M+H]^+_{mono}$ observed + DTT
7	$I_{17}CS_{19}$		**C¹**	2440.1	2440.4	322.1	n/o
	$V_{20}TCGGGVQK_{28}$		**C²**			848.4	848.3
	$I_{54}CNKQDCPILE_{64}$		**C⁵, C⁶**			1275.6	1275.5
5	$I_{17}CS_{19}$		**C¹**	2748.1	2747.8	322.1	n/o
	$V_{20}\underline{T}CGGGVQK_{28}$		**C²**			1156.4	1156.0
	$I_{54}C\underline{N}KQDCPILE_{64}$		**C⁵, C⁶**			1275.6	1275.4
8	$I_{17}CSV\underline{T}CGGGVQK_{28}$		**C¹, C²**	2730.2	2729.9	1459.5	1459.4
	$I_{54}CNK\underline{Q}DCPILE_{64}$		**C⁵, C⁶**			1275.6	1275.3
7	$V_{20}TCGGGVQKR_{29}$	◇	**C²**	1818.9	1818.9	1004.5	1004.3
	$Q_{58}DCPILE_{64}$		**C⁶**			817.4	817.1
5	$V_{20}TCGGGVQKR_{29}$	◇	**C²**	2126.9	2126.6	1312.5	1312.6
	$Q_{58}\underline{D}CPILE_{64}$		**C⁶**			817.4	816.8
8	$I_{17}CSV\underline{T}CGGGVQKR_{29}$		**C¹, C²**	2904.3	2904.0	1615.6	1615.6
	$I_{54}CNK_{57}$		**C⁵**			477.2	477.1
	$Q_{58}DCPILE_{64}$		**C⁶**			817.4	817.1
7	$V_{20}TCGGGVQK_{28}$	◇	**C²**	1662.8	1662.7	848.4	848.1
	$Q_{58}DCPILE_{64}$		**C⁶**			817.4	817.1
5	$V_{20}TCGGGVQK_{28}$	◇	**C²**	1970.8	1970.4	1156.4	1156.0
	$Q_{58}\underline{D}CPILE_{64}$		**C⁶**			817.4	816.8
8	$I_{17}CSV\underline{T}CGGGVQK_{28}$		**C¹, C²**	2748.2	2747.8	1459.5	1459.4
	$I_{54}CNK_{57}$		**C⁵**			477.2	477.1
	$Q_{58}DCPILE_{64}$		**C⁶**			817.4	817.1
5 & 7	$I_{17}CS_{19}$	•	**C¹**	796.3	796.2	322.1	n/o
	$I_{54}CNK_{57}$		**C⁵**			477.2	n/o

[a] Abbreviations: Cys #, the six conserved cysteines (as numbered in Figure 1) that are found within the peptide; T, Glc-Fuc-O-Thr; n/o, Not observed. All experimentally determined masses are within 0.05% of the predicted masses. ◇, disulfide-bonded peptides that contain the C^2-C^6 disulfide bond. •, disulfide-bonded peptides that contain the C^1-C^5 disulfide bond.

8−10 contained the glyco-forms bearing either the monosaccharide Fuc-O-Thr or the disaccharide Glc-Fuc-O-Thr (Table 3). These masses are consistent with the O-linked glycosylation previously reported for the threonine within the sequence CSVTCG of P3 expressed in the baculovirus system (*20*). In addition, masses consistent with disulfide linkage of three peptides ($I_{17} \rightarrow S_{19} + V_{20} \rightarrow R_{29} + I_{54} \rightarrow E_{64}$) via **C¹** ($Cys_{18}$), **C²** ($Cys_{22}$), **C⁵** ($Cys_{55}$), and **C⁶** ($Cys_{60}$) were found in HPLC peaks 5−7 (Table 3). The glycosylation present on the disulfide-bonded tripeptide determined the elution time from the HPLC, such that the unglycosylated form eluted later than either the Fuc-O-Thr or Glc-Fuc-O-Thr modified forms (Figure 5, Table 3).

Evidence for disulfide linkages of peptides containing **C¹** and **C²** with peptides bearing **C⁵** and **C⁶** in P1 and P2 was also obtained by LC-ESI-MS analysis of trypsin digested P123 (Figure 4). Peak 6 was determined to contain a mass of 4617 Da that is consistent with disulfide linkage of noncontiguous peptides of the first TSR ($A_1 \rightarrow R_{32} + T_{52} \rightarrow R_{61}$) via **C¹** and **C²** in the first peptide and **C⁵** and **C⁶** in the second peptide (Figure 4). The glycosylated form of the 4617 Da fragment was observed in peaks 7 and 8 (Figure 4); C-mannosylation of the central Trp in the first TSR of P123 has been previously identified (*20*). The area labeled 9 (Figure 4) is diffuse and contains several masses that are consistent with disulfide bonded peptides originating from the second (P2) and third (P3) TSRs without and with C-mannosylation. An observed mass of 10743 Da is consistent with the complete disulfide bonding between four peptides $F_{62} \rightarrow R_{88}$, $A_{113} \rightarrow K_{116}$, $D_{117} \rightarrow R_{145}$, and $S_{146} \rightarrow R_{184}$ (Figure 4). Two additional peptides with masses of 10 927 and 11 111 Da were identified in peak 9, and these are likely C-mannosylated glyco-forms of the 10 743 Da peptide (Figure 4). In support of the identity of the ($F_{62} \rightarrow R_{88} + A_{113} \rightarrow K_{116} + D_{117} \rightarrow R_{145} + S_{146} \rightarrow R_{184}$) disulfide-bonded tetrapeptide, we identified the N-terminal sequence of each of these four peptides by Edman degradation of an apparent 12 kDa fragment of trypsinized P123 purified by SDS-PAGE under nonreducing conditions, as described above (Figure 3).

To determine the disulfide bond connectivity between cysteines **C¹**, **C²**, **C⁵**, and **C⁶**, HPLC peaks 5−7 from thermolysin-digested P3, which was cleaved between **C¹** and **C²**, were subjected to trypsinization to cleave between **C⁵** and **C⁶**. These digests were then analyzed by MALDI-TOF-MS analysis in the absence and presence of reducing agent (Table 4). A disulfide bond between **C²** and **C⁶** was deduced on the basis of four separate linked sets of peptides: $V_{20} \rightarrow R_{29} + Q_{58} \rightarrow E_{64}$; glyco-form of $V_{20} \rightarrow R_{29} + Q_{58} \rightarrow E_{64}$; $V_{20} \rightarrow K_{28} + Q_{58} \rightarrow E_{64}$; glyco-form of $V_{20} \rightarrow K_{28} + Q_{58} \rightarrow E_{64}$ (Table 4, marked with ◇). A disulfide bond between **C¹** and **C⁵** was deduced within the linked peptides: $I_{17} \rightarrow S_{19} + I_{54} \rightarrow K_{57}$ (Table 4, marked with •). The glycosylated dipeptide in peak 8 ($I_{17} \rightarrow R_{29} + I_{54} \rightarrow E_{64}$), which spans the same sequence as the tripeptide ($I_{17} \rightarrow S_{19} + V_{20} \rightarrow R_{29} + I_{54} \rightarrow E_{64}$) but lacks cleavage between **C¹** and **C²**, was cleaved at the two trypsin sites, but the four cysteines remained linked (Table 4). Thus, P3 peptides modified by the various permutations of cleavages at thermolysin and trypsin sites provide strong evidence for the 1−5 and 2−6 disulfide linkages in third TSR of TSP1.

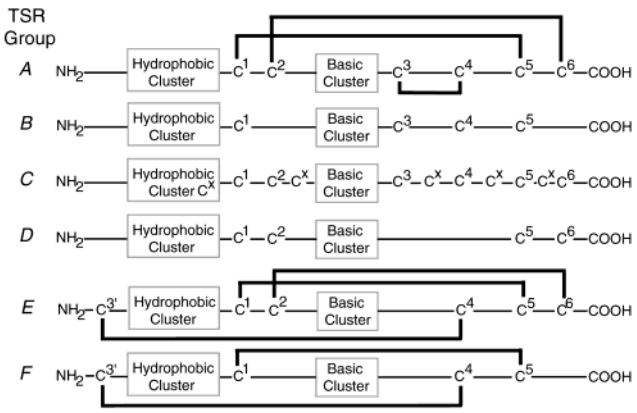


FIGURE 6: Classification and representative alignments of TSR groups A−F: Schematic of the six proposed groupings (A−F) of TSRs. NH_2 and COOH represent the amino and carboxy termini of each TSR. The hydrophobic cluster refers to the tryptophan-rich sequence while the basic cluster refers to the conserved arginine/lysine residues. The cysteines are numbered as described in the methods. The displacement of $\mathbf{C^3}$ to the amino terminus is denoted $\mathbf{C^{3'}}$. In Group *C* TSRs, the presence of cysteines that are different from the six core cysteines are denoted as $\mathbf{C^X}$. The $\mathbf{C^X}$ are placed at all the locations where additional cysteines have been observed in the primary sequence of group *C* TSRs. Thus, the diagram is a composite of locations of $\mathbf{C^X}$ rather than the location of $\mathbf{C^X}$ in a given group C TSR. The dark lines connecting cysteines represent the disulfide bonds that were determined biochemically for group *A* TSP1 TSRs (this paper) and groups E and F HB-GAM TSRs (*44*).

Taken together with the analyses of peptides containing $\mathbf{C^3}$ and $\mathbf{C^4}$, the MS data support a disulfide connectivity of 1−5, 2−6, 3−4 for the TSRs of TSP1.

DISCUSSION

Disulfide Bonds. We determined by mass spectrometry that the cysteines of the recombinant TSP1 TSRs have 1−5, 2−6, and 3−4 connectivity (Figure 6). This study is the first biochemical determination of the disulfide bond connectivity for a group 1 TSR, as defined by Smith et al. (*17*). Since the classification by Smith et al., more TSRs have emerged that do not fit this classification, such as the C-terminal TSRs of midkine and HB-GAM. To facilitate discussion and put our results in context, we performed sequence analysis of TSRs in the databases, including those of midkine and HB-GAM, that lack $\mathbf{C^3}$ but have the $\mathbf{C^{3'}}$ cysteine not present in the TSRs of TSP1 (Figure 6). We extended the work of Smith et al. (*17*) by classifying TSRs into six groups based on number and position of the conserved cysteines (Figure 6 and Supporting Information). The conserved hydrophobic residues and basic residues are denoted as the hydrophobic and basic cluster, respectively, and are positioned relative to the conserved cysteines. Our experimentally determined 1−5, 2−6, 3−4 connectivity is consistent with the absence of $\mathbf{C^3}$ and $\mathbf{C^4}$ in Group D TSRs and of $\mathbf{C^2}$ and $\mathbf{C^6}$ in Groups B and F TSRs (Supporting Information). The TSR family seems remarkable in that it can accommodate variations in the overall placement of cysteines with the only disulfide pair conserved in all TSRs being 1−5 (Figure 6). Perhaps most remarkable is that the global fold of a TSR must accommodate $\mathbf{C^4}$ in two different disulfide connectivities: 1−5, 2−6, 3−4 and 1−5, 2−6, 3′−4.

Conserved Tryptophans and Conserved Arginines. In each TSR of TSP1, the three tryptophans of the hydrophobic cluster are separated by two amino acids. Although this pattern is noted in many TSRs, there is considerable variability in the number, type, and spacing of hydrophobic residues within the cluster (see Supporting Information). This paper is the first examination of the structure of these residues by near-UV CD and fluorescence spectroscopy. The three tryptophans per TSP1 TSR are in a structured environment and the structure was enhanced for P3 relative to P123 (Figure 2). One explanation for the observed CD differences lies in the increased number of proline residues, which are conformationally limiting, within the variable region of the hydrophobic cluster of P3 relative to P1 and P2 (Figure 1). A second explanation for the difference in near-UV CD intensity is the absence of five residues from P1 that are present in P2 and P3 (Figure 1), which if involved in stabilizing the three tryptophans per TSR could lead to a decreased structure for P123 relative to P3. The enhanced structure of the hydrophobic cluster region of P3 was also associated with lower quantum yields and lower acrylamide quenching constants, relative to P123 (Table 2). These data indicate that variable packing densities exist for the hydrophobic cluster residues within TSRs in solution.

Fluorescence quenching experiments revealed the tryptophans are solvent accessible and in a positively charged environment. Two arginines separated by one amino acid are referred to as a basic cluster (Figure 6). In TSR groups A−F (see Supporting Information), the conservation of an arginine or a lysine at either of these sites of the basic cluster was greater than 70%. We found that these conserved arginines of TSP1 TSRs are accessible to trypsin at 37 °C. It is possible that within many TSRs, these conserved basic residues form a tertiary structure with the hydrophobic cluster residues, as within members of the cytokine receptor superfamily, including the prolactin and erythropoietin receptors (*35*−*37*). These proteins contain a WSxW-like motif that is part of a β-bulge secondary structure and forms a positively charged tertiary structure via amino-aromatic stacking interactions with conserved arginines and lysines (*35*−*37*). This sort of stacking of conserved tryptophans and arginines in TSP1 TSRs could account for the observed chirality and positive charge density, determined by near-UV CD and by fluorescence quenching, respectively. In addition, the stacking interaction among the conserved tryptophans and conserved arginines could reconcile previous findings of weak anti-angiogenic activity in all L-amino acid synthetic peptides, which contain either the three conserved tryptophans or arginines of the TSP1 TSRs (*7*−*10*); these synthetic peptides required micromolar concentrations for activity.

Trypsinization of P123 revealed a preferential interdomain cleavage (KRFK) between P1 and P2. Long linker sequences between P1 and P2 of mouse TSP2 (*38*) and between the TSRs of midkine (*39*) have also been found to be easily proteolyzed. Synthetic peptides containing the KRFK sequence of TSP1 bind and activate latent transforming growth factor-β (TGF-β) (*40, 41*). Our data indicate that this sequence should be readily accessible in the native protein for binding latent TGF-β.

TSRs and Angiogenesis. The purified recombinant proteins of P123 and P3, like platelet TSP1 and recombinant P123

and P2 from S2 cells (*13*), inhibited bFGF-induced EC migration at nanomolar concentrations. In addition, P123 and P3, like platelet TSP1, inhibit S1P-induced EC migration. The fact that either the second (*13*) or third TSR of TSP1 can inhibit bFGF-induced EC migration suggests that such inhibition is a general feature of TSRs from TSP1. On the basis of the data presented here, those published previously (*13*, *20*) concerning inhibition of EC migration(*20*), and the different glycosylation observed on recombinant TSRs (*20*), the common feature in P2 and P3 must not involve C-mannosylation or the glucose extension of Fuc-O−Ser/Thr within the **C**^{1}SV(S/T)**C**^{2}G sequence. The **C**^{1}SV(S/T)-**C**^{2}G sequence has been implicated in binding the receptor CD36 (*9*, *42*), which has been shown to mediate the in vitro inhibitory effects of TSP1 on stimulant-induced endothelial cell migration (*9*) as well as the in vivo inhibition by TSP1 of stimulant-induced neovascularization (*43*). Because the cysteines of **C**^{1}SV(S/T)**C**^{2}G are not disulfide bonded to each other and since this sequence bears the O-linked glycosylation, it is likely the CSVTCG-like peptides are presented to cell surface receptor(s) very differently from the native disulfide-bonded TSRs.

In the future, coordinated sets of synthetic peptides and mutagenized TSRs may be used to examine the TSR residues and tertiary structure responsible for its in vitro and in vivo anti-angiogenic activity. One problem with mutagenesis studies is deciphering whether any observed alterations in biological activity result from local changes induced by the mutagenesis or global structural changes. Of particular interest would be the role of the conserved tryptophans, arginines, and their tertiary structure(s) in the anti-angiogenic function of the TSRs. The information within this paper can serve as a solid basis for monitoring perturbations in TSR mutants.

ACKNOWLEDGMENT

We are grateful to Dr. Tina Misenheimer and Richele Watkins for critical review of the manuscript.

SUPPORTING INFORMATION AVAILABLE

The methods used for grouping and aligning 160 TSRs amino acid sequences is described. A brief description of the six TSR groups (A−F) determined and the sequence alignments for each are provided. This material is available free of charge via the Internet at http://pubs.acs.org.

CHANGES AFTER ASAP POSTING

This paper was first published on the Web on November 5, 2002, with errors in Table 3. The correct version appeared on the Web on November 26, 2002.

REFERENCES

1. Bornstein, P., and Sage, E. H. (1994) *Methods Enzymol. 245*, 62−85.
2. Huwiler, K. G., and Mosher, D. F. (2000) *Jpn. J. Thromb. Hemost.*
3. Weinstat-Saslow, D. L., Zabrenetzky, V. S., VanHoutte, K., Frazier, W. A., Roberts, D. D., and Steeg, P. S. (1994) *Cancer Res. 54*, 6504−6511.
4. Bleuel, K., Popp, S., Fusenig, N. E., Stanbridge, E. J., and Boukamp, P. (1999) *Proc. Natl. Acad. Sci. U.S.A. 96*, 2065−2070.
5. Castle, V. P., Dixit, V. M., and Polverini, P. J. (1997) *Lab. Invest. 77*, 51−61.
6. Streit, M., Velasco, P., Brown, L. F., Skobe, M., Richard, L., Riccardi, L., Lawler, J., and Detmar, M. (1999) *Am. J. Pathol. 155*, 441−452.
7. Vogel, T., Guo, N.-H., Krutzsch, H. C., Blake, D. A., Hartman, J., Mendelovitz, S., Panet, A., and Roberts, D. D. (1993) *J. Cell Biochem. 53*, 74−84.
8. Tolsma, S. S., Volpert, O. V., Good, D. J., Frazier, W. A., Polverini, P. J., and Bouck, N. (1993) *J. Cell Biol. 122*, 497−511.
9. Dawson, D. W., Pearce, S. F. A., Zhong, R., Silverstein, R. L., Frazier, W. A., and Bouck, N. P. (1997) *J. Cell Biol. 138*, 707−717.
10. Dawson, D. W., Volpert, O. V., Pearce, S. F., Schneider, A. J., Silverstein, R. L., Henkin, J., and Bouck, N. P. (1999) *Mol. Pharmacol. 55*, 332−338.
11. Good, D. J., Polverini, P. J., Rastinejad, F., Le Beau, M. M., Lemons, R. S., Frazier, W. A., and Bouck, N. P. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 6624−6628.
12. Taraboletti, G., Roberts, D., Liotta, L. A., and Giavazzi, R. (1990) *J. Cell Biol. 111*, 765−772.
13. Miao, W. M., Seng, W. L., Duquette, M., Lawler, P., Laus, C., and Lawler, J. (2001) *Cancer Res. 61*, 7830−7839.
14. Adams, J. C., and Tucker, R. P. (2000) *Dev. Dynam. 218*, 280−299.
15. Nishimori, H., Shiratsuchi, T., Urano, T., Kimura, Y., Kiyono, K., Tatsumi, K., Yoshida, S., Ono, M., Kuwano, M., Nakamura, Y., and Tokino, T. (1997) *Oncogene 15*, 2145−2150.
16. Iruela-Arispe, M. L., Vazquez, F., and Ortega, M. A. (1999) *Ann. N.Y. Acad. Sci. 886*, 58−66.
17. Smith, K. F., Nolan, K. F., Reid, K. B. M., and Perkins, S. J. (1991) *Biochemistry 30*, 8000−8008.
18. Leung-Hagesteijn, C., Spence, A. M., Stern, B. D., Zhou, Y., Su, M.-W., Hedgecock, E. M., and Culotti, J. G. (1992) *Cell 71*, 289−299.
19. Levy, G. G., Nichols, W. C., Lian, E. C., Foroud, T., McClintick, J. N., McGee, B. M., Yang, A. Y., Slemieniak, D. R., Stark, K. R., Gruppo, R., Sarode, R., Shurn, S. B., Chandrasekaran, V., Stabler, S. P., Sabio, H., Bouhassira, E. E., Upshaw, J. D., Jr., Ginsburg, D., and Tsai, H. M. (2001) *Nature 413*, 488−494.
20. Hofsteenge, J., Huwiler, K. G., Macek, B., Hess, D., Lawler, J., Mosher, D. F., and Peter-Katalinic, J. (2001) *J. Biol. Chem. 276*, 6485−6498.
21. Krieg, J., Hartmann, S., Vicentini, A., Glasner, W., Hess, D., and Hofsteenge, J. (1998) *Mol. Biol. Cell 9*, 301−309.
22. Doucey, M. A., Hess, D., Cacan, R., and Hofsteenge, J. (1998) *Mol. Biol. Cell 9*, 291−300.
23. Hofsteenge, J., Blommers, M., Hess, D., Furmanek, A., and Miroshnichenko, O. (1999) *J. Biol. Chem. 274*, 32786−32794.
24. Hartmann, S., and Hofsteenge, J. (2000) *J. Biol. Chem. 275*, 28569−28574.
25. Sun, X., Skorstengaard, K., and Mosher, D. F. (1992) *J. Cell Biol. 118*, 693−701.
26. Misenheimer, T. M., Huwiler, K. G., Annis, D. S., and Mosher, D. F. (2000) *J. Biol. Chem. 275*, 40938−40945.
27. Mosher, D. F., Huwiler, K. G., Misenheimer, T. M., and Annis, D. S. (2002) *Methods Cell Biol. 69*, 69−81.
28. Mach, H., Middaugh, C. R., and Lewis, R. V. (1992) *Anal. Biochem. 200*, 74−80.
29. Wada, Y. (1996) *Methods Mol. Biol. 61*, 101−113.
30. Isaacs, B. S., Brew, S. A., and Ingham, K. C. (1989) *Biochemistry 28*, 842−850.
31. Pain, R. (1996) Determining the CD spectrum of a protein, in *Current Protocols in Protein Science* (Coligan, J. E., et al., Eds.), Chapter 7.6, John Wiley & Sons, Inc., New York.
32. Panetti, T. S., Nowlen, J. K., and Mosher, D. F. (2000) *Arterioscler. Thromb. Vasc. Biol. 20*, 1013−1019.
33. Smith, C. A., Pangburn, M. K., Vogel, C.-W., and Muller-Eberhard, H. J. (1984) *J. Biol. Chem. 259*, 4582−4588.
34. Woody, R. W. (1994) *Eur. Biophys. J. 23*, 253−262.
35. Livnah, O., Stura, E. A., Johnson, D. L., Middleton, S. A., Mulcahy, L. S., Wrighton, N. C., Dower, W. J., Jolliffe, L. K., and Wilson, I. A. (1996) *Science 273*, 464−471.
36. Kossiakoff, A. A., Somers, W., Ultsch, M., Andow, K., Muller, Y. A., and De Vos, A. M. (1994) *Protein Sci. 3*, 1697−1705.
37. Somers, W., Ultsch, M., De Vos, A. M., and Kossiakoff, A. A. (1994) *Nature 372*, 478−481.

38. Chen, H., Sottile, J., O'Rourke, K. M., Dixit, V. M., and Mosher, D. F. (1994) *J. Biol. Chem. 269*, 32226−32232.
39. Matsuda, Y., Talukder, A. H., Ishihara, M., Hara, S., Yoshida, K., Muramatsu, T., and Kaneda, N. (1996) *Biochem. Biophys. Res. Comm. 228*, 176−181.
40. Schultz-Cherry, S., Chen, H., Mosher, D. F., Misenheimer, T. M., Krutzsch, H. C., Roberts, D. D., and Murphy-Ullrich, J. E. (1995) *J. Biol. Chem. 270*, 7304−7310.
41. Ribeiro, S. M., Poczatek, M., Schultz-Cherry, S., Villain, M., and Murphy-Ullrich, J. E. (1999) *J. Biol. Chem. 274*, 13586−13593.
42. Li, W. X., Howard, R. J., and Leung, L. L. K. (1993) *J. Biol. Chem. 268*, 16179−16184.
43. Jimenez, B., Volpert, O. V., Crawford, S. E., Febbraio, M., Silverstein, R. L., and Bouck, N. (2000) *Nat. Med. 6*, 41−48.
44. Fabri, L., Nice, E. C., Ward, L. D., Maruta, H., Burgess, A. W., and Simpson, R. J. (1992) *Biochem. Int. 28*, 1−9.

BI026463U