Assignment of the Four Disulfides in the N-terminal Somatomedin B Domain of Native Vitronectin Isolated from Human Plasma

by

Nancy A. Horn*, Gregory B. Hurst†, Anand Mayasundari*, Neil A. Whittemore*, Engin H. Serpersu* and Cynthia B. Peterson*

*Department of Biochemistry, Cellular, and Molecular Biology, and the Center of Excellence in Structural Biology
University of Tennessee, Knoxville, Tennessee 37996
and
†the Chemical Sciences Division,
Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831

This research was supported by Grants HL50676 from the Heart, Lung and Blood Institute at the National Institutes of Health (CBP) and MCB 01110741 from the National Science Foundation (EHS). AM was also supported by a Postdoctoral Fellowship, Ref. No. 0120344B, from the American Heart Association, Southeast Regional Affiliate.

*To whom correspondence should be addressed:
Department of Biochemistry, Cellular and Molecular Biology
M407 Walters Life Sciences Building
University of Tennessee
Knoxville, TN 37996

Phone: (865) 974-4083
FAX: (865)974-6306
E-mail: cynthia_peterson@utk.edu

Running Title: Disulfides in the Vitronectin SMB Domain

Keywords: Vitronectin, Somatomedin B, Mass Spectrometry, Peptide Mapping, NMR, Cystine-Stabilized Helix
ABSTRACT

The primary sequence of the N-terminal somatomedin B (SMB) domain of native vitronectin (vitronectin) contains 44 amino acids, including a framework of four disulfide bonds formed by 8 closely spaced cysteines in sequence patterns similar to those found in the cystine knot family of proteins. The SMB domain of vitronectin was isolated by digesting the protein with endoproteinase Glu-C and purifying the N-terminal 1-55 peptide by reverse-phase HPLC. Through a combination of techniques, including stepwise reduction and alkylation at acidic pH, peptide mapping with MALDI-MS and NMR, the disulfide bonds contained in the SMB domain have been determined to be C5:C9, C19:C31, C21:C32 and C25:C39. This pattern of disulfides differs from two other connectivities that have been reported previously for recombinant forms of the SMB domain expressed in E. coli. This arrangement of disulfide bonds in the SMB domain from native vitronectin forms a rigid core around the C19:C31 and C21:C32 disulfides. A small positively charged loop is created at the N-terminus by the C5:C9 cystine. The most prominent feature of this disulfide-bonding pattern is a loop between C25 and C39 similar to cystine-stabilized $\alpha$-helical structures commonly observed in cystine knots. This $\alpha$-helix has been confirmed in the solution structure determined for this domain using NMR (Mayasundari, A., Whittemore, N. A., Serpersu, E. H., and Peterson, C. B. (2004) J. Biol. Chem., in press). It confers function on the SMB domain, comprising the site for binding to plasminogen activator inhibitor type-1 and the urokinase receptor.
INTRODUCTION

Vitronectin is a large glycoprotein with wide-ranging distribution and function. The hallmark feature of vitronectin structure is a series of distinct functional domains that allow it to interact both with itself and with a number of other ligands in a variety of environments including the circulation, the extracellular matrix, and platelets (1-4). Of particular interest is the N-terminal somatomedin B (SMB) domain of vitronectin. This domain contains the high-affinity binding site for the serpin PAI-1. The interaction between PAI-1 and vitronectin is important to the function of both proteins in thrombolysis, cell adhesion and pericellular proteolysis (1,5-9). Equally important for the adhesive properties of vitronectin are binding sites for cell-surface receptors, including integrins and uPAR, that are housed within this small N-terminal domain (2,10-15).

Since it is known that the SMB domain provides a high-affinity binding site for PAI-1 and that this interaction stabilizes PAI-1 in its physiologically active form, the structure of the SMB domain of vitronectin has been hotly pursued. Computational predictions for the structure of this domain using a threading algorithm were challenging compared to other domains from vitronectin, (16), as there were no reported structures at the time for homologues of this small domain containing four disulfides. Nevertheless, there are over 100 homologues in the sequence database, suggesting that this folding motif has been conserved in evolution. Only recently have three-dimensional structures describing the SMB domain become available from two different approaches. First, an x-ray structure has been reported on a co-crystal of PAI-1 and a recombinant form of the SMB domain expressed in E. coli (17). Subsequently, we completed the determination of a solution structure for the SMB domain purified from circulating vitronectin that was isolated from human plasma (18). Although the two structures differ in overall fold, they share a common feature, a single α-helix that
contains key amino acids for PAI-1 binding. Not surprisingly, a recent report using NMR on a recombinant SMB arrived at a similar structure to that observed in the co-crystal with PAI-1 (19).

Key to understanding the structure of the SMB domain is defining the correct disulfide connectivities of the eight cysteines that exist \textit{in vivo}. The magnitude of the problem can be appreciated by comparison with classic work on ribonuclease, in which over 100 possibilities exist for unique disulfide-bonding patterns for 8 sulfhydryls, and improper oxidation of disulfides during refolding yields an ensemble of misfolded enzymes exhibiting less than 1% activity (20,21). In terms of primary sequence, the SMB region is similar to a growing class of proteins known to form cystine knots. (22-24). However, unlike typical examples of the growth factor cystine knot or the inhibitory cystine knot proteins, the SMB domain of vitronectin has four disulfide bridges rather than the more familiar three-disulfide arrangement. In this respect, the SMB domain shares sequence features with some members of cystine knot families that block ion channels, have anti-microbial activity and the bone morphogenic protein agonists (24-27).

Previous efforts to understand the structure of the SMB domain and, specifically, to define the disulfide pairings in this domain have relied upon recombinant surrogates for the native protein (17,19,28). We chose to isolate the N-terminus of vitronectin from the native monomeric form of the protein purified from human plasma for the purposes of this study, to avoid any potential pitfalls inherent in using a recombinant protein to determine native structure. Analyses of several representatives of the cystine knot family have shown that disulfide bonds may be eliminated or misfolded compared to those found in the native peptide, while retaining complete or nearly complete biological activity (29,30). Thus, there is no reliable way to assure that a peptide derived from a recombinant source truly represents the disulfide connectivities of the native protein.
Determining the cysteine linkages of proteins that contain closely spaced complex disulfide bonds is a difficult problem. A number of methods have been employed to address the analytical issues posed by these structures, but they all have limitations. Often, a combination of techniques must be recruited to completely characterize the disulfide-bonding pattern in such proteins (31-33). The primary approach used here was the selective partial reduction of the disulfide bonds in the SMB domain using Tris(2-carboxyethyl)phosphine (TCEP), followed by alkylation of the freed cysteines by N-ethyl maleimide (NEM) under acidic conditions to avoid disulfide scrambling. The SMB region was cleaved from vitronectin by endoproteinase Glu-C, purified by reverse-phase HPLC and partially reduced with TCEP under acidic conditions. The partially reduced isoforms were subjected to peptide mapping using mass spectrometry and NMR to identify the individual disulfide pairings. Our results indicate that the domain scaffold with its intricate network of disulfides is important to orient the single $[\alpha]$-helix within this domain (18) properly for binding of ligands including PAI-1 and uPAR. The SMB domain thus represents a new member of the cystine-stabilized helix (CSH) family (34).
MATERIALS AND METHODS

Materials. N-Ethylmaleimide (NEM) and Tris(2-carboxyethyl)phosphine (TCEP) hydrochloride were from Fluka (Buchs, Switzerland). Dithiothreitol (DTT) and trifluoroacetic acid (TFA) were from Pierce (Rockford, IL). Iodoacetamide (IAM) and bovine insulin were from Sigma (St. Louis, MO). Guanidine-HCl, Tris-base, dibasic sodium phosphate, ammonium bicarbonate, ammonium sulfate, acetic acid, phosphoric acid, HPLC-grade methanol and HPLC-grade acetonitrile were from Fisher Scientific (Suwanee, GA). Sequencing grades of endoproteinase Glu-C, trypsin, and chymotrypsin were from Roche Molecular Biochemicals (Indianapolis, IN).

Endoproteinase Glu-C Digestion of Vitronectin. Native vitronectin was purified from human blood plasma using a modification of the method developed by Dahlback and Podack (35) and stored as a precipitate under saturated ammonium sulfate. The ammonium sulfate slurry of vitronectin was pelleted in 2 ml Eppendorf® tubes in a Beckman microfuge at 12,000 rpm for 15 min. The pellet was dissolved in 4 M guanidine-HCl, 0.02 M iodoacetamide, 0.05 M Na$_2$HPO$_4$, pH 8.0 at a concentration of 2-4 mg/ml as determined by absorbance at 280 nm. The dissolved vitronectin (5-10 mg) was placed in the dark for 1-2 hours at approximately 25°C to block free cysteines. With the free cysteines S-carbamidomethylated, the full length vitronectin was then transferred to a Slide-A-Lyzer® dialysis cassette from Pierce with a 10,000 molecular weight cut-off, where it was dialyzed against three changes of 1 liter of 4 M guanidine-HCl, 0.1 M Na$_2$HPO$_4$, 0.125 M ammonium bicarbonate, pH 7.4, adjusted with phosphoric acid. Finally, the alkylated vitronectin was dialyzed exhaustively against 0.1 M Na$_2$HPO$_4$, 0.125 M ammonium bicarbonate, pH 7.4, adjusted with phosphoric acid. The protein was removed from the dialysis cassette and the concentration determined by absorbance at 280 nm. The concentration was adjusted to 1-2 mg/ml.
using the same buffer. The protease, endoproteinase Glu-C, was added to vitronectin at a 1:100 ratio (w/w), and the digest was performed at 37°C for 24-30 hours.

Purification of Intact and Partially Reduced Forms of the SMB Domain. The peptides generated by the endoproteinase Glu-C digestion of vitronectin were separated by reverse-phase HPLC using a linear gradient at 0.5 ml/min from 5 to 55% acetonitrile over 107 min. The gradient was made by mixing buffer A (0.1% TFA (w/v) in water) and buffer B (acetonitrile with 0.085% TFA (w/v)) in the appropriate proportions on an Agilent 1100 series quaternary pump HPLC system. A multiple wavelength detector allowed for monitoring of samples at 220 nm and 280 nm, and a fluorescent detector was used for monitoring for tryptophan fluorescence (ex 280 nm, em 356 nm). A C-18 Nucleosil column custom packed by Alltech (5 μm particle size, 100 Å pore size, 4.6 mm x 150 mm) was used. Column temperature was maintained at 50°C. Fractions were collected using a Gilson FC203 B fraction collector for identification by MALDI-MS. Additional purification prior to reduction was performed using the same column and HPLC system with a linear gradient at 0.5 ml/min from 15 to 45% B over 60 min.

Partial Reduction with TCEP and Alkylation with NEM. Partial reductions were carried out with an excess of TCEP and the degree of reduction controlled by stopping the reaction at various times by the addition of NEM. The times and concentrations were adapted from Young et al. (36). The HPLC fraction containing the peptide to be reduced was dried using a Savant Speed-Vac with a dry ice and methanol trap. The dried fractions were dissolved in freshly prepared 0.1% TFA (w/v,) pH 2.8. TCEP was prepared immediately prior to use at a concentration of 0.35 M in 0.1% TFA (w/v), pH 2.8. NEM was also freshly prepared at a concentration of 0.8 M in 0.1% TFA (w/v) pH 2.8 and 50% methanol (v/v). Solutions containing TCEP and NEM were protected from the light at
all times. A 0.05 M Tris-Base buffer, pH 7.8, was freshly prepared to titrate the partially reduced samples before alkylation with NEM. Immediately prior to performing the partial reduction, the amount of Tris buffer necessary to titrate a given amount of 0.1% TFA (w/v), pH 2.8, to pH 6.5 was determined empirically. TCEP was added to peptide to yield a solution with a final concentration of 0.02 M TCEP. Samples were generally withdrawn at 30, 60, 90 and 120 min and sufficient NEM was added to the reaction to create a 5:1 molar ratio of NEM to TCEP. Following the addition of NEM, Tris buffer was added to raise the pH to 6.5 to facilitate the alkylation reaction. Alkylation was allowed to proceed at ambient temperature (approximately 25°C) for at least 2 hours. The partially reduced and alkylated peptides were then separated by HPLC as described above and identified by MALDI-MS. Sampling times were adjusted as necessary to enrich for individual isoforms. Additional purification of partially reduced and alkylated forms was performed using a Nucleosil C-18 column and HPLC system with a linear gradient at 0.5 ml/min from 15 to 45% B over 60 min, as described above.

Digestion of Partially Reduced Forms of the SMB Domain. Partially reduced forms of the SMB region of vitronectin were digested with endoproteinase Glu-C, trypsin and/or chymotrypsin to generate peptides for the determination of individual disulfide bonds. Prior to digestion, partially reduced, alkylated and HPLC purified peptides were completely reduced with DTT, and any free cysteines were S-carbamidomethylated with iodoacetamide. Fractions containing the individual partially reduced isoforms were dried on the Savant Speed-Vac and dissolved in 0.1 M ammonium bicarbonate, pH 8.5. Subsequently, DTT (1M stock) was added to each peptide at a final concentration of 0.01 M, and the sample was incubated for 4 hr to completely reduce any remaining disulfide bonds in the partially reduced and alkylated peptide. Iodoacetamide (1M stock in 0.1 M
ammonium bicarbonate, pH 8.5) was added to the fully reduced peptide to give a final concentration of 0.1 M. S-carbamidomethylation with iodoacetamide was performed in the dark at ambient temperature (approximately 25°C) for 2 hours.

Prior to protease digestion, S-carbamidomethylated samples were desalted using a 300 mg Maxi-Clean® C18 cartridge from Alltech according to manufacturers instructions, washed on the cartridge with 3% acetonitrile with 0.1% TFA (w/v), eluted with 50% acetonitrile with 0.1% TFA (w/v), dried on the Savant Speed-Vac and dissolved in the appropriate digestion buffer. All buffers for digestion of partially reduced forms were prepared on the day of use, and digests were performed at 37°C for approximately 15 hours. Endoproteinase Glu-C digests were performed in 0.1 M ammonium bicarbonate, pH 7.4. Trypsin and chymotrypsin digests were performed in 0.1 M ammonium bicarbonate, pH 8.0. All enzymes were added to give a ratio of approximately 50:1 substrate to enzyme (w/w). In the case of dual digests, the peptide was dried after the first digest using the Savant Speed-Vac, then dissolved in the appropriate buffer for the second digest, and enzyme was added in a 50:1 substrate to enzyme (w/w) ratio. After the digestion sequence was complete, the resulting peptide mixture was dried directly on the Savant Speed-Vac and dissolved in 0.1% TFA (w/v). Alternatively, the sample was passed over a 300 mg Maxi-Clean® C18 cartridge from Alltech, washed on the cartridge with 3% acetonitrile with 0.1% TFA (w/v), eluted with 50% acetonitrile with 0.1% TFA (w/v), dried as described above and dissolved in 0.1% TFA (w/v) for MALDI-MS.

*MALDI-MS.* Matrix-assisted laser desorption/ionization (MALDI) mass spectra were obtained in positive ion mode using a PerSeptive Biosystems Voyager DE time-of-flight mass spectrometer, or a Voyager Elite DE reflectron time-of-flight mass spectrometer, each equipped with
a nitrogen laser (Applied Biosystems, Framingham MA). Spectra from the Voyager DE instrument were obtained, in linear mode with +20 kV total accelerating voltage, +18.82 kV applied to the grid, +10 V applied to the guide wire, and a 110 ns acceleration delay. Spectra from the Voyager Elite DE instrument were obtained in linear mode with +25 kV total accelerating voltage, +23 kV applied to the grid, +37.5 V applied to the guide wire, and a 100 ns acceleration delay. Spectra are averages of up to 256 individual laser pulses, obtained from several locations on each sample spot. Sample aliquots were applied to a pre-spotted thin-layer matrix, prepared by applying 0.5 μL portions of 5 g/L nitrocellulose (Immobilon NC Pure, Millipore, Bedford MA), 20 g/L α-cyano-4-hydroxy-cinnamic acid (Aldrich, Milwaukee WI; re-crystallized from ethanol before use) in 1:1 isopropanol:acetone to the sample plate and allowing to dry (37). External calibration of the m/z axis was performed using gramicidin S and bovine insulin. Expected masses to compare with the observed MALDI-MS results were calculated using Protein Prospector (38).

**Electrospray MS.** Analysis was also performed by electrospray quadrupole ion trap mass spectrometry (LCQ Deca, ThermoFinnigan, San Jose CA). Protease digestion samples that were purified by HPLC were desalted using a C18 ZipTip (Millipore, Bedford MA), and the peptides were eluted in 10μL of 50:50 acetonitrile:water. A 100 μL volume of 48:48:2 acetonitrile:water:acetic acid (48/48/2) was added to the desalted sample. The mixture was then infused directly into the electrospray source of the mass spectrometer. The +5 charge-state ion of the 1-55 digest fragment, at m/z 1257.3, was isolated using a 5 m/z isolation width and subjected to collision-induced dissociation to produce a tandem mass spectrum. Charge state deconvolution produced a plot with abscissa units of mass (Da).
NMR Spectroscopy. A sample of the SMB domain (residues 1-51 of vitronectin) was prepared for NMR measurements by cyanogen bromide digestion and HPLC purification, as described (18). NMR spectra were collected using ~90 μM SMB in a 500 μl sample volume in both D₂O and H₂O containing 10% D₂O at pH 4.4 and 298K. All NMR experiments were performed on a 600 MHz Varian INOVA instrument equipped with a single gradient axis and a triple resonance probe for the observation of proton, carbon, and nitrogen nuclei. Two-dimensional NMR data were acquired in phase-sensitive mode using the States-Haberkorn method for quadrature detection in the indirect dimension (39). Spectra of SMB in H₂O were recorded by using WET (40,41) or WATERGATE sequences (42) for water suppression. Two-dimensional homo-nuclear NOESY (43) spectra were recorded with mixing times of 150, 200, and 250 ms. TOCSY spectra (44) were recorded using DIPSI spin-lock sequence with a 8 kHz RF field and mixing times of 15, 30, 60, and 80 ms. Typically, spectra were acquired with 256 t₁ increments, 2048 data points, a relaxation delay of 1s and a spectral width of 8500 Hz. For DQF-COSY spectra (45), 512 t₁ increments were acquired. Spectra were recorded with 64-96 scans per increment for NOESY, 24-80 scans per increment for TOCSY and 56 scans per increment for DQF-COSY. In all NOESY and TOCSY spectra, the data were multiplied by a 60-90° phase shifted sin² window function in both dimensions before Fourier transformation.

NMR Data Analysis. NMR data were processed with Felix 2000 and SPARKY (T. D. Goddard and D. G. Kneller, The University of California at San Diego) software operating on a Silicon Graphics Indigo 2 workstation or on a Silicon Graphics Origin 300 server. Cross-peak intensities observed in NOE experiments were divided into four categories as strong, medium, weak, and very weak. These intensities were converted into distance restraints as follows: strong, 1.8-2.7
Å; medium, 1.8-3.4 Å; weak, 1.8-4.5 Å, and very weak 1.8-6.0 Å. An additional 1.0 Å was added to upper limits involving methyl protons. Similarly, an additional 0.5 Å for methylene protons and 2.3 Å for degenerate H_d and H_e protons of tyrosines and phenylalanines were added to upper limits. Also, a 0.2 Å was added to the upper limits of NOEs involving amide protons. Coupling constants were extracted from the DQF-COSY spectra. Backbone $\phi$ angles were restrained to $-120^\circ \pm 50^\circ$ for $^3J_{HNH_d} = 8-9$ Hz, and $-120^\circ \pm 40^\circ$ for $^3J_{HNH_e} > 9$Hz. A restraint of $-100^\circ \pm 80^\circ$ was also applied to the $\phi$ angle for residues that show stronger NH-$\phi$-$\phi_{i+1}$ NOE than the intraresidue NH-H-$\phi$ NOE (46). A total of 329 NOE restraints and 18 $\phi$ restraints were used in structure determination. All calculations were carried out using the AMBER force field interfaced with DISCOVER (Accelrys, San Diego, CA) on an Origin300 workstation. Random structures were generated by subjecting the peptide to an initial 10,000-step minimization at 298 K. The temperature was then raised gradually to 1000 K during a 1000 step dynamics simulation. The peptide was subjected to minimization and a 10 ps dynamics run at 1000 K. The NMR-derived restraints were then imposed on the peptide and the peptide was slowly annealed to 298 K in a 100 ps trajectory. Finally, the structures were subjected to further minimization at 298 K. The force constant for the distance restraints was 100 kcal/mol Å$^{-2}$ and the dielectric constant was 4.
RESULTS

*Endoproteinase Glu-C Digestion of Vitronectin and Isolation of the N-terminal Peptide.*

Determination of the disulfide-bonding pattern of the SMB domain of vitronectin began with isolation by protease digestion and reverse-phase HPLC. The first step in isolating the SMB domain was digesting full-length vitronectin with endoproteinase Glu-C. Vitronectin contains a total of 14 cysteine residues, arranged in six disulfide bonds with the remaining two free sulfhydryls occupying buried positions (47). Four of the six disulfides are found within the first 39 amino acids of the protein in the SMB domain, while the other two have been assigned as C137:C161 and C274:C453 (16). The free sulfhydryls are at positions 196 and 411 (16). Prior to digestion with endoproteinase Glu-C, these two free cysteines were carboxamidomethylated with iodoacetamide to avoid any free thiols that could promote disulfide rearrangements. The carboxamidomethylated vitronectin was then digested with endoproteinase Glu-C and the resulting peptides analyzed by reverse-phase HPLC (Figure 1A). The key to generating the optimal separation during HPLC analysis and purification of the N-terminal domain was the use of a small pore (100 Å) stationary phase. The size exclusion properties of the small pore matrix provided additional separation that was necessary to separate the N-terminus from other peptides and full-length vitronectin.

Preparative amounts of the digest were injected (1-1.5 mg/injection) and purified on the same small pore reverse-phase matrix. The resulting fractions were analyzed by MALDI-MS. A peak eluting at approximately 47-48 min was identified by mass as the fully oxidized 1-47 amino acid N-terminal fragment of vitronectin (Figure 1A). The expected average mass-to-charge ratio (m/z) was 5332.9 ([M+H]+, isotopically averaged) and the observed m/z was 5336.1. The peak eluting at approximately 53-54 min was identified by mass spectrometry as the fully oxidized 1-55 amino acid N-terminal fragment of vitronectin (Figure 1A). The expected m/z was 6281.9 ([M+H]+, isotopically...
averaged) and the observed m/z was 6284.2. Identity of these peptides also was confirmed by Edman sequencing (data not shown). The two N-terminal fragments, vitronectin 1-47 and vitronectin 1-55, were dried and re-purified by reverse-phase HPLC prior to partial reduction. The sample comprising residues 1-55, shown in Figure 1B following purification and re-chromatography by HPLC to evaluate purity, was used for the remainder of the analyses on disulfide bonds described in this work. Figure 1C shows the MALDI-MS spectrum of this sample.

Figure 1D shows an experiment performed on this sample by electrospray quadrupole ion trap MS using a Finnigan LCQ Deca instrument. From multiple charge states, the average mass of a peptide representing the SMB region (residues 1-55 with 4 disulfides) was measured as 6281.2, in excellent agreement with the calculated value of 6280.9 for the peptide. Modern MS technology allows for isolation of a peak of interest in the electrospray spectrum to confirm the peptide by sequence tag identification. The +5 charge-state ion at m/z 1257.3 was isolated using a 5 m/z isolation width, subjected to collision-induced dissociation, and the resulting MS/MS (or tandem MS) spectrum was obtained (Figure 1D). The sequence tag derived from the spectrum, corresponding to species differing in length by a single amino acid residue, agrees with the known vitronectin sequence, thus confirming that the m/z 1257.3 parent ion is due to the partial digest fragment containing residues 1-55.

Partial Reduction with TCEP and Alkylation of Free Cysteines with N-ethyl Maleimide. The second step in determining the disulfide bonding structure of the N-terminal domain of vitronectin was the partial reduction and alkylation of the liberated free thiols under acidic conditions. TCEP, the reducing agent, and NEM, the primary alkylating agent, were specifically chosen because the reactions may be performed under acidic conditions to avoid disulfide scrambling. The purified N-
terminal 1-55 fragment of vitronectin was dried and dissolved in 0.1% TFA, pH 2.8. The solution was adjusted to a final concentration of 0.02 M TCEP, and samples were withdrawn every 30 minutes over a period of 120 minutes in order to capture all of the possible reaction products. Based upon bovine insulin injected as a standard on the HPLC for comparison to the 1-55 fragment at 220 nm prior to partial reduction, the average concentration of the N-terminal peptide in the reaction mixture was estimated to be approximately 8 μM. The reduction rate of the peptide at this concentration and pH in the presence of 0.02 M TCEP was reproducible. Higher pH and/or higher concentrations of TCEP resulted in unacceptably fast reaction times, making it difficult to capture the single and doubly reduced reaction products.

The partially reduced samples were alkylated upon addition of NEM to a final concentration of 0.1 M NEM; sufficient Tris Base was added to bring the pH to 6.5 to facilitate the alkylation reaction with NEM, which effectively stops further reduction by TCEP. The 30- and 60-minute samples were analyzed by MALDI-MS to determine the partial reduction products (Figure 2). All five possible reduction products ranging from fully oxidized to fully reduced were observed, confirming that the N-terminus of native vitronectin contains four cystines that can be selectively reduced over time. Figure 3 is a schematic that depicts this strategy of partial reduction and NEM blocking, with the three reduction products of interest for further analysis, the SMB domain corresponding to one, two, or three disulfides open and alkylated (SMB1O, SMB2O, and SMB3O, respectively). These reduced and alkylated peptides generated during the partial reduction were purified by reverse-phase HPLC (Figure 4A) and the resulting fractions analyzed by MALDI-MS. Small shoulders observed on the HPLC peaks (Figure 4A) from the partial reduction were confirmed by MALDI-MS to be isoforms reflecting varying amounts of alkylation with NEM. Selected fractions containing the individual reaction products from the partial reduction of the SMB domain
corresponding to one, two, or three disulfides open and alkylated (SMB1O, SMB2O, and SMB3O) were re-purified by HPLC (Figures 4B-D) and carried forward for further analysis to determine individual disulfide pairs.

**Protease Digestion and Analysis of Alkylated Partial Reduction Products.** The strategy for determining the identity of the liberated sulfhydryls and remaining disulfides in the partially reduced and alkylated forms of vitronectin 1-55 is presented in the schemes outlined in Figures 5, 6 and 7. Two approaches were used. In some cases, the samples were digested directly and analyzed by MALDI-MS. Alternatively, the partially reduced and NEM-blocked samples were treated further with DTT to achieve complete reduction, and alkylation with IAM differentially tagged the residues that still existed as cystines after TCEP treatment. Subsequent enzymatic digestion and mass analysis was then pursued, and the presence of NEM or CAM at individual sites was used to determine which cysteine residues were oxidized and reduced in the samples with one cystine, two cystines, or three cystines reduced by TCEP (denoted SMB1O, SMB2O, and SMB3O in Figures 3 - 7).

First, the isoform of the SMB domain that was partially reduced leaving three disulfides intact and one disulfide pair reduced and alkylated with NEM (SMB1O, Figure 5) was digested directly with trypsin, and the peptides were analyzed by MALDI-MS (Table 1, Supplemental Figure 1). A peptide product was identified at \( m/z \) of 2035.5. This corresponds to the predicted \( m/z \) of 2033.8 for two peptides linked through a C5:C9 disulfide bond (SMB1O-T1/2 in Figure 5). Specifically, this product corresponds to vitronectin residues 1-6 (DQESCK) disulfide bonded to vitronectin residues 7-17 (GRCTEGFNVDK). This disulfide bond was the first identified in our analysis, and its presence was confirmed repeatedly during the course of this work.
A sample of the SMB1O form of the domain having a single pair of cysteines alkylated with NEM was fully reduced with DTT, and the newly released cysteines were carboxyamidomethylated with IAM, as outlined in Figure 5. The resulting peptide was then digested with chymotrypsin and analyzed by MALDI-MS (Supplemental Figures 2 and 3). Three peptide products were identified (Table 1). A peak of m/z 2964.5 was observed which corresponds to the expected m/z of 2965.3 for residues 1-24 with all four cysteines (C5, C9, C19, C21) carboxyamidomethylated (peptide SMB1O-Ch1, Figure 5). The presence of the CAM tag indicates that all of these cysteines were disulfide bonded in the original SMB1O peptide generated by TCEP and NEM treatment. This is consistent with the initial observation of the C5:C9 bond assignment.

Secondly, a peptide with an observed m/z of 1578.7 was also seen in the spectrum of this digest. This agrees with the expected m/z, 1575.7, of a peptide containing residues 25-35 in which two of the cysteines were carboxyamidomethylated and one was alkylated with maleimide (peptide SMB1O-Ch2, Figure 5). Also, seen in this reaction is a further cleavage product of SMB1O-Ch2, a peptide spanning residues 29-35 with both cysteines carboxyamidomethylated (peptide SMB1O-Ch3, Figure 5). This peptide yields an observed m/z of 933.5, which corresponds to the expected m/z of 934.0 for this sequence. From comparison of the results on peptides SMB1O-Ch2 and SMB1O-Ch3, it is deduced that C25 in peptide SMB1O-Ch2 must be the residue alkylated with NEM. Thus, C25 was a half cystine released during the partial reduction with TCEP that yielded the SMB1O form. All of the other 6 cysteine residues identified from this analysis were CAM labeled, indicating that these 6 residues remain paired in the initial reduction to form SMB1O. By the process of elimination, the only remaining cysteine residue, C39, should be the other half-cystine partner for C25. Unfortunately, the MALDI-MS spectrum for this digest did not show the peptide comprising residues 36-55 in whole or part. While these results, combined with those for the tryptic digest of
the 1-open form of vitronectin 1-55, suggest that C25:C39 form a disulfide pair, C39 alkylated with NEM was not recovered to confirm this inference.

Definitive evidence for the C25:C39 assignment was gained from analysis of the SMB domain that was partially reduced to yield four free cysteines alkylated with NEM (SMB2O, Figure 6). This partial reduction product was fully reduced and alkylated with IAM, so that the cystines remaining after TCEP treatment can be identified with the CAM tag. The doubly tagged peptide was digested with chymotrypsin, and the products were analyzed by MALDI-MS (Supplemental Figure 4). Three prominent peptides were observed (Table 1). The first (SMB2O-Ch1, Figure 6), with an observed m/z of 1675.9, agrees with the expected m/z of 1676.9 for a peptide containing residues 36-49 with the sole cysteine (C39) alkylated by NEM. The second peptide, SMB2O-Ch2 (Figure 6) is identified as containing residues 14-23, with both cysteines tagged with CAM. Finally, the third m/z observed was 1095.2, which corresponds to residues 28-35 with both cysteines tagged with CAM, with a calculated m/z 1097.2 (SMB2O-Ch3, Figure 6). Thus, C19, C21, C31 and C32 observed here in the CAM form were, therefore, disulfide bonded to each other in SMB2O. C39 was observed in peptide SMB2O-Ch1 to be alkylated by NEM, indicating that C39 was a half-cystine released during the partial reduction that disrupted two disulfides. Thus, the combined information derived from all three digests of the partial reduction products from both the SMB1O and SMB2O isoforms supports the conclusion that C25:C39 form a disulfide pair and are the first pair released during the partial reduction.

Having confirmed that the first two disulfide bonds disrupted by TCEP treatment are C25:C39 and C5:C9, respectively, the final assignments required discerning the arrangement of two other disulfides among residues C19, C21, C31 and C32. The partially reduced form of the SMB domain that resulted in three reduced disulfides (6 free cysteines) labeled with NEM after TCEP
treatment, and a single remaining disulfide bond was purified (SMB3O, Figure 7). The residual
disulfide bond was reduced with DTT and the newly freed cysteines were carboxamidomethylated
with IAM. This peptide was initially digested with chymotrypsin and analyzed by MALDI-MS.
Two prominent peptides were observed from this analysis. One, with a \( m/z \) of 2648.0, corresponded
to the expected core fragment comprising residues 18-35 with three cysteines tagged with NEM and
only two (from the last disulfide to remain unreduced with TCEP) tagged with CAM, with a
calculated \( m/z \) of 2648.9. This data in itself was not definitive, because there are 5 cysteines within
this core peptide and they cannot be identified without sequencing data. Although it had already
been determined that C25 was paired with C39 (outside this core peptide), the three possible
arrangements between the remaining four cysteine residues are not obvious.

Therefore, the SMB3O sample was digested sequentially with two enzymes, endoproteinase
Glu-C followed by chymotrypsin (Figure 7). This combination of enzymes is expected to cleave at
position 22 or 23, intermediate between the C19-X- C21 region and the C31- 32 sequence. Mass
identification of each of the two separate fragments (containing C19 and C21 or C31 and C32) with
both NEM and CAM labels would indicate that proximal cysteines do not pair and that disulfides
must form in a non-linear fashion. In fact, this result is observed; each of two peptides, SMB3O-
GCh1 and SMB3O-GCh2 (Figure 7, Table 1), contained one NEM-modified residue and one CAM-
modified cysteine (Supplemental Figures 5 and 6). The MALDI-MS data are therefore consistent
with C19 bonding with either C31 or C32, and, likewise, C21 bonding with the other between the
same two residues. Thus, the data indicate one of only two disulfide arrangements in the core
region, i.e. a combination of C19:C31 and C21:C32 or a pairing between C19:C32 and C21:C31.
Although final distinction between these two possible disulfide assignments was not possible from
this approach alone, the analysis ruled out the unlikely possibility that cysteines C31 and C32 are in
a vicinal disulfide. Such vicinal arrangements are extremely rare in protein structures and do not appear in the SMB domain.

*NMR Approach to Assign the Disulfides in the Core Region Spanning Residues 19-35.* As the tactic using peptide mapping and mass spectrometry was in progress, we initiated a study to determine the three-dimensional structure of the SMB domain isolated from human plasma vitronectin using $^1$H-NMR. The source material for the NMR analysis was the N-terminal 51 amino acids from vitronectin, isolated by cyanogen bromide cleavage, rather than the 55-residue fragment isolated by endoproteinase Glu-C digestion that was used for the peptide mapping/MS analyses. Since the four cystines are confined to residues 5-39, the two samples were suitable for parallel studies on the disulfide bonds. The procedure for determining the overall three-dimensional fold for the SMB domain was standard, involving sequence assignments for $^1$H resonances and two-dimensional NOESY data to assign interactions between protons that occur at medium to long range and thus dictate the fold of the domain (18). Simulated annealing calculations were used with iterative energy minimizations to arrive at the final structure (18). Initially the structural calculations were pursued without any disulfide restraints, yielding results that were consistently in agreement with the C5:C9 and C25:C39 assignments (18). Thus, an NMR approach was taken to distinguish the two possibilities for the final disulfide assignments in the 19-35 core region.

For this analysis, the two disulfides generated from the peptide mapping/MS work were imposed to refine the structure of the SMB domain, and two possible alternatives were then pursued separately in the structure determination: C19:C31 and C21:C32 pairs or C19:C32 and C21:C31 pairs. The disulfide bridge restraints were set for the two alternatives (48), and both were subjected to simulated annealing separately. Figure 8 shows the results of these analyses, which clearly assign...
the core disulfides as C19:C31 and C21:C32. Plotted in Figure 8 are two arrays that compare sets of simulated annealing calculations for the two possible arrangements of disulfide pairs. The results are color coded to represent the calculated RMSD relating pairs of structures in the matrix. For each set of calculations, the color range is yellow (0 – 2.5 Å), blue (2.5 – 5 Å), and cyan (>5 Å). From this visual display of the matrix of structures generated, it is obvious that calculations with the second alternative involving disulfides C19:C32 and C21:C31 did not yield acceptable converged structures. Thus, the disulfide arrangement for the two cystine linkages in the core of the SMB domain is C19:C31 and C21:C32.
DISCUSSION

In the study presented here, we have selectively reduced, in a step-wise fashion, the native disulfide bonds of the SMB domain of vitronectin and then alkylated the freed cysteines. Using peptide mapping, mass spectrometry and NMR, we have determined that the disulfide bonds in the SMB domain of vitronectin are C5:C9, C19:C31, C21:C32, and C25:C39. Apart from the first two cysteine residues that form the C5:C9 bond, the order of disulfides in this small domain is reminiscent of the arrangement found in a large number of proteins that fall into the “cystine knot” motif (22,24,26,49). That is, the remaining six cysteines are bonded in an “n:n+3” pattern, i.e. the first in the series forms a disulfide with the fourth (C19:C31), the second with the fifth (C21:C32), and the third with the sixth (C25:C39). Perhaps the best recognized members of this cystine knot family are the growth factor-type cystine knots, which contain six cysteines paired in a similar pattern to form a “ring” structure through which part of the peptide chain is threaded in this small domain (22,50). In spite of this interesting comparison, this analogy with cystine knots should not be over-interpreted. For one thing, many of the growth factor-type cystine knots have an extra unpaired cysteine residue. Also, a major difference between this structure and cystine knots is apparent from the solution structure of the SMB domain (18), which contains a single α-helix as the only element of secondary structure, whereas cystine knots are often rich in mini-β sheets (22).

Furthermore, the local context and amino acid spacing between cysteine residues is important in directing the overall fold of a cystine knot. A close inspection of the primary sequence of the cysteine-rich N-terminal portion of vitronectin reveals that it contains four disulfides within contexts highly similar to cystine knot proteins (27) (Figure 9A). However, the ordering of recognized sequence patterns surrounding individual cysteines along the SMB polypeptide chain differs from many of the known homologues. For example, key common patterns in cystine knot proteins are
represented in the form of CXGXC and CXC, where X may be any amino acid other than C. In six-membered and eight-membered cystine knots, such as observed with the EGF fold (27) or bone morphogenetic protein agonists (24), respectively, these four cysteines occur in a well-defined order and form the “ring” structure of the classical cystine motif. Typically, the order is C$_2$XGXC$_3$ and C$_5$XC$_6$ with disulfide pairings of C$_2$ - C$_5$ and C$_3$ - C$_6$ (24) (Figure 9A). However, in the SMB domain of vitronectin, the cysteines occur in the order C$_1$XGXC$_2$ and C$_3$XC$_4$ with disulfide pairings of C$_1$-C$_2$, C$_3$-C$_6$, and C$_4$-C$_7$ resulting in a pseudo-knot (51) (Figure 9B). Thus, this domain assumes a different bonding pattern from the typical knotted protein and the ring motif is lacking.

In this way, we were able to show that the SMB domain of vitronectin does not form a classical disulfide knot. Nevertheless, it does possess features with a high degree of similarity to cystine knot proteins. Most prominent among these is a loop formed by the disulfide pairing between C$_25$ and C$_39$, which is reminiscent of structures classified as cystine-stabilized $\alpha$-helices (CSH motifs). Such structures are known to function in extracellular ligand interactions (27). Interestingly, this loop contains residues that have been shown to be involved in binding PAI-1 (13) and uPAR (15), and the region from residues 26 to 30 forms the sole $\alpha$-helix in this small domain (17,18). This ligand-binding loop has a calculated pI of 3.7 and is oriented toward the solvent by the tightly structured core formed by the C$_{19}$:C$_{31}$ and C$_{21}$:C$_{32}$ disulfide pairings. On the far N-terminus is a small positively charged loop formed by the C$_{5}$:C$_{9}$ disulfide bond, suggesting that the SMB domain of vitronectin has a small positively charged tail opposite a large negatively charged ligand-binding nose.

This sequence of disulfide pairings is in agreement with the solution structure from NMR determined recently by our laboratory (18) (Figure 9B). These results are, however, at odds with previously published results for recombinant counterparts of the SMB domain of vitronectin (17,28).
Kamikubo *et al.* employed similar partial reduction methods with TCEP to discern the disulfide connectivities of a form of the SMB domain isolated from a fusion protein expressed in *E. coli* (28). They identified C5:C9, C19:C21, C25:C31, and C32:C39 as the cystine pattern in their recombinant protein (rSMB) (28) (*Figure 9C*). The underlying assumption made by Kamikubo *et al.* was that rSMB contained the native disulfide framework of vitronectin because it was functionally active in their *in vitro* assays and it had comparable PAI-1 binding activity to multimeric vitronectin. Also, mutagenesis of half-cystines within two of their proposed disulfides resulted in a complete loss of biological activity. However, work on other proteins displaying a cystine knot motif has shown that biological activity may often be misleading, in that forms with one or more disulfides missing or disrupted can maintain full biological activity (30,52-54).

While not previously observed as a native structure among proteins harboring the familiar cystine knot motif, the structure proposed by Kamikubo *et al.* in 2002 (28) and further elaborated upon in 2004 (19) is familiar from the literature on cystine knots where it has been characterized as a highly stable folding intermediate (55). In a study describing non-native disulfide bonding patterns in oxidative folding intermediates of the *Amaranthus* α-amylase inhibitor, a member of the inhibitory cystine knot family, a major folding intermediate (MFI) is identified as a peptide structure with a “linear uncrossed pattern” (55). The NMR structure for the MFI of the α-amylase inhibitor showed that it appears as compact as the native knotted protein, presumably because of conformational restraints imposed by a vicinal disulfide bridge. Other disulfide intermediates are short-lived compared to the MFI with its linear bonding pattern (56). Another example of a linear arrangement of disulfides observed en route to the correct final fold comes from work by Wilken and Bedows on chorionic gonadotropin β-subunit (57). In this work, a transient disulfide forms *in vivo* between cysteines C3 and C4, (nearest neighbors, although not vicinal, in the sequence). This
disulfide is rearranged to form a typical C₃-C₆ linkage in the secreted, mature form of the protein. Thus, it appears that the linear disulfide assignments reported by Kamikubo et al. (2002) on rSMB represent a stable folding intermediate rather than the native arrangement of cystines.

The first report of a three-dimensional structure for the SMB domain by Zhou et al. (2003) also used a recombinant SMB domain that was co-crystallized with the active recombinant mutant PAI-1 (17). The disulfide-bonding pattern observed in their study was C5:C21, C9:C39, C19:C32, and C25:C31 (Figure 9D). This matches neither our pattern for native SMB (Figure 9B), nor the original linear pattern reported for a recombinant SMB by Kamikubo et al. (28) (Figure 9C). More recently Kamikubo et al. have determined a three-dimensional structure for their recombinant SMB and have suggested that there are numerous disulfide patterns that are consistent with their fold (19). We emphasize by contrast the unique assignment of disulfides for the SMB domain isolated in our work from human plasma vitronectin, for which we have no evidence for heterogeneity or multiple disulfide arrangements.

Although the solution structure from our NMR work (18) and the structures on the recombinant SMB domains (17,19) differ in overall fold, they are similar in one critical feature—the presence of the single α-helix that houses PAI-1- and uPAR-binding residues (17,28). In spite of different disulfide scaffolds, both structures form what appears to be a conserved helix reminiscent of the CSH motif observed in traditional cystine knot proteins. This curious result may stem from the demonstrated malleability of CSH structures to accommodate cysteine substitutions, yet retain biological activity (29,30). Nevertheless, the different disulfide pairs in the recombinant forms of the SMB domain do not represent the proper arrangement in the physiological scenario. In similar types of knotted structures, folding studies have shown that the energetic differences between native and non-native folds is not large (55,56,58). In a study on disulfides in vascular endothelial growth
factor, it was observed that removal of disulfides leads to folding anomalies and a lower $T_m$ for unfolding, although the overall thermodynamic stability for the various mutants was comparable (59). Also, other work has suggested that a native disulfide arrangement may be more significant for the biological half-life and proper secretion of the polypeptide, as opposed to activity (60). In this regard, comparisons among several disulfide arrangements in gonadotropin $\alpha$-subunit showed that, in spite of the fact that all were nominally active in a bioassay, only the native disulfide cross-links lend normal kinetics of secretion from cells (53). Such studies suggest that the native fold is more kinetically favored, regardless of small energetic differences between native and misfolded disulfide knots.

The SMB domain of vitronectin possesses sequence homologies that are tantalizing in their similarity to the cystine knot motif. In comparison to classic cystine knots, it appears that a permutation in the arrangement of sequence micro-domains surrounding the cysteines in the SMB domain nevertheless results in a structure that mimics these models in terms of folding and stability. Adopting the fold of a pseudo-knot, the SMB domain of vitronectin shares common attributes with true cystine knot proteins. While studies on recombinant SMB domains have contributed valuable information regarding some features that are required for ligand binding to vitronectin, the SMB domain isolated from native monomeric vitronectin, having disulfide bonds between C5:C9, C19:C31, C21:C32, and C25:C39, represents the biologically relevant form of the domain with a CSH motif.
FOOTNOTES

Abbreviations include: SMB, somatomedin B; PAI-1, rSMB, recombinant somatomedin B: plasminogen activator inhibitor-type 1, uPAR, urokinase-type plasminogen activator receptor; TCEP, Tris(2-carboxyethyl)phosphine; HPLC, high performance liquid chromatography; TFA, trifluoroacetic acid; IAM, iodoacetamide; NEM, N-ethylmaleimide; CAM, carbamidomethyl; NMR, nuclear magnetic resonance; MALDI-MS, matrix-assisted laser desorption/ionization-mass spectrometry; CSH, cystine stabilized helix

ACKNOWLEDGMENTS

We thank Nathan VerBerkmoes, a graduate student in the Program in Genome Science and Technology at the University of Tennessee and Oak Ridge National Laboratory, for assistance with the tandem MS experiments.
REFERENCES

1. Schroeck, F., Arroyo de Prada, N., Sperl, S., Schmitt, M., and Viktor, M. (2002) Biol Chem 383, 1143-1149. "Interaction of plasminogen activator inhibitor type-1 (PAI-1) with vitronectin (Vn): mapping the binding sites on PAI-1 and Vn"
2. Mousa, S. A. (2002) Curr Opin Investig Drugs 3, 1191-1195. "Vitronectin receptors in vascular disorders"
3. Wechsler-Reya, R. J. (2001) Trends Neurosci 24, 680-682. "Caught in the matrix: how vitronectin controls neuronal differentiation"
4. Preisssner, K. T., and Seiffert, D. (1998) Thromb Res 89, 1-21. "Role of vitronectin and its receptors in haemostasis and vascular remodeling"
5. Minor, K. H., and Peterson, C. B. (2002) J Biol Chem 277, 10337-10345. "Plasminogen activator inhibitor type 1 promotes the self-association of vitronectin into complexes exhibiting altered incorporation into the extracellular matrix"
6. Podor, T. J., Shaughnessy, S. G., Blackburn, M. N., and Peterson, C. B. (2000) J Biol Chem 275, 25402-25410. "New insights into the size and stoichiometry of the plasminogen activator inhibitor type-1.vitronectin complex"
7. Minor, K. H., Wilkins-Port, C. E., McKeown-Longo, P. J., and Peterson, C. B. (2004) Mol. Biol. Cell in press "Plasminogen Activator Inhibitor-1 Alters the Binding and Degradation of Vitronectin by Cultured Human Fibroblasts"
8. Stefansson, S., and Lawrence, D. A. (1996) Nature 383, 441-443. "The serpin PAI-1 inhibits cell migration by blocking integrin alpha V beta 3 binding to vitronectin"
9. Loskutoff, D. J., Curriden, S. A., Hu, G., and Deng, G. (1999) Apmis 107, 54-61. "Regulation of cell adhesion by PAI-1"
10. Sidenius, N., and Blasi, F. (2003) Cancer Metastasis Rev 22, 205-222. "The urokinase plasminogen activator system in cancer: recent advances and implication for prognosis and therapy"
11. Sigurdardottir, O., and Wiman, B. (1994) Biochim Biophys Acta 1208, 104-110. "Identification of a PAI-1 binding site in vitronectin"
12. Deng, G., Royle, G., Seiffert, D., and Loskutoff, D. J. (1995) Thromb Haemost 74, 66-70. "The PAI-1/vitronectin interaction: two cats in a bag?"
13. Deng, G., Royle, G., Wang, S., Crain, K., and Loskutoff, D. J. (1996) J Biol Chem 271, 12716-12723. "Structural and functional analysis of the plasminogen activator inhibitor-1 binding motif in the somatomedin B domain of vitronectin"
14. Deng, G., Curriden, S. A., Hu, G., Czekay, R. P., and Loskutoff, D. J. (2001) J Cell Physiol 189, 23-33. "Plasminogen activator inhibitor-1 regulates cell adhesion by binding to the somatomedin B domain of vitronectin"
15. Deng, G., Curriden, S. A., Wang, S., Rosenberg, S., and Loskutoff, D. J. (1996) J Cell Biol 134, 1563-1571. "Is plasminogen activator inhibitor-1 the molecular switch that governs urokinase receptor-mediated cell adhesion and release?"
16. Xu, D., Baburaj, K., Peterson, C. B., and Xu, Y. (2001) Proteins 44, 312-320. "Model for the three-dimensional structure of vitronectin: Predictions for the multi-domain protein from threading and docking"
Zhou, A., Huntington, J. A., Pannu, N. S., Carrell, R. W., and Read, R. J. (2003) *Nature Structural Biology* **10**, 541-544. "How vitronectin binds PAI-1 to modulate fibrinolysis and cell migration"

Mayasundari, A., Whittemore, N. A., Serpersu, E. H., and Peterson, C. B. (2004) *J Biol Chem* *in press* "The Solution Structure of the N-terminal Domain of Human Vitronectin: Proximal Sites that Regulate Fibrinolysis and Cell Migration"

Kamikubo, Y., De Guzman, R., Kroon, G., Curriden, S., Neels, J. G., Churchill, M. J., Dawson, P., Oldqiej, S., Jaglielska, A., Sheraga, H. A., Loskutoff, D. J., and Dyson, H. J. (2004) *Biochemistry* *on-line citation* "Disulfide bonding arrangements in active forms of the somatomedin B domain of human vitronectin"

Anfinsen, C. B., Haber, E., Sela, M., and White, F. H., Jr. (1961) *Proc Natl Acad Sci U S A* **47**, 1309-1314. "The kinetics of formation of native ribonuclease during oxidation of the reduced polypeptide chain"

Anfinsen, C. B. (1973) *Science* **181**, 223-230. "Principles that govern the folding of protein chains"

Isaacs, N. W. (1995) *Curr Opin Struct Biol* **5**, 391-395. "Cystine knots"

Sun, P. D., and Davies, D. R. (1995) *Annu Rev Biophys Biomol Struct* **24**, 269-291. "The cystine-knot growth-factor superfamily"

Avsian-Kretchmer, O., and Hsueh, A. J. (2004) *Mol Endocrinol* **18**, 1-12. "Comparative genomic analysis of the eight-membered ring cystine knot-containing bone morphogenetic protein antagonists"

Escoubas, P., and Rash, L. (2004) *Toxicon* **43**, 555-574. "Tarantulas: eight-legged pharmacists and combinatorial chemists"

Craik, D. J., Daly, N. L., and Waine, C. (2001) *Toxicon* **39**, 43-60. "The cystine knot motif in toxins and implications for drug design"

Vitt, U. A., Hsu, S. Y., and Hsueh, A. J. (2001) *Mol Endocrinol* **15**, 681-694. "Evolution and classification of cystine knot-containing hormones and related extracellular signaling molecules"

Kamikubo, Y., Okumura, Y., and Loskutoff, D. J. (2002) *J Biol Chem* **277**, 27109-27119. "Identification of the disulfide bonds in the recombinant somatomedin B domain of human vitronectin"

Hiro'oka, T., Maassen, D., Berger, P., and Boime, I. (2000) *Endocrinology* **141**, 4751-4756. "Disulfide bond mutations in follicle-stimulating hormone result in uncoupling of biological activity from intracellular behavior"

Zhu, Q., Liang, S., Martin, L., Gasparini, S., Menez, A., and Vita, C. (2002) *Biochemistry* **41**, 11488-11494. "Role of disulfide bonds in folding and activity of leiurotoxin I: just two disulfides suffice"

Gray, W. R. (1993) *Protein Sci* **2**, 1749-1755. "Echistatin disulfide bridges: selective reduction and linkage assignment"

Wu, J., and Watson, J. T. (2002) *Methods Mol Biol* **194**, 1-22. "Assignment of disulfide bonds in proteins by chemical cleavage and peptide mapping by mass spectrometry"

Yen, T. Y., Yan, H., and Macher, B. A. (2002) *J Mass Spectrom* **37**, 15-30. "Characterizing closely spaced, complex disulfide bond patterns in peptides and proteins by liquid chromatography/electrospray ionization tandem mass spectrometry"
34. Tamaoki, H., Miura, R., Kusunoki, M., Kyogoku, Y., Kobayashi, Y., and Moroder, L. (1998) *Protein Eng* **11**, 649-659. "Folding motifs induced and stabilized by distinct cystine frameworks"

35. Dahlback, B., and Podack, E. R. (1985) *Biochemistry* **24**, 2368-2374. "Characterization of human S protein, an inhibitor of the membrane attack complex of complement. Demonstration of a free reactive thiol group"

36. Young, Y., Zeni, L., Rosenfeld, R. D., Stark, K. L., Rohde, M. F., and Haniu, M. (1999) *J. Pept. Res.* **54**, 514-521. "Disulfide assignment of the c-terminal cysteine knot of agouti-related protein (AGRP) by direct sequencing analysis."

37. Vorm, O., Roepstorff, P., and Mann, M. (1994) *Anal. Chem.* **66**, 3281-3287. "Improved Resolution and Very High Sensitivity in MALDI TOF of Matrix Surfaces Made by Fast Evaporation"

38. Karl R. Clauser, P. B., and Alma L. Burlingame. (1999) *Anal Chem* **71**, 2871 -2882. "Role of Accurate Mass Measurement (±10 ppm) in Protein Identification Strategies Employing MS or MS/MS and Database Searching"

39. States, D. J., Haberkorn, R. A., and Ruben, D. J. (1982) *Journal of Magnetic Resonance* **48**, 286-292. "A Two-Dimensional Nuclear Overhauser Experiment with Pure Absorption Phase in 4 Quadrants"

40. Ogg, R. J., Kingsley, P. B., and Taylor, J. S. (1994) *Journal of Magnetic Resonance Series B* **104**, 1-10. "Wet, a T-1-Insensitive and B-1-Insensitive Water-Suppression Method for in-Vivo Localized H-1-Nmr Spectroscopy"

41. Smallcombe, S. H., Patt, S. L., and Keifer, P. A. (1995) *Journal of Magnetic Resonance Series A* **117**, 295-303. "WET solvent suppression and its applications to LC NMR and high-resolution NMR spectroscopy"

42. Piotto, M., Saudek, V., and Sklenar, V. (1992) *Journal of Biomolecular NMR* **2**, 661-665. "Gradient-Tailored Excitation for Single-Quantum NMR-Spectroscopy of Aqueous-Solutions"

43. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) *Journal of Chemical Physics* **71**, 4546-4553. "Investigation of Exchange Processes by 2-Dimensional NMR-Spectroscopy"

44. Shaka, A. J., Lee, C. J., and Pines, A. (1988) *Journal of Magnetic Resonance* **77**, 274-293. "Iterative Schemes for Bilinear Operators - Application to Spin Decoupling"

45. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wuthrich, K. (1983) *Biochemical and Biophysical Research Communications* **117**, 479-485. "Improved Spectral Resolution in Cosy H-1-Nmr Spectra of Proteins Via Double Quantum Filtering"

46. Clubb, R. T., Ferguson, S. B., Walsh, C. T., and Wagner, G. (1994) *Biochemistry* **33**, 2761-2772. "Three-Dimensional Solution Structure of Escherichia coli Periplasmic Cyclophilin."

47. Zhuang, P., Li, H., Williams, J. G., Wagner, N. V., Seiffert, D., and Peterson, C. B. (1996) *J Biol Chem* **271**, 14333-14343. "Characterization of the denaturation and renaturation of human plasma vitronectin. II. Investigation into the mechanism of formation of multimers"

48. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *Febs Letters* **229**, 317-324. "Determination of 3-Dimensional Structures of Proteins from Interproton Distance Data by Hybrid Distance Geometry-Dynamical Simulated Annealing Calculations"

49. Craik, D. J., Daly, N. L., Bond, T., and Waine, C. (1999) *J Mol Biol* **294**, 1327-1336. "Plant cyclotides: A unique family of cyclic and knotted proteins that defines the cyclic cystine knot structural motif"
Muller, Y. A., Christinger, H. W., Keyt, B. A., and de Vos, A. M. (1997) *Structure* **5**, 1325-1338. "The crystal structure of vascular endothelial growth factor (VEGF) refined to 1.93 Å resolution: multiple copy flexibility and receptor binding"

Taylor, W. R., Xiao, B., Gamblin, S. J., and Lin, K. (2003) *Comput Biol Chem* **27**, 11-15. "A knot or not a knot? SETting the record 'straight' on proteins"

Calabro, V., Sabatier, J. M., Blanc, E., Lecomte, C., Van Rietschoten, V., and Darbon, H. (1997) *J Pept Res* **50**, 39-47. "Differential involvement of disulfide bridges on the folding of a scorpion toxin"

Sato, A., Perlas, E., Ben-Menahem, D., Kudo, M., Pixley, M. R., Furuhashi, M., Hsueh, A. J., and Boime, I. (1997) *J Biol Chem* **272**, 18098-18103. "Cystine knot of the gonadotropin alpha subunit is critical for intracellular behavior but not for in vitro biological activity"

Hymowitz, S. G., Filvaroff, E. H., Yin, J. P., Lee, J., Cai, L., Risser, P., Maruoka, M., Mao, W., Foster, J., Kelley, R. F., Pan, G., Gurney, A. L., de Vos, A. M., and Starovasnik, M. A. (2001) *Embo J* **20**, 5332-5341. "IL-17s adopt a cystine knot fold: structure and activity of a novel cytokine, IL-17F, and implications for receptor binding"

Cemazar, M., Zahariev, S., Lopez, J. J., Carugo, O., Jones, J. A., Hore, P. J., and Pongor, S. (2003) *Proc Natl Acad Sci U S A* **100**, 5754-5759. "Oxidative folding intermediates with nonnative disulfide bridges between adjacent cysteine residues"

Cemazar, M., Zahariev, S., Pongor, S., and Hore, P. J. (2004) *J Biol Chem* "Oxidative folding of Amaranthus alpha-amylase inhibitor: Disulfide bond formation and conformational folding"

Wilken, J. A., and Bedows, E. (2004) *Biochemistry* **43**, 5109-5118. "Disulfide Bond Rearrangement during Formation of the Chorionic Gonadotropin beta-Subunit Cystine Knot in Vivo"

Daly, N. L., Clark, R. J., and Craik, D. J. (2003) *J Biol Chem* **278**, 6314-6322. "Disulfide folding pathways of cystine knot proteins. Tying the knot within the circular backbone of the cyclotides"

Muller, Y. A., Heiring, C., Misselwitz, R., Welfle, K., and Welfle, H. (2002) *J Biol Chem* **277**, 43410-43416. "The cystine knot promotes folding and not thermodynamic stability in vascular endothelial growth factor"

Darling, R. J., Ruddon, R. W., Perini, F., and Bedows, E. (2000) *J Biol Chem* **275**, 15413-15421. "Cystine knot mutations affect the folding of the glycoprotein hormone alpha-subunit. Differential secretion and assembly of partially folded intermediates"
FIGURE LEGENDS

**Figure 1.** Purification and MALDI-MS Analysis of N-terminal SMB Domain Produced from Endoproteinase Glu-C Digestion. **Figure 1A** shows a preparative injection of the endoproteinase Glu-C digest of native vitronectin with N-terminal peptides, 1-47 and 1-55, indicated (arrows). Vitronectin was S-carbamidomethylated with iodoacetamide to block free cysteines and digested with endoproteinase Glu-C. The resulting peptides were separated by reverse phase HPLC on a C18 column as described in Materials and Methods, and individual peaks were collected for further analysis. **Figure 1B** shows the reverse phase HPLC profile of VN 1-55 re-purified from the initial HPLC separation (**Panel 1A**) of the endoproteinase Glu-C digest containing the SMB domain of vitronectin. **Figure 1C** shows the MALDI-MS of the reverse-phase purified VN 1-55, with an observed average mass of 6284.2. The spectra were smoothed with a 19-point Savitsky-Golay filter. **Figure 1D** shows an electrospray MS-MS analysis on the purified VN 1-55. From multiple charge states in the mass spectrum (not shown), the average mass was measured as 6281.2, in agreement with the MALDI-MS (**Panel 1C**) and the expected mass of 6280.9. The +5 charge-state ion at m/z 1257.3 was isolated using a 5 m/z isolation width, and subjected to collision-induced dissociation. The masses of fragment ions are given, along with the amino acid of molecular mass that corresponds to the difference in mass between adjacent species.

**Figure 2.** MALDI-MS Analysis of the Partial Reduction and Alkylation of VN 1-55. VN 1-55 was partially reduced with TCEP at pH 2.8 followed by alkylation with NEM at pH 6.5. **Figure 2A** shows the progress of the reduction and alkylation after 30 minutes. The mass increases by 252 Da with the reduction and alkylation of each disulfide bond with three open bonds clearly shown at 30
Figure 2B shows the progress at 60 minutes where all of the isoforms are present including the fully reduced and alkylated form.

**Figure 3. Strategy for Partial Reduction of SMB, Followed by Blocking of Freed Cysteines with NEM.** Shown schematically are the starting material, fully oxidized SMB, and partially reduced species that have either one, two, or three cystine bonds opened. NEM was used to block the freed cysteines that are liberated in each of these species. The partially reduced, blocked intermediates that were analyzed further in this work are labeled as SMB1O, SMB2O, and SMB3O for the partially reduced and NEM-blocked species with either one, two, or three cystines opened with TCEP, respectively. Gray circles represent cysteines in a disulfide linkage, aqua circles represent TCEP-reduced and NEM-blocked cysteines, and red lines represent disulfide bridges in the SMB domain.

**Figure 4. Purification of Individual Partially Reduced Isoforms of VN 1-55.** Figure 4A shows the reverse phase HPLC separation of the isoforms of VN 1-55 created during the partial reduction and alkylation. SMB is the fully oxidized SMB domain of vitronectin. As outlined schematically in Figure 3, SMB10 is the singly reduced isoform, shown here in both singly and doubly alkylated forms. SMB20 is the doubly reduced isoform that has been fully alkylated. SMB30 is the triply reduced isoform, and SMB40 is the fully reduced isoform shown here with one free cysteine as well as the fully alkylated form. A linear gradient was used at 0.5 ml/min from 5 to 55% buffer B over 107 min, as described in Materials and Methods. Figures 4B-D show the reverse phase HPLC profiles of the partially purified isoforms representing 1, 2 and 3 open disulfides, respectively. Figure 4B, the re-purification of SMB1O, used a gradient from 15 to 45% B over 50 min at a flow
rate of 0.5 ml/min. Figure 4C, the re-purification of SMB2O, used a gradient from 18 to 45% B over 50 min at a flow rate of 0.5 ml/min. Figure 4D, the re-purification of SMB3O, used a gradient of 20 to 50% B over 50 min at a flow rate of 0.5 ml/ml. Shown beside Panels B through D are schematic figures depicting SMB1O, SMB2O and SMB3O, respectively (see Figure 3).

**Figure 5. Strategy for Proteolysis and MALDI-MS Analysis on Digestion Products from SMB1O.** This scheme outlines the experimental procedures that were performed to identify free cysteines and disulfide bridges in the SMB domain with one cystine opened with TCEP, and the two liberated cysteines blocked with NEM (SMB1O). This SMB isoform was treated in two ways. In one strategy, the SMB1O was digested with trypsin, and the digest products were identified by mass analysis using MALDI-MS. A unique peak was identified that corresponded to two cleavage fragments connected by a disulfide, denoted as SMB1O-T1/2. In the second approach, the SMB1O fragment was fully reduced with DTT and the newly liberated cysteines were blocked with IAM, as described in Materials and Methods. This fully reduced, and differentially blocked SMB isoform, was then digested with chymotrypsin, and digest products were analyzed by MALDI-MS. Three fragments were identified, denoted SMB1O-Ch1, SMB1O-Ch2, SMB1O-Ch3. Table 1 lists expected masses and observed masses for all fragments depicted in this scheme.

**Figure 6. Strategy for Proteolysis and MALDI-MS Analysis on Digestion Products from SMB2O.** This scheme outlines the experimental procedures that were performed to identify free cysteines and disulfide bridges in the SMB domain with two cystines opened with TCEP, and the four liberated cysteines blocked with NEM (SMB2O). This SMB2O fragment was fully reduced with DTT and the newly liberated cysteines were blocked with IAM, as described in Materials and Methods. This fully reduced, and differentially blocked SMB isoform, was then digested with
chymotrypsin, and digest products were analyzed by MALDI-MS. Three fragments were identified, denoted SMB2O-Ch1, SMB2O-Ch2, and SMB2O-Ch3. **Table 1** lists expected masses and observed masses for all fragments depicted in this scheme.

**Figure 7. Strategy for Proteolysis and MALDI-MS Analysis on Digestion Products from SMB3O.** This scheme outlines the experimental procedures that were performed to identify free cysteines and disulfide bridges in the SMB domain with three cystines opened by TCEP, and the six liberated cysteines blocked with NEM (SMB1O). This SMB3O fragment was fully reduced with DTT and the newly liberated cysteines were blocked with IAM, as described in Materials and Methods. This fully reduced, and differentially blocked isoform, was then digested with two enzymes, endoproteinase Glu-C and chymotrypsin, and digest products were analyzed by MALDI-MS. Two fragments were identified, denoted SMB3O-GCh1 and SMB3O-GCh2. **Table 1** lists expected masses and observed masses for the fragments depicted in this scheme.

**Figure 8. Distinguishing Two Disulfide Pairing Possibilities in the Core of the SMB Domain by NMR Measurements and Refinement.** Simulated annealing calculations were used to distinguish two possibilities for the disulfide-bonding patterns between C19, C21, C31 and C32. Calculations were performed as described in Materials and Methods using proton assignments, measured NOEs, and the C5:C9 and C25:C39 disulfide bonds as restraints. Each calculated structure was compared to all other calculated structures by means of a two-dimensional array. The resulting grid was color coded to represent the calculated RMSD relating pairs of structures in the matrix. For each set of calculations, the color range is yellow (0 – 2.5 Å), blue (2.5 – 5 Å), and cyan (>5 Å). **Panel A** is a comparison of 61 structures calculated with C19:C31 and C21:C32 disulfides in the core region. **Panel B** is a comparison of 38 structures with C19:C32 and C21:C31 disulfides in the core region.
Figure 9. Pattern of Disulfides in Somatomedin B from Human Plasma Vitronectin. A schematic of the sequence of the vitronectin SMB domain with assigned disulfides is shown for comparison with the typical disulfide pattern for the growth factor type cystine knots (Figure 9A, (27)). Local sequence motifs that are common between the growth factor type cystine knots and the SMB domain are color coded, revealing a different ordering between the two sequences. The assignment of disulfides determined in this work for the SMB domain from human plasma vitronectin is depicted in Figure 9B. This assignment of cystines agrees with the solution structure determined on the native domain from human plasma vitronectin (18).¹The arrangement of disulfides originally proposed for a recombinant form of the SMB domain by Kamikubo et al. (28) is shown in Figure 9C. This group has recently published work that proposes that at least 4 other disulfide patterns for rSMB are compatible with their NMR structure (19).²The disulfide pattern observed by Zhou and coworkers for a recombinant SMB in a co-crystal with PAI-1 is shown in Figure 9D (17).
| Peptide$^a$ | Identification$^b$ | Observed Mass$^c$<br>($m/z$ avg) | Expected Mass$^c$<br>($m/z$ avg) | Disulfide Assignment |
|---|---|---|---|---|
| $^{1}$DQESCK$^6$<br>$^{7}$GRCTEGFNVDDK$^{18}$ | SMB10-T1/2 | 2063.4 +/- 1.1 (n=2) | 2063.3 | $^{C5}$:$^{C9}$ |
| $^{1}$DQESC$^{CAM}$KGRC$^{CAM}$TEGFNV$^{15}$<br>$^{16}$DKKC$^{CAM}$QC$^{CAM}$DEL$^{24}$ | SMB10-Ch1 | 2964.6 +/- 0.2 (n=3) | 2965.3 | $^{C25}$:$^{C39}$ |
| $^{25}$C$^{NEM}$SYYQSC$^{CAM}$C$^{CAM}$TDY$^{35}$ | SMB10-Ch2 | 1576.1 +/- 0.3 (n=3) | 1575.7 |
| $^{29}$QSC$^{CAM}$C$^{CAM}$TDY$^{35}$ | SMB10-Ch3 | 933.9 +/- 0.1 (n=3) | 934.0 |
| $^{36}$TAEC$^{NEM}$KPVTRGDVF$^{49}$ | SMB20-Ch1 | 1677.4 +/- 0.6 (n=3) | 1676.9 | $^{C25}$:$^{C39}$ |
| $^{14}$NVDKKC$^{CAM}$QC$^{CAM}$DEL$^{24}$ | SMB20-Ch2 | 1409.2 +/- 0.2 (n=3) | 1409.6 |
| $^{28}$YQSC$^{CAM}$C$^{CAM}$TDY$^{35}$ | SMB20-Ch3 | 1097.0 +/- 0.4 (n=2) | 1097.2 |
| $^{14}$NVDKKKCQCDE$^{23}$<br><br><span>$^a$CAM + NEM</span> | SMB30-GCh1 | 1364.0 +/- 0.6 (n=3) | 1364.5085 | $^{C19}$:$^{C31}$ + $^{C21}$:$^{C32}$<br>or<br>$^{C19}$:$^{C32}$ + $^{C21}$:$^{C31}$ |
| $^{29}$QSCCTDYTAEE$^{38}$<br><br><span>$^a$CAM + NEM</span> | SMB30-GCh2 | 1304.4 +/- 0.4 (n=3) | 1303.4 |

$^a$Sequences of peptides from the SMB domain are given, with the corresponding numbers in the amino acid sequence. Cysteines labeled with CAM (carbamidomethyl) or NEM (N-ethylmaleimide) are shown by subscripts. Boxes in the bottom row indicate that one cysteine in the peptide is labeled with CAM and the other with NEM, but it is not possible to distinguish these possibilities from MALDI data alone.

$^b$Nomenclature for identified peptides correlates with the schemes shown in Figures 5-7.

$^c$Observed m/z ratios by MALDI are given mean +/- standard deviations. The number of replicate samples measured is shown in parentheses.
Figure 1
Figure 2

A

B

Counts

Mass (m/z)

Counts

Mass (m/z)
Figure 6
Figure 7
A. Consensus growth factor cystine knot

\[ C_1-(X)_n-C_2-X-G-X-C_3-(X)_n-[C]-C_4-(X)_n-C_5-X-C_6 \]

B. SMB domain from plasma vitronectin

\[ \text{NH}_2-X_4-C_1-X-G-X-C_2-X_9-C_3-X-C_4-X_3-C_5-X_5-C_6-C_7-X_6-C_8-X_11-\text{COOH} \]

C. rSMB\(^1\)

\[ \text{NH}_2-X_4-C_1-X-G-X-C_2-X_9-C_3-X-C_4-X_3-C_5-X_5-C_6-C_7-X_6-C_8-X_11-\text{COOH} \]

D. rSMB\(^2\)

\[ \text{NH}_2-X_4-C_1-X-G-X-C_2-X_9-C_3-X-C_4-X_3-C_5-X_5-C_6-C_7-X_6-C_8-X_11-\text{COOH} \]
SUPPLEMENTAL DATA

Manuscript: “Assignment of the Four Disulfides in the N-terminal Somatomedin B Domain of Native Vitronectin Isolated from Human Plasma,” by N. A. Horn, G. B. Hurst, A. Mayasundari, N. A. Whittemore, E. H. Serpersu and C. B. Peterson

The supplemental data consist of 6 figures.
Supplemental Figure 1. Identification of SMB1O-T1/2 by MALDI. Mass spectrometry was performed as described in experimental procedures. The spectra were smoothed with a 19-point Savitsky-Golay filter.
**Supplemental Figure 2. Identification of SMB2O-Ch1 and SMB2O-Ch2 by MALDI.**
Mass spectrometry was performed as described in experimental procedures. The spectra were smoothed with a 19-point Savitsky-Golay filter.
Supplemental Figure 3. Identification of SMB2O-Ch3 by MALDI. Mass spectrometry was performed as described in experimental procedures. The spectra were smoothed with a 19-point Savitsky-Golay filter.
Supplemental Figure 4. Identification of SMB2O-Ch1, SMB2O-Ch2, and SMB2O-Ch3 by MALDI. Mass spectrometry was performed as described in experimental procedures. The spectra were smoothed with a 19-point Savitsky-Golay filter.
Supplemental Figure 5. Identification of SMB3O-GCh2 by MALDI. Mass spectrometry was performed as described in experimental procedures. The spectra were smoothed with a 19-point Savitsky-Golay filter.
Supplemental Figure 6. Identification of SMB3O-GCH1 by MALDI. Mass spectrometry was performed as described in experimental procedures. The spectra were smoothed with a 19-point Savitsky-Golay filter.
Assignment of the four disulfides in the N-terminal somatomedin B domain of native vitronectin isolated from human plasma
Nancy A. Horn, Gregory B. Hurst, Anand Mayasundari, Neil A. Whittemore, Engin H. Serpersu and Cynthia B. Peterson

J. Biol. Chem. published online June 1, 2004

Access the most updated version of this article at doi: 10.1074/jbc.M405716200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2004/08/26/M405716200.DC1