Model for the Three-Dimensional Structure of Vitronectin: Predictions for the Multi-Domain Protein from Threading and Docking

Dong Xu,[1]* Kunnumal Baburaj,[2] Cynthia B. Peterson,[2] and Ying Xu[1]
[1]*Computational Biology Section, Life Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee*
[2]*Department of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, Tennessee*

ABSTRACT **The structure of vitronectin, an adhesive protein that circulates in high concentrations in human plasma, was predicted through a combination of computational methods and experimental approaches. Fold recognition and sequence–structure alignment were performed using the threading program PROSPECT for each of three structural domains, i.e., the N-terminal somatomedin B domain (residues 1–53), the central region that folds into a four-bladed β-propeller domain (residues 131–342), and the C-terminal heparin-binding domain (residues 347–459). The atomic structure of each domain was generated using MODELLER, based on the alignment obtained from threading. Docking experiments between the central and C-terminal domains were conducted using the program GRAMM, with limits on the degrees of freedom from a known inter-domain disulfide bridge. The docked structure has a large inter-domain contact surface and defines a putative heparin-binding groove at the inter-domain interface. We also docked heparin together with the combined structure of the central and C-terminal domains, using GRAMM. The predictions from the threading and docking experiments are consistent with experimental data on purified plasma vitronectin pertaining to protease sensitivity, ligand-binding sites, and buried cysteines. Proteins 2001;44:312–320.**

Key words: vitronectin; protein structure prediction; domains; threading; docking; heparin

INTRODUCTION

Human plasma vitronectin (also known as serum spreading factor, S-protein, or epibolin) is a multifunctional glycoprotein found in blood and in the extracellular matrix. It interacts with a wide variety of structurally dissimilar ligands. Examples include glycosaminoglycans (e.g., heparin), serine protease inhibitors (e.g., plasminogen activator inhibitor type-1 (PAI-1), or the antithrombin–thrombin complex), cell surface receptors (including several integrins and the urokinase-receptor), extracellular matrix constituents (e.g., collagen), and components of the complement system. The interactions between these biomolecules and vitronectin control several important physiological processes. These include blood coagulation and fibrinolysis, cell adhesion and cellular migration, tumor metastasis and tissue remodeling, and modulation of the immune system.

Extensive studies have been carried out on vitronectin since its discovery in 1967 (reviewed by Preissner[1] and Schvartz et al.[2]). Many of the studies have focused on localizing binding sites for the various ligands. From sequence comparison, limited proteolysis, peptide mapping, immunochemical studies, and work with recombinant fragments, a general idea about the organization of vitronectin has emerged. The structure of the 459-amino acid protein is thought to be organized into three domains that provide the necessarily broad repertoire of binding epitopes for target ligands.

The N-terminal domain (residues 1–53) consists of 44 amino acids that are identical to the circulating protein somatomedin B, whose structure is unknown. This domain does not match any protein of known structure with significant sequence similarity using sequence–sequence comparisons, such as BLAST and PSI-BLAST.[3] The N-terminal somatomedin B domain is likely the key binding site for integrins,[4] PAI-1,[5–8] and the urokinase/plasminogen activator receptor.[9] A central domain is thought to mediate binding to some bacteria.[10–12] There is a long linker between the somatomedin B domain and the central domain, i.e., residues 54–130. The secondary structure prediction program PHD[13] predicts it in the loop conforma-

Abbreviations: PAI-1, plasminogen activator inhibitor type-1; DABIA, 4-dimethylaminoazo-benzene-4′-iodoacetamide; HPLC, high-pressure liquid chromatography.

Grant sponsor: Office of Health and Environmental Research, U.S. Department of Energy/UT-Battelle, LLC; Grant number: DE-AC05-000R22725; Grant sponsor: National Institutes of Health; Grant number: HL50676.

The coordinates of structural models referred to in this article can be found at http://compbio.ornl.gov/~nxy/vitronectin/ as Supplementary Material.

Kunnumal Baburaj's present address is Laboratory of Cell Biology, Building 3-Room 302, National Institutes of Health, Bethesda, MD 20891.

*Correspondence to: Dong Xu, Computational Biology Section, Life Sciences Division, 1060 Commerce Park Drive, Oak Ridge National Laboratory, Oak Ridge, TN 37831-6480. E-mail: xud@ornl.gov

Received 10 November 2000; Accepted 15 March 2001

tion. A highly charged sequence near the C-terminus mediates binding to heparin[14–16] and collagen.[17,18] The C-terminal domain containing the heparin-binding sequence also may serve as a binding site for plasminogen,[19] complement factors,[20] and perforin,[20,21] and as a secondary binding site for PAI-1[22–24] and/or the urokinase receptor.[9] Both the central domain (residues 131–342) and the C-terminal domain (residues 347–459) have the signature of hemopexin,[25] whose structure is known as a four-bladed β-propeller fold. However, sequence–sequence comparisons did not produce an unambiguous sequence alignment across the whole domain in either case.

In spite of the keen interest in structure/function issues with vitronectin, the three-dimensional structure of the protein has not been solved experimentally. One obstacle to crystallographic studies on vitronectin comes from the conformational lability of the protein that has been observed both in vitro and in vivo. Indeed, although most circulating vitronectin is monomeric, the protein also adopts an altered, oligomeric conformation. The self-associating behavior of vitronectin is observed upon in vitro unfolding/refolding either by heat or chemical denaturation.[26–29] This propensity for the altered conformation of vitronectin to form oligomers presumably has biological relevance since such higher-ordered multimeric forms of the glycoprotein have been detected in the extracellular matrix,[28] platelet releasates,[30] and within the α-granules of platelets.[30] Another hindrance to structural work stems from the heterogeneity that arises in the protein due to extensive posttranslational modification (e.g., N-linked glycosylation, phosphorylation, and sulfation). The lack of a structural model is a major bottleneck in further understanding the function of vitronectin and its associated domains.

In this article, we propose an atomic model of vitronectin using computational methods, and compare the model with known experimental information. We have adopted the domain assignment described above for the N-terminal, central, and heparin-binding domains; we predict the structure of each domain separately. The primary method used for the prediction is fold recognition through sequence–structure alignment (threading). We employed the program PROSPECT (PROtein Structure Prediction and Evaluation Computer Toolkit) developed at Oak Ridge National Laboratory.[31,32] PROSPECT has demonstrated its good performance through extensive tests using known structures,[32] and a blind test in the third communitywide experiments on the Critical Assessment of Techniques for Protein Structure Prediction (CASP-3).[33,34] In addition, we have docked the central and C-terminal domains together to gain insight into the overall fold of the protein. We also docked heparin together with the combined structure of the central and C-terminal domains. The docking experiments provide clues to the understanding of the binding of heparin to vitronectin. All the structural models that we derived for vitronectin can be downloaded from http://compbio.ornl.gov/~nxy/vitronectin/.

METHODS

Labeling of Free Sulfhydryls on Vitronectin

Vitronectin was isolated from human plasma as described previously,[26] using a method adapted from the original purification scheme of Dahlback and Podack.[35] For labeling, vitronectin (1.5 mg) was dialyzed (overnight) against Tris buffer (0.1 M, pH 8.3), and guanidinium chloride (solid) was added to make up a concentration of 6 M, followed by the addition of 0.3 mg DABIA (Pierce Chemical Company) in dimethylformamide (200 μl). The mixture was stirred overnight at room temperature in the dark. The precipitated reagent was removed by centrifugation followed by filtration using a 0.45-μm Acrodisc filter. The solution of the labeled vitronectin and the dissolved unlabeled DABIA was then passed through a desalting column (Sephadex G25, 10 × 250 mm) to remove the excess reagent and guanidinium chloride. The chromatography was monitored at 280 nm.

Trypsin Cleavage

The DABIA-labeled protein was dialyzed into ammonium bicarbonate buffer (50 mM, pH 8.5), and 2% trypsin (w/w) was added. Proteolysis was allowed to proceed for 24 h at 37°C.

HPLC Separation of Trypsin-Cleaved Polypeptide Fragments

The trypsin-cleaved fragments of DABIA-labeled protein were subjected to high-performance liquid chromatography (HPLC) (Beckman system Gold) on a reversed-phase column (Novopac C_{18}, 3.9 × 150 mm). The solvent system consists of 0.1% trifluoroacetic acid (TFA) in water (A) and 0.1% TFA in 80% acetonitrile:water (B). The proportion of solvent B was increased from 0–70% in 80 min, then from 70–100% within 85 min and kept 100% during the next 10 min. The flow rate was 1 ml/min and the peptides were monitored at 220 nm. The DABIA-labeled fragments were identified by monitoring the absorbance at 436 nm. The peaks were collected manually; the purity of the labeled fractions was determined by rechromatography. The positions of free sulfhydryls were identified by sequencing the DABIA-labeled fragments, using the Edman degradation method. Peptide sequencing was performed at the Microchemical Facility at Emory University, Atlanta, GA.

Computational Modeling

The threading using PROSPECT was performed on each sequence segment of the three domains. The knowledge-based energy function used in PROSPECT consists of three additive terms: (1) a singleton term, which measures the fitness of aligning a particular amino acid to a particular template-fold environment defined by secondary structure and solvent accessibility; (2) a pairwise interaction term, which measures the contact preference between a pair of amino acids on the query sequence assigned to nearby template positions; and (3) gap penalties for unaligned amino acids on the query sequence and template positions. PROSPECT guarantees to find a globally opti-

mal alignment under the given energy function.[31] Two sets of the known structure templates are used for threading: protein chains (defined by FSSP nonredundant set[36]) and compact domains (defined by the DALI nonredundant domain library[37]). A neural network assessment for the reliability of the fold recognition is given by PROSPECT based on individual energy terms and the threading energy distribution between the template and all the protein sequences in FSSP. The threading alignments were also compared with PSI-BLAST[3] alignment for the second and the C-terminal domains. The atomic structures were generated using MODELLER[38] based on the alignments obtained by threading.

Docking was carried out between the central and C-terminal domains, as well as between the heparin and the combined structure of the central and C-terminal domains using the computer package GRAMM.[39] The program performs an exhaustive six-dimensional search through the relative translations and rotations of the molecules. GRAMM is particularly suitable in our case, as it allows inputs of low-resolution structures like our predicted domain structures. We have used a nuclear magnetic resonance (NMR) structure for heparin [Protein Data Bank (PDB) code 1hpn] [40] in docking. The structures have been visualized and presented in this article, using VMD.[41]

RESULTS

Identification of Free Sulfhydryls in Vitronectin

Important information for the structural predictions on vitronectin will be gained from the complete assignment of free sulfhydryls and disulfide bonds in vitronectin from experimental approaches. A first step toward this goal was an effort to determine the sequence identities of the two free sulfhydryls that are buried in the native folded protein.[27] To provide a chromophoric tag on these residues, and to prevent disulfide rearrangement from occurring during the digestion, the free sulfhydryls were labeled with DABIA.[42] The labeled vitronectin was then digested with trypsin, and the peptides were separated chromatographically. An HPLC profile from a reversed-phase separation of the digest of DABIA-labeled vitronectin on a C_{18} column is presented in Figure 1. Note that there are two prominent peaks with absorbance at 430 nm, indicating the presence of the DABIA label. Both peaks have been sequenced with the following results:

Peak 1: a. DVWGI; b. GQYLY . . .

Peak 2: MDWLVPAT?EPIQS

The missing residue in the sequence of peak 2 corresponds to Cys-411. The sequencing results from peak 1 show that it is actually a mixture of two trypsin digest fragments. Peak 1a begins at amino acid 179, corresponding to a fragment that contains Cys-196. Peak 1b corresponds to the sequence of vitronectin beginning at position 158 (although the residue corresponding to Cys-161 was recognized by the instrument as a leucine, perhaps because of modification of the cysteine residue). From these data, we can firmly assign one of the free sulfhydryls to Cys-411, and the second residue is assumed to be 196, as there was an amino acid present in the cycle corresponding to residue 161. In support of this assignment is a previous report that has identified one of the six disulfides in vitronectin as forming a covalent interaction between Cys-137 and Cys-161.[43]

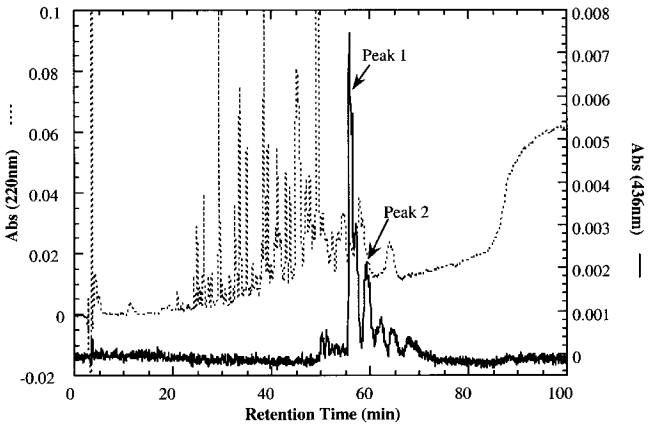

Fig. 1. High-performance liquid chromatography (HPLC) separation of trypsin digest of DABIA-labeled vitronectin. Separation of the peptides on a reversed-phase C_{18} column using an acetonitrile gradient, as described in Methods, with detection of all peptides in the digest at 220 nm (hatched line) and with specific detection of DABIA-labeled peptides at 436 nm (solid line).

Folding Predictions for the Three Domains

PROSPECT was used to determine predicted structures for each of the three domains of vitronectin. The threading results are summarized in Table I, and the structural models are shown in Figure 2a–c. As shown in the linear schematic of vitronectin in Figure 2d, the structural predictions for the domains do not include the connecting region between the N-terminal and central domains, which is predicted to be highly unstructured as discussed above. Another short linker between the central and C-terminal domains, comprising residues 324–353, was not modeled. Results for the individual domains are summarized below.

Somatomedin B domain (Fig. 2a)

We threaded the sequence segment residues 1–53 against the FSSP library and the DALI domain library, using PROSPECT. The best hit in terms of the most favorable total energy (as described in Methods) in both libraries is the PDB entry 1afp (antifungal protein), an open β-barrel fold with five strands and contains three disulfide bonds. The alignment, as shown in Figure 3, has a sequence identity of 19% between the query protein and the template. The neural network assessment for the threading reliability shows that the probability for 1afp to be the correct fold is >50%. There is no hit to the sequence of any known structure using PSI-BLAST. We also performed fold recognition using tools other than PROSPECT, but no better template than 1afp was found to cover the domain of residues 1–53. For example, PDB-Blast[44] found the template of anticoagulant phospholipase (1b4w with 122 residues). But the structure model for vitronectin based on the

TABLE I. Summary of Threading Results for the Three Domains

Domain	N-Terminal Domain	Central Domain	C-Terminal Domain
Residue range	1–53	131–342	347–459
Template	1afp	1gen	1gen
Total score	−827.4	−1402.1	−238.8
Mutation score	−935.0	−2617.0	−1286.0
Singleton score	−122.4	−1050.7	−670.8
Pairwise score	−319.9	−17.7	−106.2
Secondary structure score	155.1	−216.2	305.5
Gap penalty	394.8	2499.5	1518.7
Sequence identity	19%	21%	17%
Confidence level	>50%	>99%	>90%

PDB-Blast alignment does not have a globular shape, as expected for a protein domain. Furthermore, GenTHREADER[45] found the template of metallothionein (1dmc) as the best hit, but this structure has only 31 residues and thus does not encompass the entire N-terminal domain of vitronectin.

Central four-bladed β-propeller domain (Fig. 2b)

The sequence segment of residues 131–342 was threaded against the FSSP database. The best hit in terms of total energy for this domain is the PDB entry 1gen (C-terminal domain of gelatinase A), which has a full four-bladed β-propeller fold. The alignment, as shown in Figure 4, has a sequence identity of 21% between the query protein and the template. It is basically in agreement with one of several PSI-BLAST alignments for vitronectin (detected using a cutoff expectation value of 0.001). The regions that have slightly different alignments between PROSPECT and PSI-BLAST are indicated in lowercase in Figure 4. The variation may indicate that the structure prediction for these regions has low confidence. The sequence segment (residues 131–342) also has a similar number of residues as a template of the four-bladed β-propeller fold (~200 residues). The neural network assessment for the threading reliability indicates that it is certain that this domain adopts a fold of 1gen (the probability for 1gen to be the correct fold is >99%). A theoretical model derived from an automated pipeline is available for this domain (residues 128–323) in ModBase.[46] It was based on the template 1ck7, which has the same fold of 1gen. The root-mean-square deviation (RMSD) of all heavy atoms for residues 129–323 between our model and the model in ModBase is 6.47 Å, attributable primarily to some minor alignment differences. We believe that our model is better because the alignment in ModBase was automatically retrieved from one of many top PSI-BLAST hits that have similar scores, while our alignment (also among top PSI-BLAST hits) is supported by PROSPECT.

C-terminal heparin-binding domain (Fig. 2c)

The sequence segment of residues 347–459 was threaded. Again, the best hit in terms of total energy for this domain is the PDB entry 1gen. However, the sequence segment covers only one-half of the four-bladed β-propeller fold. Figure 5 gives the alignment, which has a sequence identity of 17% between the query protein and the template. The alignment is basically in agreement with one of several PSI-BLAST alignments for vitronectin (detected using a cutoff expectation value of 0.001). One region has slightly different alignments between PROSPECT and PSI-BLAST, and it is indicated in lowercase in Figure 5. Again, the differences resulting from the two algorithms are not coincidental and presumably reflect low confidence in the predictions for these regions. The neural network assessment for the threading reliability indicates that the probability that 1gen is the correct fold in this domain is >90%.

Docking of the Central and C-Terminal Heparin-Binding Domains

Vitronectin normally circulates as a mixture of single-chain and two-chain, disulfide cross-linked forms. The two-chain form of vitronectin is produced by proteolysis within a region of the protein that lies on the C-terminal side of the heparin-binding sequence, producing a two-chain form of the protein with a heavy chain and a C-terminal light chain of approximately 10K molecular weight. This C-terminal 10K fragment contains two cysteines at positions 411 and 453 and is disulfide cross-linked to the heavy chain of the protein. Since our study has demonstrated that Cys-411 is free, we conclude that buried sulfhydryl residue, Cys-453, is the amino acid that makes the disulfide bridge from the light chain to the remainder of the protein. The most important constraint used for the docking was this inter-domain disulfide cross-link that connects heavy and light chains in the two-chain form of vitronectin. The cysteine residue involved in the inter-domain disulfide pair was chosen as residue 274 on the basis of the work presented in this article on free sulfhydryls that indicates that Cys-196 is free. This assignment of the 274–453 disulfide bond is also supported by the assignment of an intra-domain disulfide bond in the central region between cysteines 137 and 161.[43]

Docking studies on the two β-propeller domains with the intra-domain disulfide were conducted using GRAMM. The top 100 docking conformations were generated. There is no prediction reliability assessment in GRAMM. The conformations that satisfy the constraints were similar, and we picked the one with the best docking score (in terms of good contact and few penetrations between the two domains) from them. The docked conformation be-

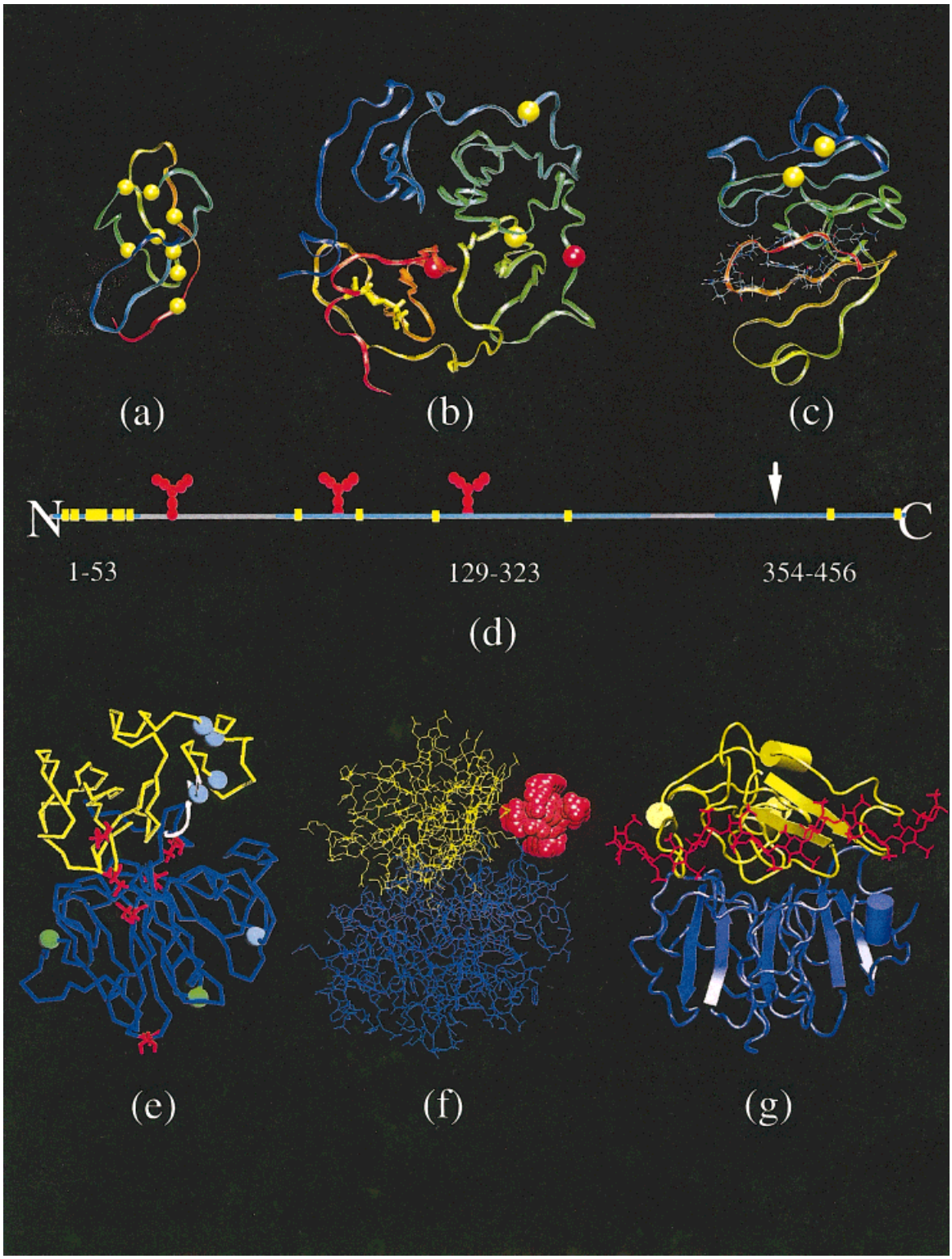

Fig. 2. Structural models of vitronectin. (a–c) show the structural models of the N-terminal somatomedin B domain, the central domain with a four-bladed β-propeller fold, and the C-terminal heparin-binding domain, respectively. The color from the red to blue shows the sequence order from the N-terminus to the C-terminus. The yellow solid spheres indicate cysteines. White spheres (**a**) show the integrin attachment site; red spheres (**b**) show glycosylation sites (residues 150 and 223); and thin lines (**c**) show the heparin-binding site. **d:** Linear representation of the sequence of vitronectin, with blue highlighting corresponding to domains modeled in (a–c), yellow dots indicating cysteines, red structures representing carbohydrate attachment sites, and the white arrow pointing to the known site of protease cleavage to the two-chain form. **e:** Docking structure between the central (blue) and C-terminal (yellow) domains, with cysteines (red lines), presumed heparin-binding residues (354–363 in white ribbons), sites susceptible to protease (residues 305, 361, 370, 379 and 383 with light blue spheres), and N-linked glycosylation sites (residues 150 and 223 with green spheres). **f,g:** Predicted docking conformation between the heparin (red) and the combined structure of the central (blue) and C-terminal (yellow) domains from two different perspectives.

```
Query: 1 DQESCKGRCTEGFNVDKKCQCD------ELCSYYQSCCTDYTAECKPQVTRGDVFTMPE 53
               |:|       |  |           :|  |    |     | |      :|  :
Templ: 1 --ATYNGKCYKK---DNICKYKAQSGKTAICKCYVKKCPRDGAKCEFDSYKGKCYC--- 51
```

Fig. 3. Threading alignment between the somatomedin B domain (indicated by Query) and the template 1afp (indicated by Templ), where | and : show the two aligned residues are identical and similar, respectively. Numbering for the template is the same as that used in the PDB file.

```
Query: 123 LHPGRPQPPAEEELCSGKP-FDAFTDLKNGSLFAFRGQYCYELDEKAVRPGYPKLIRDVW 181
                    |   |:|     ||    :: | :| |: :: :       :|  | |:   :
Templ: 461 ------LGPVTPEICKQDIVFDGIAQIR-GEIFFFKDRFIWRTVTPRDKPMGPLLVATFW 513

Query: 182 G-iegpidaaftrincQGKTYLFKGSQYWRFEDGVLDPDYPRNI-SDGFDGIPDNVDAAL 239
             :   ||| :     : |   | |::|| :    |:  ||: : | |       ||||
Templ: 514 PELPEKIDAVYEAPQ-EEKAVFFAGNEYWIYSASTLERGYPKPLTSLGLPPDVQRVDAAF 572

Query: 240 ALPAHSYSGRERVYFFKGKQYWEYQFQHQPSQEECEGSSLSAVFE-----hfammqrdsw 294
                ::|  :: | | | ::| |    :            |           :  :
Templ: 573 -----NWSKNKKTYIFAGDKFWRYNEVKKKMDPGFPKLIADAWNAIPDNLDAVVDLQGGG 627

Query: 295 edife----LLFWGRTSAGTRQPQFISRDWHGV 323
               |       |      |  : :   |  || |
Templ: 628 HSYFFKGAYYLKLENQSLKSVKFGSIKSDWLGC 660
```

Fig. 4. Threading alignment between the four-bladed β-propeller domain (indicated by Query) and the template 1gen (indicated by Templ), where | and : show the two aligned residues are identical and similar, respectively. Region shown in lowercase on the query sequence indicates different alignments between PROSPECT and PSI-BLAST. Numbering for the template is the same as that used in the PDB file.

```
Query: 354 HRNRKGYRSQRGHSRGRNQNSRRPSRATWLSlfsseesnlgannYDDYRMDWLVPATCE 412
            :  |             : ::       :    :|           |   :|
Sbjct: 527 PQEEKAVFFAGNEYWIYSASTL---ERGYPKPLTSLGLPPDVQRV-DAAFNW------S 575

Query: 413 PIQSVFFFSGDKYYRVNLRTRRVDTVDPPYPRSIAQYWLGCPAP 456
              :  : |:|||::| |   ::   :|| :|: ||  |   |
Sbjct: 576 KNKKTYIFAGDKFWRYNEVKKK---MDPGFPKLIADAWNAIPDN 616
```

Fig. 5. Threading alignment between the heparin-binding domain (indicated by Query) and the template 1gen (indicated by Templ), where | and : show the two aligned residues are identical and similar, respectively. Regions shown in lowercase on the query sequence indicate different alignments between PROSPECT and PSI-BLAST. Numbering for the template is the same as that used in the PDB file.

tween the domains is shown in Figure 2e. Interestingly, it has a groove lined with known heparin-binding residues.[47] We then predicted the complex between heparin and the combined two-domain structure using GRAMM. Again, the top 100 docking conformations were generated. Among them, some conformations fit into the groove with good docking scores, and they essentially share the same conformation. Figure 2f,g shows the structure complexed with heparin from two different perspectives.

DISCUSSION

We have predicted the structures of the three domains in vitronectin. To our knowledge, the structure of the somatomedin B domain is the first model ever proposed. It has been suggested that the other two domains contain the hemopexin domain signature, but the alignment based on sequence comparison is ambiguous, and detailed study was not reported in the literature. The structure quality of our models is good based on the assessment using WHATIF[48] and PROCHECK.[49] Among the three domains, the models of the four-bladed β-propeller domain and the heparin-binding domain have very high confidence levels (with probabilities of >99% and >90%, respectively). Their alignments are in agreement with PSI-BLAST alignments, which use the sequence profile to significantly increase the underlying signal, while reducing noise in multiple sequence alignments. The secondary structures of the each domain basically agree with the ones predicted by several tools such as PHD[13] and the PSA Server.[50] The fold of the somatomedin B domain is thought to be correct with some degree of confidence (>50%), but it is indeed less reliable than the other two domains. The identification of the four disulfide bonds in this short region of sequence will be needed for further refinement of the model. The secondary structure predictions for this domain vary significantly between different methods, and therefore cannot be used to assess the model.

The two docking structures (one between the central and C-terminal domains, and one between the heparin and the combined structure of the central and C-terminal domains) agree with known experimental data. They also have good packing in terms of few penetration and significant contacts between the two parts in a complex. These show that our docking models are reasonable.

SUMMARY

Although there have been no direct structural determinations on vitronectin or its domains, substantial informa-

tion from biochemical characterization of the protein can be used to constrain and/or test the model that is predicted from the computational approach. Indeed, any viable model should agree with information that has been gained experimentally. The most useful information for these purposes comes from disulfide and sulfhydryl identities identified in this study and others,[43] limited proteolysis studies used to define domains within the long polypeptide,[14,15,51] and heparin binding to vitronectin.[47,52]

Information on Sulfhydryl and Disulfide Identification

Information regarding the sequence assignment of the disulfide bonds and free sulfhydryls in vitronectin is an important determinant of the three-dimensional organization of the molecule. Previous work had demonstrated that 2 of the 14 cysteines in vitronectin are free and that they occupy buried positions,[27] while the rest form 6 disulfide bonds in the protein. The docked structures shown in Figurer 2d–f contain two of the six disulfides in vitronectin, as well as the two free sulfhydryls. The inter-domain disulfide bridge between residues 274 and 453 was one of the main constraints on the docked model, helping define the extensive contact surface between the two domains. The intra-domain 137–161 disulfide bridge (the only experimentally identified disulfide in the protein) is compatible with the structure predicted for the four-bladed propeller structure of the central domain. Also, the known burial of Cys-411 within the folded structure of vitronectin is in agreement with the predicted model shown here. It should be noted that a buried orientation of Cys-196 is not as obvious from the model; however, the two-domain docked structure does not represent the full-length protein, so contacts with the somatomedin B or linking region could affect the exposure of this residue.

Regions of Vitronectin That Are Susceptible to Proteases

Early studies aimed at defining domain boundaries and functions of vitronectin made extensive use of limited proteolysis. From these efforts, regions of high sensitivity to protease cleavage have been defined. The protease-sensitive sites previously identified within vitronectin are localized to two regions, one in the inter-domain linker from residues 44–90, and a second site near the C-terminus spanning residues in the vicinity of the heparin-binding site. These include thrombin cleavage sites on the C-terminal side of Arg-305 and Arg-370, a plasmin cleavage site following Arg-361, and elastase-sensitive sites following Ala-330 and Leu-383.[14,15,51] Note also that this C-terminal region contains the bond that is targeted for processing of vitronectin to the two-chain form upon cleavage following Arg-379 by an unidentified protease in vivo.[53] The known sites of protease sensitivity in this C-terminal region are shown in the docked model in Figure 2e with the blue spheres. These regions are exposed and accessible in the three-dimensional model for the fold of vitronectin, lending additional credence to this model as a valid one.

Details Regarding Heparin-Binding Sites on Vitronectin

Cleavage with cyanogen bromide was used in the mid 1980's to isolate a peptide and localize heparin binding to a highly basic region near the C-terminus of protein[16] that has been subsequently narrowed down to residues 340–359 in synthetic peptide studies.[14,54] Although this region of the protein was originally proposed to be buried or "cryptic" in the circulating, monomeric form of vitronectin, solution binding studies have since shown that this region of the protein is accessible.[47,52] Two-dimensional NMR studies on a model peptide derived from this region of the protein indicate that at least one of two arginine residues at positions 351 and 353 makes direct contacts with heparin.[47] The sequence of this peptide (from 354–363) is shown in the white ribbon structure in Figure 2e, defining a charged heparin-binding surface at the interface of the central and C-terminal domains. The coordinates of residues 351 and 353, which are not aligned to the template, are not included in the model. But it is feasible for them to lie at the base of the heparin-binding pocket between the domains, in a suitable position to make direct contact with heparin without violating the distance constraint to residue 354.

More recent studies using recombinant methods have suggested that sequences from the central domain of vitronectin exhibit binding of heparin, albeit weaker than the binding demonstrated for the C-terminal binding site.[55] The positioning of the heparin-binding pocket in this model at the interface between domains demonstrates the possibility that some specific contacts are made from the central domain to heparin. Thus, sequences from both domains together appear to define the heparin-binding region on vitronectin. Indeed, detailed studies of binding affinity and stoichiometry have shown that the stoichiometry of heparin binding to monomeric vitronectin is 1:1, indicating that any proposed binding sequences in other regions of the protein are not functional as "secondary" sites.[47,52] In this regard, it should be noted that the polypeptide sequences identified in recent phage display approaches (residues 82–137 and 175–219)[11] are remote from the heparin-binding pocket in our model and do not appear to contribute to the primary binding site on vitronectin.

Functions of the N-Terminal Domain

Site-directed mutagenesis has been used to evaluate important residues within the somatomedin B domain that are required for structure and PAI-1 binding.[9] For these alanine-scanning studies, a truncated form of vitronectin that corresponds to the somatomedin B domain was expressed in *Escherichia coli*. The results of this study demonstrated the importance of the four disulfide bonds for proper function of the protein. Furthermore, the results pointed to the involvement of other residues including Gly-12, Asp-22, Leu-24, Tyr-27, Tyr-28, and Asp-34 in binding to PAI-1. From the modeled structure, all these residues are found on the surface of the folded domain. The other prominent function associated with this region of the

protein is the integrin attachment site, corresponding to the RGD sequence at residues 45–47 (white spheres in Fig. 2a).[51,56] These three amino acids are found at the end of a tight turn in the modeled structure, also oriented on the surface of the molecule in a favorable orientation for binding to the cell-surface receptors.

CONCLUSIONS

Threading has been used successfully to provide the first three-dimensional working model for the structural domains of plasma vitronectin. These structural models are consistent with a large set of known experimental data, including positioning of ligand binding sites, accessibility of protease cleavage sites, and orientation of free sulfhydryls and disulfide bonds, as discussed above. Other practical considerations of the folded structure have been fulfilled with the model; for example, the central four-bladed β-propeller domain has two N-linked glycosylation sites (Asn-150 and Asn-223, shown in green in Fig. 2e),[57] both of which are on the surface of the domain. This positioning would be expected so that the attached carbohydrate side chains would protrude from the surface in a solvent-exposed orientation.

The β-propeller structure observed for the central domain of vitronectin characterizes the largest, but least well characterized domain of the protein in terms of function. One of the hallmarks of vitronectin is its ability to interact with a broad set of structurally distinct ligands. Several important ligands are thought to bind to this central region of the protein, although their positions for binding have not been well localized. Interestingly, β-sheet structures are often involved in intermolecular contacts between proteins, and the β-propeller structures would be ideal for these types of interactions. Many of the β-propeller structures that have been identified to date (several on the basis of homology) have known functions in governing protein–protein interactions. Some noteworthy examples are lectins[58] and the WD-40 domains in G-protein signaling molecules.[59]

The self-association of vitronectin into a multimeric form that functions in tissues and the extracellular matrix is one way in which the activity and localization of the protein is apparently regulated. Although there has been much speculation about the region(s) of the protein involved in this oligomerization, the process has not been well characterized. It has been proposed that heparin or the heparin-binding sequence in vitronectin mediate the association, but this has been disproved in biochemical and biophysical studies evaluating the effects of the ligand on urea denaturation.[26,27] In addition, it has been shown that the C-terminal 10K fragment is not required for self-association.[60] With the flexibility observed among β-propeller structures in incorporating four to eight β blades,[61] and with the example of domain–domain association via the β-sheet structures in our computed model for vitronectin, it can be postulated that this central β-propeller region may be involved in self-association. Support for this line of thinking comes from β-propeller structures predicted in other matrix proteins[62] and also demonstrated in lectins,[58] both examples in which oligomerization and multivalency have a long track record of biological importance.

ACKNOWLEDGMENTS

The authors thank Dr. Engin H. Serpersu for helpful discussions. The computational research was sponsored by the Office of Health and Environmental Research, U.S. Department of Energy, under contract DE-AC05-000R22725 managed by UT-Battelle, LLC. Funding also was provided by National Institutes of Health, grant HL50676 (to C.B.P.).

REFERENCES

1. Preissner KT. Structure and biological role of vitronectin. Annu Rev Cell Biol 1991;7:275–310.
2. Schvartz I, Seger D, Shaltiel S. Vitronectin. Int J Biochem Cell Biol 1999;31:539–544.
3. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25: 3389–3402.
4. Cherny RC, Honan MA, Thiagarajan P. Site-directed mutagenesis of the arginine–glycine–aspartic acid in vitronectin abolishes cell adhesion. J Biol Chem 1993;268:9725–9729.
5. Deng G, Royle G, Wang S, Crain K, Loskutoff DJ. Structural and functional analysis of the plasminogen activator inhibitor-1 binding motif in the somatomedin B domain of vitronectin. J Biol Chem 1996;271:12716–12723.
6. Seiffert D, Loskutoff DJ. Evidence that type 1 plasminogen activator inhibitor binds to the somatomedin B domain of vitronectin. J Biol Chem 1991;266:2824–2830.
7. Gibson AD, Baburaj K, Day DE, Verhamme I, Shore JD, Peterson CB. The use of fluorescent probes to characterize conformational changes in the interaction between vitronectin and plasminogen activator inhibitor-1. J Biol Chem 1997;272:5112–5121.
8. Seiffert D, Ciambrone G, Wagner NV, Binder BR, Loskutoff DJ. The somatomedin B domain of vitronectin. Structural requirements for the binding and stabilization of active type 1 plasminogen activator inhibitor. J Biol Chem 1994;269:2659–2666.
9. Deng G, Curriden SA, Wang S, Rosenberg S, Loskutoff DJ. Is plasminogen activator inhibitor-1 the molecular switch that governs urokinase receptor-mediated cell adhesion and release? J Cell Biol 1996;134:1563–1571.
10. Liang OD, Maccarana M, Flock JI, Paulsson M, Preissner KT, Wadstrom T. Multiple interactions between human vitronectin and *Staphylococcus aureus*. Biochim Biophys Acta 1993;1225:57–63.
11. Liang OD, Preissner KT, Chhatwal GS. The hemopexin-type repeats of human vitronectin are recognized by *Streptococcus pyogenes*. Biochem Biophys Res Commun 1997;234:445–449.
12. Liang OD, Rosenblatt S, Chhatwal GS, Preissner KT. Identification of novel heparin-binding domains of vitronectin. FEBS Lett 1997;407:169–172.
13. Rost B, Sander C. Prediction of protein secondary structure at better than 70% accuracy. J Mol Biol 1993;232:584–599.
14. Kost C, Stuber W, Ehrlich HJ, Pannekoek H, Preissner KT. Mapping of binding sites for heparin, plasminogen activator inhibitor-1, and plasminogen to vitronectin's heparin-binding region reveals a novel vitronectin-dependent feedback mechanism for the control of plasmin formation. J Biol Chem 1992;267:12098–12105.
15. Gechtman Z, Belleli A, Shaltiel S. The cluster of basic amino acids in vitronectin contributes to its binding of plasminogen activator inhibitor-1: evidence from thrombin-, elastase- and plasmin-cleaved vitronectins and antipeptide antibodies. Biochem J 1997; 325:339–349.
16. Suzuki S, Pierschbacher MD, Hayman EG, Nguyen K, Ohgren Y, Ruoslahti E. Domain structure of vitronectin. Alignment of active sites. J Biol Chem 1984;259:15307–15314.
17. Ishikawa M, Hayashi M. Activation of the collagen-binding of endogenous serum vitronectin by heating, urea and glycosaminoglycans. Biochim Biophys Acta 1992;1121:173–177.

18. Ishikawa-Sakurai M, Hayashi M. Two collagen-binding domains of vitronectin. Cell Struct Funct 1993;18:253–259.
19. Preissner KT. Specific binding of plasminogen to vitronectin. Evidence for a modulatory role of vitronectin on fibrin(ogen)-induced plasmin formation by tissue plasminogen activator. Biochem Biophys Res Commun 1990;168:966–971.
20. Tschopp J, Masson D, Schafer S, Peitsch M, Preissner KT. The heparin binding domain of S-protein/vitronectin binds to complement components C7, C8 and C9 and perforin from cytolytic T-cells and inhibits their lytic activities. Biochemistry 1988;27: 4103–4109.
21. Milis L, Morris CA, Sheehan MC, Charlesworth JA, Pussell BA. Vitronectin-mediated inhibition of complement: evidence for different binding sites for C5b-7 and C9. Clin Exp Immunol 1993;92:114–119.
22. Preissner KT, Grulich-Henn J, Ehrlich HJ, Declerck P, Justus C, Collen D, Pannekoek H, Muller-Berghaus G. Structural requirements for the extracellular interaction of plasminogen activator inhibitor-1 with endothelial cell matrix-associated vitronectin. J Biol Chem 1990;265:18490–18498.
23. Chain D, Korc-Grodzicki B, Kreizman T, Shaltiel S. Endogenous cleavage of the Arg-379–Ala-380 bond in vitronectin results in a distinct conformational change which buries Ser-378, its site of phosphorylation by protein kinase A. Biochem J 1991;274:387–394.
24. Gechtman Z, Sharma R, Kreizman T, Fridkin M, Shaltiel S. Synthetic peptides derived from the sequence around the plasmin cleavage site in vitronectin. Use in mapping the PAI-1 binding site. FEBS Lett 1993;315:293–297.
25. Hunt LT, Barker WC, Chen HR. A domain structure common to hemopexin, vitronectin, interstitial collagenase, and a collagenase homolog. Protein Seq Data Anal 1987;1:21–26.
26. Zhuang P, Blackburn MN, Peterson CB. Characterization of the denaturation and renaturation of human plasma vitronectin. I. Biophysical characterization of protein unfolding and multimerization. J Biol Chem 1996;271:14323–14332.
27. Zhuang P, Li H, Williams JG, Wagner NV, Seiffert D, Peterson CB. Characterization of the denaturation and renaturation of human plasma vitronectin. II. Investigation into the mechanism of formation of multimers. J Biol Chem 1996;271:14333–14343.
28. Stockmann A, Hess S, Declerck P, Timpl R, Preissner KT. Multimeric vitronectin. Identification and characterization of conformation-dependent self-association of the adhesive protein. J Biol Chem 1993;268:22874–22882.
29. Bittorf SV, Williams EC, Mosher DF. Alteration of vitronectin. Characterization of changes induced by treatment with urea. J Biol Chem 1993;268:24838–24846.
30. Seiffert D, Schleef RR. Two functionally distinct pools of vitronectin (Vn) in the blood circulation: identification of a heparin-binding competent population of Vn within platelet alpha-granules. Blood 1996;88:552–560.
31. Xu Y, Xu D, Uberbacher EC. An efficient computational method for globally optimal threading. J Comp Biol 1998;5:597–614.
32. Xu Y, Xu D. Protein threading using PROSPECT: design and evaluation. Proteins 2000;40:343–354.
33. Protein structure prediction issue. Proteins 1999;suppl. 3.
34. Xu Y, Xu D, Crawford OH, Einstein JR, Larimer F, Uberbacher EC, Unseren MA, Zhang G. Protein threading by PROSPECT: a prediction experiment in CASP3. Protein Eng 1999;12:101–109.
35. Dahlback B, Podack ER. Characterization of human S protein, an inhibitor of the membrane attack complex of complement. Demonstration of a free reactive thiol group. Biochemistry 1985;24:2368–2374.
36. Holm L, Sander C. Mapping the protein universe. Science 1996;273: 595–602.
37. Holm L, Sander C. Dictionary of recurrent domains in protein structures. Proteins 1998;33:88–96.
38. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol 1993;234:779–815.
39. Vakser IA. Low-resolution docking: prediction of complexes for underdetermined structures. Biopolymers 1996;39:455–464.
40. Mulloy B, Forster MJ, Jones C, Davies DB. N.M.R. and molecular-modelling studies of the solution conformation of heparin. Biochem J 1993;293:849–858.
41. Humphrey W, Dalke A, Schulten K. VMD: visual molecular dynamics. J Mol Graph 1996;14:33–38.
42. Krieglstein KA, Henschen A, Weller U, Habermann E. Arrangement of disulfide bridges and positions of sulfhydryl groups in tetanus toxin. Eur J Biochem 1990;188:39–45.
43. Skorstengaard K, Halkier T, Hojrup P, Mosher D. Sequence location of a putative transglutaminase cross-linking site in human vitronectin. FEBS Lett 1990;262:269–274.
44. Li W, Pio F, Pawlowski K, Godzik A. Saturated BLAST: an automated multiple intermediate sequence search used to detect distant homology. Bioinformatics 2000;16:1105–1110.
45. Jones DT. GenTHREADER: an efficient and reliable protein fold recognition method for genomic sequences. J Mol Biol 1999;287: 797–815.
46. Sánchez R, Pieper U, Mirkovic N, DeBakker PIW, Wittenstein E, Sali A. ModBase: a database of annotated comparative protein structure models. Nucleic Acids Res 2000; 28:250–253.
47. Gibson AD, Lamerdin JA, Zhuang P, Baburaj K, Serpersu EH, Peterson CB. Orientation of heparin-binding sites in native vitronectin. Analyses of ligand binding to the primary glycosaminoglycan-binding site indicate that putative secondary sites are not functional. J Biol Chem 1999;274:6432–6442.
48. Vriend G. WHAT IF: a molecular modeling and drug design program. J Mol Graphics 1990;8:52–56.
49. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Crystallogr 1993;26:283–291.
50. Stultz CM, White JV, Smith TF. Structural analysis based on state-space modeling. Protein Sci 1993;2:305–314.
51. Seiffert D, Loskutoff DJ. Kinetic analysis of the interaction between type 1 plasminogen activator inhibitor and vitronectin and evidence that the bovine inhibitor binds to a thrombin-derived amino-terminal fragment of bovine vitronectin. Biochim Biophys Acta 1991;1078:23–30.
52. Zhuang P, Chen AI, Peterson CB. Native and multimeric vitronectin exhibit similar affinity for heparin. Differences in heparin binding properties induced upon denaturation are due to self association into a multivalent form. J Biol Chem 1997;272:6858–6867.
53. Tollefsen DM, Weigel CJ, Kabeer MH. The presence of methionine or threonine at position 381 in vitronectin is correlated with proteolytic cleavage at Arginine 379. J Biol Chem 1990;265:9778–9781.
54. Preissner KT, Muller-Berghaus G. Neutralization and binding of heparin by S protein/vitronectin in the inhibition of Factor Xa by antithrombin. Involvement of an inducible heparin-binding domain of S-protein/vitronectin. J Biol Chem 1987;262:12247–12253.
55. Yoneda A, Ogawa H, Kojima K, Matsumoto I. Characterization of the ligand binding activities of vitronectin: interaction of vitronectin with lipids and identification of the binding domains for various ligands using recombinant domains. Biochemistry 1998;37: 6351–6360.
56. Seiffert D, Ciambrone G, Wagner NV, Binder BR, Loskutoff DJ. The somatomedin B domain of vitronectin. Structural requirements for the binding and stabilization of active type 1 plasminogen activator inhibitor. J Biol Chem 1994;269:2659–2666.
57. Seiffert D, Wagner NV. Evidence for a specific interaction of vitronectin with arginine: effects of reducing agents on the expression of functional domains and immunoepitopes. Biochimie 1997;79:205–210.
58. Beisel HG, Kawabata S, Iwanaga S, Huber R, Bode W. Tachylectin 2: crystal structure of a specific GlcNAc/GalNAc-binding lectin involved in the innate immunity host defense of the Japanese horseshoe crab *Tachypleus tridentatus*. EMBO J 1999;18:2313–2322.
59. Wall MA, Posner BA, Sprang SR. Structural basis of activity and subunit recognition in G protein heterotrimers. Structure 1998;6: 1169–1183.
60. Gibson AD, Peterson CB. Full length and truncated forms of vitronectin provide insight into effects of proteolytic processing on function. Biochem Biophys Acta 2001;1545:389–403.
61. Fulop V, Jones DT. β Propellers: structural rigidity and functional diversity. Curr Opin Struct Biol 1999;9:715–721.
62. Springer TA. An extracellular β-propeller predicted in lipoprotein and scavenger receptors, tyrosine kinases, epidermal growth factor precursor, and extracellular matrix components. J Mol Biol 1998;283:837–862.