PROTEINS: Structure, Function, and Bioinformatics 64:665–672 (2006)

Solution Structures of the First and Fourth TSR Domains of F-Spondin

Kimmo Pääkkönen,[1]† Helena Tossavainen,[2]† Perttu Permi,[2] Harri Rakkolainen,[2] Heikki Rauvala,[2] Erkki Raulo,[2] Ilkka Kilpeläinen,[2,3] and Peter Güntert[1]*

[1]*Tatsuo Miyazawa Memorial Program, RIKEN Genomic Sciences Center, Yokohama, Japan*
[2]*Institute of Biotechnology, NMR Laboratory, University of Helsinki, Helsinki, Finland*
[3]*Department of Chemistry, University of Helsinki, Helsinki, Finland*

ABSTRACT **F-spondin is a protein mainly associated with neuronal development. It attaches to the extracellular matrix and acts in the axon guidance of the developing nervous system. F-spondin consists of eight domains, six of which are TSR domains. The TSR domain family binds a wide range of targets. Here we present the NMR solution structures of TSR1 and TSR4. TSR domains have an unusual fold that is characterized by a long, nonglobular shape, consisting of two β-strands and one irregular extended strand. Three disulfide bridges and stack of alternating tryptophan and arginine side-chains stabilize the structure. TSR1 and TSR4 structures are similar to each other and to the previously determined TSR domain X-ray structures from another protein, TSP, although TSR4 exhibits a mobile loop not seen in other structures. Proteins 2006;64:665–672.** © 2006 Wiley-Liss, Inc.

Key words: F-spondin; TSR; extracellular matrix; neuron development; NMR structure

INTRODUCTION

F-spondin[1,2] is an extracellular matrix protein that comprises six C-terminal domains belonging to the TSR superfamily. F-spondin genes are found in both vertebrates and invertebrates. The N-terminus is conserved in F-spondins, but there is some variation of the number and type of TSR repeats between vertebrates and invertebrates.[3] The TSR superfamily is usually associated with different roles in neural development. Over 40 different human proteins have been found to contain TSR domains.[4] They bind to various targets: collagen V,[5] fibronectin,[6] CD36,[7] TGFβ,[8] and heparin.[9,10]

F-spondin is expressed especially in the floor plate of the developing embryo. It patterns axonal trajectory in the spinal cord by promoting the outgrowth of commissural axons[11] and inhibiting the outgrowth of motor axons.[12] A role in neuronal regeneration has also been suggested.[13] Studies on bovine F-spondin interactions with rat aortic vascular smooth muscle cells show that F-spondin is also expressed outside the brain, where it stimulates the growth of smooth muscle cells.[14] Recently it was shown that F-spondin binds the amyloid-β precursor protein and prevents its cleavage.[15]

F-spondin has two N-terminal domains, reelin and spondin, which are related to the corresponding domains found in the proteins Reelin and Mindin, and six C-terminal TSR domains. The total length of rat F-spondin is 807 amino acids. The first two domains consist of approximately 200 amino acids each, and the TSR domains are about 60 amino acids in length.[16,17] Mindin, which also has a spondin domain, is a secreted protein that binds to the extracellular matrix.[18,19] Mindin has recently been suggested to also act as a pattern-recognition molecule for microbial pathogens.[20] Reelin is an extracellular matrix protein associated with neuron development.[21]

F-spondin contains several glycosylation sites. Both C-mannosylation and O-fucosylation have been reported for F-spondin.[22,23] The first five TSR domains in F-spondin contain conserved tryptophans that have been identified as sites for C-mannosylation.[23] The glycosylation sites for the last domain are not yet known.

TSR4 contains the possible CD36-binding motif CSVTCG, which is also found in TSR2 and TSR3.[7] This motif is a potential binding site for the malarial pathogen *Plasmodium falciparum*.[24]

TSP-1 contains three TSR domains. The structure of TSR2 and TSR3, along with the linker between the two domains, have been determined previously by X-ray crystallography.[25] A novel, antiparallel, three-stranded fold was revealed that is characterized by a stabilizing stacked array of tryptophans, arginines, and cysteines. To widen the structural knowledge within the TSR domain family, we have embarked on the structure determination of the TSR domains in F-spondin by liquid-state NMR spectros-

Abbreviations: 2D, two-dimensional; ApoEr2, apolipoprotein E receptor type 2; APP, amyloid precursor protein; cDNA, complementary DNA; COSY, correlation spectroscopy; DTT, dithiothreitol; HSQC, heteronuclear single quantum coherence; LA, low density lipoprotein receptor type A ligand-binding repeats; MALDI-TOF, matrix-assisted laser desorption/ionization time-of-flight; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser enhancement spectroscopy; PDB, Protein Data Bank; RMSD, root-mean-square deviation; RT-PCR, reverse transcription polymerase chain reaction; TSP, thrombospondin; TSP-1, thrombospondin-1; TSR, thrombospondin type 1 repeat; TSR1, first TSR domain of F-spondin; TSR4, fourth TSR domain of F-spondin.

†These authors contributed equally to this work.

*Correspondence to: Peter Güntert, 1-7-22 Suehiro-cho, Tsurumi, Yokohama, 230-0045, Japan. E-mail: guentert@gsc.riken.jp

Received 28 September 2005; Revised 6 February 2006; Accepted 15 March 2006

Published online 30 May 2006 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/prot.21030

copy. In this work the structures of TSR1 and TSR4 domains of rat F-spondin were determined, corresponding to residues 441–499 and 613–666 of the rat F-spondin sequence (Swiss-Prot code: P35446).

MATERIALS AND METHODS

Protein Preparation

The F-spondin cDNA was obtained by RT-PCR from embryonic (E16) rat brain total RNA. The primers were designed to incorporate BamHI- and EcoRI restriction sites at the 5′ and 3′ ends, respectively, of the selected domain corresponding sequence in F-spondin cDNA. The primers are ACGGATCCGAAACCTGCATCTACTCC and GTGAATTCCTAGCCATCTTCATCGCTGC for TSR1, and GCGGATCCATCCCGTGCTTGCTGTCTC and TCGAATTCCTAGGGGCACTCTGGCAGCATA for TSR4. The PCR-produced domain constructs were sequenced and subsequently produced in *Escherichia coli* using the glutathione-S-transferase fusion vector pGEX-2T (Pharmacia-Amersham) as described previously.[26] The constructs for both domains contain two extra amino acids, Gly and Ser, at the amino terminus. The disulphide bridges were determined by digesting both native and DTT-reduced samples with trypsin, and analyzing the resulting peptides with Bruker Dantonics Ultraflex MALDI-TOF mass spectrometry.[27]

NMR Experiments

The NMR samples contained 0.8–1.2 m*M* uniformly ^{15}N, ^{13}C-labeled protein, 2 m*M* NaN_3 and 20 m*M* bisTris in 95% H_2O/5% D_2O at pH 6.5 (TSR1) and 6.8 (TSR4). No carbohydrates were present in the proteins. All NMR spectra were acquired at 10°C on Varian Inova 600 and 800 MHz spectrometers. The spectra were processed with VNMR 6.1C (Varian Inc., Palo Alto, CA) and NMRPipe,[28] and analyzed with Sparky 3.106.[29]

In the case of TSR1, a set of 10 triple-resonance spectra was used for the assignment of the chemical shifts: HNCA, iHNCA, HN(CO)CA, HNCO, HN(CA)CO, HNCACB, HN(CO)CACB, CC(CO)NH, H(CCO)NH, and HCCH-COSY.[30,31] In the case of F-spondin TSR4, we utilized a smaller set of six spectra, excluding the spectra involving exclusively Cα correlations and HN(CA)CO. NOE peaks were identified and integrated in a ^{15}N-edited NOESY-HSQC,[32] and in a ^{13}C-edited NOESY-HSQC[33] spectrum modified to simultaneously excite aliphatic and aromatic carbon resonances. NOESY mixing times were 100 ms for both TSR1 NOESY spectra and 120 ms for both TSR4 NOESY spectra. Additional structural restraints were derived from $^3J_{HN\alpha}$ coupling constants[34] in the form of ϕ angle restraints.

We determined T_1 and T_2 relaxation times for the backbone amide nitrogens from series of 2D ^{15}N-^1H-HSQC spectra we acquired, with relaxation delays ranging from 0.01 to 1.9 s for T_1 and from 0.01 to 0.25 s for T_2, using eight time points per series. We used the relaxation data fitting routine of Sparky 3.106 to analyze the relaxation data. Peak heights in the series of spectra were fitted to a decaying exponential. Peak positions were adjusted separately in each spectrum of the series.

Resonance Assignment and Structure Calculation

Backbone resonance assignments were made with the help of the program Autoassign.[35,36] Except for the first two residues, complete backbone assignments were obtained for TSR1. In the case of TSR4, two additional residues remained unassigned by Autoassign. The assignments made by Autoassign for TSR1 and TSR4 were confirmed to be correct. The backbone assignments were checked and completed manually. The side-chain assignments were done manually.

The structure calculations were made automatically by the program CYANA.[37–39] Peaks from the ^{15}N-edited NOESY-HSQC and ^{13}C-edited NOESY-HSQC were picked and integrated manually, and the peak lists, together with the chemical shift assignments, were used as input for the iterative NOE assignment and structure calculation with CYANA. We generated 200 conformers in each of the 7 cycles of the combined automated NOESY assignment and structure calculation algorithm.[38] We subjected the 20 conformers with lowest final target function values to restrained energy minimization with respect to the AMBER force field[40] using the program OPALp.[41,42] The protein was immersed in a shell of water molecules with a thickness of 8 Å. We applied a maximum of 3000 steps of restrained conjugate gradient minimization, using, in addition to the standard AMBER force field, a pseudopotential for NOE upper distance bounds that was proportional to the sixth power of the restraint violation. The force constant was chosen such that a restraint violation of 0.1 Å contributed 0.3 kcal/mol to the potential energy. Structure figures were prepared with the program MOLMOL, which was also used to calculate ring current shifts.[43]

RESULTS

The 3D structures of TSR1 and TSR4 were determined by heteronuclear NMR spectroscopy combined with automated NOESY assignment and structure calculation. To assess the possible differences in the dynamical behavior of the domains, we measured T_1 and T_2 relaxation times of the backbone amide nitrogens.

Except for the two non-native N-terminal residues, complete backbone assignments and nearly complete side-chain assignments were obtained. Unusual chemical shifts were observed for side-chain protons of residues Arg466 and Arg468 of TSR1 and the corresponding Arg637 and Arg639 of TSR4. These residues are in the cores of the domains, and close to aromatic rings. Ring current shift calculations for the final structures confirmed qualitatively the strong chemical shift changes of these protons. The chemical shifts have been deposited in the BioMagResBank with accession numbers 6175 for TSR1 and 10002 for TSR4. The three disulfide bonds [Fig. 1(A)] were determined by trypsin digestion and mass spectrometry (see Materials and Methods section). Furthermore, the C^α and C^β chemical shifts of the cysteines in both domains are consistent with the criteria for the oxidized form of cysteines as described by Sharma and Rajarathnam.[44]

The resonance assignments, together with NOE peak positions and volumes and ϕ dihedral angle restraints

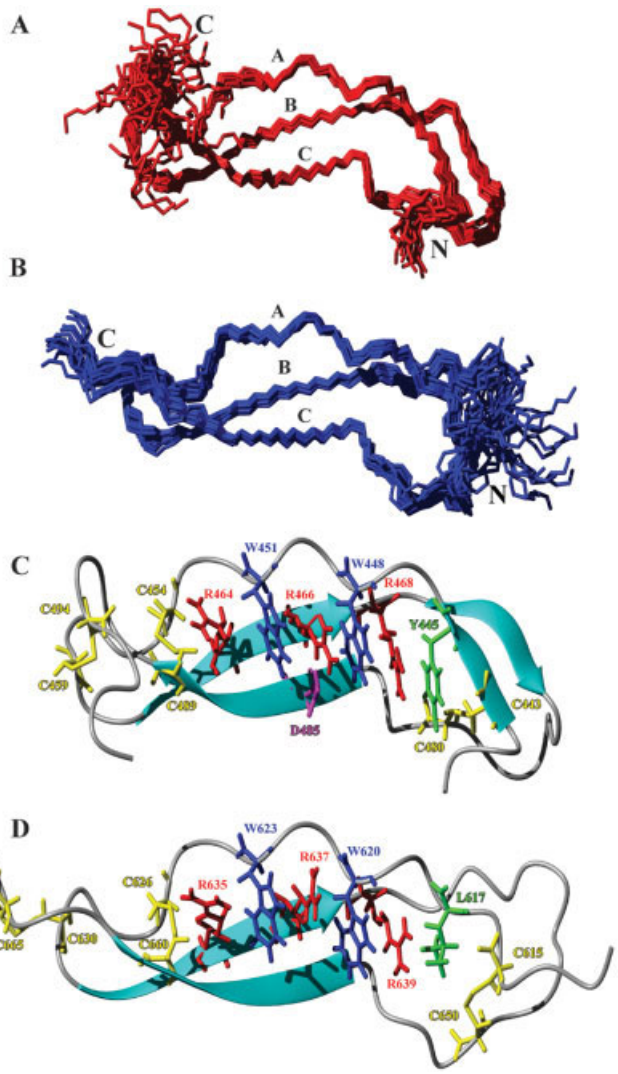

Fig. 1. The NMR solution of structures of (**A**) the F-spondin TSR domain 1 and (**B**) the F-spondin TSR domain 4. Twenty energy-refined conformers are shown for each domain. The first strand has a rippled conformation, which is characteristic for this fold. The two other strands form an antiparallel β-sheet (residues 462–467 and 484–489 for TSR1, and 634–640 and 657–663 for TSR4). TSR1 has a short additional β-sheet in the region where TSR4 has a less well-defined loop region (residues 443–445 and 471–473). The secondary structure and core residues of a representative conformer of the solution structures of the F-spondin (**C**) TSR1 and (**D**) TSR4 domains are shown. Tryptophans are drawn in blue, arginines in red, and cysteines in yellow. The green residue in TSR1 is tyrosine, which may further stabilize the structure through interactions with the nearby arginine side-chain. In TSR4, this residue is leucine. Aspartate 485 is shown in magenta.

TABLE I. Statistics for the NMR Solution Structures of the F-Spondin TSR Domains 1 and 4

	TSR1	TSR4
Number of residues	61	56
Distance restraints		
All	994	825
Short-range, $\|i-j\| \leq 1$	527	508
Medium-range, $1 \leq \|i-j\| \leq 5$	100	55
Long-range, $\|i-j\| \geq 5$	367	262
ϕ Dihedral angle restraints	43	25
Average number of restraints per residue	16.3	14.7
Average CYANA target function value, Å^2	0.66 ± 0.10	0.57 ± 0.10
Average AMBER energy, kcal/mol	−2000 ± 83	−1819 ± 70
RMSD		
All residues, backbone, Å	1.46 ± 0.27	1.46 ± 0.31
All residues, heavy atoms, Å	1.68 ± 0.23	1.71 ± 0.23
Ordered structure, backbone,[a] Å	0.50 ± 0.11	0.92 ± 0.21
Ordered structure, heavy atoms, Å	0.79 ± 0.08	1.41 ± 0.16
Number of hydrogen bonds[b]	36	23
Ramachandran plot statistics, %		
Most favored regions	81.3	76.0
Additional allowed regions	18.1	23.0
Generously allowed regions	0.6	0.9
Disallowed regions	0.0	0.0

The solution structure is represented by the 20 energy-refined conformers with lowest CYANA target function values. No distance violation over 0.2 Å or dihedral angle violations over 5° were observed in the structures.

[a]The disordered N- and C-terminal residues are excluded (i.e., residues 438–442 and 495–499 for TSR1, and 610–614 and 663–666 for TSR4).

[b]After energy minimization with OPALp.

from $^3J_{HN\alpha}$ coupling constants were used as input for the automated iterative NOE assignment and structure calculation protocol incorporated in the program CYANA version 2.0.[37–39] Well-defined structures in good agreement with the experimental restraints resulted for both domains, as evidenced by the structural statistics of Table I. Neither distance restraint violations above 0.2 Å nor dihedral angle violations over 5° were observed in the two ensembles, and less than 1% of the residues were found in generously allowed or disallowed regions of the Ramachandran plot (Table I). The structures have been deposited in the PDB with accession codes 1SZL for TSR1 and 1VEX for TSR4.

TSR1 is structurally slightly better defined than TSR4, with an average backbone RMSD to the mean coordinates of 0.50 Å for the well-defined part of the molecule (residues 443–494; Table I, Fig. 1). This results from a higher number of assigned peaks in the NOESY spectra and reflects the generally superior quality of the TSR1 spectra. Especially the number of long-range distance restraints is considerable higher in TSR1 than in TSR4. In TSR1 the backbone RMSD shows no large variations along the sequence, whereas in TSR4 the second loop of the structure (residues 643–654) is less well defined, raising the average backbone RMSD of residues 615–662 to 0.92 Å.

The overall fold is very similar for the two F-spondin TSR domains (Fig. 1). This is not surprising given that the alignment of the amino acid sequences [Fig. 2(A)] shows 48% sequence identity. The structural details are somewhat different, resulting in a backbone RMSD between the mean coordinates of TSR1 and TSR4 of 1.3 Å, excluding the flexible termini and the loops and taking into account the one-residue insertion in the TSR1 sequence. In the

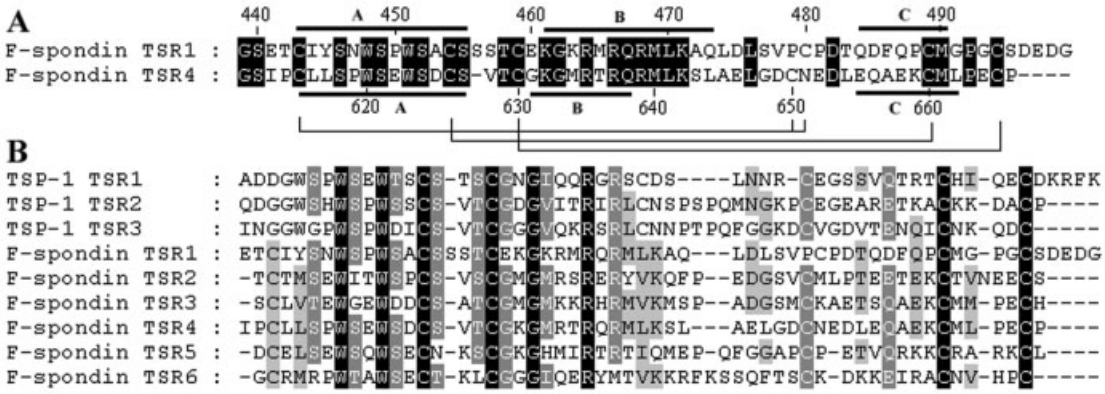

Fig. 2. (**A**) Sequence alignment of the F-spondin TSR domains 1 and 4. Disulfide bonds are indicated by connecting lines, and the strands A, B, and C with solid lines. (**B**) Sequence alignment of F-spondin TSR domains and TSP-1 TSR domains. The cloning artifacts (residues Gly and Ser in the start of the sequences) have been omitted from the alignment.

following, the structure of TSR1 is presented in detail. Because this description is valid to a large degree also for TSR4, we comment on only the significant differences between the two structures.

The overall fold is antiparallel three-stranded. Strand A (residues 443–455 in TSR1 and 615–627 in TSR4) is irregular and has a rippled form, whereas strands B (461–473 in TSR1 and 631–638 in TSR4) and C (484–490 in TSR1 and 655–662 in TSR4) are regular β-strands and form a β-sheet (Fig. 1). In strand A, the two conserved tryptophans, Trp448 and Trp451, are located at the ridges of the rippled backbone pointing toward strand C. Strand B provides three arginines, Arg464, Arg466, and Arg468, that form the RWRWR stacked array with the two tryptophans. This arrangement is energetically favorable, as the cation–π interaction between an aromatic ring and a positively charged side-chain group contributes significantly to protein stability.[45] This interaction is common especially between arginines and tryptophans. The array is preceded by Tyr445 from strand B, whose side-chain is favorably oriented to be involved in an additional cation–π interaction with Arg468. The RWRWR base structure is capped on both ends by disulfide bonds, one at the N-terminal side of the domain formed by Cys443 in strand A and Cys480 in the loop between strands B and C, and two at the C-terminal side, formed by Cys454–Cys489 and Cys459–Cys494, where Cys454 and Cys459 are in the loop between strands A and B, and Cys489 and Cys494 in the C-terminal following strand C. In TSR4, the aforementioned tyrosine is replaced by a leucine, whose orientation is also toward the inside of the domain. Altogether the tryptophan, arginine, and cysteine side-chains form an eight-layered stacked structure. The $C^{\delta 1}$ atoms of the two tryptophan residues point outward from the core, enabling the C-mannosylation of these moieties.[46]

A regular hydrogen bond network is formed between the antiparallel β-strands B and C. Additional lateral contacts are likely to contribute to the stability of this β-sheet. For instance, in the majority of structures a salt bridge is observed between the side-chains of Arg466 and Asp485. Strand A, although not conforming to the requirements of a regular β-strand, is intimately linked to strand B via hydrogen bonds. Amide protons from the four serine residues Ser446, Ser449, Ser452 and Ser455 of the rippled backbone of strand A form hydrogen bonds to carbonyl carbons of residues Met469, Gln467, Met465 and Lys463 in strand B, respectively. Serine side-chain hydroxyl groups on strand A make additional hydrogen bonds to the backbone of strand B. In over one third of the conformers, we observed a hydrogen bond between Ser446 Oγ and Met469 H, Ser452 Oγ and Met465 H, and Ser449 Hγ and Gln467 O. Overall the hydrogen bond network in TSR4 is similar to TSR1, albeit slightly less complete.

The A-B loop conformation is stabilized by a disulfide bond to the C-terminal loop. In the B-C loop, residues 470–473 are involved in an additional β-sheet with the N-terminal residues 443–445. This short β-sheet is distorted by a classical β-bulge: hydrogen bonds are formed between Ile444H^N-Ala472C′, Ile444C′-Ala472H^N and Ile444C′-Lys471H^N, and the ϕ,ψ angles of Lys471 conform to those of a right-handed α-helix. A type I β-turn follows the short β-strand in the loop, and at the end, an irregular stretch containing the cysteine that makes the disulfide bond to the N-terminal part of the structure, surrounded by two *trans* prolines.

The B-C loop of TSR4 is not well defined. A smaller number of NOEs was obtained for this segment compared to the structured regions. The relaxation time measurements (Fig. 3) show that on the average TSR4 has shorter T_2 relaxation times than TSR1, and several residues in TSR4 show increased T_1/T_2 ratios, indicative of chemical exchange. This is especially noticeable for residues Leu616, Leu617, Met640, Asp649, Cys650, and Asn651 (corresponding TSR1 numbering in Fig. 3: Ile444, Tyr445, Met469, Val478, Pro479, and Cys480). The latter three residues are in the B-C loop. Cys650 is connected by a disulfide bond to Cys615, which also has a low T_2 value and slightly higher T_1/T_2 ratio, as do the neighboring residues of both cysteines. Residues Ala645 and Glu646 have short T_2 times, and they too are in the disordered loop region.

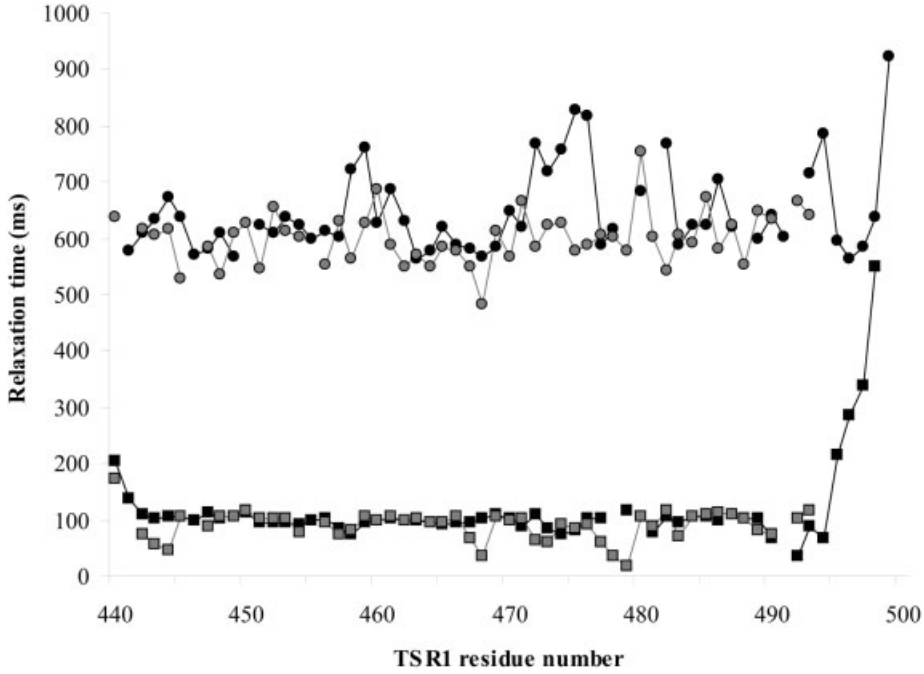

Fig. 3. T_1 (circles) and T_2 (boxes) relaxation times for the backbone amide nitrogens of TSR1 (black) and TSR4 (gray). The residue numbering of TSR1 is used.

DISCUSSION

The structures of the TSR1 and TSR4 domains are markedly similar to the previously solved structures of TSR domains of TSP-1,[25] with a sequence homology on the order of 30%. The backbone atoms of the structure families can be superimposed on the TSP-1 structures, excluding the loops and the chain termini, with backbone RMSDs of 1.7 Å for F-spondin TSR1 and TSP-1 domains 2 and 3. For F-spondin TSR4 and TSP-1 domains 2 and 3 the corresponding RMSDs are 1.2 Å and 1.5 Å, respectively. The disulfide bond network is slightly different in F-spondin TSRs and TSP-1 TSR domains. In the structures of F-spondin TSRs, there is a disulfide bond between the beginning of strand A and the beginning of strand C, whereas in TSP-1 TSR domains the corresponding disulfide bond is formed between strands B and C. Some variation in the amino acids forming the eight-layered stacked array is observed, too. Only the four central residues, corresponding to Arg468, Trp448, Arg466, and Trp451 of TSR1, are strictly conserved. The outer layers are formed by Tyr and Arg in TSR1, Leu and Arg in TSR4, Trp and Ile in TSP-1 domain 2, and Trp and Glu in TSP-1 domain 3. Irrespective of their nature, all these sidechains maintain a planar and parallel orientation.

Based on disulfide bond pattern analysis from amino acid sequences the TSR domains have tentatively been divided into two major groups by Tan et al.,[25] with TSP-1 TSRs being representatives of group 1 and F-spondin TSRs representing group 2. The present structure determination of F-spondin TSRs shows that although the disulfide bond pattern observed in the N-terminal parts of F-spondin and TSP-1 TSR domains is indeed different, this difference does not significantly influence the overall fold. The RMSD values between the structures of the two TSR types remain small.

The electrostatic surface potentials of F-spondin TSR1 and TSR4 show a similar positive face as the TSP-1 TSR2 domain (Fig. 4). Two arginines are conserved both in the sequence alignment and by their positions in the structures of all three domains. In addition, two positive and two negative charges are present in reasonably close positions on the same side of the molecule, although they originate from sequentially different residues (K464, K467, E462, and E459 in TSP-1 TSR2; R464, K463, D485, and D482 in F-spondin TSR1; and R637, K659, E658, and D653 in F-spondin TSR4). The charges on the surface of F-spondin TSR1 are mainly concentrated on one side of the domain, while F-spondin TSR4 has several charged residues on both sides of the domain.

Recent studies have shown that F-spondin binds the APP[15] and the ApoEr2.[47] F-spondin binds APP through its N-terminal reelin and spondin domains. The F-spondin C-terminal TSR domains 1–4 bind the LA repeats 3, 7, or 8 of ApoEr2. These interactions may play a role in regulating Aβ peptide formation, which is important for Alzheimer's disease.[47]

The LA repeats are small and ubiquitous domains that bind calcium. Several structures of these domains have been solved.[48–53] They often show regions of negative surface charges,[51] and it has been postulated that electrostatics are important for the interactions between LA repeats and its ligands.[54] Indeed, the interaction between LA repeat 3 of the very low density lipoprotein receptor and the minor group human rhinovirus was postulated to

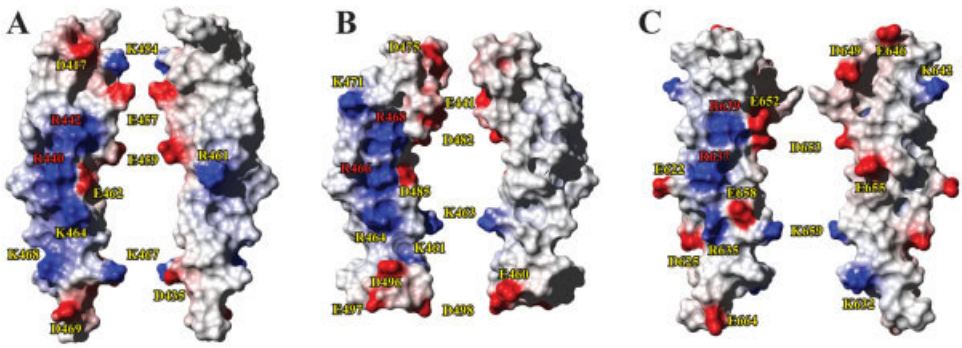

Fig. 4. Electrostatic surface potentials of (**A**) TSP TSR2, (**B**) F-spondin TSR1, and (**C**) F-spondin TSR4. Each domain is shown from two sides, which are rotated by 180° around a vertical axis relative to each other. For the F-spondin TSR domains, the structures with the lowest total energy are shown. Charged residues are labeled. Red labels identify residues that are conserved in the sequence alignment and have similar positions on the surfaces of all three domains. The charged side-chains D496, E497, and D498 in F-spondin TSR1 and D625, K632, K642, E646, D649, and D653 in F-spondin TSR4 have larger than 2 Å side-chain RMSDs.

be based on electrostatic interactions between the molecules.[52] The positive charges on the surfaces of the F-spondin TSR1 and TSR4 domains (Fig. 4) makes it plausible that the interactions between F-spondin TSRs and LA repeats could also be mediated by electrostatic interactions. However, the C-mannosylations and O-fucosylation will change the charge distributions on the F-spondin TSR domain surfaces, and it has not been shown yet which of the first TSR domains of F-spondin are interacting with LA repeats. Further binding and structural studies will be needed to resolve this question.

The flexibility of the F-spondin TSR4 B-C loop region when compared to TSR1 is intriguing. Differences in amino acid sequence and relaxation times may suggest some functional difference, whose nature remains to be understood. The last two TSR domains of F-spondin have been reported to bind to the extracellular matrix, whereas the domains 1–4 do not,[12] but domains 1–4 may be involved in mediating the repulsive effect that inhibits motor neuron outgrowth.[3] It is also possible that the difference in the loop structure is related to the single-domain constructs used, and that a multidomain protein may have different characteristics. In our experience, the TSR domains of F-spondin differed markedly in their stability when expressed individually, some of them being either poorly folded or manifesting significant conformational exchange (unpublished data). As there is a linker of only two residues between F-spondin TSR3 and TSR4, the two domains come into close proximity, and it is conceivable that the C-terminal end of TSR3 together with the A-B loop, stabilize the B-C loop of TSR4. Furthermore, fucosylation has been reported to stabilize protein structure through the formation of a hydrophobic cluster.[55] Thr573 of TSR3 has been identified as a site for O-fucosylation.[23] Assuming that the structure of TSR3 is similar to that of the two other TSR domains, this residue is located in the TSR3 A-B loop. By formation of hydrophobic contacts or hydrogen bonds, the B-C loop region of TSR4 may additionally be stabilized by this post-translational modification. Further studies with multidomain constructs and binding experiments will be needed to fully understand the differences between the domains and their interactions.

CONCLUSIONS

The structures of the TSR domains 1 and 4 from F-spondin reveal an antiparallel, three-stranded fold in which one of the strands has an irregular rippled form, whereas the two others form a regular β-sheet. The structure of this small domain is stabilized by multiple interactions between backbone and side-chain atoms, as well as covalent interactions in the form of disulfide bonds. Interleaved side-chains of two tryptophans, three arginines, and three disulfide-bonded cysteines form an eight-layer, stacked array (Fig. 1). The structures of TSR1 and TSR4 are very similar to each other, with the major difference being a mobile loop in TSR4 that is not present in TSR1. Its significance for the biological function remains to be elucidated in detail.

REFERENCES

1. Klar A, Baldassare M, Jessell TM. F-spondin: a gene expressed at high levels in the floor plate encodes a secreted protein that promotes neural cell adhesion and neurite extension. Cell 1992;69: 95–110.
2. Ruiz i Altaba A, Cox C, Jessell TM, Klar A. Ectopic neural expression of a floor plate marker in frog embryos injected with the midline transcription factor Pintallavis. Proc Natl Acad Sci USA 1993;90:8268–8272.
3. Feinstein Y, Klar A. The neuronal class 2 TSR proteins F-spondin and Mindin: a small family with divergent biological activities. Int J Biochem Cell Biol 2004;36:975–980.
4. Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA,

Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X. The sequence of the human genome. Science 2001;291:1304–1351.

5. Takagi J, Fujisawa T, Usui T, Aoyama T, Saito Y. A single chain 19-kDa fragment from bovine thrombospondin binds to type V collagen and heparin. J Biol Chem 1993;268:15544–15549.
6. Sipes JM, Guo N, Negre E, Vogel T, Krutzsch HC, Roberts DD. Inhibition of fibronectin binding and fibronectin-mediated cell adhesion to collagen by a peptide from the second type I repeat of thrombospondin. J Cell Biol 1993;121:469–477.
7. Asch AS, Silberger S, Heimer E, Nachman RL. Thrombospondin sequence motif (CSVTCG) is responsible for CD36 binding. Biochem Biophys Res Commun 1992:1208–1217.
8. Schultz-Cherry S, Chen H, Mosher DF, Misenheimer TM, Krutzsch HC, Roberts DD, Murphy-Ullrich JE. Regulation of transforming growth factor-beta activation by discrete sequences of thrombospondin 1. J Biol Chem 1995;270:7304–7310.
9. Guo NH, Krutzsch HC, Negre E, Zabrenetzky VS, Roberts DD. Heparin-binding peptides from the type I repeats of thrombospondin: structural requirements for heparin binding and promotion of melanoma cell adhesion and chemotaxis. J Biol Chem 1992;267:19349–19355.
10. Gantt SM, Clavijo P, Bai X, Esko JD, Sinnis P. Cell adhesion to a motif shared by the malaria circumsporozoite protein and thrombospondin is mediated by its glycosaminoglycan-binding region and not by CSVTCG. J Biol Chem 1997;272:19205–19213.
11. Burstyn-Cohen T, Tzarfaty V, Frumkin A, Feinstein Y, Stoeckli E, Klar A. F-Spondin is required for accurate pathfinding of commissural axons at the floor plate. Neuron 1999;23:233–246.
12. Tzarfati-Majar V, Burstyn-Cohen T, Klar A. F-spondin is a contact-repellent molecule for embryonic motor neurons. Proc Natl Acad Sci USA 2001;98:4722–4727.
13. Burstyn-Cohen T, Frumkin A, Xu YT, Scherer SS, Klar A. Accumulation of F-spondin in injured peripheral nerve promotes the outgrowth of sensory axons. J Neurosci 1998;18:8875–8885.
14. Miyamoto K, Morishita Y, Yamazaki M, Minamino N, Kangawa K, Matsuo H, Mizutani T, Yamada K, Minegishi T. Isolation and characterization of vascular smooth muscle cell growth promoting factor from bovine ovarian follicular fluid and its cDNA cloning from bovine and human ovary. Arch Biochem Biophys 2001;390:93–100.
15. Ho A, Südhof TC. Binding of F-spondin to amyloid-β precursor protein: A candidate amyloid-β precursor protein ligand that modulates amyloid-β precursor protein cleavage. Proc Natl Acad Sci USA 2004;101:2548–2553.
16. Lawler J, Hynes RO. The structure of human thrombospondin, an adhesive glycoprotein with multiple calcium-binding sites and homologies with several different proteins. J Cell Biol 1986;103:1635–1648.
17. Bornstein P, O'Rourke K, Wikstrom K, Wolf FW, Katz R, Li P, Dixit VM. A second, expressed thrombospondin gene (Thbs2) exists in the mouse genome. J Biol Chem 1991;266:12821–12824.
18. Higashijima S, Nose A, Eguchi G, Hotta Y, Okamoto H. Mindin/F-spondin family: novel ECM proteins expressed in the zebrafish embryonic axis. Dev Biol 1997;192:211–227.
19. Umemiya T, Takeichi M, Nose A. M-spondin, a novel ECM protein highly homologous to vertebrate F-spondin, is localized at the muscle attachment sites in the *Drosophila* embryo. Dev Biol 1997;186:165–176.
20. He YW, Li H, Zhang J, Hsu CL, Lin E, Zhang N, Guo J, Forbush KA, Bevan MJ. The extracellular matrix protein mindin is a pattern-recognition molecule for microbial pathogens. Nat Immunol 2004;5:88–97.
21. D'Arcangelo G, Miao GG, Chen SC, Soares HD, Morgan JI, Curran T. A protein related to extracellular matrix proteins deleted in the mouse mutant reeler. Nature 1995;374:719–723.
22. Hofsteenge J, Huwiler KG, Macek B, Hess D, Lawler J, Mosher DF, Peter-Katalinic J. C-mannosylation and O-fucosylation of the thrombospondin type 1 module. J Biol Chem 2001;276:6485–6498.
23. Gonzalez de Peredo A, Klein D, Macek B, Hess D, Peter-Katalinic J, Hofsteenge J. C-mannosylation and O-fucosylation of thrombospondin type 1 repeats. Mol Cell Proteomics 2002;1:11–18.
24. Rich KA, George FWt, Law JL, Martin WJ. Cell-adhesive motif in region II of malarial circumsporozoite protein. Science 1990;249:1574–1577.
25. Tan K, Duquette M, Liu JH, Dong Y, Zhang R, Joachimiak A, Lawler J, Wang JH. Crystal structure of the TSP-1 type 1 repeats: a novel layered fold and its biological implication. J Cell Biol 2002;159:373–382.
26. Kilpeläinen I, Kaksonen M, Avikainen H, Fath M, Linhardt RJ, Raulo E, Rauvala H. Heparin-binding growth-associated molecule contains two heparin-binding β-sheet domains that are homologous to the thrombospondin type I repeats. J Biol Chem 2000;275:13564–13570.
27. Navale V, Kaushal P, Hunt S, Burducea I, Gentz R, Khan F, Vertes A. Peptide mapping and disulfide bond analysis of myeloid progenitor inhibitory chemokine and keratinocyte growth factor by matrix-assisted laser desorption ionization mass spectrometry. Anal Biochem 1999;267:125–134.
28. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: a multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 1995;6:277–293.
29. Goddard TD, Kneller DG. Sparky. Version 3. San Francisco: University of California; 1997.
30. Sattler M, Schleucher J, Griesinger C. Heteronuclear multidimensional NMR experiments for the structure determination of proteins in solution employing pulsed field gradients. Progr NMR Spectr 1999;34:93–158.
31. Permi P. Intraresidual HNCA: an experiment for correlating only intraresidual backbone resonances. J Biomol NMR 2002;23:201–209.
32. Zhang O, Kay LE, Olivier JP, Forman-Kay JD. Backbone ^{1}H and ^{15}N resonance assignments of the N-terminal SH3 domain of drk in folded and unfolded states using enhanced-sensitivity pulsed field gradient NMR techniques. J Biomol NMR 1994;4:845–858.
33. Muhandiram DR, Farrow NA, Guang-yi X, Smallcombe SH, Kay LE. A gradient ^{13}C NOESY-HSQC experiment for recording NOESY spectra of ^{13}C-labeled proteins dissolved in H_2O. J Magn Reson 1993;B102:317–321.
34. Aitio H, Permi P. Semi-constant-time HMSQC (SCT-HMSQC-HA) for the measurement of $^{3}J_{HNH\alpha}$ couplings in ^{15}N-labeled proteins. J Magn Reson 2000;143:391–396.
35. Zimmerman DE, Kulikowski CA, Huang Y, Feng W, Tashiro M, Shimotakahara S, Chien C, Powers R, Montelione GT. Automated analysis of protein NMR assignments using methods from artificial intelligence. J Mol Biol 1997;269:592–610.
36. Moseley HN, Monleon D, Montelione GT. Automatic determination of protein backbone resonance assignments from triple resonance nuclear magnetic resonance data. Methods Enzymol 2001;339:91–108.
37. Güntert P, Mumenthaler C, Wüthrich K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol 1997;273:283–298.

38. Herrmann T, Güntert P, Wüthrich K. Protein NMR structure determination with automated NOE assignment using the new software CANDID and the torsion angle dynamics algorithm DYANA. J Mol Biol 2002;319:209–227.
39. Güntert P. Automated NMR protein structure calculation. Prog NMRes Spectr 2003;43:105–125.
40. Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA. A 2nd generation force-field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 1995;117:5179–5197.
41. Luginbühl P, Güntert P, Billeter M, Wüthrich K. The new program OPAL for molecular dynamics simulations and energy refinements of biological macromolecules. J Biomol NMR 1996;8: 136–146.
42. Koradi R, Billeter M, Güntert P. Point-centered domain decomposition for parallel molecular dynamics simulation. Comput Phys Commun 2000;124:139–147.
43. Koradi R, Billeter M, Wüthrich K. MOLMOL: a program for display and analysis of macromolecular structures. J Mol Graph 1996;14:51–55, 29–32.
44. Sharma D, Rajarathnam K. ^{13}C NMR chemical shifts can predict disulfide bond formation. J Biomol NMR 2000;18:165–171.
45. Gallivan JP, Dougherty DA. Cation-π interactions in structural biology. Proc Natl Acad Sci USA 1999;96:9459–9464.
46. Furmanek A, Hofsteenge J. Protein C-mannosylation: facts and questions. Acta Biochim Pol 2000;47:781–789.
47. Hoe HS, Wessner D, Beffert U, Becker AG, Matsuoka Y, Rebeck GW. F-spondin interaction with the apolipoprotein E receptor ApoEr2 affects processing of amyloid precursor protein. Mol Cell Biol 2005;25:9259–9268.
48. Daly NL, Scanlon MJ, Djordjevic JT, Kroon PA, Smith R. Three-dimensional structure of a cysteine-rich repeat from the low-density lipoprotein receptor. Proc Natl Acad Sci USA 1995;92: 6334–6338.
49. Daly NL, Djordjevic JT, Kroon PA, Smith R. Three-dimensional structure of the second cysteine-rich repeat from the human low-density lipoprotein receptor. Biochemistry 1995;34:14474–14481.
50. Fass D, Blacklow S, Kim PS, Berger JM. Molecular basis of familial hypercholesterolaemia from structure of LDL receptor module. Nature 1997;388:691–693.
51. Simonovic M, Dolmer K, Huang W, Strickland DK, Volz K, Gettins PG. Calcium coordination and pH dependence of the calcium affinity of ligand-binding repeat CR7 from the LRP: comparison with related domains from the LRP and the LDL receptor. Biochemistry 2001;40:15127–15134.
52. Verdaguer N, Fita I, Reithmayer M, Moser R, Blaas D. X-ray structure of a minor group human rhinovirus bound to a fragment of its cellular receptor protein. Nat Struct Mol Biol 2004;11:429–434.
53. Rudenko G, Henry L, Henderson K, Ichtchenko K, Brown MS, Goldstein JL, Deisenhofer J. Structure of the LDL receptor extracellular domain at endosomal pH. Science 2002;298:2353–2358.
54. Jeon H, Blacklow SC. Structure and physiologic function of the low-density lipoprotein receptor. Annu Rev Biochem 2005;74:535–562.
55. Mer G, Hietter H, Lefevre JF. Stabilization of proteins by glycosylation examined by NMR analysis of a fucosylated proteinase inhibitor. Nat Struct Biol 1996;3:45–53.