Published in final edited form as:

J Mol Biol. 2010 July 16; 400(3): 589–604. doi:10.1016/j.jmb.2010.04.043.

Structure, evolutionary conservation, and conformational dynamics of *Homo sapiens* fascin-1, an F-actin crosslinking protein

Reza Sharifi Sedeh¹, Alexander A. Fedorov², Elena V. Fedorov², Shoichiro Ono³, Fumio Matsumura³, Steven C. Almo^{2,*}, Mark Bathe^{4,*}

¹Department of Mechanical Engineering, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, USA,

²Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, USA,

³Department of Molecular Biology and Biochemistry, Rutgers University, Piscataway, NJ 08855, USA,

⁴Department of Biological Engineering, MIT, 77 Massachusetts Avenue, Cambridge, MA 02139, USA

Abstract

Eukaryotes employ several highly conserved actin-binding proteins to crosslink filamentous actin into compact ordered bundles present in distinct cytoskeletal processes including microvilli, stereocilia, and filopodia. Fascin is one such actin-binding protein that is present predominantly in filopodia, which are believed to play a central role in normal and aberrant cell migration. An important outstanding question regards the structural basis for the unique localization and functional properties of fascin compared with other actin crosslinking proteins. Here, we present the crystal structure of full length *Homo sapiens* fascin-1, as well as examine its packing, conformational flexibility, and evolutionary sequence conservation. The structure reveals a novel arrangement of four tandem β -trefoil domains that form a bi-lobed structure with approximate pseudo-twofold symmetry. Each lobe has internal approximate pseudo-two-fold and pseudo-three-fold symmetry axes that are approximately perpendicular, with β -hairpin triplets located symmetrically on opposite sides of each lobe that mutational data suggest are actin-binding domains. Sequence conservation analysis confirms the importance of hydrophobic core residues that stabilize the β -trefoil fold, as well as interfacial residues that likely stabilize the overall fascin molecule. Sequence conservation also indicates highly conserved surface patches near the putative actin-binding domains of fascin, which conformational dynamics analysis suggest to be coupled

*To whom correspondence should be addressed: Mark Bathe, Department of Biological Engineering, MIT, 77 Massachusetts Avenue, NE47-223, Cambridge, MA 02139, USA, T +1-617-324-5685 | F +1-617-324-7554 | mark.bathe@mit.edu, or Steven C. Almo, Department of Biochemistry, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, New York 10461, USA, T +1-718-430-2746 | F +1-718-430-8565 | almo@aecon.yu.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

via an allosteric mechanism that might have important functional implications for F-actin crosslinking by fascin.

Keywords

actin-binding protein; cell motility; cancer; molecular simulation; F-actin bundle

INTRODUCTION

The actin cytoskeleton of eukaryotic cells is centrally involved in a range of cellular functions including migration, endocytosis, and division. In each case, the cell employs a host of accessory proteins to regulate actin dynamics spatiotemporally to achieve cellular function. In the case of migration, the leading edge of the cell consists predominantly of two types of protrusive filamentous actin (F-actin) structures: the dendritic sheetlike lamellipodium^{1–5} and dynamic cortical spike-like filopodia.^{6–8} Filopodia consist of compact ordered bundles of unipolar actin filaments that are nucleated at the leading edge by formins^{9,10} and crosslinked tightly into highly ordered bundles by the actin-binding protein fascin, which is both highly conserved evolutionarily and tissue- and cell-type-specific.^{7,11} While fascin dominates in cortical actin bundles such as filopodia, oocyte microvilli, and the dendrites of dendritic cells—which play direct roles in cell migration, cell-matrix adhesion, and cell-cell interactions—fascin is also associated to a lesser extent with cytoplasmic actin bundles that participate in the maintenance of internal cell architecture.¹² Fascin is additionally known to be upregulated in a number of highly motile cell phenotypes including invasive cancer cells.^{13–17}

It is well established that the molecular size and conformational flexibility of actin crosslinking proteins are highly correlated with the morphological cytoskeletal structures with which they are associated.^{18–21} For example, human filamin is a relatively large homodimer (160 nm) consisting of 24 tandem immunoglobulin repeats that is associated primarily with dendritic cytoskeletal networks,²² α -actinin is a smaller (35 nm) anti-parallel homodimer that is associated with both networks and bundles such as the contractile ring in dividing cells and stress fibers in adherent cells,^{23,24} and the compact ABP fimbrin (10 nm) (also called plastin) is found nearly exclusively in highly ordered unipolar actin bundles such as microvilli. The X-ray structures of several of these and related actin-crosslinking proteins have been determined. They consist of dual calponin homology actin-binding domains that are suggested to stabilize actin filaments.^{25–31} Fimbrin additionally has two N-terminal calcium-binding EF-hand motifs that confer calcium regulation of its F-actin crosslinking activity in human isoforms.^{32,33}

Homo sapiens fascin-1 also crosslinks actin filaments into compact unipolar bundles, but in a manner that is regulated by phosphorylation of serine 39 by protein kinase C,³⁴ which inhibits its actin-bundling activity without affecting its localization to the filopodial tip complex.⁷ *H. sapiens* fascin-1 is a compact, 55 kDa (493 residues) globular monomer with actin-binding domains that differ in primary sequence. This makes fascin distinct from other families of actin crosslinking proteins, which are composed of multiple domains that have a

high degree of sequence homology (e.g., fimbrin, α -actinin, and ABP120). Fascin was originally discovered in extracts from sea urchin eggs³⁵ and subsequently found in other invertebrates as well as vertebrates including *Drosophila*,³⁶ starfish sperm,³⁷ *Xenopus laevis*,³⁸ rodents,³⁹ and humans.⁴⁰ *Homo sapiens* fascin-1 is the original vertebrate fascin discovered in humans, whereas retinal and testis fascin were later discovered and named fascin-2 and fascin-3, respectively.¹¹ Sequence alignment shows no similarities between fascins and other known actin binding proteins in humans, but strong similarities within the fascin family itself.⁴¹ Atomic models for actin-fascin bundles have been proposed based on optical diffraction studies of negatively stained reconstituted material.⁴² These suggest an 11 nm transverse banding pattern and uniformly polarized actin filaments organized in a hexagonal array with an interfilament distance of 11.5 nm. Assuming a single fascin monomer per crosslink, the predicted fascin-actin stoichiometry of 1:4.5 is in good agreement with experimentally measured ratios.⁴³ This maximal stoichiometry is a result of the helical twist of the actin filament that limits crosslinking sites between pairs of filaments to every fourth or fifth actin monomer. Fascin is additionally suggested to impart unusual mechanical stiffness to actin bundles in cells^{7,44–46} and reconstituted actin systems^{24,47,48} as compared with fimbrin, despite similar actin-binding affinities and bundle structure.^{42,49} Despite its importance to cell function, the structural basis for the unique localization and actin-bundling properties of fascin remain unknown. Here, we report the 2.9 Å resolution crystal structure of full-length recombinant *H. sapiens* fascin-1 and examine its packing, evolutionary sequence conservation, and conformational flexibility. The results of normal mode analysis of fascin suggest potential functional consequences of its distinct molecular structure and flexibility in crosslinking F-actin. Mutational and other experimental studies are needed to test these predictions.

RESULTS

Overall structure

The two fascin molecules in the asymmetric unit are highly similar, with a root-mean-square deviation (RMSD) of 0.64 Å for 474 C α atom pair equivalences. Fascin is composed of four tandem repeat β -trefoil domains (Figure 1) with pseudo-three-fold symmetry, consisting of 12 β -strands that form the barrel (B) and cap (C) regions in the sequence: BCCBBCCBBCCB (Figure 2a). The domains pack to form a distorted tetrahedron composed of two lobes: domains F1 (residues 8–139) and F2 (residues 140–260) form the first lobe and domains F3 (residues 261–381) and F4 (residues 382–493) form the second lobe. These two lobes can be superimposed with an RMSD of 2.3 Å for 224 C α atom pair equivalences. The two lobes are related by an approximate two-fold axis (rotation angle 163°, screw distance 8.1 Å). This axis passes between the two lobes and is approximately perpendicular to the plane of the image in Figure 1a. Despite the similarity of the two fascin lobes, the difference in relative internal orientations of their domains is 9°. Overall, 6,676 Å² of accessible surface area is buried at domain-domain interfaces, with approximately one-third of this surface area buried in the interface between domains F1 and F2, one-quarter between domains F3 and F4, 18% between domains F2 and F4, 11% between domains F1 and F4, and 11% between domains F2 and F3 (Table S1).

Within the first lobe, the two β -trefoil domains F1 and F2 are related by an approximate pseudo-two-fold axis (rotation 170° , screw $6.2 \text{ \AA}$) passing between these domains and lying approximately vertically in the plane of the image in Figure 1a, with the C-terminus of the first domain linked directly to the N-terminus of the second. Within the second lobe, the two β -trefoil domains F3 and F4 are also related by an approximate two-fold axis (rotation 169° , screw $7.3 \text{ \AA}$) passing between these domains and lying approximately vertically in the plane of the image in Figure 1a. Domains F3 and F4 in the second lobe are also packed head to tail like domains F1 and F2 in the first lobe. The interface between the domains in each lobe is formed by interactions between the β -barrels, with their β -hairpin triplets at the extreme ends of the lobes. This association results in an almost coaxial alignment of the pseudo-three-fold axes of the two domains within each end-capped β -barrel lobe (Figures 1c and 1d), with the pseudo-three-fold axis lying approximately perpendicular to the inter-domain dyad.

β -trefoil domain structure

Each β -trefoil domain consists of 12 β -strands that form three structurally homologous subunits (green, red, and blue in Figure 2a). These three subunits are related by a pseudo-three-fold axis and each contains a $\beta\beta\beta$ -loop- β motif⁵⁰ that forms two pairs of anti-parallel β -strands. The first and fourth β -strand of every unit coalesce to form a barrel structure (strands $\beta 1$, $\beta 4$, $\beta 5$, $\beta 8$, $\beta 9$ and $\beta 12$), with the strands inclined approximately 56° to the barrel axis. Three additional pairs of β -strands ($\beta 2$, $\beta 3$, $\beta 6$, $\beta 7$, $\beta 10$, and $\beta 11$) form three β -hairpins that cover the barrel from one side, with the exception of domains F1 and F4 based on the DSSP (Dictionary of Protein Secondary Structure).⁵¹ Structural superposition of the β -trefoil domains of fascin highlights their structural similarity (Figure 2b), with pairwise RMSDs ranging from 1.3 – $2.2 \text{ \AA}$ (Table S2).

The β -trefoil fold is present in over 40 distinct proteins deposited in the PDB (Protein Data Bank)⁵² according to the SCOP (Structural Classification of Proteins) database.⁵³ Members of the family include the mammalian cytokines interleukin-1 α ⁵⁴ and -1β ,⁵⁵ fibroblast growth factor,⁵⁶ soybean trypsin inhibitor,⁵⁷ the lectin B-chains from ricin⁵⁸ and abrin,⁵⁹ the galactose-specific lectin amaranthin,⁶⁰ and the actin-binding protein hisactophilin.⁶¹ The β -trefoil domains of fascin exhibit considerably higher sequence identity and structural similarity to each other than to other members of the β -trefoil family, suggesting evolution of the fascin molecule via multiple gene duplication events from an ancestral β -trefoil fold⁵⁶ (Figure S1a, and Tables S3 and S4).

The core of the β -trefoil fold is stabilized by hydrophobic interactions between bulky hydrophobic residues contributed by each β -strand, extending nearly to the barrel axis to form a tightly packed hydrophobic core.⁵⁰ Residues that stabilize the core of the β -trefoil fold are identified from *Erythrina caffra* trypsin inhibitor (ETI: PDB ID 1TIE),⁶² *H. sapiens* interleukin-1 β (IL-1 β : PDB ID 1I1B),⁵⁵ and *Bos torus* acidic fibroblast growth factor in Murzin et al.⁵⁰* Here, structure-based multiple sequence alignment of 59 β -trefoil domains obtained from 45 distinct proteins available in the PDB demonstrates high variability at all

* *Homo sapiens* acidic fibroblast growth factor is used instead of *Bos taurus* acidic fibroblast growth factor because structural coordinates for the latter are not currently available in the PDB.

amino acid positions other than these stabilizing hydrophobic core residues, which are highly conserved across the β -trefoil domains analyzed (Figures 2d and S2a). 90% of the sequence pairs of β -trefoil domains are less than 28% identical (Figure S3a). This indicates the phylogenetic diversity among the 59 β -trefoil domains. Residues responsible for hydrophobic packing in the β -hairpin triplet are mostly Leu, Val, and Ile in all 59 β -trefoil domains, where each β -hairpin donates two residues to stabilize the triplet.

It is of interest to compare the single β -trefoil domain protein hisactophilin from the motile slime mold *Dictyostelium discoideum* with fascin because it is also an actin-binding protein.^{61,63} Hisactophilin functions in osmoprotection of *D. discoideum*, enhancing the structural integrity of the cell cortex by crosslinking F-actin to the plasma membrane in a pH-dependent manner via protonation of its numerous histidine residues,⁶⁴ which compose 31 out of a total of 118 amino acids. Interestingly, while eight analogous positions are occupied by the charged residues His, Arg, and Lys in the first β -trefoil domain of fascin, seven in the second, four in the third, and three in the fourth, only three of these are His residues: the first domain has two His residues, the fourth domain has one, and domains two and three do not have any. Additionally, the 109 residues in fascin-1 that are at analogous positions to the histidines in hisactophilin appear to be distributed randomly throughout the molecule, including 29 interfacial positions between β -trefoil domains (Figure S1b). Thus, despite the structural and functional similarities of fascin to hisactophilin, regulation of its actin-binding via pH-dependent protonation of its histidine residues is unlikely to be present to the same extent as in hisactophilin, if at all.

The stabilizing hydrophobic core residues present in the β -trefoil fold are also highly conserved across homologous fascin molecules, although they are relatively more variable than compared with different β -trefoil proteins (30% versus 16% mean conservation grades) (Figures 3, 4a, and S2a). The conservation grade of the core-stabilizing hydrophobic residues in fascin and its homologues reduces to 19% when residues are classified by their physicochemical properties using a seven-letter alphabet for the entropy-based conservation measure, suggesting that their hydrophobic nature is functionally more important than their specific amino acid type.⁶⁵

β -trefoils associate to form two lobes in fascin

The β -trefoil domains of fascin associate at their exposed bases to form the (F1 and F2) first and (F3 and F4) second lobes of fascin with flanking polar interactions on the solvent-exposed surface of each lobe (Figures 4b, 4c, and S2b). A single ion pair, D98–R185, is present between domains F1 and F2 in the first lobe (Figure 4b), and another ion pair, R343–D450, is present between domains F3 and F4 in the second lobe (Figure 4c).^{*} Two proline residues, P140 and P384, are highly conserved in the fascin family with respective conservation grades of 11 and seven percent. The first induces a bend in the linker segment between domains 1 and 2 in the first lobe, and the second does the same in the linker between domains 3 and 4 in the second lobe. 27% of all interfacial residues between

^{*}These salt bridges are obtained for the oxygen-nitrogen distance cut-off of 3.2 Å. Increasing the cut-off to 4 Å results in one (D97–R224) and two (D342–K464 and R344–D420) more salt bridges between domains F1 and F2 and F3 and F4, respectively (see Physical property analysis).

domains F1 and F2 and between F3 and F4 reside in the first quartile of all residues in the molecule as ranked by conservation grade, suggesting that their inter-trefoil interactions might be important to the structural stability of each lobe. A detailed analysis of the relative contributions of electrostatic and hydrophobic interactions to lobe and fold stability would be of interest, but are beyond the scope of the present work.

Lobes associate to form the full-length fascin molecule

The full-length fascin molecule is formed by association of the lobes at approximately a 56 degree skew angle (Figure 1). The central region of fascin is stabilized by multiple polar interactions connecting all four β -trefoil domains together, where the residues H139, Q141, S259, R383, and R389 that contribute to these interactions are also highly conserved in the fascin family with respective conservation grades of nine, eight, ten, 0.5, and 17 percent (Figures 3 and 5). 43% of all interfacial residues between the lobes of fascin reside in the first quartile of residues in the molecule as ranked by conservation grade, also suggesting their importance to the overall structural stability of the full-length molecule.

Putative actin-binding sites of fascin

In the absence of any structure of a fascin-F-actin complex, the precise actin-binding sites of fascin remain unknown. However, two point mutations in *Drosophila melanogaster* fascin (the singed gene) are known to impact bundling: G409E (equivalent to G393 in *H. sapiens* fascin-1) that leads to partial inactivation of fascin *in vivo* and S289N (equivalent to S274 in *H. sapiens* fascin-1) that almost completely inactivates fascin, suggesting that one actin-binding domain may be in the region near S274 that is also highly conserved (14% mean conservation grade) and solvent exposed (Figures 3 and 6).⁶⁶ Interestingly, mutation of S251 to phenylalanine (equivalent to P236 in *H. sapiens* fascin-1) partially restores fascin function lost in the S289N mutation for reasons that are unknown.⁶⁶

A second actin-binding site is thought to be in the highly conserved protein-kinase-Ca domain consisting of residues 29–43 in the first β -trefoil domain of fascin (Figure 3). It has high sequence similarity to an actin-binding site of myristoylated alanine-rich C-kinase substrate (MARCKS) FGFKVNASASSLKKK (residues 29–43 in *H. sapiens* fascin-1).¹¹ Mutation of S39 to alanine renders *in vitro* actin-binding by fascin-1 insensitive to regulation by phosphorylation by PKC α ,^{34,67} and is therefore a constitutively active mutant as shown also in mouse B16F1 cells.⁷ There, the constitutively active mutant fascin leads to a significant increase in the number and length of cortical filopodia, whereas the inactive phosphomimetic fascin mutant S39E is shown to reduce filopodial frequency.⁷

The majority of the highly conserved residues of fascin have functional importance that is rationalized based either on contribution to fold/structural stability or to actin-binding. Examples include the putative actin-binding site between residues 29–43, hydrophobic core stabilizing residues in the β -trefoil fold, interfacial residues between the lobes, and to lesser extent interfacial residues between β -trefoil domains within each lobe. This is supported by the distribution of conservation grades for each of these sub-sets of residues, which is biased towards high conservation when compared with all residues in the molecule (Figure S4 and

S5a). The remaining subset of highly conserved residues in the molecule consists of 51 (41%) residues (Figure S5 and Tables S5 and S6).

Half of these residues consist either of residues that are highly conserved across fascin molecules and the β -trefoil fold, suggesting their functional importance to the fold but not necessarily to the specific biological function of fascin, or glycines, which likely confer local flexibility to the fascin molecule. Of the remaining residues, approximately ten percent are located proximal to the putative actin-binding site consisting of residues 29–43, and the remainder are located proximal to the known functional sites S274 and P236, discussed previously (Figure 6b and Table S6). The fact that these residues are highly conserved across homologous fascins but not generally across the β -trefoil fold suggests that these have functional importance to the molecule, such as in binding to F-actin or stabilizing the fascin fold.

Conformational dynamics

Analysis of the conformational dynamics of fascin illustrates that end-association of the β -trefoil domains confers structural integrity within each cylindrical lobe that is not present across lobes (Figure S6). The generalized linear correlation coefficient (r_{LMI}) calculated for residue pairs using NMA may be used to infer the dynamical correlation between different regions of the protein (Figure 7). Interestingly, correlations in the motions of domains F1 and F3 are the highest among all pairs of β -trefoil domains despite the fact that they are not in direct contact (Figure 7b and Table 1, and see also Figure S7). This suggests a potential allosteric coupling between these domains, which contain the putative actin-binding sites of the molecule. The structural origin of this long-range coupling is in the tight binding between the β -trefoil domains of fascin to form two β -barrel lobes that interact at an interfacial region that forms a hinge about which the lobes may bend (Figure S6). This lobe-structure is in stark contrast to the similarly sized crosslinking protein fimbrin, which consists of four calponin homology domains that each consist of triple α -helix bundles.⁶⁸ Further analysis of fimbrin is required to understand potential functional implications of the distinct architectures of these molecules. Identification of the precise actin-binding sites of fascin via mutagenesis and structural studies are needed to elucidate implications of the observed structural and dynamical properties of the molecule on its actin-crosslinking function.

DISCUSSION

The crystal structure of *H. sapiens* fascin-1 reveals that its β -trefoil domains associate via internal hydrophobic interactions and external ionic interactions at their bases to form cylindrical β -barrel lobes. NMA indicates that these interactions confer structural integrity to each cylindrical lobe that is maintained in the full-length molecule. Conservation analysis confirms the functional importance of the cores of the β -trefoil domains, which contain bulky hydrophobic residues that are highly conserved both across fascins and the entire β -trefoil family of proteins, which are otherwise divergent in sequence and biological function. Indeed, the majority of highly conserved residues in fascin are suggested either to contribute to its structural stability or to binding to F-actin. These residues are located in and near the

putative actin-binding site between residues 29–43 and the previously identified functional residues S274 and P236, hydrophobic core stabilizing residues of the β -trefoil fold, interfacial residues between the lobes, and to lesser extent interfacial residues between β -trefoil domains within each lobe (Figures S4 and S5, and Table S6).

Association of the cylindrical lobes of fascin at their largely polar interface results in a hinge that facilitates large-scale collective motions of each lobe, whereby the opposing putative actin-binding domains of fascin are correlated in their dynamical motions. This distinct molecular design of fascin might have important implications on its biological function via a lever-like mechanism: action at one end of the molecule is transmitted in a lever-like fashion to the other end. Indeed, fascin binding to F-actin at one end (e.g., between F1 and F4) may induce a direct conformational change at the opposing end of the molecule (between F2 and F3) that could alter its binding affinity for F-actin, resulting in cooperative binding that may explain its unique ability to form tightly packed and ordered actin bundles in vitro and in vivo. Future mutational studies guided in part by the results of this work, together with cryo-EM-based reconstructions of the actin-bound fascin crosslink, will eventually facilitate a detailed understanding of the molecular basis for the unique structural and mechanical properties that these ubiquitous and highly conserved actin-binding proteins confer to actin bundles.

EXPERIMENTAL AND COMPUTATIONAL PROCEDURES

Purification and crystallization

Recombinant *H. sapiens* fascin was prepared using a procedure similar to that described earlier.³⁴ *H. sapiens* fascin was expressed as a GST fusion using pGEX2T in the *E. coli* strain JR600. Cultures were grown to an OD₆₀₀ of 0.6 at 37 °C and cooled to 25 °C. Protein expression was induced by the addition of 1mM IPTG and growth continued overnight at 25 °C. Cells were harvested by centrifugation, washed with PBS buffer, resuspended in PBS containing 1 mM DTT, 0.2 mM PMSF and disrupted by sonication. The lysate was clarified by centrifugation in a SS-34 rotor at 15,000 rpm for 20 minutes and applied to a glutathione–Sepharose column equilibrated in PBS containing 1 mM DTT (PBS/DTT). The fusion protein was eluted with 10 mM glutathione, 50 mM Tris, 1 mM DTT, 0.1 M NaCl, pH 8.0 and dialyzed against PBS/DTT. Fascin was liberated from GST by cleavage with thrombin, followed by glutathione-Sepharose chromatography. Fascin was subsequently dialyzed against 10 mM Tris, 1 mM DTT, pH 8.0 and applied to a DE52 column that was developed with a linear gradient of 0 to 70 mM NaCl in 10 mM Tris, 1 mM DTT, pH 8.0. The purified molecule was dialyzed against 10 mM Tris, 10 mM NaCl, 30 mM KCl, 0.1 mM EDTA, 1 mM DTT, pH 8.0 and stored at –70 °C. The expressed protein consisted of 507 residues, including a 14-residue N-terminal extension that results from the cloning strategy. Selenomethionyl-substituted (Se-Met) fascin was prepared by transforming *E. Coli* strain B834, a methionine auxotroph and purified as described for native fascin. Native and selenomethionyl fascin crystals were grown at room temperature by hanging drop vapor diffusion, with a protein concentration of 20 mg/ml in 10 mM Tris HCl, pH 8.0, 10 mM NaCl, 30 mM KCl, 0.1 mM EDTA and 1 mM DTT. 6 μ l drops composed of equal volumes of protein solution and a well solution were suspended over 1 ml reservoirs containing 20%

PEG 4000, 0.1 M HEPES, pH 7.5 and 1 mM DTT. Rod-like crystals grew to a typical size 0.6×0.2×0.2 mm in five to ten days and exhibited diffraction consistent with space group C2 ($a = 165.43$, $b = 71.69$, $c = 116.92$ Å, $\beta = 132.17^\circ$; with two fascin molecules per asymmetric unit). Isomorphous derivatives were prepared by soaking native crystals in mother liquor containing 1 mM KAu(CN)₂, 0.1 mM Kag(CN)₂ and 1 mM Pt(Cl)₂-tetrapyridine. Se-Met fascin was used as fourth derivative and a fifth derivative was prepared by soaking Se-Met crystals in 1 mM KAu(CN)₂.

Structure determination and refinement

All data were collected at room temperature on a Siemens X-1000 area detector coupled to an RU200 rotating anode and were reduced and scaled using XDS and XSCALE.⁶⁹ Heavy atom sites were located in difference Patterson maps and refined with the PHASES package.⁷⁰ Data collection and phasing statistics are given in Table 2. Non-crystallographic symmetry (NCS) was determined by comparing pairs of NCS-related heavy atom positions. The NCS transformation was further refined in iterative solvent flattening and NCS averaging procedure implemented in program DM.⁷¹ Initial maps were fit with polyalanine using TOM⁷² and allowed for the entire fold to be determined. A representative part of experimental electron density map is shown in Figure 8. Subsequent phase combination allowed for the placement of the correct amino acid sequence. The structure was refined using X-PLOR 3.1,⁷³ employing tight NCS restraints during each step of the refinement. The N-terminal extension (14 residues) and the first 7 fascin residues are not visible in two non-crystallographically related molecules. In addition, two loops composed of residues 50–57 and 276–279 have very high temperature factors and were not included in the refinement. The final fascin model consists of two non-crystallographically related molecules and includes 7431 nonhydrogen atoms. The model has an R-value of 22.3 %, a free R-factor of 26.8% in the resolution range 25.0–2.9 Å, and no Ramachandran outliers. Final refinement statistics are provided in Table 2.

Sequence analysis

Three independent procedures for computing evolutionary sequence conservation are employed: the Conservation Surface Mapping method (ConSurf),^{74,75} the Real-valued Evolutionary Trace method (ET),⁷⁶ and a simple entropy-based method using a 21-letter alphabet.⁶⁵ The correlation coefficients between each of these three conservation analysis procedures are greater than 80%. In the absence of any information regarding the superiority of any one approach, the grades are weighted such that each contributes equally to the final conservation grade, except for positions that are missing residues in the original PDB file, for which only the entropy-21-based grade is used. The “conservation grade” presented in Results is defined as percent relative to all residues in the protein. For example, a residue conservation grade of 5% means that 5% of residues in the protein are at least as conserved as this residue, and 95% are less conserved. Highly conserved residues are defined as residues residing in the first quartile of residues in the molecule as ranked by conservation grade.

ConSurf is applied on the set of homologous *H. sapiens* fascin-1 sequences stored in the ConSurf Databank (ConSurf-DB).⁷⁵ ConSurf-DB stores results of ConSurf analyses on all

structures in the PDB.⁵² Conservation grades are computed using the Rate4Site algorithm⁷⁷ on 22 homologous sequences selected by the ConSurf-DB protocol for *H. sapiens* fascin-1.⁷⁵

The Real-valued ET procedure is a hybrid method that incorporates an entropy grade into a phylogenetic analysis of a multiple sequence alignment.⁷⁶ The phylogenetic tree is created using hierarchical clustering based on the Unweighted Pair Group Method with Arithmetic Mean.⁷⁸ The tree is divided into distinct subfamilies using different partitionings, where an entropy grade is calculated for each position in a sub-family-specific manner.⁷⁶ Conservation analysis is performed on *H. sapiens* fascin-1 (SwissProt ID Q16658) using 31 sequences extracted from the HSSP database.⁷⁹

In the third conservation analysis procedure, a simple entropy measure with a 21-letter alphabet is used. Sequences with high homology to *H. sapiens* fascin-1 are retrieved from the NCBI⁸⁰ using PSI-BLAST⁸¹ with an E-value cutoff of 10^{-6} and one iteration. Sequences that differ by more than 5% in length from the query sequence are removed together with redundant sequences, resulting in 61 sequences (Table S7). Sequences are aligned using ClustalX⁸² with default parameters. The entropy of each position in the sequence alignment is computed using the standard measure for entropy, $S_i = -\sum_{k=1}^{21} p_{ik} \log_{21} p_{ik}$, where p_{ik} is the probability of occurrence of amino acid type k at sequence position i .

Structure-based sequence alignment of β -trefoil domains is performed using STAMP.⁸³

Physical property analysis

Interfacial residues between fascin domains are identified by the change in their relative solvent accessibility upon domain-domain association,⁸⁴ where relative solvent accessibility is defined as the ratio of the actual solvent accessibility of the residue to its solvent accessibility in the extended Gly-x-Gly tripeptide.⁸⁵ Interfacial residues are defined as those residues that change their relative solvent accessibility by more than 1% between the dissociated and associated states.⁸⁴ Salt bridges are found using VMD 1.8.6 salt bridge plugin⁸⁶ with default parameters. Increasing the oxygen-nitrogen distance cut-off of 3.2 Å (VMD default value) to 4 Å⁸⁷ is also considered.

Conformational dynamics of fascin are calculated using Normal Mode Analysis (NMA).^{88,89} The crystal structures of chains A and B of fascin (PDB ID 1DFC) are analyzed with missing loops built using Swiss-PDB viewer.⁹⁰ All-atom energy minimization and subsequent NMA are performed using CHARMM v33a1⁹¹ using the implicit solvent force-field EEF1.⁹² Hydrogen atoms and amino acid side chains are first minimized sequentially with all remaining protein atoms fixed. Repeated rounds of Steepest Descent minimization followed by Adopted Basis Newton-Raphson minimization are subsequently performed in the presence of successively reduced harmonic constraints applied to backbone chain atoms. The final energy-minimized structures of chains A and B have RMS energy gradients of 4×10^{-4} and 2.5×10^{-4} kcal/mol/Å, respectively, with RMS coordinate differences between the crystal and energy-minimized structures of 1.9 and 1.7 Å. The Block Normal Mode method^{93,94} is used with one residue per block to compute the 207 lowest frequency normal modes. The equilibrium thermal fluctuation of α -carbon atom i is computed using

$\langle \Delta r_i^2 \rangle = k_B T \sum_k y_{ik}^2 / \lambda_k m_i$ m_i is atomic mass, $k_B T$ is thermal energy, is the eigenvalue corresponding to normal mode k , and y_{ik}^2 is the squared magnitude of the projections of the k^{th} normal mode on the Cartesian components of the displacement vector of the i^{th} atom.⁹⁵

Correlations in atomic fluctuations are computed using the generalized linear correlation coefficient, $r_{\text{LMI}}[\mathbf{x}_i, \mathbf{x}_j] = \left(1 - \exp\left(-2I_{\text{lin}}[\mathbf{x}_i, \mathbf{x}_j]/3\right)\right)^{1/2}$, based on the linear mutual information metric, $I_{\text{lin}}[\mathbf{x}_i, \mathbf{x}_j] = \frac{1}{2}[\ln \det \mathbf{C}_{(i)} + \ln \det \mathbf{C}_{(j)} - \ln \det \mathbf{C}_{(ij)}]$, where $\mathbf{C}_{(i)} = \langle \mathbf{x}_i^T \mathbf{x}_i \rangle$ and $\mathbf{C}_{(ij)} = \langle \mathbf{x}_i, \mathbf{x}_j^T \mathbf{x}_i, \mathbf{x}_j \rangle$ and are marginal-covariances of atom i and the pair-covariance matrix of atoms i and j , respectively (Supplementary Material).⁹⁶ Results are compared with the more commonly employed Pearson correlation coefficient, $C_{ij} = \langle \mathbf{x}_i \cdot \mathbf{x}_j \rangle / \langle \mathbf{x}_i^2 \rangle \langle \mathbf{x}_j^2 \rangle^{1/2}$, where $\mathbf{x}_i$ denotes displacement of atom i in Figure S8.

Accession numbers

X-ray coordinates and structure factors for *H. sapiens* fascin-1 have been deposited in the Protein Data Bank under accession code 1DFC.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

The authors are grateful to Martin Karplus and Danijela Vignjevic for extensive comments on the manuscript. Funding from the Samuel A. Goldblith Professorship to M.B. is gratefully acknowledged.

Abbreviations used:

F-actin	filamentous actin
NMA	normal mode analysis
RMSD	root-mean-square deviation
DSSP	dictionary of protein secondary structure
PDB	protein data bank
SCOP	structural classification of proteins
ETI	<i>Erythrina caffra</i> trypsin inhibitor
IL-1β	interleukin-1 β
H	hisactophilin
FGF	fibroblast growth factor
MARCKS	myristoylated alanine-rich C-kinase substrate

Se-Met	Selenomethionyl-substituted
NCS	Non-crystallographic symmetry
ConSurf	conservation surface mapping method
ET	evolutionary trace method
ConSurf-DB	ConSurf databank

REFERENCES

1. Abercrombie M (1980). The Croonian Lecture, 1978 - The crawling movement of metazoan cells. *Proc. R. Soc. Lond. B. Biol. Sci* 207, 129–147.
2. Small JV, Stradal T, Vignal E & Rottner K (2002). The lamellipodium: where motility begins. *Trends Cell Biol.* 12, 112–120. [PubMed: 11859023]
3. Pollard TD & Borisy GG (2003). Cellular motility driven by assembly and disassembly of actin filaments. *Cell* 112, 453–465. [PubMed: 12600310]
4. Pantaloni D, Le Clainche C & Carlier MF (2001). Mechanism of actin-based motility. *Science* 292, 1502–1506. [PubMed: 11379633]
5. Matsudaira P (1994). Actin crosslinking proteins at the leading edge. *Semin. Cell Biol* 5, 165. [PubMed: 7919230]
6. Faix J & Rottner K (2005). The making of filopodia. *Curr. Opin. Cell Biol* 18, 18–25. [PubMed: 16337369]
7. Vignjevic D, Kojima S, Aratyn Y, Danciu O, Svitkina T & Borisy GG (2006). Role of fascin in filopodial protrusion. *J. Cell Biol* 174, 863–875. [PubMed: 16966425]
8. Mallavarapu A & Mitchison T (1999). Regulated actin cytoskeleton assembly at filopodium tips controls their extension and retraction. *J. Cell Biol* 146, 1097–1106. [PubMed: 10477762]
9. Romero S, Le Clainche C, Didry D, Egile C, Pantaloni D & Carlier MF (2004). Formin is a processive motor that requires profilin to accelerate actin assembly and associated ATP hydrolysis. *Cell* 119, 419–429. [PubMed: 15507212]
10. Goode BL & Eck MJ (2007). Mechanism and function of formins in the control of actin assembly. *Annu. Rev. Biochem* 76, 593–627. [PubMed: 17373907]
11. Kureishy N, Sapountzi V, Prag S, Anilkumar N & Adams JC (2002). Fascins, and their roles in cell structure and function. *Bioessays* 24, 350–361. [PubMed: 11948621]
12. Adams JC (2004). Roles of fascin in cell adhesion and motility. *Curr. Opin. Cell Biol* 16, 590–596. [PubMed: 15363811]
13. Hashimoto Y, Skacel M & Adams JC (2005). Roles of fascin in human carcinoma motility and signaling: Prospects for a novel biomarker? *Int J Biochem Cell Biol* 37, 1787–1804. [PubMed: 16002322]
14. Jawhari AU, Buda A, Jenkins M, Shehzad K, Sarraf C, Noda M, Farthing MJG, Pignatelli M & Adams JC (2003). Fascin, an actin-bundling protein, modulates colonic epithelial cell invasiveness and differentiation in vitro. *Am. J. Pathol* 162, 69–80. [PubMed: 12507891]
15. Vignjevic D, Schoumacher M, Gavert N, Janssen KP, Jih G, Lae M, Louvard D, Ben-Ze'ev A & Robine S (2007). Fascin, a novel target of beta-Catenin-TCF signaling, is expressed at the invasive front of human colon cancer. *Cancer Res.* 67, 6844–6853. [PubMed: 17638895]
16. Peraud A, Mondal S, Hawkins C, Mastronardi M, Bailey K & Rutka JT (2003). Expression of fascin, an actin-bundling protein, in astrocytomas of varying grades. *Brain Tumor Pathol.* 20, 53–58. [PubMed: 14756441]
17. Hashimoto Y, Parsons M & Adams JC (2007). Dual actin-bundling and protein kinase c-binding activities of fascin regulate carcinoma cell migration downstream of Rac and contribute to metastasis. *Mol. Biol. Cell* 18, 4591–4602. [PubMed: 17855511]
18. Matsudaira P (1991). Modular organization of actin cross-linking proteins. *Trends Biochem. Sci* 16, 87–92. [PubMed: 2058002]

19. Revenu C, Athman R, Robine S & Louvard D (2004). The co-workers of actin filaments: From cell structures to signals. *Nat. Rev. Mol. Cell Biol* 5, 635–646. [PubMed: 15366707]
20. Lodish H, Berk A, Zipursky SL, Matsudaira P, Baltimore D & Darnell J (1999). *Molecular Cell Biology*. 4th edit, W.H. Freeman and Company, New York.
21. Bartles JR (2000). Parallel actin bundles and their multiple actin-bundling proteins. *Curr. Opin. Cell Biol* 12, 72–78. [PubMed: 10679353]
22. Stossel TP, Condeelis J, Cooley L, Hartwig JH, Noegel A, Schleicher M & Shapiro SS (2001). Filamins as integrators of cell mechanics and signalling. *Nat. Rev. Mol. Cell Biol* 2, 138–145. [PubMed: 11252955]
23. Liu J, Taylor DW & Taylor KA (2004). A 3-D reconstruction of smooth muscle alpha-actinin by CryoEM reveals two different conformations at the actin-binding region. *J. Mol. Biol* 338, 115–125. [PubMed: 15050827]
24. Tseng Y, Fedorov E, McCaffery JM, Almo SC & Wirtz D (2001). Micromechanics and ultrastructure of actin filament networks crosslinked by human fascin: A comparison with alpha-actinin. *J. Mol. Biol* 310, 351–366. [PubMed: 11428894]
25. Carugo KD, Banuelos S & Saraste M (1997). Crystal structure of a calponin homology domain. *Nat. Struct. Biol* 4, 175–179. [PubMed: 9164454]
26. Goldsmith SC, Pokala M, Shen WY, Fedorov AA, Matsudaira P & Almo SC (1997). The structure of an actin-crosslinking domain from human fimbrin. *Nat. Struct. Biol* 4, 708–712. [PubMed: 9302997]
27. Keep NH, Norwood FLM, Moores CA, Winder SJ & Kendrick-Jones J (1999). The 2.0 angstrom structure of the second calponin homology domain from the actin-binding region of the dystrophin homologue utrophin. *J. Mol. Biol* 285, 1257–1264. [PubMed: 9887274]
28. Keep NH, Winder SJ, Moores CA, Walke S, Norwood FLM & Kendrick-Jones J (1999). Crystal structure of the actin-binding region of utrophin reveals a head-to-tail dimer. *Struct. Fold. Des* 7, 1539–1546.
29. Fucini P, Renner C, Herberhold C, Noegel AA & Holak TA (1997). The repeating segments of the F-actin cross-linking gelation factor (ABP-120) have an immunoglobulin-like fold. *Nat. Struct. Biol* 4, 223–230. [PubMed: 9164464]
30. McCoy AJ, Fucini P, Noegel AA & Stewart M (1999). Structural basis for dimerization of the *Dictyostelium* gelation factor (ABP120) rod. *Nat. Struct. Biol* 6, 836–841. [PubMed: 10467095]
31. Djinić-Carugo K, Young P, Gautel M & Saraste M (1999). Structure of the alpha-actinin rod: Molecular basis for cross-linking of actin filaments. *Cell* 98, 537–546. [PubMed: 10481917]
32. Namba Y, Ito M, Zu YL, Shigesada K & Maruyama K (1992). Human T-cell L-plastin bundles actin-filaments in a calcium-dependent manner. *J. Biochem* 112, 503–507. [PubMed: 1491005]
33. Dearruda MV, Watson S, Lin CS, Leavitt J & Matsudaira P (1990). Fimbrin is a homolog of the cytoplasmic phosphoprotein plastin and has domains homologous with calmodulin and actin gelation proteins. *J. Cell Biol* 111, 1069–1079. [PubMed: 2391360]
34. Ono S, Yamakita Y, Yamashiro S, Matsudaira PT, Gnarr JR, Obinata T & Matsumura F (1997). Identification of an actin binding region and a protein kinase C phosphorylation site on human fascin. *J. Biol. Chem* 272, 2527–2533. [PubMed: 8999969]
35. Kane RE (1975). Preparation and purification of polymerized actin from sea urchin egg extracts. *J. Cell Biol* 66, 305–315. [PubMed: 1095598]
36. Cant K, Knowles BA, Mooseker MS & Cooley L (1994). *Drosophila Singed*, a fascin homolog, is required for actin bundle formation during oogenesis and bristle extension. *J. Cell Biol* 125, 369–380. [PubMed: 8163553]
37. Maekawa S, Endo S & Sakai H (1982). A protein in starfish sperm head which bundles actin-filaments *in vitro*: purification and characterization. *J. Biochem* 92, 1959–1972. [PubMed: 7161268]
38. Holthuis JCM, Schoonderwoert VTG & Martens GJM (1994). A vertebrate homolog of the actin-bundling protein fascin. *Biochim. Biophys. Acta Gene Struct. Expression* 1219, 184–188.
39. Edwards RA, Herrerasosa H, Otto J & Bryan J (1995). Cloning and expression of a murine fascin homolog from mouse brain. *J. Biol. Chem* 270, 10764–10770. [PubMed: 7738015]

40. Duh FM, Latif F, Weng YK, Geil L, Modi W, Stackhouse T, Matsumura F, Duan DR, Linehan WM, Lerman MI & Gnarr JR (1994). cDNA cloning and expression of the human homolog of the sea urchin fascin and *Drosophila singed* genes which encodes an actin-bundling protein. *DNA Cell Biol.* 13, 821–827. [PubMed: 8068206]
41. Edwards RA & Bryan J (1995). Fascins, a family of actin bundling proteins. *Cell Motil. Cytoskeleton* 32, 1–9. [PubMed: 8674129]
42. DeRosier DJ & Tilney LG (1982). How actin filaments pack into bundles. *Cold Spring Harb. Symp. Quant. Biol* 46, 525–540. [PubMed: 6955098]
43. Bryan J & Kane RE (1978). Separation and interaction of major components of sea urchin actin gel. *J. Mol. Biol* 125, 207–224. [PubMed: 731692]
44. Aratyn YS, Schaus TE, Taylor EW & Borisy GB (2007). Intrinsic dynamic behavior of fascin in filopodia. *Mol. Biol. Cell* 18, 3928–3940. [PubMed: 17671164]
45. Bathe M, Heussinger C, Claessens M, Bausch AR & Frey E (2008). Cytoskeletal bundle mechanics. *Biophys. J* 94, 2955–2964. [PubMed: 18055529]
46. Howard J (2008). Molecular mechanics of cells and tissues. *Cell. Mol. Bioeng* 1, 24–32.
47. Claessens MMAE, Bathe M, Frey E & Bausch AR (2006). Actin-binding proteins sensitively mediate F-actin bundle stiffness. *Nature Mater.* 5, 748–753. [PubMed: 16921360]
48. Honda M, Takiguchi K, Ishikawa S & Hotani H (1999). Morphogenesis of liposomes encapsulating actin depends on the type of actin-crosslinking. *J. Mol. Biol* 287, 293–300. [PubMed: 10080893]
49. Volkmann N, DeRosier D, Matsudaira P & Hanein D (2001). An atomic model of actin filaments cross-linked by fimbrin and its implications for bundle assembly and function. *J. Cell Biol* 153, 947–956. [PubMed: 11381081]
50. Murzin AG, Lesk AM & Chothia C (1992). Beta-trefoil fold patterns of structure and sequence in the kunitz inhibitors interleukins-1-beta and 1-alpha and fibroblast growth factors. *J. Mol. Biol* 223, 531–543. [PubMed: 1738162]
51. Kabsch W & Sander C (1983). Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577–2637. [PubMed: 6667333]
52. Bernstein FC, Koetzle TF, Williams GJB, Meyer EF, Brice MD, Rodgers JR, Kennard O, Shimanouchi T & Tasumi M (1977). The Protein Data Bank: A computer-based archival file for macromolecular structures. *J. Mol. Biol* 112, 535–542. [PubMed: 875032]
53. Murzin AG, Brenner SE, Hubbard T & Chothia C (1995). SCOP: A structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol* 247, 536–540. [PubMed: 7723011]
54. Graves BJ, Hatada MH, Hendrickson WA, Miller JK, Madison VS & Satow Y (1990). Structure of interleukin 1-alpha at 2.7 Angstrom resolution. *Biochemistry* 29, 2679–2684. [PubMed: 2346741]
55. Finzel BC, Clancy LL, Holland DR, Muchmore SW, Watenpaugh KD & Einspahr HM (1989). Crystal structure of recombinant human interleukin-1-beta at 2.0 Angstroms resolution. *J. Mol. Biol* 209, 779–791. [PubMed: 2585509]
56. Ponting CP & Russell RB (2000). Identification of distant homologues of fibroblast growth factors suggests a common ancestor for all beta-trefoil proteins. *J. Mol. Biol* 302, 1041–1047. [PubMed: 11183773]
57. McLachlan AD (1979). 3-fold structural pattern in the soybean trypsin-inhibitor (kunitz). *J. Mol. Biol* 133, 557–563. [PubMed: 537058]
58. Rutenber E & Robertus JD (1991). Structure of ricin B-chain at 2.5-A resolution. *Proteins: Struct. Funct. Genet* 10, 260–269. [PubMed: 1881882]
59. Tahirov TH, Lu TH, Liaw YC, Chen YL & Lin JY (1995). Crystal structure of abrin-A at 2.14 Angstrom. *J. Mol. Biol* 250, 354–367. [PubMed: 7608980]
60. Transue TR, Smith AK, Mo HQ, Goldstein IJ & Saper MA (1997). Structure of benzyl T-antigen disaccharide bound to *Amaranthus caudatus* agglutinin. *Nat. Struct. Biol* 4, 779–783. [PubMed: 9334739]
61. Habazettl J, Gondol D, Wilschek R, Otlewski J, Schleicher M & Holak TA (1992). Structure of hisactophilin is similar to interleukin-1-beta and fibroblast growth factor. *Nature* 359, 855–858. [PubMed: 1436061]

62. Onesti S, Brick P & Blow DM (1991). Crystal-structure of a Kunitz-type trypsin-inhibitor from *Erythrina caffra* seeds. *J. Mol. Biol* 217, 153–176. [PubMed: 1988676]
63. Scheel J, Ziegelbauer K, Kupke T, Humbel BM, Noegel AA, Gerisch G & Schleicher M (1989). Hisactophilin, a histidine-rich actin-binding protein from *Dictyostelium discoideum*. *J. Biol. Chem* 264, 2832–2839. [PubMed: 2914932]
64. Hanakam F, Gerisch G, Lotz S, Alt T & Seelig A (1996). Binding of hisactophilin I and II to lipid membranes is controlled by a pH-dependent myristoyl-histidine switch. *Biochemistry* 35, 11036–11044. [PubMed: 8780505]
65. Valdar WSJ (2002). Scoring residue conservation. *Proteins: Struct. Funct. Genet* 48, 227–241. [PubMed: 12112692]
66. Cant K & Cooley L (1996). Single amino acid mutations in *Drosophila* fascin disrupt actin bundling function in vivo. *Genetics* 143, 249–258. [PubMed: 8722779]
67. Yamakita Y, Ono S, Matsumura F & Yamashiro S (1996). Phosphorylation of human fascin inhibits its actin binding and bundling activities. *J. Biol. Chem* 271, 12632–12638. [PubMed: 8647875]
68. Klein MG, Shi WX, Ramagopal U, Tseng Y, Wirtz D, Kovar DR, Staiger CJ & Almo SC (2004). Structure of the actin crosslinking core of fimbrin. *Structure* 12, 999–1013. [PubMed: 15274920]
69. Kabsch W (1988). Evaluation of single-crystal X-ray-diffraction data from a position-sensitive detector. *J. Appl. Crystallogr* 21, 916–924.
70. Furey W & Swaminathan S (1990). PHASES: A program package for the processing and analysis of diffraction data from macromolecules. American Crystallographic Association Meeting Abstracts, Series 2 18, 73.
71. Bailey S (1994). The CCP4 Suite-programs for protein crystallography. *Acta Crystallogr. Sect. D. Biol. Crystallogr* 50, 760–763. [PubMed: 15299374]
72. Jones TA (1985). Interactive computer graphics: FRODO. *Methods Enzymol.* 115, 157–171. [PubMed: 3841179]
73. Brunger A (1992). X-PLOR: version 3.1: A system for X-ray crystallography and NMR, Yale University Press, New Haven, Connecticut.
74. Armon A, Graur D & Ben-Tal N (2001). ConSurf: An algorithmic tool for the identification of functional regions in proteins by surface mapping of phylogenetic information. *J. Mol. Biol* 307, 447–463. [PubMed: 11243830]
75. Goldenberg O, Erez E, Nimrod G & Ben-Tal N (2009). The ConSurf-DB: pre-calculated evolutionary conservation profiles of protein structures. *Nucleic Acids Res.* 37, D323–D327. [PubMed: 18971256]
76. Mihalek I, Res I & Lichtarge O (2004). A family of evolution-entropy hybrid methods for ranking protein residues by importance. *J. Mol. Biol* 336, 1265–1282. [PubMed: 15037084]
77. Mayrose I, Graur D, Ben-Tal N & Pupko T (2004). Comparison of site-specific rate-inference methods for protein sequences: Empirical Bayesian methods are superior. *Mol. Biol. Evol* 21, 1781–1791. [PubMed: 15201400]
78. Lichtarge O, Bourne HR & Cohen FE (1996). An evolutionary trace method defines binding surfaces common to protein families. *J. Mol. Biol* 257, 342–358. [PubMed: 8609628]
79. Schneider R, deDaruvar A & Sander C (1997). The HSSP database of protein structure-sequence alignments. *Nucleic Acids Res.* 25, 226–230. [PubMed: 9016541]
80. Wheeler DL, Church DM, Federhen S, Lash AE, Madden TL, Pontius JU, Schuler GD, Schriml LM, Sequeira E, Tatusova TA & Wagner L (2003). Database resources of the National Center for Biotechnology. *Nucleic Acids Res.* 31, 28–33. [PubMed: 12519941]
81. Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W & Lipman DJ (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389–3402. [PubMed: 9254694]
82. Thompson JD, Higgins DG & Gibson TJ (1994). Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. [PubMed: 7984417]
83. Russell RB & Barton GJ (1992). Multiple protein sequence alignment from tertiary structure comparison: Assignment of global and residue confidence levels. *Proteins: Struct. Funct. Genet* 14, 309–323. [PubMed: 1409577]

84. Caffrey DR, Somaroo S, Hughes JD, Mintseris J & Huang ES (2004). Are protein-protein interfaces more conserved in sequence than the rest of the protein surface? *Protein Sci.* 13, 190–202. [PubMed: 14691234]
85. Miller S, Lesk AM, Janin J & Chothia C (1987). The accessible surface area and stability of oligomeric proteins *Nature* 328, 834–836. [PubMed: 3627230]
86. Humphrey W, Dalke A & Schulten K (1996). VMD: Visual molecular dynamics. *J. Mol. Graph* 14, 33–38. [PubMed: 8744570]
87. Kumar S & Nussinov R (1999). Salt bridge stability in monomeric proteins. *J. Mol. Biol* 293, 1241–1255. [PubMed: 10547298]
88. Brooks B & Karplus M (1983). Harmonic dynamics of proteins: Normal modes and fluctuations in bovine pancreatic trypsin inhibitor. *Proc. Natl. Acad. Sci. U. S. A* 80, 6571–6575. [PubMed: 6579545]
89. Go N, Noguti T & Nishikawa T (1983). Dynamics of a small globular protein in terms of low frequency vibrational modes. *Proc. Natl. Acad. Sci. U. S. A* 80, 3696–3700. [PubMed: 6574507]
90. Guex N & Peitsch MC (1997). SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* 18, 2714–2723. [PubMed: 9504803]
91. Brooks BR, Bruccoleri RE, Olafson BD, States DJ, Swaminathan S & Karplus M (1983). CHARMM—A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem* 4, 187–217.
92. Lazaridis T & Karplus M (1999). Effective energy function for proteins in solution. *Proteins: Struct. Funct. Genet* 35, 133–152. [PubMed: 10223287]
93. Tama F, Gadea FX, Marques O & Sanejouand YH (2000). Building-block approach for determining low frequency normal modes of macromolecules. *Proteins: Struct. Funct. Genet* 41, 1–7. [PubMed: 10944387]
94. Li GH & Cui Q (2002). A coarse-grained normal mode approach for macromolecules: An efficient implementation and application to Ca²⁺-ATPase. *Biophys. J* 83, 2457–2474. [PubMed: 12414680]
95. Brooks BR, Janezic D & Karplus M (1995). Harmonic analysis of large systems. 1. Methodology. *J. Comput. Chem* 16, 1522–1542.
96. Lange OF & Grubmüller H (2006). Generalized correlation for biomolecular dynamics. *Proteins: Struct. Funct. Bioinform* 62, 1053–1061.
97. Ichiye T & Karplus M (1991). Collective motions in proteins: a covariance analysis of atomic fluctuations in molecular dynamics and normal mode simulations. *Proteins: Struct. Funct. Genet* 11, 205–217. [PubMed: 1749773]
98. DeLano W (2002). The PyMOL molecular graphics program (<http://www.pymol.org>).
99. Kyte J & Doolittle RF (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol* 157, 105–132. [PubMed: 7108955]
100. Clamp M, Cuff J, Searle SM & Barton GJ (2004). The Jalview Java alignment editor. *Bioinformatics* 20, 426–427. [PubMed: 14960472]
101. Evans SV (1993). SETOR Hardware-lighted 3-dimensional solid model representations of macromolecules. *J. Mol. Graph* 11, 134–138. [PubMed: 8347566]
102. Grant BJ, McCammon JA, Caves LSD & Cross RA (2007). Multivariate analysis of conserved sequence-structure relationships in kinesins: Coupling of the active site and a tubulin-binding sub-domain. *J. Mol. Biol* 368, 1231–1248. [PubMed: 17399740]

Figure 1. Overall structure of *H. sapiens* fascin-1.

(a) View I of *H. sapiens* fascin-1 showing lobe one that consists of β -trefoil domains F1 (residues 8–139, yellow) and F2 (residues 140–260, green) and lobe two that consists of β -trefoil domains F3 (residues 261–381, blue) and F4 (residues 382–493, red). The pseudo-two-fold axis is located between the lobes and is oriented approximately normal to the plane of the image. (b) View of fascin-1 from the top in comparison with View I in (a), obtained by a 90 degree rotation of fascin about the X-axis in View I. The pseudo-two-fold axis is approximately vertical in the plane of the image. (c) View of fascin-1 along the pseudo-three-fold axis of lobe one, obtained by rotation of fascin-1 in View I by 156 degrees about the X-axis, -119 degrees about the Y-axis, and -156 degrees about the Z-axis. (d) View of fascin-1 along the pseudo-three-fold axis of lobe two. This view is obtained by rotation of fascin-1 in View I by -150 degrees about the X-axis, 125 degrees about the Y-axis, and -151 degrees about the Z-axis. All structural figures are generated using PyMol.⁹⁸

Figure 2. Structure and sequence analyses of the β -trefoil fold.

(a) View along the internal pseudo-three-fold axis of β -trefoil domain F2 of *H. sapiens* fascin-1, containing 12 β -strands. The hairpin triplet ($\beta 2$, $\beta 3$, $\beta 6$, $\beta 7$, $\beta 10$, and $\beta 11$) is located proximal and the six-stranded barrel (strands $\beta 1$, $\beta 4$, $\beta 5$, $\beta 8$, $\beta 9$ and $\beta 12$) distal. (b) Structural superposition of domains F1 (yellow), F2 (green), F3 (blue), and F4 (red) of fascin-1. (c) Structural superposition of homologous β -trefoil folds trypsin inhibitor (violet), interleukin-1 β (orange), human acidic fibroblast growth factor (brown), hisactophilin (black), and domain F1 (yellow) of fascin-1. (d) Structure-based sequence alignment of 59 homologous β -trefoil folds:⁵⁰ the four domains of *H. sapiens* fascin-1 (F1–F4), trypsin inhibitor (ETI), interleukin-1 β (IL-1B), human acidic fibroblast growth factor (FGF), and hisactophilin (H). Stars denote residues that stabilize the core of the β -trefoil fold. Green denotes hydrophobic and white denotes hydrophilic residues using the hydrophobicity table of Kyte *et al.*⁹⁹ Secondary structure assignment is based on.⁵⁰ All multiple sequence alignment figures are generated using Jalview.¹⁰⁰

Figure 3. Multiple sequence alignment of homologous fascins.

Multiple sequence alignment of a subset of 61 homologous fascin sequences with NCBI accession codes and %-sequence identities to *H. sapiens* fascin-1 provided in parentheses: *H. sapiens* fascin-1 (NP_003079; 100%), *H. sapiens* fascin-2 (NP_036550; 57%), *H. sapiens* fascin-3 (NP_065102; 28%), *Strongylocentrotus purpuratus* fascin (NP_999701; 37%), and *Drosophila melanogaster* singed (isoform A) (NP_727226; 41%). Green denotes strictly conserved residues, yellow denotes strong conservation, and colors in between interpolate conservation grade linearly. A schematic of the fascin sequence is shown at bottom right. Point mutations in *D. melanogaster* singed (the singed gene) that disrupt actin-bundling are S289N (S274 in *H. sapiens* fascin-1) (blue five-point star) and G409E (G393 in *H. sapiens* fascin-1) (green five-point star). S251F (P236 in *H. sapiens* fascin-1) (solid black circle) restores fascin function in the S289N mutant.⁶⁶ Mutation of S39 (also S39 in *H. sapiens*

fascin-1) (red triangle) to alanine renders *in vitro* actin-binding by fascin-1 insensitive to regulation by phosphorylation by PKC α .^{34,67} See text for details.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 4. Residues suggested to stabilize the β -trefoil cores and lobes of fascin-1.

(a) Conservation grade and schematic of hydrophobic core residues (spheres) measured across homologous fascin molecules. Views are along the three-fold axis of the β -trefoil domain, where “top” and “bottom” refer to viewing the domain from its cap and barrel, respectively. Conservation grades vary from blue to red, denoting highly variable and conserved residue positions, respectively. (b) Ionic and hydrophobic interactions between β -trefoil domains of lobe one. (c) Ionic and hydrophobic interactions between the β -trefoil domains of lobe two. Hydrophobic residues that lose more than $40 \text{ \AA}^2$ of their solvent accessible surface upon domain-domain association and residues participating in domain-domain salt bridges are labeled in (b) and (c).

Figure 5. Conservation grade and solvent accessible surface burial of surface residues of the lobes of fascin-1.

Interior and exterior views of the lobes of fascin-1 show residue conservation and solvent accessible surface burial due to lobe-lobe association. White and orange denote low (<1%) and high (>75%) solvent accessible surface burial, respectively. Conservation grades vary from blue to red, denoting highly variable and conserved residue positions, respectively. Interior and exterior lobe views are related by 180 degree rotations about the horizontal axis, where each lobe is aligned horizontally with the axis of its lowest principal moment of inertia. Point mutations affecting fascin function and residues H139, Q141, S259, R383, and R389 are indicated. Residue numbers in parentheses indicate the analogous positions in *D. melanogaster* fascin. (Inset) Close-up view of highly conserved interfacial residues H139, Q141, S259, R383, and R389 in stick representation.

Figure 6. Residue conservation near putative actin-binding sites of fascin-1.

(a) Solvent-accessible surface residue conservation grades in fascin-1. Conservation grades vary from blue to red, denoting highly variable and conserved residue positions, respectively. Point mutations affecting fascin function and residues 29 and 43 are indicated. Residue numbers in parentheses indicate the analogous positions in *D. melanogaster* fascin. (b) Residues near residues 29–43 (putative actin-binding site), P236, and S274 that are highly conserved across homologous fascins but not across homologous β -trefoils. Residues 23, 25, 27, and 44 that are near residues 29–43, residues 149, 151, 154–155, 166, 174, 235, and 251–252 that are near P236, and residues 274, 283–285, 292, and 360–361 that are near S274 are represented as spheres. Remaining residues are shown in cartoon representation. D166 and N284 are considered variable because the corresponding columns in the structure-based multiple sequence alignment of β -trefoil domains consist mostly of gaps (90%).

Figure 7. Dynamically correlated domains of fascin-1.

(a) Average-link hierarchical clustering is used to identify residue clusters that are highly correlated based on the magnitude of their generalized linear mutual information coefficient. Upper and lower triangles show inter-residue correlations based on sequence position and after clustering, respectively. High dynamical correlations are found within individual β -trefoil domains, as expected, but also between distant domains F1 and F3 (also see Table 1). (b) View I of fascin-1 colored according to the five clusters shown in (a). Point mutations affecting fascin function and residues 29 and 43 are indicated. Residue numbers in parentheses indicate analogous positions in *D. melanogaster* fascin.

Figure 8. Side-by-side stereo view of fascin-1.

Side-by-side stereo view of the experimental electron density map for *H. sapiens* fascin-1 displayed at 1σ . The 2.9 Å map is based on phases from five derivatives, density modification and NCS averaging. The Cα trace is shown for the β-hairpin triplet in domain 4 of chain B. Figure generated using SETOR.¹⁰¹

Table 1

Average generalized linear mutual information coefficient and fraction of residues that are in contact (in parentheses) for the five clusters in fascin-1 shown in Figure 7. Clusters (1) red-cyan; (2) cyan-green; and (3) green-blue have mean generalized linear correlation coefficients in the 73rd, 46th, and 13th percentiles, respectively, excluding intra-domain interactions and interactions between residues in direct contact (i.e., residues that have at least two heavy atoms within 5Å of one another).¹⁰²

	yellow	red	cyan	green	Blue
yellow	-	0.26 (11%)	0.29 (11%)	0.26 (31%)	0.21 (16%)
red	-	-	0.31 (0%)	0.26 (23%)	0.27 (9%)
cyan	-	-	-	0.25 (0%)	0.27 (13%)
green	-	-	-	-	0.17 (0%)
blue	-	-	-	-	-

Table 2

Fascin data collection, phasing, and refinement statistics.

Data Collection	NATIVE*	DER 1**	DER 2**	DER3**	DER 4**	DER5**
Resolution Range (Å)	2.9	2.9	3.0	3.2	3.2	3.0
Unique Reflections	21421	20944	19737	15882	15763	18422
Completeness (%)	94.2	92.1	96.1	93.5	92.8	89.7
Rmerge (%) [#]	4.8	5.8	5.3	5.7	6.9	6.2
Riso (%) ^{##}		12.0	19.9	14.6	11.4	15.6
Phasing						
RCullis (%) ^{###}		57.2	60.3	61.6	68.6	63.8
Phasing Power ^{####}		1.62	1.31	1.32	1.27	1.79
No. of heavy atom sites		4	2	5	6	10
Refinement						
Resolution Range (Å)	25–2.9					
No. of unique reflections	20541					
Completeness (%)	92.9					
Rcryst (%)	22.3					
Rfree (%)	26.8					
No. of protein atoms	7431					
RMSD bonds (Å)	0.009					
RMSD angles (°)	1.4					

* Native crystals were grown from 20% PEG 4000, 0.1 M HEPES pH 7.5, 1mM DTT. Space group C2, unit cell $a = 165.4$, $b = 71.7$, $c = 116.9$, $\beta = 132.2^\circ$, two molecules in the asymmetric unit.

** Derivatives 1, 2, and 3 were produced by soaking native crystals in 1mM solutions of Kau(CN)₂, KAg(CN)₂ and PT(Cl)₂-tetrapiridine. Derivative 4 is the selenomethionyl fascin and derivative 5 is the selenomethionyl fascin crystal soaked in 1mM Kau(CN)₂.

[#]R_{merge} = $\sum |I_{\text{obs}} - \langle I \rangle| / \sum \langle I \rangle$ summed over all observations of all reflections

^{##}R_{iso} = $\sum |F_{\text{ph}} - F_{\text{p}}| / \sum F_{\text{p}}$ summed over all unique reflections

^{###}R_{Cullis} = $\sum ||F_{\text{ph,obs}} - F_{\text{p,obs}}| - F_{\text{h,calc}}| / \sum ||F_{\text{ph,obs}} - F_{\text{p,obs}}||$

^{####}Phasing power = $\sum |F_{\text{h}}| / \sum ||F_{\text{ph,obs}}| - |F_{\text{ph,calc}}||$