Structure, Vol. 11, 197–203, February, 2003, ©2003 Elsevier Science Ltd. All rights reserved. PII S0969-2126(03)00002-9

Novel Fold Revealed by the Structure of a FAS1 Domain Pair from the Insect Cell Adhesion Molecule Fasciclin I

Naomi J. Clout, Dominic Tisi,[1] and Erhard Hohenester*
Biophysics Section
Department of Biological Sciences
Imperial College
London SW7 2AZ
United Kingdom

Summary

Fasciclin I is an insect neural cell adhesion molecule consisting of four FAS1 domains, homologs of which are present in many bacterial, plant, and animal proteins. The crystal structure of FAS1 domains 3 and 4 of *Drosophila* fasciclin I reveals a novel domain fold, consisting of a seven-stranded β wedge and a number of α helices. The two domains are arranged in a linear fashion and interact through a substantial polar interface. Missense mutations in the FAS1 domains of the human protein βig-h3 cause corneal dystrophies. Many mutations alter highly conserved core residues, but the two most common mutations, affecting Arg-124 and Arg-555, map to exposed α-helical regions, suggesting reduced protein solubility as the disease mechanism.

Introduction

The fasciclins are a family of insect cell adhesion molecules (CAMs) expressed on different and overlapping subsets of fasciculating axons and growth cones [1, 2]. Fasciclins II and III contain immunoglobulin-like (IG) and fibronectin type III (FN3) domains linked to transmembrane helices, reminiscent of other CAMs. Fasciclin I, in contrast, is a GPI-anchored protein [3] containing a tandem of four homologous domains of ≈150 residues, termed FAS1 domains, which are not related to any other protein domain of known structure [4]. Mutations in the *Drosophila fasI* gene result in a mild phenotype characterized by altered nerve terminal arborization [5], whereas embryos doubly mutant in *fasI* and *abl*, the fly ortholog of the Abelson tyrosine kinase, have major axon guidance defects [6]. Although the extracellular portion of grasshopper fasciclin I is monomeric in solution [7], *Drosophila* fasciclin I is capable of mediating homophilic cell adhesion [8].

FAS1 domains are present in many other secreted and membrane-anchored proteins (for a complete list, see the SMART [smart.embl-heidelberg.de] and Pfam [www.sanger.ac.uk/Software/Pfam] databases). Most of these proteins contain multiple FAS1 domains, either in tandem or interspersed with other domains, but proteins consisting of single FAS1 domains also exist. Despite their low overall sequence conservation (typically <20% identity for any pairwise alignment), FAS1 domains are recognized readily due to the presence of two conserved sequence motifs. Unusual for extracellular domains, the FAS1 superfamily includes members from all phyla: mycobacterial secreted proteins [9]; plant proteins, such as the major CAM of *Volvox*, Algal-CAM [10], and *Arabidopsis* arabinogalactans [11]; several *C. elegans* proteins of unknown function; and a number of vertebrate proteins. In humans, FAS1 domains are found in a family of very large hyaluronan and scavenger receptors [12, 13], as well as in the adhesive proteins βig-h3/RGD-CAP/kerato-epithelin [14, 15] and periostin/OSF-2 [16]. The latter two proteins most closely resemble fasciclin I, consisting of uninterrupted tandems of four FAS1 domains. Importantly, mutations in the FAS1 domains of βig-h3 result in corneal dystrophies, due to the deposition of insoluble protein aggregates [17, 18].

Although the biological functions of many FAS1 domains remain to be elucidated, it has been suggested that the FAS1 domain represents an evolutionarily ancient cell adhesion domain common to plants and eukaryotes [10]. To provide a template for the FAS1 superfamily, we have undertaken a structural analysis of the founding member of the family, fasciclin I. Our structure reveals a novel domain fold, which is distinguished from other CAM domains by the presence of several α helices, and gives new insight into the disease mechanism of βig-h3 mutations in corneal dystrophies.

Results and Discussion

Structure Determination

Previous attempts to solve the crystal structure of full-length grasshopper fasciclin I, comprising all four FAS1 domains but lacking the GPI anchor, were unsuccessful [7]. To increase our chances of obtaining well-ordered crystals, we expressed *Drosophila* fasciclin I fragments corresponding to FAS1 domains 1-2 and 3-4, respectively, in a eukaryotic system (293-EBNA cells). Two constructs of the 1-2 pair were expressed only at low levels. A first construct of the 3-4 pair spanning residues 328–628 was highly expressed, but cleavage of the His_6 tag used for purification proved inefficient. A slightly longer construct starting at residue 314 was highly expressed and the tag was readily cleaved by enterokinase. This glycosylated fasciclin I fragment, termed FasI 3-4, could be crystallized. The FasI 3-4 structure was determined by multiple isomorphous replacement (MIR) and refined to $R_{free} = 0.263$ at 2.6 Å resolution (Table 1).

Description of the Structure

The FasI 3-4 structure consists of two globular FAS1 domains of approximately 30 Å in diameter interacting through a substantial domain interface (Figure 1). Two well-ordered N-acetylglucosamine moieties attached to Asn-441 are located in the interface region. The N and

*Correspondence: e.hohenester@ic.ac.uk
[1]Present address: Astex Technology Ltd, Cambridge CB4 0WE, United Kingdom.

Key words: cell adhesion; axon guidance; extracellular module; genetic disorder; corneal dystrophy; X-ray crystallography

Table 1. Crystallographic Statistics

Data collection	Native	$KAu(CN)_2$	K_2PtCl_4
Soaking conditions		0.5 mM, 2 hr	0.15 mM, 2 hr
Resolution range (Å)	20–2.6	20–2.8	20–3.0
Unique reflections	17,949	15,047	12,190
Average multiplicity	8.4 (8.3)[a]	9.6	12.8
Completeness (%)	97.1 (97.5)[a]	99.7	99.0
R_{merge}[b]	0.058 (0.309)[a]	0.087	0.109
Phasing (20–3.0 Å)			
Number of sites		1	1
R_{deriv}[c]		0.079	0.268
R_{Cullis}[d] (centric/acentric)		0.83/0.72	0.90/0.89
Phasing power[e] (centric/acentric)		1.04/0.86	0.90/0.60
Refinement (20–2.6 Å)	**Native**		
Reflections (working set/test set)	16,153/1,758		
Protein atoms	2,380		
Asn-linked glycan	2 N-acetylglucosamines		
Solvent sites	76 H_2O, 4 SO_4^{2-}		
R_{cryst}[f]	0.225		
R_{free}[f]	0.263		
Rmsd bond lengths (Å)	0.008		
Rmsd bond angles (°)	1.2		
Ramachandran plot (%)[g]	86.2/12.6/0.7/0.4		

[a] Values in parentheses are for reflections in the highest resolution shell (2.74–2.60 Å).
[b] $R_{merge} = \Sigma_h \Sigma_i |I_i(h) - \langle I(h) \rangle| / \Sigma_h \Sigma_i I_i(h)$, where $I_i(h)$ is the i-th measurement of reflection h and $\langle I(h) \rangle$ is the weighted mean of all measurements of h.
[c] $R_{deriv} = \Sigma_h ||F_{PH}| - |F_P|| / \Sigma_h |F_P|$, where F_P and F_{PH} are the native and derivative structure factors, respectively.
[d] $R_{Cullis} = \Sigma_h |||F_{PH}| - |F_P|| - |F_H|| / \Sigma_h ||F_{PH}| - |F_P||$, where F_H is the calculated heavy atom structure factor.
[e] The phasing power is defined as rms F_H/rms lack of closure.
[f] $R = \Sigma_h |F_{obs} - F_{calc}| / \Sigma_h F_{obs}$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively.
R_{cryst} and R_{free} were calculated using the working and test set, respectively.
[g] Residues in most favored, additionally allowed, generously allowed, and disallowed regions [28].

C termini are located at opposite poles of the elongated FasI 3-4 structure, and the two FAS1 domains are related by a rotation of 124° and a translation of 26 Å along an axis roughly parallel to the long molecular dimension.

The FAS1 domain fold is organized around a wedge-shaped arrangement of two orthogonal β sheets, which share a curved strand, β6 (Figure 1). The mixed β2-β1-β7-β6 sheet is central to the domain: one face is covered by the N-terminal helices α1 and α2, and the other packs against the antiparallel β3-β4-β5-β6 sheet. The outer face of the latter β sheet is fully exposed to solvent. Strands β1 and β2 are connected by a series of α helices, which pack between the V-shaped arrangement of N-terminal helices and the central β sheet. Helix α5, together with the extended segment preceding it, closes the open end of the β wedge. The C-terminal segment following β7 runs alongside the curved β6 strand in an irregular manner, with the kink in β6 defining the acute end of the β wedge. The FAS1 domain contains two extensive hydrophobic cores, one within the β wedge and the other between the β2-β1-β7-β6 sheet and helices α1 to α4. The FAS1 domain fold has no similarities to any other domain of known structure. The best match in DALI [19] had a Z-score of 2.0 and matched only one third of the FAS1 domain.

The two FAS1 domains in FasI 3-4 are rather similar and can be superimposed with an rms deviation of 1.7 Å for 105 Cα atoms (Figure 2). The only striking difference is the presence of an additional helix, αL, in FAS1 domain 4, which, together with α1 and α2, creates a triangular arrangement of helices packing against the β2-β1-β7-β6 sheet. The αL helix is an integral part of the linker/interface between the FAS1 domains in FasI 3-4. Domain 3 contains residues equivalent to αL, but they are disordered in the crystal, presumably because stabilizing interactions with the preceding FAS1 domain are missing.

The interface between FAS1 domains 3 and 4 buries a total of 1700 Å^2 of solvent-accessible surface area, with the two N-acetylglucosamine moieties at Asn-441 accounting for 140 Å^2 of this area. Notably, there are five direct hydrogen bonds and several close van der Waals contacts between the glycan and the protein. The domain interface is relatively flat and almost exclusively hydrophilic. Apart from two direct hydrogen bonds, all contacts occur via water molecules in the interface, of which there are many. In domain 3, the interface region involves the loop between α5 and β3, the entire β4 strand, and the β5-β6 hairpin. In domain 4, the αL helix and the following loop, as well as β6, contribute to the interface. The domain linker (Pro-466/Tyr-467/Thr-468) packs tightly against Gly-439 in the β5-β6 hairpin of domain 3, with Tyr-467 additionally stacking with His-461.

Neither the crystal packing nor an inspection of the FasI 3-4 surface offer any clues about the mechanism of homophilic adhesion mediated by fasciclin I [8]. Adhesion may require the membrane-distal domains 1 and 2, or may be due to a multiplicity of weak interactions

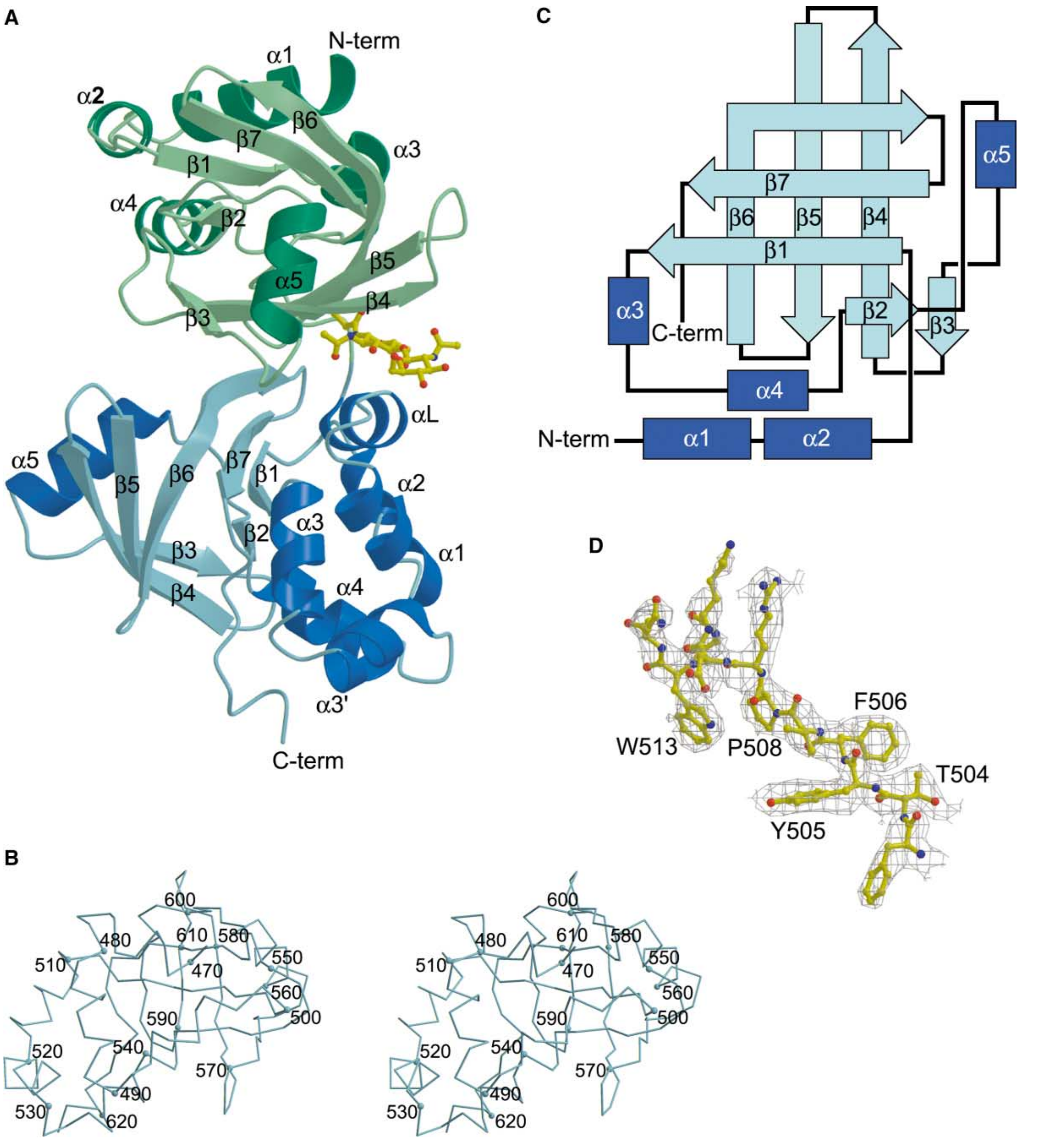

Figure 1. Structure of FasI 3-4

(A) Cartoon representation of the FasI 3-4 structure. Domain 3 is in green and domain 4 is in blue. The polypeptide chain termini and secondary structure elements are labeled. Two N-acetylglucosamine moieties attached to Asn-441 are shown in atomic detail.

(B) Stereo view of a Cα trace of FAS1 domain 4 of FasI 3-4. Every tenth residue is shown as a small sphere and is labeled, starting at residue 470.

(C) Schematic representation of the FAS1 fold, in the orientation of (B).

(D) Representative portion of the experimental electron density at 3.0 Å resolution after density modification (1.2 σ contouring). The final atomic model is superimposed on the map. The region shown corresponds to strand β1 and helix α3 in domain 4.

on the cell surface, as indicated by the observation that soluble full-length fasciclin I is monomeric in solution [7].

There exist a number of developmentally regulated splice variants of fasciclin I, which differ by insertions of a few residues near the boundary of FAS1 domains 2 and 3 [20]. The site of insertion is located in the N-terminal region of the FasI 3-4 construct, which is disordered in the crystal. From sequence analysis, we predict the fasciclin I splice variants to differ in the αL-α1 linker of FAS1 domain 3 (Figure 3). Insertion of extra residues in this region will, at the very least, affect the structure of the surface-exposed αL-α1 loop. However, because

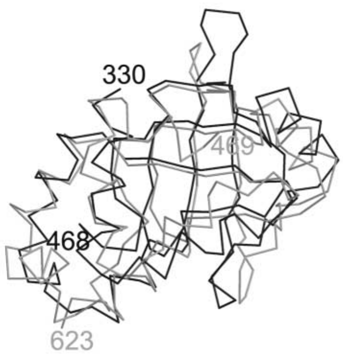

Figure 2. Comparison of FAS1 Domains 3 and 4

Stereo view of a superposition of FAS1 domains 3 (dark gray) and 4 (light gray) of FasI 3-4. A total of 105 Cα atoms were superimposed with an rms deviation of 1.7 Å.

the αL helix is likely to be involved in the interface with domain 2, the insertions could have a more dramatic effect, changing the overall conformation of fasciclin I. In either scenario, it is easy to see how alternative splicing could modulate the adhesive properties of fasciclin I.

The FAS1 Domain Superfamily

We have determined the first structure of a FAS1 domain, and it is of interest to consider the relevance of our structure for the entire FAS1 domain superfamily. Previous sequence analysis revealed two regions of high conservation, termed H1 and H2 [21]. Our structure shows that H1 corresponds to β1 and α3, whereas H2 corresponds to the C-terminal half of β6, as well as β7 and part of the C-terminal segment (Figure 3). The conservation of H1 is almost absolute, whereas H2 is poorly conserved in a small number of superfamily members, such as the second FAS1 domain of Algal-CAM. A third region that is particularly well conserved is the short strand β2. The most variable region is the protruding and partly solvent-exposed α3-α4 linker (domain 4 of FasI 3-4 contains an extra helix, α3′, in this region). As expected, most other loops connecting secondary structure elements are also quite variable, with the notable exception of the β6-β7 hairpin.

In addition to the expected conservation of apolar core residues, the FAS1 domain consensus contains a large number of polar residues, the location of which is revealing. The Oγ atom of the invariant threonine at the start of β1 forms hydrogen bonds to the backbone of the extended segment preceding α5. The conserved aspartic acid/asparagine in α3 interacts with the back-

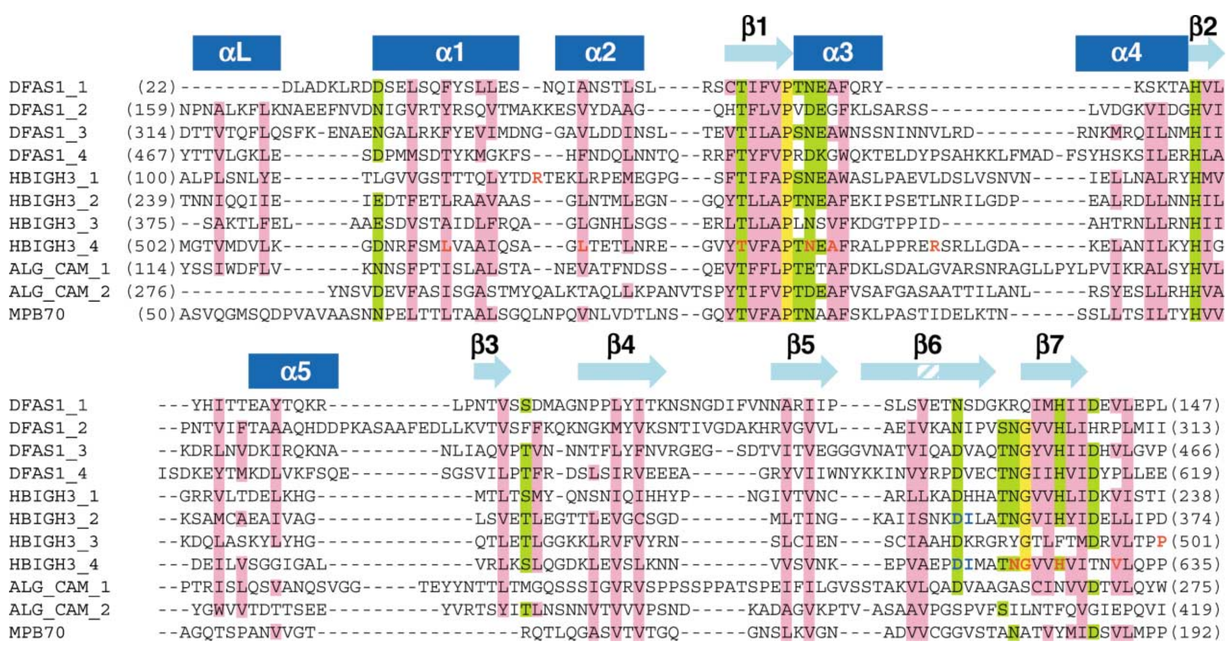

Figure 3. Structure-Based Sequence Alignment of Selected FAS1 Domains

The sequences are of *Drosophila melanogaster* fasciclin I (SwissProt P10674; splice variant II [20]), human βig-h3 (Q15582), *Volvox carteri* Algal-CAM (Q41644), and *Mycobacterium tuberculosis* MPB70 (Q50769). The sequence numbering includes the signal peptide. Structurally important FAS1 residues were assigned based on a superposition of the two FAS1 domains of FasI 3-4 (see Figure 2); other sequences were aligned accordingly. The secondary structure elements of the FAS1 domain are indicated above the alignment. Conserved residues are shaded as follows: pink, apolar; green, polar; yellow, glycine and proline. βig-h3 residues mutated in corneal dystrophies [17, 18] are in red, and residues implicated in integrin binding to βig-h3 [15] are in blue.

bone of the N terminus of α1. The buried conserved histidine at the start of β2 makes hydrogen bonds to the backbone of β1 and α4, and also packs against the conserved proline at the end of β1, which in turn interacts with a conserved aromatic residue in α3. The extended segment between β2 and α5 makes a conserved interaction, with an aspartic acid/asparagine in β6. Finally, the buried conserved histidine in β7 hydrogen bonds to the extended segment preceding α5, and the conserved aspartic acid at the end of β7 initiates the irregular C-terminal segment. Thus, conserved hydrogen-bonding interactions are used to reinforce the FAS1 fold at key positions, in particular in regions lacking in regular secondary structure, such as the β2-α5 segment plugging the open end of the β wedge. Another notable feature of the FAS1 domain is the absence of stabilizing disulfide bridges.

Most FAS1 domains occur in pairs or as a four-domain tandems. Whether the domain arrangement seen in the FasI 3-4 structure is conserved in other proteins is unclear. Some features of the linker/interface region, such as the αL helix and a proline residue equivalent to Pro-466, are present in other family members. On the other hand, the paucity of direct interactions in the FasI 3-4 interface, as well as the general lack of conservation of interface residues, makes it likely that other arrangements exist. For this reason, it is impossible to confidently predict the overall structure of proteins containing multiple FAS1 domains. If the relative domain arrangement seen in the FasI 3-4 structure is representative of other pairs, the four FAS1 domains of fasciclin I [4], βig-h3 [14, 15], and periostin [16] would be arranged in a near-linear fashion, with a long dimension of approximately 120 Å.

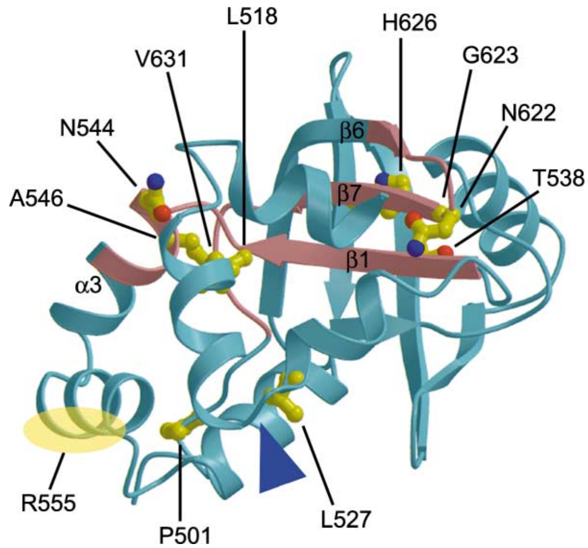

Figure 4. Location of Missense Mutations in βig-h3 Causing Corneal Dystrophies in Humans

Shown is a cartoon representation of domain 4 of fasciclin I, in which selected residues [17, 18], drawn in atomic detail, were changed to their equivalents in FAS1 domain 4 of βig-h3 for illustration purposes. The local structure around Arg-555 in βig-h3 cannot be modeled reliably, and the general location of this residue is indicated by yellow shading. The blue arrowhead indicates the location of Arg-124 in FAS1 domain 1 of βig-h3. The conserved sequence motifs, H1 and H2 [21], are colored brown.

Implications for βig-h3 Function and Disease

Of the human proteins containing FAS1 domains, βig-h3 has received the most attention, owing to causative mutations in the *BIGH3* gene in a group of corneal dystrophies [17, 18]. The precise physiological function of βig-h3 is still unclear, but is likely to involve cell adhesion via integrins α1β1 [14] or α3β1 [15]. Mutagenesis has implicated an Asp-Ile sequence motif in βig-h3 FAS1 domains 2 and 4 in α3β1 integrin binding [15]. This motif maps to the conserved kink in strand β6 (Figure 3). The 2 residues are not buried in the FAS1 domain core and are, in principle, available for receptor binding. However, in the relative domain arrangement of the FasI 3-4 structure, the motif in domain 4 is buried in the domain interface. Whether the Asp-Ile motifs in βig-h3 are directly involved in integrin binding, or whether their mutation abrogated binding because of long-range structural perturbations, remains to be seen. The high conservation of the motif across functionally unrelated molecules (Figure 3) and the apparently crucial hydrogen bond formed by the aspartic acid would seem to argue for the latter alternative.

The corneal dystrophies due to *BIGH3* mutations are believed to result from the accumulation of pathological βig-h3 amyloid deposits, suggesting that the mutations affect the stability and/or solubility of the βig-h3 protein. Altered proteolytic processing has been observed in some forms of the disease [22]. More than half of the *BIGH3* mutations affect one of two arginine residues, Arg-124 in FAS1 domain 1, and Arg-555 in FAS1 domain 4 [18]. Arg-124 is located in the turn between α1 and α2, whereas Arg-555 is located in the α3-α4 linker, which is the most variable region of the FAS1 fold (Figures 3 and 4). Although our structure does not allow the reliable modeling of Arg-555, it seems unlikely that mutation of either arginine, in these solvent-exposed regions, would result in misfolding in the endoplasmic reticulum. Presumably, mutant βig-h3 is secreted, but aggregates with time in the extracellular space because of a subtle decrease in protein stability or solubility.

The remaining *BIGH3* missense mutations [18] present a very different picture. They target conserved core residues in FAS1 domain 4 of βig-h3 (Figures 3 and 4). Pro-501 is predicted to be in the linker between FAS1 domains 3 and 4 (the equivalent residue in the FasI 3-4 pair is also a proline). Leu-518 and Leu-527 are important for the packing of helices α1 and α2, respectively, onto the body of the FAS1 domain. The remaining seven mutations intriguingly cluster in the two regions of highest sequence conservation: Thr-538, Asn-544, and Ala-546 map to the highly conserved β1-α3 region, whereas Asn-622, Gly-623, His-626, and Val-631 map to β7. It appears likely that some of these rarer *BIGH3* mutations cause misfolding inside the cell. However, a possible folding/secretion defect would be difficult to reconcile with the apparent extracellular deposition of mutant βig-h3 protein. Perhaps, the FAS1 domain is more resilient to individual mutations than the high conservation of certain regions might suggest. Knowledge of the FAS1 domain structure will be valuable in defining the physiological activities of βig-h3, as well as the disease mechanism(s) of corneal dystrophies.

Biological Implications

Most cell adhesion molecules (CAMs) and secreted extracellular matrix proteins have a modular architecture: they are composed of a limited set of domains, which frequently form tandem repeats. The great majority of CAMs consist of immunoglobulin-like and fibronectin type III domains, which share a characteristic β sandwich fold. Fasciclin I is a *Drosophila* neural CAM consisting of a repeat of four unique domains, termed FAS1 domains. Related FAS1 domains are found not only in other animal proteins, but also in bacteria and plants, leading to the suggestion that the FAS1 domain may represent (one of) the most ancient adhesive extracellular domain(s). Of note, missense mutations in the FAS1 domains of the human adhesive protein βig-h3 cause corneal dystrophies by an ill-defined mechanism involving amyloid deposition.

To provide a structural template for the FAS1 superfamily, we have determined the crystal structure of a FAS1 domain pair from fasciclin I. The structure consists of a near-linear arrangement of two compact FAS1 domains, each composed of a seven-stranded β wedge and a number of α helices in a novel arrangement. The discovery of an αβ fold in a CAM is unprecedented.

Knowledge of the FAS1 domain structure has allowed us to generate a reliable sequence alignment of the divergent FAS1 superfamily and provides new insight into the functions of FAS1-containing proteins. In particular, we have analyzed the location of missense mutations in human βig-h3 causing corneal dystrophies. The two most common mutations affect arginine residues in exposed α-helical regions of the FAS1 domain, suggesting reduced βig-h3 stability or solubility as the cause of amyloid deposition. Many of the rarer βig-h3 mutations, however, affect structurally important FAS1 core residues and are predicted to cause disease by compromising βig-h3 folding and secretion.

Experimental Procedures

Protein Expression and Purification

A full-length cDNA encoding *Drosophila* fasciclin I was used as a template to amplify FAS1 domains 3 and 4 (FasI 3-4; residues 314–628 in the numbering scheme used here, which includes the signal peptide and refers to splice variant II [20]) with Pfu polymerase following the manufacturer's instructions (New England Biolabs). The primers introduced NheI and XhoI sites at the 5′ and 3′ ends, respectively, which were used to ligate the PCR product into a modified pCEP-Pu vector [23]. This episomal vector codes for a fusion protein containing the BM-40 secretion signal, a His_6 tag, a myc antigen, and an enterokinase cleavage site, followed by the FasI 3-4 fragment. After enterokinase cleavage, a single vector-derived leucine residue remains at the N terminus of FasI 3-4. The vector was used to transfect human 293-EBNA cells (maintained in Dulbecco's modified Eagle medium [DMEM F12, Invitrogen] containing 10% fetal bovine serum) using Fugene reagent (Roche) according to the manufacturer's instructions. Cells containing the episomal vector were selected with 1 μg/ml puromycin. The recombinant protein was purified from serum-free culture medium using TALON metal affinity resin (Clontech), yielding ≈10 mg of protein from 1 L of conditioned medium. The tag was cleaved with enterokinase (EKMax, Invitrogen) and the enzyme removed with EKAway affinity resin (Invitrogen). Final purification was achieved on a ResourceQ ion exchange chromatography column (50 mM Na-HEPES [pH 7.5], 0–0.5 M NaCl). This yielded FasI 3-4 protein, which migrated on SDS-PAGE as a single band of approximately 50 kDa M_r (calculated M_r 37 kDa). Hence, the FasI 3-4 construct is modified at one or several of the four consensus sites for Asn-linked glycosylation.

Crystallization and Data Collection

For crystallization trials by the hanging drop vapor diffusion method, FasI 3-4 was concentrated to 15 mg/ml in 0.01 M Na-HEPES (pH 7.5). Crystals were obtained using 0.1 M Tris-HCl (pH 8.5), 2.0 M ammonium sulfate as precipitant. The crystals took about 1 month to grow to full size (0.2 × 0.2 × 0.2 mm^3). For diffraction data collection, crystals were cryoprotected with 20% glycerol and flash-frozen to 100 K. Data were collected at station 9.6 at the SRS Daresbury (λ = 0.87 Å) using an ADSC Quantum-4 CCD detector. The data were integrated with MOSFLM [24] and reduced with programs of the CCP4 suite [25]. The crystals belong to space group $P4_332$ with a = 150.81 Å. There is one FasI 3-4 molecule in the asymmetric unit, resulting in a solvent content of 68% (not taking the glycan into account).

Structure Determination and Refinement

A native data set to 2.6 Å resolution and two heavy atom derivative data sets were used for phasing by the MIR method (Table 1). The heavy atom sites were found by standard Patterson and Fourier techniques and refined with MLPHARE [25]. Phases to 3.0 Å resolution were calculated including anomalous scattering (mean figure of merit 0.354) and dramatically improved by density modification with DM [25]. About 90% of the FasI 3-4 molecule could be built into the first electron density map with O [26]. Refinement was carried out with CNS [27] at 2.6 Å resolution and converged at R = 0.225 and R_{free} = 0.263. Continuous and clear density was observed for residues 328–624, excluding the exposed 369–375 loop. There is no density for the first 16 and last 5 residues of FasI 3-4, which are presumed to be disordered. Clear density was observed for the first two N-acetylglucosamine moieties attached to Asn-441. Patchy extra density at Asn-416 and Asn-497 suggests further modifications, but these were not modeled. Solvent-accessible surface areas were calculated with NACCESS (wolf.bms.umist.ac.uk/naccess/). Figures 1A, 1B, 1D, 2, and 4 were made with BOBSCRIPT [29] and RASTER3D [30].

Acknowledgments

We thank Dr. Kai Zinn (California Institute of Technology) for fasciclin I cDNA samples, Dr. Patrik Maurer (University of Cologne, Germany) for the pCEP-Pu expression vector, the staff at SRS Daresbury beamline 9.6 for help with data collection, and Dr. Peter Brick for critical reading of the manuscript. This work was supported by a Wellcome Trust Senior Research Fellowship in Basic Biomedical Science to E.H.

Received: September 5, 2002
Revised: October 30, 2002
Accepted: October 31, 2002

References

1. Bastiani, M.J., Harrelson, A.L., Snow, P.M., and Goodman, C.S. (1987). Expression of fasciclin I and II glycoproteins on subsets of axon pathways during neuronal development in the grasshopper. Cell *48*, 745–755.
2. Patel, N.H., Snow, P.M., and Goodman, C.S. (1987). Characterization and cloning of fasciclin III: a glycoprotein expressed on a subset of neurons and axon pathways in *Drosophila*. Cell *48*, 975–988.
3. Hortsch, M., and Goodman, C.S. (1990). *Drosophila* fasciclin I, a neural cell adhesion molecule, has a phosphatidylinositol lipid membrane anchor that is developmentally regulated. J. Biol. Chem. *265*, 15104–15109.
4. Zinn, K., McAllister, L., and Goodman, C.S. (1988). Sequence

analysis and neuronal expression of fasciclin I in the grasshopper and *Drosophila*. Cell *53*, 577–587.
5. Zhong, Y., and Shanley, J. (1995). Altered nerve terminal arborization and synaptic transmission in *Drosophila* mutants of cell adhesion molecule fasciclin I. J. Neurosci. *15*, 6679–6687.
6. Elkins, T., Zinn, K., McAllister, L., Hoffmann, F.M., and Goodman, C.S. (1990). Genetic analysis of a *Drosophila* neural cell adhesion molecule: interactions of fasciclin I and Abelson tyrosine kinase mutations. Cell *60*, 565–575.
7. Wang, W.-C., Zinn, K., and Bjorkman, P.J. (1993). Expression and structural studies of fasciclin I, an insect cell adhesion molecule. J. Biol. Chem. *268*, 1448–1455.
8. Elkins, T., Hortsch, M., Bieber, A.J., Snow, P.M., and Goodman, C.S. (1990). *Drosophila* fasciclin I is a novel homophilic adhesion molecule that along with fasciclin III can mediate cell sorting. J. Cell Biol. *110*, 1825–1832.
9. Ulstrup, J.C., Jeansson, S., Wiker, H.G., and Harboe, M. (1995). Relationship of secretion pattern and MPB70 homology with osteoblast-specific factor 2 to osteitis following *Mycobacterium bovis* BCG vaccination. Infect. Immun. *63*, 672–675.
10. Huber, O., and Sumper, M. (1994). Algal-CAMs: isoforms of a cell adhesion molecule in embryos of the alga *Volvox* with homology to *Drosophila* fasciclin I. EMBO J. *13*, 4212–4222.
11. Schultz, C.J., Johnson, K.L., Currie, G., and Bacic, A. (2000). The classical arabinogalactan protein gene family of *Arabidopsis*. Plant Cell *12*, 1751–1767.
12. Pollitz, O., Gratchev, A., McCourt, P.A.G., Schledzewski, K., Guillot, P., Johansson, S., Svineng, G., Franke, P., Kannicht, C., Kzhyshkowska, J., et al. (2002). Stabilin-1 and -2 constitute a novel family of fasciclin-like hyaluronan receptor homologues. Biochem. J. *362*, 155–164.
13. Adachi, H., and Tsujimoto, M. (2002). FEEL-1, a novel scavenger receptor with in vitro bacteria-binding and angiogenesis-modulating activities. J. Biol. Chem. *277*, 34264–34270.
14. Ohno, S., Noshiro, M., Makihira, S., Kawamoto, T., Shen, M., Tan, W., Kawashima-Ohya, Y., Fujimoto, K., Tanne, K., and Kato, Y. (1999). RGD-CAP (βig-h3) enhances the spreading of chondrocytes and fibroblasts via integrin α1β1. Biochim. Biophys. Acta *1451*, 196–205.
15. Kim, J.-E., Kim, S.-J., Lee, B.-H., Park, R.-W., Kim, K.-S., and Kim, I.-S. (2000). Identification of motifs for cell adhesion within the repeated domains of transforming growth factor-β-induced gene, βig-h3. J. Biol. Chem. *275*, 30907–30915.
16. Takeshita, S., Kikuno, R., Tezuka, K., and Amann, E. (1993). Osteoblast-specific factor-2: cloning of a putative bone adhesion protein with homology to the insect protein fasciclin I. Biochem. J. *294*, 217–278.
17. Munier, F.L., Korvatska, E., Djemai, A., Le Paslier, D., Zografos, L., Pescia, G., and Schorderet, D.F. (1997). Kerato-epithelin mutations in four 5q31-linked corneal dystrophies. Nat. Genet. *15*, 247–251.
18. Munier, F.L., Frueh, B.E., Othenin–Girard, P., Uffer, S., Cousin, P., Wang, M.X., Heon, E., Black, G.C.M., Blasi, M.A., Balestrazzi, E., et al. (2002). *BIGH3* mutation spectrum in corneal dystrophies. Invest. Ophthalmol. Vis. Sci. *43*, 949–954.
19. Holm, L., and Sander, C. (1997). DALI/FSSP classification of three-dimensional protein folds. Nucleic Acids Res. *25*, 231–234.
20. McAllister, L., Rehm, E.J., Goodman, C.S., and Zinn, K. (1992). Alternative splicing of micro-exons creates multiple forms of the insect cell adhesion molecule fasciclin I. J. Neurosci. *12*, 895–905.
21. Kawamoto, T., Noshiro, M., Shen, M., Nakamasu, K., Hashimoto, K., Kawashima-Ohya, Y., Gotoh, O., and Kato, Y. (1998). Structural and phylogenetic analyses of RGD-CAP/βig-h3, a fasciclin-like adhesion protein expressed in chick chondrocytes. Biochim. Biophys. Acta *1395*, 288–292.
22. Korvatska, E., Henry, H., Mashima, Y., Yamada, M., Bachmann, C., Munier, F.L., and Schorderet, D.F. (2000). Amyloid and non-amyloid forms of 5q31-linked corneal dystrophy resulting from kerato-epithelin mutations at Arg-124 are associated with abnormal turnover of the protein. J. Biol. Chem. *275*, 11465–11469.
23. Kohfeldt, E., Sasaki, T., Göhring, W., and Timpl, R. (1998). Nidogen-2: a new basement membrane protein with diverse binding properties. J. Mol. Biol. *282*, 99–109.
24. Leslie, A.G.W. (1994). MOSFLM Users Guide (Cambridge, UK: MRC-LMB).
25. CCP4 (Collaborative Computing Project 4) (1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.
26. Jones, T.A., Zhou, J.-Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.
27. Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography and NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.
28. Laskowski, R.A., MacArthur, M.W., Moss, D.S., and Thornton, J.M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. *26*, 283–291.
29. Esnouf, R.M. (1997). An extensively modified version of MOLSCRIPT which includes greatly enhanced colouring facilities. J. Mol. Graph. *15*, 132–134.
30. Merritt, E.A., and Bacon, D.J. (1997). Raster3D: photorealistic molecular graphics. Methods Enzymol. *277*, 505–524.

Accession Numbers

The coordinates of the FasI 3-4 structure have been deposited in the Protein Data Bank under ID code 1o70.