The structure of the FERM domain of merlin, the neurofibromatosis type 2 gene product

Beom Sik Kang,† David R. Cooper,† Yancho Devedjiev, Urszula Derewenda and Zygmunt S. Derewenda*

Department of Molecular Physiology and Biological Physics, University of Virginia, Charlottesville, Virginia 22908-0736, USA

† These authors contributed equally to the project.

Correspondence e-mail: zsd4n@virginia.edu

Neurofibromatosis type 2 is an autosomal dominant disorder characterized by central nervous system tumors. The cause of the disease has been traced to mutations in the gene coding for a protein that is alternately called merlin or schwannomin and is a member of the ERM family (ezrin, radixin and moesin). The ERM proteins link the cytoskeleton to the cell membrane either directly through integral membrane proteins or indirectly through membrane-associated proteins. In this paper, the expression, purification, crystallization and crystal structure of the N-terminal domain of merlin are described. The crystals exhibit the symmetry of space group $P2_12_12_1$, with two molecules in the asymmetric unit. The recorded diffraction pattern extends to 1.8 Å resolution. The structure was solved by the molecular-replacement method and the model was refined to a conventional R value of 19.3% (R_{free} = 22.7%). The N-terminal domain of merlin closely resembles those described for the corresponding domains in moesin and radixin and exhibits a cloverleaf architecture with three distinct subdomains. The structure allows a better rationalization of the impact of selected disease-causing mutations on the integrity of the protein.

1. Introduction

Neurofibromatosis type 2 (NF2), first described in 1822 by the Scottish surgeon Wishart, is an often devastating autosomal dominant disorder affecting one in every 40 000–90 000 potential births, depending on geographic factors (Evans *et al.*, 1992, 2000; Gutmann, 2001; Martuza & Eldridge, 1988). Until about 1985, NF2 was often linked with neurofibromatosis type 1, also a dominant inherited disorder, and the two were collectively referred to as von Recklinhausen disease. Individuals affected by NF2 develop central nervous system tumors such as Schwann cell tumors of the eighth cranial nerve (bilateral vestibular schwannomas), meningiomas and ependymomas, which although classified as cancers are typically slow-growing and non-malignant. The clinical symptoms vary profoundly from a mild to a very severe phenotype, with diagnostic prevalence of the disease significantly lower than birth incidence (Evans *et al.*, 2000; Gutmann, 2001).

Neurofibromatosis type 2 is associated with a homozygous inactivation of the NF2 gene. Located within 17 exons in the long arm of chromosome 22, this gene encodes a 595-residue protein denoted as schwannomin or merlin (Rouleau *et al.*, 1993; Trofatter *et al.*, 1993). Alternative splicing of exon 16 results in the presence of another isoform, which differs only in the C-terminal 11 residues, with important functional consequences (Sherman *et al.*, 1997). There is convincing evidence that mutations inactivating some or all of the biological functions of merlin, which acts as a tumor

suppressor protein, are the causal factor behind the etiology of NF2. For example, in schwannoma, overexpression of wild-type NF2 gene but not of a mutant leads to growth suppression, impaired cell motility, adhesion and spreading (Gutmann *et al.*, 1998, 1999; Sherman *et al.*, 1997). Furthermore, mice with targeted mutations in the NF2 gene develop malignant tumors (McClatchey *et al.*, 1998).

Merlin is a member of a larger group of proteins, which includes protein 4.1, talin and three closely homologous proteins known collectively as ERM, *i.e.* ezrin, radixin and moesin (Mangeat *et al.*, 1999; Tsukita *et al.*, 1994, 1997; Tsukita & Yonemura, 1997). The ERM proteins have no known catalytic function, but are believed to participate in signaling phenomena by providing a link between the actin cyto-skeleton and the membrane (Tsukita *et al.*, 1994). Like other ERM proteins, merlin contains three domains: the N-terminal domain (also denoted as the FERM domain) comprising approximately the first 300 residues, a central coiled-coil fragment and a C-terminal polypeptide containing the last 120 residues. The C-terminal polypeptide of merlin is unique among the ERM family members in that it does not contain an actin-binding motif (Mangeat *et al.*, 1999; Turunen *et al.*, 1998). The molecular physiology of merlin and of the ERM proteins in general involves intermolecular or intramolecular head-to-tail interaction between the FERM domain and the C-terminal polypeptide (Meng *et al.*, 2000; Nguyen *et al.*, 2001; Sherman *et al.*, 1997; Tsukita *et al.*, 1997). The FERM domain of merlin has been implicated in intermolecular interactions with such proteins as CD44 (Herrlich *et al.*, 2000), EBP50 (NHE-RF; Murthy *et al.*, 1998), SCHIP-1 (Goutebroze *et al.*, 2000), HRS (Scoles *et al.*, 2000), β1-integrin (Obremski *et al.*, 1998) and RhoGDI (Maeda *et al.*, 1999). Whether or not all of these interactions are physiologically relevant remains to be validated, as are the specific signaling pathways relevant to merlin. However, the regulated association of the FERM domain of merlin with the C-terminal polypeptide (also denoted C-ERMAD) mediates tumor-growth suppression in normal cells (Sherman *et al.*, 1997). Under normal conditions the association between the two domains is regulated by phosphorylation of the C-terminal polypeptide, although it is not clear what induces this process.

Recently, the molecular architecture of the ERM proteins has become better understood owing to X-ray diffraction analyses of the FERM domains of radixin and moesin. The moesin domain structure was solved at 1.9 Å resolution in complex with its partner C-terminal polypeptide, but with the intervening coiled-coil fragment removed by recombinant methods (Pearson *et al.*, 2000), and was also studied independently in a form which includes an extension into the coiled-coil region at 2.7 Å resolution (Edwards & Keep, 2001). The radixin FERM domain was solved with and without bound inositol-(1,4,5)-triphosphate (IP3) at 2.8 and 2.9 Å resolution, respectively (Hamada *et al.*, 2000). In addition, a more distantly related domain from protein 4.1 was also solved by X-ray diffraction at 2.8 Å resolution (Han *et al.*, 2000). These studies revealed that the FERM domains are structurally very similar, with a cloverleaf-like architecture

consisting of three distinct subdomains. The N-terminal subdomain has a ubiquitin-like fold and is followed by a subdomain resembling an acyl-CoA binding protein and a third subdomain reminiscent of a phosphotyrosine-binding domain (PTB) or pleckstrin homology domain (PH). In the structure of the moesin intramolecular complex (Pearson *et al.*, 2000), the C-terminal polypeptide adopts an extended meandering conformation, which suggests that without its FERM partner it is unable to form a stable tertiary fold.

In spite of significant progress in the studies of ERM proteins, the structure of merlin, the specific molecule associated with NF2, has not been described. It is important to stress that there are critical functional differences between merlin and its homologs and that only merlin mutations are associated with the neurofibromatosis phenotype. Efforts to design therapeutic agents able to interact with the FERM domain of merlin in a way that could relieve NF2 symptoms would certainly benefit from an accurate knowledge of the molecular structure of merlin itself. Here, we report the structure of the human merlin FERM domain (residues 1–313) at 1.8 Å resolution. The structure reveals the expected conserved cloverleaf architecture of the FERM domain and provides an additional rationale for the pathological effects of the known NF2-associated missense mutations. It also suggests regions of the protein that are critical for the interactions with effectors and/or activators of merlin.

2. Materials and methods

2.1. Construction of FERM domain expression plasmids

A merlin clone was purchased from American Tissue and Culture Collection (ATCC 106908). It contained the nucleotide sequence corresponding to the merlin N-terminal 341 amino acids. To express this sequence using the Gateway gene-expression system (Life Technologies) and to introduce the recombinant TEV protease (rTEV) cleavage site between the glutathione S-transferase (GST) tag and the target protein, we designed three primers, attB1-rTEV primer (5′-GGGGAC-AAGTTTGTACAAAAAAGCAGGCTCCGAAAACCTG-TATTTTCAGGGC-3′), rTEV-merlin primer (5-TCCGAA-AACCTGTATTTTCAGGGCATGGCCGGGGCCATCGC-TTCCCGC-3′) and attB2-merlin primer (5′-GGGGACCAC-TTTGTACAAGAAAGCTGGGTTCATCGAGCGAGGCC-ACGCTGCCGCTCCATCTGCTTTCTATCC-3′).

A PCR product generated by two-step PCR (rTEV-merlin primer and attB2-merlin first and then attB1-rTEV primer and attB2-merlin primer) was cloned into pDEST15, a GST fusion protein vector, according to the manufacturer's instructions and this clone was named pDEST15:merlin341. To improve the efficiency of purification, we modified the vector to include a hexa-His (His$_6$) tag at the *Nde*I site in front of GST sequence using the primers 5′-TATGTCAGGGCACCATCACCAT-CACCATTCTGGGGCTGC-3′ and 5′-TAGCAGCCCCAG-AATGGTGATGGTGATGGTGCCCTGACA-3′. This vector was denoted pHisDEST15:merlin341. Finally, we introduced a stop codon after Ala113 to eliminate the amino acids

extraneous to the FERM domain, using the primers 5′-GAGGAGAAGGAAAGCCTAGTCTTTGGAAGTTCAG-CAG-3′ and 5′-CTGCTGAACTTCCAAAGACTAGCCTT-TCCTTCTCCTC-3′. This resulted in the clone pHisDEST15:merlin313, which was used in all subsequent protein-expression experiments.

2.2. Protein purification and crystallization

To overexpress the double-tagged merlin FERM domain, pHisDEST15:merlin313 was introduced into *Escherichia coli* BL21 (DE3) RIL strain (Stratagene). LB medium containing ampicillin (50 mg ml^{-1}) was inoculated using 5%(*v/v*) of overnight seed culture. After cultivation at 310 K for 3 h, 1 m*M* IPTG was added and cells were cultivated at 295 K for a further 12 h. Cells were harvested by centrifugation at 5000*g* for 20 min, resuspended with 50 m*M* Tris–HCl pH 7.5 (buffer *A*) and disrupted by sonication (Sonifier 450, Branson) for 30 s ml^{-1}. The cell lysate was centrifuged at 26 000*g* for 45 min and the soluble supernatant was applied to a glutathione-Sepharose 4B column (Amersham Pharmacia Biotech). After washing the column with 50 m*M* Tris–HCl pH 8.5, 50 m*M* NaCl, the recombinant protein was eluted with buffer *B* (10 m*M* glutathione). The eluent was subjected to a HiPrep 26/10 Desalting column (Amersham Pharmacia Biotech) equilibrated with buffer *A* to remove NaCl and glutathione. The recombinant protein was digested using rTEV protease (Life Technologies) at 283 K in the presence of 0.5 m*M* EDTA and 1 m*M* DTT. After digestion, 300 m*M* NaCl was added to the digested recombinant protein solution and it was passed through a glutathione Sepharose 4B column again to remove uncut full-length fusion protein and the His$_6$-GST tag. To remove rTEV protease and residual tag, 10 m*M* imidazole was added to the flowthrough from the glutathione Sepharose 4B column and this solution was loaded onto an Ni–NTA column (Qiagen) equilibrated with buffer *A* containing 300 m*M* NaCl and 10 m*M* imidazole. The flowthrough of this column was concentrated using a Centriprep YM30 (Amicon), loaded onto a Superdex G75 column (Amersham Pharmacia Biotech) and eluted with buffer *A* containing 300 m*M* NaCl. The fractions containing the merlin FERM domain were collected and concentrated using a Centriprep YM30 for crystallization screening. All the purification steps, except the rTEV digestion, were performed at 277 K. The purified FERM domain contains an additional glycine at the N-terminus arising from the rTEV recognition sequence. After the purification, about 30 mg of pure protein was obtained from 2.8 l of culture.

After screening for crystallization conditions using Crystal Screen and ammonium sulfate Grid Screen (Hampton Research), crystallization conditions were optimized around 0.1 *M* sodium cacodylate pH 6.5 containing ammonium sulfate and dioxane. The sitting-drop vapor-diffusion method was used for all crystallization trials. Drops were formed of 3 µl of protein solution and 3 µl of reservoir buffer and were overlaid with a 1:1 mixture of silicone and mineral oils. Crystallization trays were stored at 294 K. The best crystals were obtained

Table 1
Data-collection and refinement statistics.

Values in parentheses refer to the highest resolution shell.

Experimental data	
Space group	$P2_12_12_1$
Unit-cell parameters (Å)	
a	87.02
b	89.33
c	96.77
Resolution (Å)	30–1.80 (1.86–1.80)
Mosaicity (°)	0.69
Unique reflections	68222 (6875)
Redundancy	3.6 (3.4)
Completeness (%)	95.4 (97.2)
R_{sym}†	0.065 (0.622)
Average $I/\sigma(I)$	16.8 (2.68)
Reflections with $I > 3\sigma$ (%)	70.8 (30.2)
Refinement details	
Resolution (Å)	5.0–1.8 (1.847–1.8)
Reflections (working)	66303 (4917)
Reflections (test)	985 (77)
R_{work}† (%)	19.3 (26.0)
R_{free}† (%)	22.7 (26.8)
No. of waters	862
R.m.s. deviation from ideal geometry	
Bonds (Å)	0.011
Angles (°)	1.39
Average *B* factor (Å^2)	
Main chain	23.5
Side chain	26.1
Waters	40.6
Sulfate	44.3

† $R_{sym} = \sum_{hkl} |I - \langle I \rangle|/\sum_{hkl} I$, R_{work} or $R_{free} = \sum ||F_{obs}(hkl)| - |F_{calc}(hkl)||/\sum_{hkl} |F_{obs}(hkl)|$.

using a 5 mg ml^{-1} protein solution and a buffer containing 56% saturated ammonium sulfate, 2% dioxane and 0.1 *M* sodium cacodylate.

2.3. Data collection, structure solution and refinement

The crystal used for data collection was briefly soaked in a solution containing 12.5%(*v/v*) glycerol and 56% ammonium sulfate before being transferred to 24% glycerol and 30% ammonium sulfate and frozen by immersion in liquid nitrogen. The data were collected at beamline X9B at NSLS at a wavelength of 0.920 Å under cryoconditions using an ADSC Quantum 4 CCD. The data were indexed and scaled using *HKL*2000 (Otwinowski & Minor, 1997).

The structure was solved by molecular replacement using *AMoRe* (Navaza, 1994). The program *SEAMAN* (Kleywegt, 1996*a*) was used to create a search model based on the radixin structure (PDB code 1gc6), with serines substituted for all non-conserved residues larger than alanine. Manual model rebuilding was performed in *O* (Jones *et al.*, 1991). A combination of *CNS* (Brünger *et al.*, 1998) and *REFMAC* from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) was used for refinement, with the final refinement performed using *REFMAC*5 with default values for target stereochemistry (Murshudov *et al.*, 1997). Waters were added using *ARP/wARP* (Perrakis *et al.*, 1999).

3. Results and discussion

3.1. Crystallization, data collection and structure solution

The FERM domain crystals belong to space group $P2_12_12_1$, with unit-cell parameters $a = 87.02$, $b = 89.33$, $c = 96.76$ Å. After two weeks, the average size of the crystals was $0.2 \times 0.2 \times 0.1$ mm (Fig. 1a). The volume of the asymmetric unit (188 040 Å^3) suggested the presence of two molecules, with a resultant Matthews coefficient of 2.5 Å^3 Da^{-1} and a solvent content of 51.5%. Data with overall completeness of 95.4% were collected from a single frozen crystal and were merged and scaled with an R_{merge} of 0.065 (Table 1).

In order to assess which of the two existing atomic models of FERM domains is closer to merlin, parallel molecular-replacement calculations were performed with a model based on the radixin structure at 2.8 Å without IP3 (PDB code 1gc7; Hamada *et al.*, 2000) and a model based on the moesin structure (PDB code 1ef1; Pearson *et al.*, 2000). Although the levels of sequence identity of merlin with radixin or moesin are high (64 and 63%, respectively), a model was constructed from each, with non-conserved residues larger than alanine truncated to serines. These two polyserine models yielded molecular-replacement solutions that were marginally better than those obtained from the complete structures (data not shown). With each model, the molecular-replacement calculations gave two solutions, in agreement with expectations based on crystal density considerations. By most statistical criteria, the moesin-based model provided the best solution for the rotation and translation function, but the radixin-based model provided a better solution after rigid-body refinement. Based on this and the slightly higher sequence similarity with radixin, the radixin-based model was used as the starting point for refinement. During the refinement, non-crystallographic symmetry restraints were not applied, given the relatively high resolution of the data. A combination of *CNS* and *REFMAC* was used to refine the structure, with the final rounds of refinement performed in *REFMAC*5. Maximum-likelihood refinement of the model converged with the statistics reported in Table 1. To determine the extent of model bias, several rounds of refinement were also performed using the moesin-based model. This refinement was discontinued when we were satisfied that the model was not significantly biased by the initial model choice.

3.2. Quality of the refined atomic model

The final model consists of two molecules of the FERM domain, 861 water molecules and six sulfate ions. The first 19 amino acids in each of the FERM domain molecules are not visible in the electron density. The refined structure conforms to standard protein stereochemistry, with an r.m.s. deviation from ideal bond lengths of 0.011 Å and only five of the 588 residues of the structure falling into generously allowed regions of the Ramachandran plot (Laskowski *et al.*, 1993). Only a few side chains are not entirely contained within the electron density of a $2F_o - F_c$ σ_A-weighted map at 1σ (Figs. 1c and 1d). Each monomer contains one *cis*-proline. A limited number of residues exhibit static disorder, but alternate conformations were not refined at this point.

The main-chain temperature factors range from 12.0 to 54.3 Å^2, with average values of 22.6 and 24.4 Å^2 for chains A and B, respectively. This similarity is easily rationalized by similar packing of both molecules in the crystal lattice. The temperature factors are generally higher, as expected, in external loops. The exception to this is strand $\beta5C$, a terminal strand in a β-sheet, which is stabilized by hydrogen bonds on one side only. The low B values reported in this study reflect the superior quality of the atomic model. This is particularly
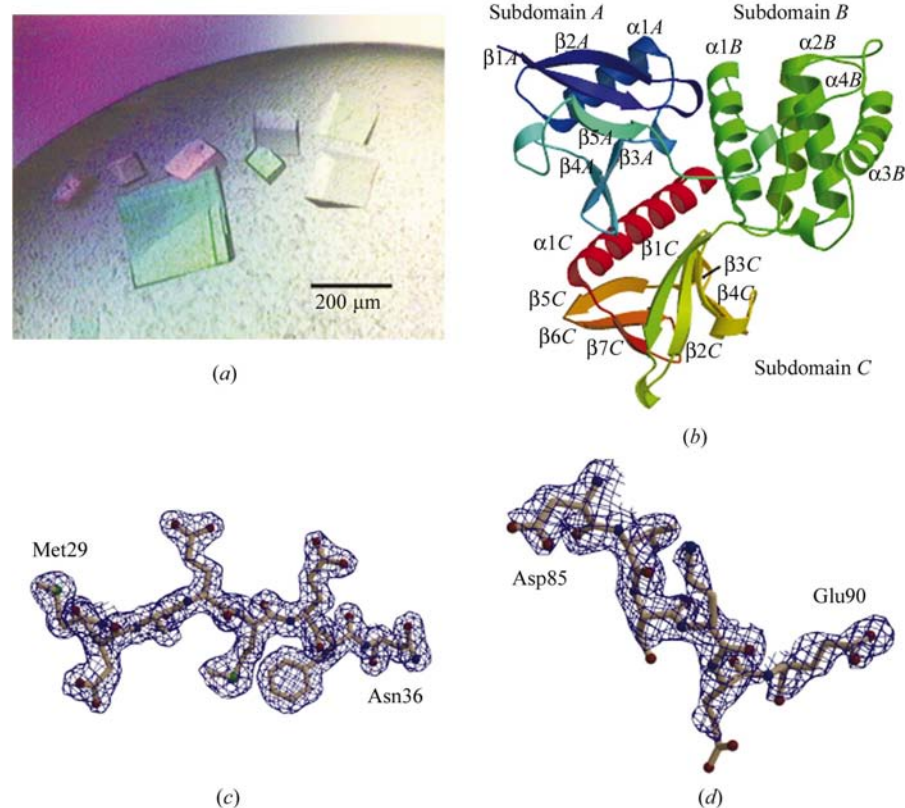

Figure 1
(a) Typical crystals of the FERM domain of merlin. (b) The FERM domain of merlin is shown in a ribbon representation color-ramped from blue to red. The subdomains are labeled as they are described in the text and the main secondary-structural elements are labeled. (c) Typical electron density is shown contoured at 1.2σ in a $2mF_{obs} - DF_{calc}$ map. (d) The electron density for the region with the highest B factors is shown contoured at 1.0 σ in a $2mF_{obs} - DF_{calc}$ map. Figs. 1(b), 2, 3, 4 and 5 were produced using *MOLSCRIPT* (Kraulis, 1991) and *Raster3D*. Figs. 1(c) and 1(d) were produced with *BOBSCRIPT* (Esnouf, 1997) and *Raster3D* (Merritt & Bacon, 1997).

striking when these values are compared with those found in the radixin models ($\sim$64 Å^2 for the main chain) or the isolated FERM domain of moesin (PDB code 1e5w; $\sim$70 Å^2 for main chain). While this discrepancy may stem from limited resolution in those studies, it is also likely that the refinement protocols may not have been optimal. For example, some of the loops in the moesin FERM domain (PDB code 1e5w) have B values in excess of 140 Å^2, which corresponds to an unrealistic value of the mean-square displacement $\langle r^2 \rangle$ of nearly 2.0 Å^2.

The r.m.s. distance between the C$^\alpha$ atoms of the two molecules, following least-squares overlap, is 0.77 Å. Only a handful of residues, mostly solvent exposed, have different side-chain conformations in the two monomers. The segments with larger differences correlate with areas of higher temperature factors and areas that are involved in crystallographic contacts. The region with the highest discrepancy is the N-terminus of $\alpha 3B$, with the preceding loop and the 3$_{10}$-helix. The average r.m.s. coordinate error derived from the program *SIGMAA* in the *CCP*4 suite is 0.11 Å^2.

3.3. The overall tertiary architecture and comparisons with moesin and radixin

The tertiary structure of the FERM domain of merlin is very close to that of the homologous domains of moesin and radixin (Edwards & Keep, 2001; Hamada *et al.*, 2000; Pearson *et al.*, 2000). The polypeptide chain folds into three clearly identifiable subdomains, each with similarities to known single-domain proteins. These three structural elements were denoted differently for the moesin and radixin structures and we here choose to follow the latter convention, according to which the merlin fragment encompassing residues 20 to approximately 100 is defined as *A*, that including residues 101–215 is denoted *B*, and the third fragment, residues 216–313, is denoted *C* (Fig. 1*b*). As noted by others (Hamada *et al.*, 2000; Pearson *et al.*, 2000), the *A* subdomain has a fold reminiscent of ubiquitin, *B* is similar to the acyl-CoA binding protein and *C* exhibits a fold found in such signaling domains as PTB, PH and EVH1. Merlin is unique in that it has an additional N-terminal extension of 19 amino acids compared with both radixin and moesin. It has been suggested recently based on limited proteolysis experiments that this fragment is disordered (Brault *et al.*, 2001) and our structure fully confirms this prediction. The first amino acid clearly

defined in the electron density is Lys20. It is natural to speculate that this portion of merlin is disordered in solution. However, since this region has been shown to be necessary for the proper functioning of merlin and is implicated in actin binding (Brault *et al.*, 2001), it is also possible that it becomes ordered as merlin binds to some effector target.

Least-squares fitting of the merlin FERM domain onto radixin and moesin reveals that the mutual disposition of the three subdomains is relatively well preserved in all three proteins, although concerted shifts of entire subdomains are noticeable albeit small. Such rearrangements affect global comparisons of r.m.s. positional differences, as the latter are
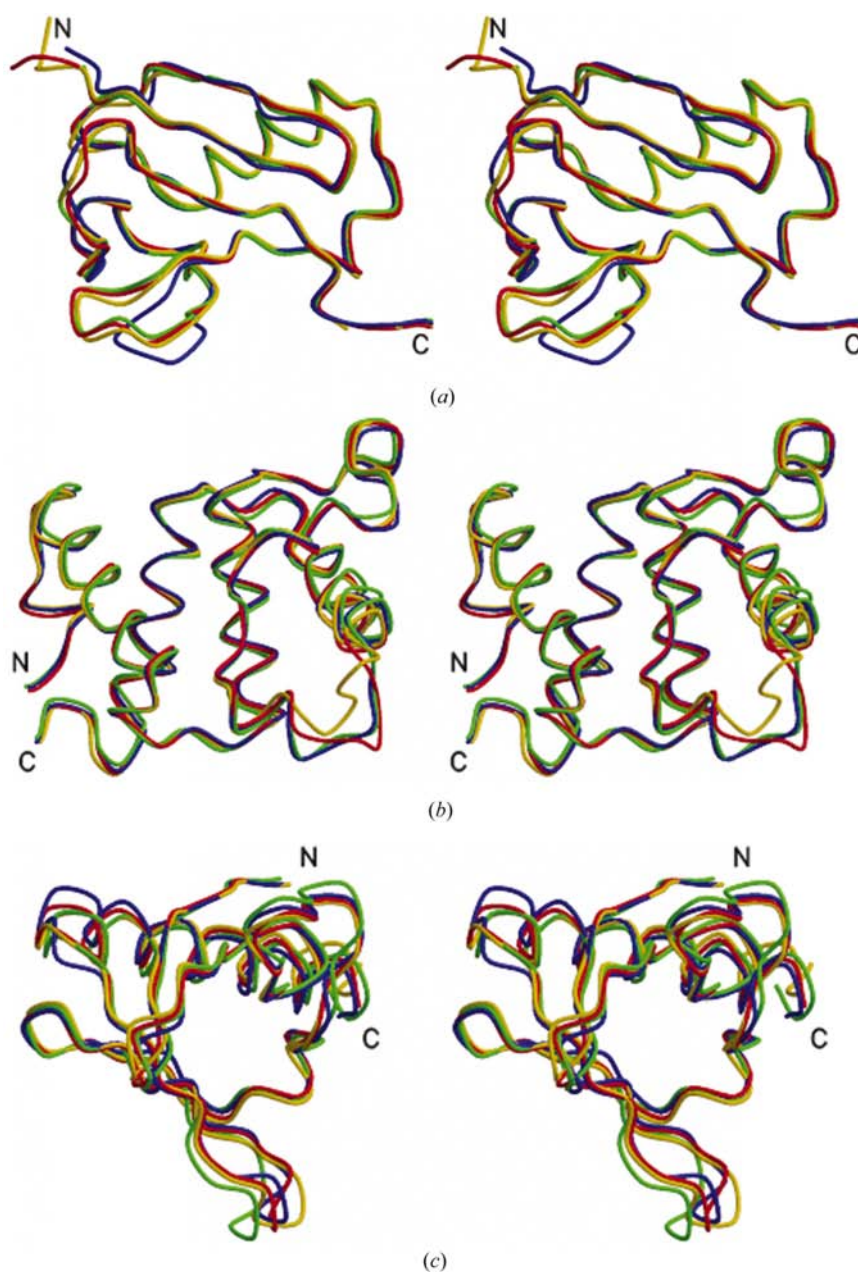

Figure 2
Stereoviews of the superpositions of the individual FERM subdomains of merlin, radixin and moesin: subdomain *A* (top), subdomain *B* (center) and subdomain *C* (bottom). In all figures, merlin is blue, radixin is red, 1ef1 moesin is green and 1e5w moesin is gold.

likely to reflect both local changes and global rearrangements. Superpositions of the individual subdomains of the FERM domains of merlin, moesin and radixin are shown in Fig. 2.

To gain a better understanding of the similarities and differences between the FERM domains, the program *LSQMAN* (Kleywegt, 1996*b*) was used to superpose the FERM domains and their individual subdomains. Owing to the high degree of similarity of the two monomers of merlin and for simplicity, values for the *A* monomer are presented here. When the entire ordered portion of the FERM domain of merlin (294 residues) is fitted onto radixin (PDB code 1gc7) or either of the deposited FERM domains of moesin (PDB codes 1ef1, which corresponds to the 1.9 Å study, and 1e5w, the 2.7 Å resolution model), the corresponding values of the r.m.s. positional differences between the C^α atoms range from 1.8 to 2.0 Å. However, when the subdomains are fitted onto the targets separately and when a few outliers with distances above 3.5 Å are excluded, the values fall dramatically to approximately 0.92 Å for moesin and 0.8 Å for radixin. Subdomain *A*, when fitted onto 1gc7, 1ef1 and 1e5w, showed r.m.s. differences of 0.65, 0.55 and 0.65 Å, respectively, with 70,

64 and 69 C^α atoms within a 1.5 Å distance. The discrepancies occurred consistently around residue 31, the loop comprising amino acids 67–72 and residue 88. For domain *B*, the results were 0.45, 0.65 and 0.51 Å, with 102, 91 and 103 atoms included, respectively. There is a consistent discrepancy, which includes the loop around residue 177. Finally, subdomain *C* exhibits r.m.s. differences of 0.61, 0.67 and 0.70 Å, respectively, with 86, 87 and 77 atoms included and a consistent departure of residues 288–291 from the average. None of the differences are of a magnitude which would suggest a significant biological effect and some can be easily rationalized in terms of crystal contacts.

The high resolution of the present study permits a detailed analysis of the interfaces between the three subdomains, including contributions from the ordered solvent. Two large interfaces contribute to the integrity of the tertiary structure of the FERM domain. The first involves residues from subdomains *A* and *C*. The C-terminal long helix of the *C* subdomain (residues 289–313) packs against two loops of subdomain *A* containing residues 69–76 and 99–103. The face of the helix involved in this interface is largely non-polar and contains Leu306, Leu299 and Leu295; the *A* subdomain contributes Phe100, Trp74 and Val72. Numerous water molecules flank this interface; however, they do not seem to be an integral part of the interface but rather form a typical hydration shell. This specific interface is different in both radixin and moesin because of the single amino-acid deletion which is found in merlin in the loop comprising residues 66–72 and confers a conformational change. As a result, the loop packs significantly closer to the N-terminus of the helix in the *C* subdomain, probably because of a salt bridge formed between Asp70 on one side and Arg291 and Lys289 on the other. Both moesin and radixin lack an aspartate in this position and instead contain bulky aromatics (Phe or Tyr) which push the loop away from the *C* subdomain. The significant difference in the local structure of this loop, as well as the dramatically different amino-acid sequence in this region, suggest that this epitope may be involved in protein–protein interactions unique to merlin.

Another interface is found between subdomains *A* and *B* which, in addition to the intervening loop, interact *via* the first helix of the *B* subdomain, which is wedged between the two subdomains and contributes several hydrophophic side chains such as Ile126, Val122, Phe118 and Phe119. There are also direct hydrogen bonds between the subdomains. This interface is closely packed and lacks any internal water molecules.
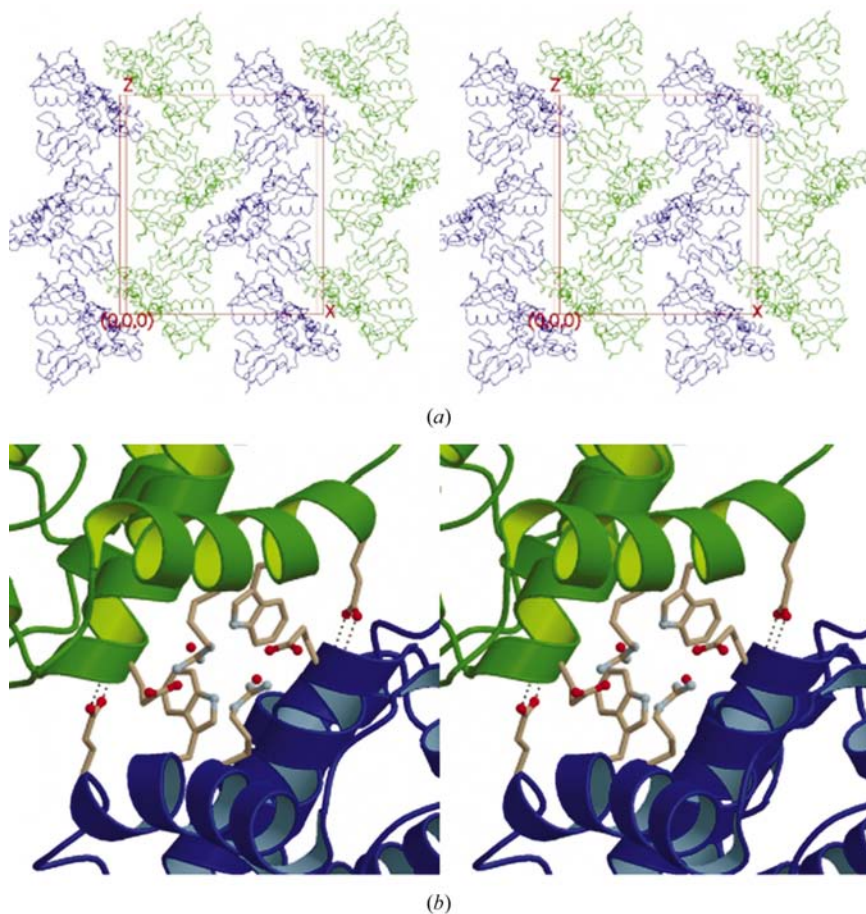

Figure 3
Crystal packing of merlin. (*a*) Stereoview of the packing in the unit cell showing the similarities of the packing of the *A* and *C* subdomains for the two merlin monomers. (*b*) Stereoview of the dimer interface. A salt bridge positions the side chains of Glu136 and Arg187 such that they pack against Trp191. Glu194 hydrogen bonds with amide N atoms at the N-terminus of a central helix of subdomain *B*. In both figures monomer *A* is green and monomer *B* is blue.

There is only a small interface between subdomains *C* and *B*. Leu250, located in the loop between *β3C* and *β4C*, fits into a small pocket formed by Tyr132, Glu215 and Met216. Overall, however, the relative positions of these two subdomains and the entire FERM 'cloverleaf' are defined by the contacts described above and the covalent linkages.

The crystals of the FERM domain of merlin contain two molecules in the asymmetric unit related by a non-crystallographic twofold axis running nearly parallel to the crystallographic *b* axis and between the *B* subdomains of adjacent molecules (Fig. 3*a*). This packing is consistent with a strong maximum in the native Patterson seen at 0.0, 0.5, 0.094 (data not shown), indicating translational non-crystallographic symmetry. The interface between these two molecules, although small (422 Å^2), is quite intricate and involves helices *α2B* and *α4B* in a symmetric arrangement (Fig. 3*b*). The two helices in each molecule interact *via* a salt bridge involving Glu136 and Arg187. Furthermore, the side chains of Arg187 from the two molecules pack tightly against each other and are flanked on each side by the indole rings of the two Trp191 residues. These in turn pack against Glu136 in the neighboring molecule, further stabilizing this contact. At each end of this interface, Glu194 caps otherwise non-bonded backbone amides of residues 136 and 137 at the N-terminus of an *α*-helix, so that each O$^\varepsilon$ atom accepts a hydrogen bond from one amide N atom. This elegant cap stabilizes the incipient helix immediately downstream of a diprolyl peptide. Finally, two symmetrical pairs of water molecules, each coordinated by at least three hydrogen-bonding partners, add to the stability

Figure 4
Missense mutations of merlin. The distribution of NF2-associated missense mutations is shown by the presence of a sphere at the C$^\alpha$. Red spheres represent a subtitution mutation, purple a deletion and green an insertion. Sites of mutations are labeled.

of this contact. We note that in moesin this general area is involved in the binding of the C-terminal polypeptide and that many of the residues participating in this interface are conserved among the FERM domains, all of which suggests a functional significance.

Other crystal contacts also contribute to the stability of the lattice. Residues 30–36 in subdomain *A* of molecule 1 interact with two loops of subdomain *C* in an adjacent molecule 1, *i.e.* residues 280–282 and 252–255. As both subdomains are roughly at the same *x* and *y* coordinates, this arrangement forms a chain which runs along the *c* axis of the crystal. Interestingly, molecule 2 shows similar contacts and thus the symmetry of the two molecules is broken by the different packing of *B* subdomains against *C* subdomains of molecules in the next layer along the *c* axis. The carboxyl terminus of the *C* subdomain of molecule 1 is buried in the loop which includes residues 169–180 of subdomain *B* of molecule 2, whereas the carboxyl terminus of molecule 2 is wedged between the two *B* subdomains. The structural differences observed between *B* subdomains in moesin, radixin and merlin may reflect, at least in part, the impact of this crystal contact.

3.4. The NF2-associated missense mutations

The structure of moesin has been used previously to analyze the structural consequences of NF2-associated mutations in merlin. However, given that the current study focuses on the NF2 causal gene product itself, it is proper to address this issue again. Although the most devastating mutations of merlin are nonsense mutations that cause a premature termination of merlin, a number of missense mutations are associated with milder cases of the disease (Gutmann *et al.*, 1998). As can be seen in Fig. 4 and Table 2, 20 of these mutations are distributed throughout the FERM domain, with a slightly higher frequency of mutations in the *A* subdomain. Most of the NF2-assocated mutations occur at sites that are conserved between merlin and other FERM domains, with six of these (Leu46, Phe62, Leu64, Lys79, Phe96 and Ile273) completely conserved among the ERM proteins and protein 4.1. While a number of the NF2-associated mutations are likely to cause critical disruption in the packing of the respective subdomain, the majority of the mutations may impact the subdomain interfaces. This suggests that the specific architecture of the cloverleaf is crucial for the normal function of the protein. None of the mutations occur at the surface interacting with the C-terminal polypeptide of merlin, as predicted from the structure of the moesin complex.

The subdomain interface that is affected by the largest number of NF2-associated mutations is the *AB* interface. Phe62 is directly involved in the *AB* interface and the mutation of this residue to a serine removes part of the hydrophobic interaction between these two subdomains. The L117I mutation in subdomain *B* is also found at the hydrophobic interface between these two subdomains. The insertion of a leucine in subdomain *A* after residue 49 may also affect the *AB* interface by altering the conformation of the subsequent
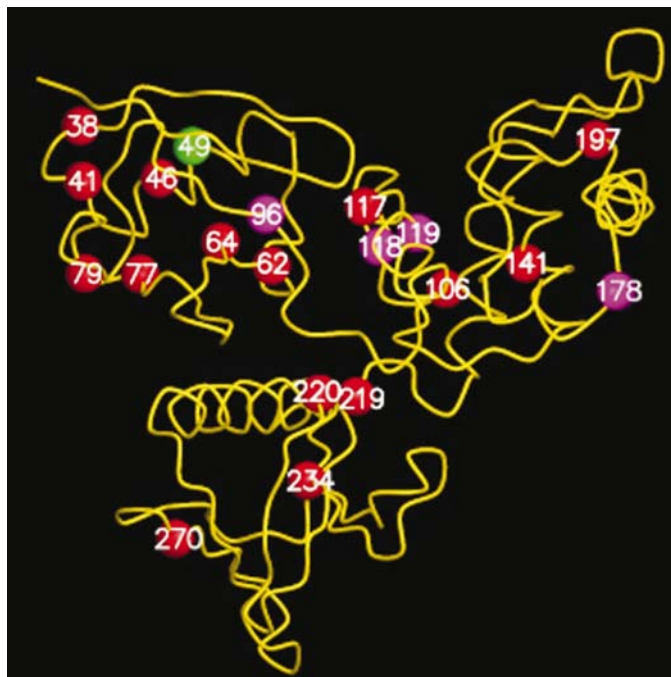
Table 2
NF2-associated mutations.

HC, hydrophobic core. Sub*A*, subdomain *A*; sub*B*, subdomain *B*; sub*C*, subdomain *C*; ins, insertion; del, deletion.

Mutation	Structural consequence	Phenotype	Reference
Subdomain *A* mutations			
E38V	Decreased solubility or impaired interactions	Mild NF2	Parry *et al.* (1996)
W41C	Side-chain packing in sub*A*	Mild NF2	Welling *et al.* (1996)
L46R	HC of sub*A*	Meningiomas	Mérel *et al.* (1995)
ins 49L	HC of sub*A* + *AB* hydrogen-bond loss	Mild NF2	Ruttledge *et al.* (1996)
F62S	*AB* hydrophobic interface	Mild or severe NF2	Scoles *et al.* (1996)
L64P	HC of sub*A*	Not reported	Xu & Gutmann (1998)
M77V	See text	Intermediate NF2	Evans *et al.* (2000)
K79E	See text	Schwannomas	Sainz *et al.* (1994)
del F96	HC of sub*A*	Severe NF2	MacCollin *et al.* (1994)
Subdomain *B* mutations			
E106G	*BC* packing (salt-bridge loss in sub*B*)	Severe NF2	Bourn *et al.* (1994)
L117I	*AB* hydrophobic interface	Meningiomas	De Vitis *et al.* (1996)
del F118 and/or del F119	*AB* hydrophobic interface + HC of sub*B*	Severe NF2	Bourn *et al.* (1995)
L141P	Breaks helix and disrupts fold of sub*B*	Not reported	Unpublished†
del Q178	See text	Severe NF2	Kluwe *et al.* (2000)
G197C	Unfavorable conformation in loop	Mild NF2	Welling *et al.* (1996)
Subdomain *C* mutations			
V219M	*BC* hydrophobic interface	Meningiomas	Mérel *et al.* (1995)
N220Y	*AC* hydrophobic interface	Mild NF2	Ruttledge *et al.* (1996)
L234R	HC of sub*C*	Severe NF2	Jacoby *et al.* (1999)
E270G	Loss of salt bridge disrupts sub*B*	Severe NF2	Kluwe *et al.* (1998)

† Unpublished mutation found at http://neuro-trials1.mgh.harvard.edu/nf2.

loop. This loop facilitates four inter-subdomain hydrogen bonds, with the extended side chains of Glu58 in subdomain *A* and Gln111 in subdomain *B* hydrogen bonding with the backbone of the adjacent subdomain. Furthermore, deletions of Phe118 and Phe119, together or individually, were also found in NF2 patients. These residues are contained in α1*B* and further underscore the functional importance of the *AB* interface.

Several NF2-assocated mutations affect the interfaces between domain *C* and the other two subdomains. Two mutations, V219M and N220Y, are located at the N-terminal strand of subdomain *C*. Their side chains point in opposite directions, with the side chain of 219 pointing towards the *BC* interface and the side chain of 220 in the direction of the *AC* interface. Introduction of a bulkier side chain at either site may disrupt the respective interface. The mutation E106G removes a side chain involved in a salt bridge with Lys209, possibly allowing subdomain *B* to rotate closer to subdomain *C*. In other ERM proteins a serine or alanine is found in this position, making this interaction unique to merlin.

The majority of the other NF2-assocated mutations are most likely to disrupt local packing within their respective subdomains. L46R and L234R introduce a large charged residue in the hydrophobic core of subdomains *A* and *C*, respectively. The L64P substitution and the deletion of Phe96 would create cavities in the hydrophobic core of subdomain *A*, as well as disrupt the local secondary structure surrounding these residues. The G197C substitution occurs at the loop between α2*B* and α3*B*, which requires a backbone conformation that is unfavorable for cysteine ($\varphi = 100$, $\psi = -13°$) and may decrease the solubility of the protein by exposing the side chain to solvent. L141P may destabilize subdomain *B* by

inserting a proline into the middle of one of the central helices (α2*B*). Although this residue is not on the side of the helix that points towards α4*B*, it is approximately at the position where these two central helices cross each other. The E270G mutation is likely to destabilize the *C* subdomain and alter one of the potential effector-binding sites (see below).

Several of the mutations do not clearly fall into the categories of disrupting subdomain interfaces or subdomain tertiary structure. One such mutation is K79E, which is at the end of α4*A*. This charge-reversing mutation is very likely to cause the formation of a salt bridge with the neighboring Lys76. Both of these lysines are conserved among ERM proteins. In the merlin structure Lys76 is hydrogen bonded to Tyr66 in α3*A*, which is also conserved in the ERM family. However, in the radixin structure the homologous lysine extends outward and interacts with the IP3. Although as yet there is no direct evidence that merlin binds inositol phosphates, almost all of the residues responsible for the binding of IP3 in the radixin structure are either conserved in merlin or replaced with functionally equivalent amino acids. The charge reversal caused by the K79E mutation would most likely prevent any inositol phosphates from binding to this pocket.

The potential effect of mutations of residues with solvent-exposed side chains is less clear. The side chain of Met77, mutated to a valine in at least one NF2 case, packs against Phe47 and the mutation may create a destabilizing solvent-accessible depression. Similarly, the substitution of Glu38 for a valine is in a solvent-exposed region. Although this substitution is sterically accommodated, it would place a hydrophobic residue on the surface of the protein, possibly substantially decreasing the solubility of merlin. It is noteworthy that the same type of substitution in hemoglobin causes sickle-cell anemia. The nearby mutation of W41C would affect the local packing of side chains in the area surrounding Glu38. The deletion of Gln178 is discussed below.

3.5. Other functional implications

The apparent differences in the biological properties of the various members of the ERM family call for a careful analysis of their respective molecular models. It has been suggested recently that the FERM domain in complex with the C-terminal polypeptide is in a 'dormant' state and that its biological inertness is a product of the occlusion of the relevant epitopes and conformational differences (Edwards & Keep, 2001). This suggestion is based on the 2.7 Å analysis of

the structure of the uncomplexed FERM domain of moesin and on its comparison with the structure of the complexed moesin at 1.9 Å resolution (Pearson *et al.*, 2000). In particular, the loop encompassing residues 260–264 (276–280 in merlin) was found to differ significantly between the two models. We note, however, that the cloverleaf-like fold has some intrinsic flexibility made possible by the interfaces between subdomains. Crystal contacts are sufficient to force minor distortions, but individual domains remain nearly identical within experimental error in their tertiary fold. Although the 276–280 loop in our structure resembles the conformation described by Edwards & Keep (2001), we believe that this does not necessarily constitute proof that the difference is caused by the binding of the C-terminal fragment.

The distribution of residues conserved in moesin, radixin and ezrin but not in merlin can shed light onto the origin of the functional differences between merlin and the ERM proteins. There are 72 such residues and an additional 19 are found in two of the three ERM proteins but not in merlin. Almost all are located at the surface of the protein, although they are not evenly distributed over the surface (Fig. 5). Of these 91 residues unique to merlin, 31 result in a change in the surface electrostatics. Relatively few affect epitopes involved in the binding of the C-terminal polypeptide. The majority of the 91 residues are clustered in three patches, two of which are roughly at a tip of the cloverleaf. One patch is located in each subdomain and therefore the patches will be described as patches *A*, *B* and *C*. These patches are likely to interact with effectors or activators of merlin.

Patch *C* is at the C-terminal end of subdomain *C* and includes resides $\beta5C$–$\beta7C$ and the beginning of $\alpha1C$. All of the merlin-specific residues in this area have their side chains exposed to the solvent and four of them, located on the face of the second β-sheet in this subdomain, involve a charge change from the ERM consensus sequence. Glu270 and Lys284, mentioned above in the context of the E270G mutations, both

constitute a charge change from the other ERM members and are located in patch *C*.

The second patch of residues unique to merlin is found near the tip of subdomain *A* and includes residues found in the distal ends of $\beta1A$ and $\beta2A$, in the following loop and in the N-terminal end of $\alpha1A$. Ezrin has been shown to contain an actin-binding site in this area (Martin *et al.*, 1997). Although the overall net charge of the region is unchanged from the ERM consensus sequence, the local electrostatic footprint is altered by the addition of two acidic and two basic residues, making it unlikely that this serves as an actin-binding site in merlin. Moreover, merlin has been shown to contain an actin-binding site within the first 27 residues, 19 of which are not found in the ERM proteins (Brault *et al.*, 2001). The fact that the E38V and W41C mutations are included in patch *A* makes it more likely that the effects of this protein are manifested by impairing the ability of merlin to bind to effectors or activators.

The subdomain *B* patch contains the beginning of $\alpha4B$ and the loop that precedes it. This region has been called the 'Blue Box' in the *Drosophila* homolog of merlin and has been shown to be vital for the protein's function (LaJeunesse *et al.*, 1998). A comparison of merlin with moesin complexed with its C-terminal fragment reveals that the Blue Box is adjacent to the loop between the *A* and *B* helices of the C-terminal polypeptide. The residues of that fragment that contact the Blue Box are not conserved between merlin and other ERM proteins; thus, the molecular surface that covers the most extended part of the C-terminal polypeptide and the flanking region of merlin is different from the corresponding regions of ezrin, radixin and moesin. This could explain why the activation of merlin is not coincident with any of these proteins. This
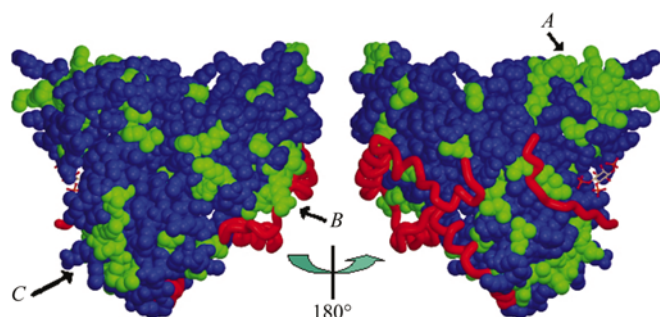

Figure 5
The locations of the residues unique to merlin are shown in green on a blue space-filled model. The arrows point to the patches described in the text. In this figure, the C-terminal polypeptide of moesin has been roughly positioned on merlin to indicate where the FERM and C-terminal polypeptide interaction is most likely to occur in merlin. The IP3 of radixin is also included. The image on the left is in the same orientation as Fig. 1(*b*).
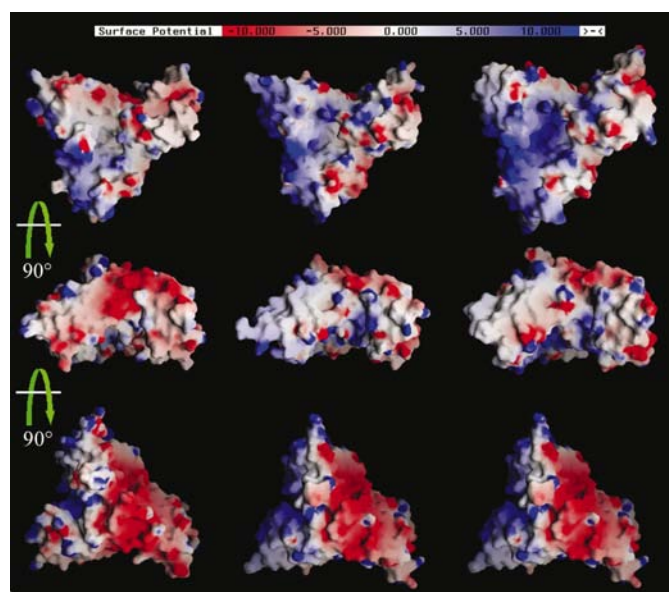

Figure 6
Electrostatic potentials generated in *GRASP* (Nicholls *et al.*, 1991) are shown for merlin (left), radixin (middle) and moesin (1ef1) (right). The top views are in the same orientation as in Fig. 1(*b*) and each successive image down the figure has been rotated 90° forward.

is also supported by the NF2 phenotype associated with the deletion of Gln178 located in the Blue Box region. Gln178 is conserved among the ERM members but not in merlin and is adjacent to where the most extended part of the C-terminal polypeptide would be as judged by the moesin complex structure. The nature of this loop would lead one to believe that this loop could rearrange itself to accommodate this deletion without too much difficulty; however, this mutation leads to a severe NF2 phenotype (Kluwe *et al.*, 2000).

Although the patches of residues unique to merlin create local epitopes, the overall molecular surface and electrostatic potential of merlin is similar to that of other members of the ERM family. The largest exception to this is the *AB* interface (Fig. 6). This cleft is much more electronegative than in the other ERM proteins. It is interesting to note that this is the surface that is affected by many of the NF2-associated missense mutations and is roughly flanked by patches *A* and *B*. This leads one to speculate that this region is crucial for the interaction of merlin with effectors or activators.

4. Conclusions

We have described the structure of the FERM domain of merlin at 1.8 Å resolution, the highest resolution to date for any of the FERM proteins. As expected, the structure is similar to those of the respective domains in radixin and moesin, but also exhibits interesting differences which may have functional implications. This work sets the stage for more detailed analysis of structure–function relationships in merlin, with the aim of designing ways of either subduing or eliminating the devastating symptoms of neurofibromatosis type 2.

References

Bourn, D., Carter, S. A., Mason, S., Gareth, D., Evans, R. & Strachan, T. (1994). *Hum. Mol. Genet.* **3**, 813–816.

Bourn, D., Evans, G., Mason, S., Tekes, S., Trueman, L. & Strachan, T. (1995). *Hum. Genet.* **95**, 572–574.

Brault, E., Gautreau, A., Lamarine, M., Callebaut, I., Thomas, G. & Goutebroze, L. (2001). *J. Cell Sci.* **114**, 1901–1912.

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., Read, R. J., Rice, L. M., Simonson, T. & Warren, G. L. (1998). *Acta Cryst.* D**54**, 905–921.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

De Vitis, L. R., Tedde, A., Vitelli, F., Ammannati, F., Mennonna, P., Bigozzi, U., Montali, E. & Papi, L. (1996). *Hum. Genet.* **97**, 632–637.

Edwards, S. D. & Keep, N. H. (2001). *Biochemistry*, **40**, 7061–7068.

Esnouf, R. M. (1997). *J. Mol. Graph.* **15**, 132–143.

Evans, D. G., Huson, S. M., Donnai, D., Neary, W., Blair, V., Newton, V. & Harris, R. (1992). *Q. J. Med.* **84**, 603–618.

Evans, D. G., Sainio, M. & Baser, M. E. (2000). *J. Med. Genet.* **37**, 897–904.

Goutebroze, L., Brault, E., Muchardt, C., Camonis, J. & Thomas, G. (2000). *Mol. Cell Biol.* **20**, 1699–1712.

Gutmann, D. H. (2001). *Hum. Mol. Genet.* **10**, 747–755.

Gutmann, D. H., Geist, R. T., Xu, H., Kim, J. S. & Saporito-Irwin, S. (1998). *Hum. Mol. Genet.* **7**, 335–345.

Gutmann, D. H., Sherman, L., Seftor, L., Haipek, C., Hoang Lu, K. & Hendrix, M. (1999). *Hum. Mol. Genet.* **8**, 267–275.

Hamada, K., Shimizu, T., Matsui, T., Tsukita, S. & Hakoshima, T. (2000). *EMBO J.* **19**, 4449–4462.

Han, B. G., Nunomura, W., Takakuwa, Y., Mohandas, N. & Jap, B. K. (2000). *Nature Struct. Biol.* **7**, 871–875.

Herrlich, P., Morrison, H., Sleeman, J., Orian-Rousseau, V., Konig, H., Weg-Remers, S. & Ponta, H. (2000). *Ann. NY Acad. Sci.* **910**, 106–120.

Jacoby, L. B., MacCollin, M., Parry, D. M., Kluwe, L., Lynch, J., Jones, D. & Gusella, J. F. (1999). *Neurogenetics*, **2**, 101–108.

Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. (1991). *Acta Cryst.* A**47**, 110–119.

Kleywegt, G. J. (1996a). *Jnt CCP4/ESF–EACBM Newsl. Protein Crystallogr.* **32**, 32–36.

Kleywegt, G. J. (1996b). *Acta Cryst.* D**52**, 842–857.

Kluwe, L., MacCollin, M., Tatagiba, M., Thomas, S., Hazim, W., Haase, W. & Mautner, V. F. (1998). *Am. J. Med. Genet.* **77**, 228–233.

Kluwe, L., Mautner, V., Parry, D. M., Jacoby, L. B., Baser, M., Gusella, J., Davis, K., Stavrou, D. & MacCollin, M. (2000a). *Neurogenetics*, **3**, 17–24.

Kraulis, P. J. (1991). *J. Appl. Cryst.* **24**, 946–950.

LaJeunesse, D. R., McCartney, B. M. & Fehon, R. G. (1998). *J. Cell Biol.* **141**, 1589–1599.

Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). *J. Appl. Cryst.* **26**, 283–291.

McClatchey, A. I., Saotome, I., Mercer, K., Crowley, D., Gusella, J. F., Bronson, R. T. & Jacks, T. (1998). *Genes. Dev.* **12**, 1121–1133.

MacCollin, M., Ramesh, V., Jacoby, L. B., Louis, D. N., Rubio, M. P., Pulaski, K., Trofatter, J. A., Eldridge, R., Parry, D., Bove, C., Short, M. P. & Gusella, J. (1994). *Am. J. Hum. Genet.* **55**, 314–320.

Maeda, M., Matsui, T., Imamura, M. & Tsukita, S. (1999). *Oncogene*, **18**, 4788–4797.

Mangeat, P., Roy, C. & Martin, M. (1999). *Trends Cell Biol.* **9**, 187–192.

Martin, M., Roy, C., Montcourrier, P., Sahuquet, A. & Mangeat, P. (1997). *Mol. Biol. Cell*, **8**, 1543–1557.

Martuza, R. L. & Eldridge, R. (1988). *N. Engl. J. Med.* **318**, 684–688.

Meng, J. J., Lowrie, D. J., Sun, H., Dorsey, E., Pelton, P. D., Bashour, A. M., Groden, J., Ratner, N. & Ip, W. (2000). *J. Neurosci. Res.* **62**, 491–502.

Mérel, P., Hoang-Xuan, K., Sanson, M., Moreau-Aubry, A., Bijlsma, E. K., Lazaro, C., Moisan, J. P., Resche, F., Nishisho, I., Estivill, X., Delattre, J. Y., Poisson, M., Theillet, C., Hulsebos, T., Delattre, O. & Thomas, G. (1995). *Genes Chromosomes Cancer*, **13**, 211–216.

Merritt, E. A. & Bacon, D. J. (1997). *Methods Enzymol.* **277**, 505–524.

Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). *Acta Cryst.* D**53**, 240–255.

Murthy, A., Gonzalez-Agosti, C., Cordero, E., Pinney, D., Candia, C., Solomon, F., Gusella, J. & Ramesh, V. (1998). *J. Biol. Chem.* **273**, 1273–1276.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Nguyen, R., Reczek, D. & Bretscher, A. (2001). *J. Biol. Chem.* **276**, 7621–7629.

Nicholls, A., Sharp, K. & Honig, B. (1991). *Proteins*, **11**, 281–296.

Obremski, V. J., Hall, A. M. & Fernandez-Valle, C. (1998). *J. Neurobiol.* **37**, 487–501.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Parry, D. M., MacCollin, M. M., Kaiser-Kupfer, M. I., Pulaski, K., Nicholson, H. S., Bolesta, M., Eldridge, R. & Gusella, J. F. (1996). *Am. J. Hum. Genet.* **59**, 529–539.

Pearson, M. A., Reczek, D., Bretscher, A. & Karplus, P. A. (2000). *Cell*, **101**, 259–270.

Perrakis, A., Morris, R. & Lamzin, V. S. (1999). *Nature Struct. Biol.* **6**, 458–463.

Rouleau, G. A., Merel, P., Lutchman, M., Sanson, M., Zucman, J., Marineau, C., Hoang-Xuan, K., Demczuk, S., Desmaze, C., Plougastel, B., Pulst, S. M., Lenoir, G., Bijsma, E., Fashold, R., Dumanski, J., de Jong, P., Parry, D., Eldridge, R., Aurias, A., Delattre, O. & Thomas, G. (1993). *Nature (London)*, **363**, 515–521.

Ruttledge, M. H., Andermann, A. A., Phelan, C. M., Claudio, J. O., Han, F. Y., Chretien, N., Rangaratnam, S., MacCollin, M., Short, P., Parry, D., Michels, V., Riccardi, V. M., Weksberg, R., Kitamura, K., Bradburn, J. M., Hall, B. D., Propping, P. & Rouleau, G. A. (1996). *Am. J. Hum. Genet.* **59**, 331–342.

Sainz, J., Huynh, D. P., Figueroa, K., Ragge, N. K., Baser, M. E. & Pulst, S. M. (1994). *Hum. Mol. Genet.* **3**, 885–891.

Scoles, D. R., Baser, M. E. & Pulst, S. M. (1996). *Neurology*, **47**, 544–546.

Scoles, D. R., Huynh, D. P., Chen, M. S., Burke, S. P., Gutmann, D. H. & Pulst, S. M. (2000). *Hum. Mol. Genet.* **9**, 1567–1574.

Sherman, L., Xu, H. M., Geist, R. T., Saporito-Irwin, S., Howells, N., Ponta, H., Herrlich, P. & Gutmann, D. H. (1997). *Oncogene*, **15**, 2505–2509.

Trofatter, J. A., MacCollin, M. M., Rutter, J. L., Murrell, J. R., Duyao, M. P., Parry, D. M., Eldridge, R., Kley, N., Menon, A. G., Pulaski, K., Haase, V., Ambrose, C., Munroe, D., Bove, C., Haines, J., Martuza, R., McDonald, M., Seizinger, N., Short, M. P., Buckler, A. & Guzella, J. (1993). *Cell*, **72**, 791–800.

Tsukita, S., Oishi, K., Sato, N., Sagara, J. & Kawai, A. (1994). *J. Cell. Biol.* **126**, 391–401.

Tsukita, S. & Yonemura, S. (1997). *Curr. Opin. Cell. Biol.* **9**, 70–75.

Tsukita, S., Yonemura, S. & Tsukita, S. (1997). *Trends Biochem. Sci.* **22**, 53–58.

Turunen, O., Sainio, M., Jaaskelainen, J., Carpen, O. & Vaheri, A. (1998). *Biochim. Biophys. Acta*, **1387**, 1–16.

Welling, D. B., Guida, M., Goll, F., Pearl, D. K., Glasscock, M. E., Pappas, D. G., Linthicum, F. H., Rogers, D. & Prior, T. W. (1996). *Hum. Genet.* **98**, 189–193.

Xu, H. M. & Gutmann, D. H. (1998). *J. Neurosci. Res.* **51**, 403–415.