Structural basis for neurofibromatosis type 2: crystal structure of the merlin FERM domain

Toshiyuki Shimizu‡, Azusa Seto‡, Nobuo Maita‡, Keisuke Hamada‡, Shoichiro Tsukita¶, Sachiko Tsukita¶||, and Toshio Hakoshima‡§  ‡‡

From the ‡‡Structural Biology Laboratory, Nara Institute of Science and Technology and §CREST, Japan Science and Technology Corporation, 8916-5 Takayama, Ikoma, Nara 630-0101, Japan, and the ¶ Department of Cell Biology, Faculty of Medicine, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan, and ||College of Medical Technology, Kyoto University, Sakyo-ku, Kyoto 606-8507, Japan.

‡‡ To whom correspondence should be addressed (e-mail: hakosima@bs.aist-nara.ac.jp)

Toshio Hakoshima
Phone: +81-0743-72-5570
Fax: +81-0743-72-5579
E-mail: hakosima@bs.aist-nara.ac.jp

Running title: crystal structure of Merlin
SUMMARY
Neurofibromatosis type 2 (NF2) is a dominantly inherited disease associated with the central nervous system. The NF2 gene product merlin is a tumor suppressor and its mutation or inactivation causes this disease. We report here the crystal structure of the merlin FERM domain containing a 22-residue α-helical segment. The structure reveals that the merlin FERM domain consists of three subdomains displaying notable features of the electrostatic surface potentials, although the overall surface potentials similar to those of ezrin/radixin/moesin (ERM) proteins indicate its electrostatic membrane association. The structure also suggests the inactivation mechanisms caused by the pathogenic mutations associated with NF2.

INTRODUCTION
Neurofibromatosis 2 (NF2) is an autosomal dominant disease that causes a predisposition to the development of nervous system tumors, mainly vestibular schwannomas involving the eighth pair of cranial nerves and schwannomas of other locations, meningiomas and ependymomas. Mutations in the NF2 gene have been identified in individuals affected with NF2, suggesting that the NF2 gene might encode a tumor suppressor (1-3). Extensive studies of this gene for the presence of inactivating mutations in NF2 patients has revealed different mutation types including nonsense, frameshift, missense, insertion, deletion and splicing-site mutations. The NF2 protein, called merlin or schwannomin, has two predominant isoforms which lack residues corresponding to exon 16 (isoform I) or exon 17 (isoform II), leading to variant C-terminal ends of the proteins. Merlin (2, 3) shows significant sequence similarity to members of the membrane-organizing proteins band 4.1 superfamily known as ezrin, radixin and moesin (ERM proteins), which link plasma membranes and actin filaments and are involved in the formation of microvilli, cell adhesion sites, ruffling membranes and cleavage furrows (4-6). Merlin contains a conserved FERM domain (7) (protein 4.1 (“four point one”), ERM) followed by a helical domain and a hydrophilic tail (Figure 1). Biochemical and cell biological data have suggested that merlin may share some molecular functions with ERM proteins (6, 8-9), that is, it is suggested that merlin may function as part of a signal transduction pathway regulating cell-cell and cell-matrix interactions, although exact mechanism of tumor suppressor activity of merlin remains unknown. Interestingly, most mutations found in NF2 patients fall into the FERM domain (Figure 2), which has been known to directly bind plasma membranes and adhesion proteins.

Recently, crystal structures of ERM proteins have been determined. These are the dormant complex of the moesin FERM domain with the C-terminal tail domain (10), and the activated form of radixin with and without inositol 1,4,5-triphosphate (IP3) (11). These structures provide the structural basis of masking/unmasking and membrane-targeting mechanisms of the ERM proteins. Moreover, the protein 4.1 has been shown to display the FERM domain structure similar to those of ERM proteins (12). However, it is also important to elucidate the effects of the various pathogenic mutations in the NF2 gene in terms of the protein structure, as well as to gain insights into the structural characterization of merlin. Here, we report the three-dimensional structure of merlin (residues 1-346), which contains the FERM domain (1-311) plus a
part of the helical region, as revealed by X-ray crystallography.

**EXPERIMENTAL PROCEDURES**

The FERM domain (1-346 residues) of mouse merlin was expressed as a fusion protein with glutathione S-transferase (GST). The protein was purified using a GST affinity column with glutathione-Sepharose 4B (Amersham Pharmacia Biotech) and ion-exchange columns HiTrapQ and SP (Amersham Pharmacia Biotech). The purified protein was verified with matrix-assisted laser desorption/ionization time-of-flight mass spectroscopy (JMS Elite, PerSeptive Inc.) and N-terminal analysis (M492, Applied Biosystems). Crystals were obtained at 4°C by the hanging drop vapor-diffusion method from solutions containing 13.6 mg/ml protein, 10 mM Mes, pH 6.8, 7.5% polyethylene glycol 6000, 150 mM NaCl, 250 mM LiCl, and 0.5 mM dithiothreitol (DTT) equilibrated against 50 mM Hapes, pH 7.0, 15% polyethylene glycol 6000 (PEG 6K), and 500 mM LiCl. All diffraction data were collected from crystals flash-frozen to 100 K after cryoprotection in a solution containing 18% PEG 6K, 25% PEG 400, and 500 mM LiCl. The crystals belong to space group \( P4_32_12 \), which was determined by the following structure analysis with unit cell parameters \( a = b = 67.4 \ \text{Å}, \ c = 179.1 \ \text{Å} \). A 2.9-Å data set was collected using a charge-coupled area detector (Mar CCD) installed on the beam line BL41XU at the Spring-8 synchrotron facility, Harima, Japan. The wavelength was set to 0.9198 Å with a crystal-to-detector distance of 150 mm. The data were collected with angular ranges of 180° with step sizes of 1° for an exposure time of 4 sec. Intensity data were processed using the programs DPS-Mosfilm (13).

The initial phases were calculated by molecular replacement using a search model based on the free state structure of the radixin FERM domain (11). Unambiguous rotation and translation solutions were obtained from several searches using different ranges of intensity data and integration radii with the program AMoRe (14). The resultant initial map shows clear electron densities for most of the FERM domain and \( \alpha \)-helical segment. An initial model of this segment was built into the electron density map using the graphics program O (15). The model was refined by the method of simulated annealing using the program CNS (16). After several cycles of rebuilding and refinement, the model was refined to an \( R \) value of 22.9% (free \( R \) value of 28.4%) for intensity data at a 2.9-Å resolution. The current model includes 323 residues of the protein, and no water molecules are included. The data collection and refinement statistics are shown in Table 1. Estimated coordinated error from the Luzzatti plot (17) and density correlation coefficient (16) is 0.40 Å and 0.937, respectively. No model was built for the 17 N-terminal and 6 C-terminal residues of the peptide, which are poorly defined in the current map. As defined in PROCHECK (18), all main-chain dihedral angles fall within allowed regions of the Ramachandran plot except those for Asp 268, which was modeled with unfavorable angles as observed in moesin (10) and radixin (11). In the peptide model, the side chains of 19 residues were poorly defined, and in the current structure they are replaced with alanines. Ribbon representations of the main-chain folding of the molecule were drawn using the program Molscript (19) and Raster3D (20), while molecular surface representations were drawn using the program GRASP (21). Coordinates have been deposited in the
Protein Data Bank of the Research Collaboratory for Structural Bioinformatics (PDB id: 1ISN).

RESULTS

Overall structure of the merlin FERM domain

The fold of the merlin FERM domain consists of three subdomains A, B, and C, each showing similarity to known structures: ubiquitin, acyl-CoA-binding protein, and the PH (plechstrin homology) / PTB (phosphotyrosine binding) / EVH1 (enabled, VASP, homology 1) domain, respectively (Figure 3A). As expected from the relatively high sequence homology with ERM proteins (~64% sequence identity), the overall structure is similar to those of the FERM domains that have been reported previously. For simplicity, we describe here a quantitative structural comparison with the IP3-bound form of the radixin FERM domain, which shows a small root-mean-square (rms) deviation (0.3 Å) with the free form. The rms deviation in Cα-carbon atoms obtained from the pair-wise superposition of each subdomain in the merlin and the IP3-bound radixin FERM domains were found to be 1.87 Å, 0.59 Å, and 0.94 Å for subdomains A, B, and C, respectively. Subdomain A shows an exceptionally large deviation due to the different conformations in a segment containing strands β3A and β4A, in which the connecting loop has a one-residue deletion between Thr 67 and Ile 68 (Figure 3B). When this segment and the flexible N-terminal region are excluded, the rms deviation between the subdomains A becomes reasonably small (0.78 Å) and the overall rms deviation between the two FERM domains is 1.00 Å, which is a relatively small deviation in each subdomain as well. The large conformational change allows the loop with the deletion site to interact with helix α1C in the subdomain C more tightly than in ERM proteins. It is possible that the fixation of helix α1C by the loop may weaken the masking effect of merlin, since the displacement of helix α1C and subdomain C controls the masking of the FERM domain with C-terminal tail domain (11). In fact, the masking effect of merlin is weaker than that of the ERM proteins (22). Merlin has an additional 17-residue segment at the N-terminus, compared with ERM proteins. These 17 residues are disordered in the current model. The structural feature in the current model is consistent with the proteinase digestion experiment (23), which has shown a high sensitivity of the merlin N-terminal residues (1-18) but a resistance of the merlin FERM domain (19-314) against α-chymotrypsin digestion.

C-terminal extended α-helical segment

The current merlin (1-346 residues) contains an additional 33-residue segment linked to the C-terminus of the FERM domain. These residues form a loop followed by a C-terminal extended helix of 22 residues (Figure 3A). We suspect that this extended helix may correspond to part of the helical domain, which consists of 160 residues. Interestingly, a recent study using low-angle shadowing electron microscopy has suggested that the open form of the full-length radixin contains a filamentous structure of 200-250 Å between two globular structures, which would correspond to the N-terminal FERM and the C-terminal tail domains (24). This filamentous structure is highly consistent with 160 residues of the helical domain. Moreover, a 47-residue extended helix linked to the C-terminal of the FERM domain has been observed in the
recent crystal structure of moesin (residues 1-346), which contains the FERM domain and part of the helical domain (25). As reported in this moesin structure, the extended $\alpha$-helical segment of merlin seems to be stabilized by the hydrogen bond formation of Arg 57 from subdomain A with Glu 317 and Gln 324. In addition, the helix makes contacts with the FERM domain through the hydrogen bond formation among Glu 107, Asp 314, and Lys 322, and hydrophobic contacts between hydrophobic residues from the $\alpha$-helical segment (Val 318 and Met 321), subdomain A (Trp 60), and linker A-B (Val 110). It should be noted that these residues are highly conserved or substituted with a homologous residue in ERM proteins and merlin (Figure 2).

Blue box
The merlin FERM domain contains a seven amino acid stretch (residues 177-183; YQMTPEM), referred to as the Blue box or BB, that is identical in human and Drosophila but divergent in other ERM proteins (26) (Figure 2). The Blue box mutants produce a dominant negative phenotype in Drosophila. This Blue box stretch, which is important for merlin function, is located around the N-terminal region of helix $\alpha$3B in subdomain B and displays a conformation similar to that of radixin with a small rms deviation (0.84 Å) (Figure 3B). However, this region of merlin may be more flexible than that of radixin: the side chains of two merlin residues (Gln 178 and Met 183) are disordered. It is possible that physico-chemical, but not structural, properties may be important for the function of this region in merlin.

DISCUSSION
Pathogenic mutations in merlin
The severe disease NF2 tends to be associated with merlin mutations that lead to premature protein termination as a result of frameshift mutation, or the creation of a premature termination codon. These nonsense mutations, which are found in all three subdomains of the FERM domain, result in an unstable, truncated protein that fails to function properly (Figure 4A). Several missense mutations have also been reported to date, and these mutations produce full-length molecules that are defective as growth regulators. The missense mutations in the FERM domain are shown in Figure 2 in addition to deletion mutants.

The missense mutations that lead to non-conservative, single amino acid changes are likely to impair function in a number of ways. Most mutations result in a clear disruption of the hydrophobic core or secondary structure elements (F62S, $\Delta$96, $\Delta$118, L234R, I273K in Figure 4B). For instance, Phe 118 is located at helix $\alpha$1B to form the hydrophobic core surrounded by Phe 61, Phe 62, and Leu 109, which are highly conserved among ERM proteins. The deletion $\Delta$118 causes severe NF2, probably due to the disruption of the hydrophobic core, and it results in diffusion in the cytoplasmic and an increased tendency to detach from the plate surface (23, 27). Leu 141 is located at the middle of helix $\alpha$2B, and the L141P mutation would disrupt this helix.

The L64P and E106G mutations have been reported to impair protein folding, resulted in an unfolded conformation (28). In the crystal structure, Leu 75 from strand $\beta$4A is packed against Leu 64 located at strand $\beta$3A, indicating that the L64P mutation affects the $\beta$-sheet structure of subdomain A. Glu 106 located at the linker A-B ($\alpha$1B)
participates in an inter-subdomain interaction by forming a hydrogen bond with Lys 289 from subdomain C.

Some mutations occur in the C-terminal extended α-helical segment from the helical domain rather than the FERM domain. The Q324L mutation on the α-helical segment would disrupt a hydrogen bond of Glu 324 with Arg 57 from subdomain A. Since the side chain of Ser 315 makes a hydrogen bond with the main chain of Glu 317, the S315F mutant probably affects the hydrogen bonding interaction, resulting in the destabilization of the extended α-helical segment.

The binding site of IP3 has been defined in the radixin crystal structure (11) and also suggested by mutagenesis in ezrin (29). The crystal structure reveals that the IP3-binding site is located at the basic cleft between subdomains A and C (indicated by a red arrow in Figure 4B), where the IP3 molecule has been recognized by direct hydrogen bonding interactions with Lys 60, Asn 62, Lys63, and Lys 278 in radixin. Therefore, the K79E mutant (equivalent to Lys 63 in radixin) would have low affinity for IP3 and PIP2 (phosphatidylinositol 4,5-bisphosphate). In fact, the K79E missense mutant has been observed to be strongly localized in the peri-nuclear cytoplasm (27) instead of the membrane. The M77V mutation, which corresponds to residue 61 of radixin, may induce small but significant changes in the fine structure of the stretch containing the IP3-recognizing residues.

The V219M mutant gave rise to the same proteinase resistance as the full-length merlin (23). The space around Val 219 seems to accommodate Met, and this mutation causes no significant changes. It was reported that the N220Y mutant showed reduced binding to EBP50 (27, 30). In the current model, most of Asn 220 is buried and it is unlikely to bind to EBP50 directly. The N220Y mutant would make short contact with residues 99-101 of the linker region between subdomains A and B, and change subdomain orientation, resulting in decreased binding to EBP50.

There remain mutations (E38V, G197C, E270G, L316F, L316W, and L339F) that are difficult to explain. These residues are exposed to solvent regions, suggesting that the mutations may cause no significant perturbation on the molecular structure. A few of these mutations may be explained in terms of interactions with possible merlin-binding partners. Further experiments are needed.

Molecular surface properties
Several merlin-binding cellular proteins have been identified, and they have been designated p165, p145, p125, p85, and p70 (31). The proteins p125 and p85 are poly (ADP-ribose) polymerase and Ku-antigen p85, respectively. SCHIP-1 has also been reported as a merlin-binding protein (32). The deletion study has revealed that most of the FERM domain is necessary for p165 and p125 binding, and a long α-helical segment in addition to the FERM domain is required for p85 and p70 binding. Some of these proteins (p165, p145, p125) do not bind ERM proteins. Although merlin and ERM proteins share significant homology in the FERM domain, electrostatic potential surfaces of the merlin FERM domain show notable features different from those of radixin (Figure 5). One characteristic surface of merlin is located at subdomain A (indicated by a yellow arrow in Figure 5A) and is negatively charged with Glu 38, Asp
45, Asp 78, and Glu 90, while radixin has a positively charged surface. Another characteristic surface of merlin is located at subdomain B (indicated by a yellow arrow in Figure 5C) and is also negatively charged with Asp 128 and Glu 167, while radixin is positively charged. It is possible that merlin-binding partners make use of these properties, but further examination is required to determine the binding surface.

The radixin FERM domain has a positively charged flat surface formed by subdomains A and C. This flat surface contains a basic cleft (indicated by a red arrow in Figure 5B) for the IP$_3$ binding, indicating an electrostatic membrane association of the FERM domain. Like radixin, the corresponding surface of merlin is positively charged, suggesting a similar electrostatic membrane association and PIP$_2$ binding by the merlin FERM domain.

Sequence analysis predicts that merlin lacks a site corresponding to the C-terminal F-actin binding site, which exists in all ERM proteins (33). Instead, merlin may contact actin through a low-affinity binding site in the C-terminal end of its FERM domain (178-367) (34), or indirectly via the actin-binding protein βII spectrin/fodrin (35). Brault et al. (23) have suggested that the amino acids 1-27 and 280-323 of merlin are sufficient to bind F-actin and are likely to form independent actin-binding units. In the current structure, these two segments are apart from each other. N-terminal residues (1-17) are disordered, and residues 21-27 form a β-strand (β1A) of subdomain A. Residues 280-323 form helix α1C and part of the C-terminal extended α-helical segment.

**ACKNOWLEDGEMENTS**

We thank Drs. K. Okada and M. Kawamoto for their technical assistance in data collection at the SPring-8 Facility, Harima. We are grateful to S. Muraguchi for preparation of figures. This work was supported by a Grant-in-Aid for Scientific Research to T. H. (12490204) from JSPS and a Grant-in-Aid for Scientific Research on Priority Area to S. T. and T. H. (10179104) from the MECSST of Japan. K. H. was supported by a JSPS research fellowship for Young Scientists. T. H. is a member of the Structural Biology Sakabe project of the Foundation for Advancement of International Science.
REFERENCES


**Figure legends**

Figure 1. Diagram of merlin domains. The N-terminal FERM domain is divided into three subdomains (A-C). Helical domain and C-terminal domain are also indicated. Three subdomains in the FERM domain are defined by this crystallographic analysis and other two domains are indicated according to Ref. 6.

Figure 2. Sequence alignment of the merlin and ERM proteins. The secondary structure of merlin from the structure reported here is shown. Thick and thin black boxes show $\alpha$-helix and $\beta$-strand, respectively. The pathogenic mutations in merlin are also shown; they were obtained from the Massachusetts General Hospital Neuro-genetic Surgery web site (http://neurosurgery.mgh.harvard.edu/NFresrch.htm) and the Human Gene Mutation Database at the Institute of Medical Genetics in Cardiff (http://archive.uwcm.ac.uk/uwcm/mg/lgmd0.html) (36). $\Delta$ and # indicate the deletion and the nonsense mutations, respectively.

Figure 3. Overall structure and superposition of the FERM domains. (A) Ribbon representation of the merlin FERM domain. Each subdomain is colored green (A), blue green (B), and blue (C), respectively. The C-terminal $\alpha$-helical segment is gray, and the linker region is brown. (B) Superposition of the FERM domain in merlin (blue) and radixin (brown). The $\alpha$-helical segment is omitted for clarity. The Blue box (residues 177-183) of merlin is located at the N-terminal region of helix $\alpha$3B and is indicated with a label. An arrow indicates the one residue deletion site of merlin.

Figure 4. Mutations mapped on the tertiary structure of merlin. (A) Nonsense mutations that result in truncated proteins are shown in red. (B) Missense mutations discussed in the manuscript. Buried and exposed residues are colored red and blue, respectively. Residues in deletion mutations are colored green. Lys 79 involved in IP$_3$ recognition is colored violet. A red arrow indicates the IP$_3$-binding site found in radixin.

Figure 5. Electrostatic potentials surface. Positive (blue) and negative (red) potentials are mapped on the van der Waals surfaces in the range $-10 \, k_B T$ (red) to $+10 \, k_B T$ (blue), where $k_B$ is Boltzmann’s constant and $T$ is the absolute temperature. A (merlin) and B (radixin) are viewed in the same direction as Figure 3. C (merlin) and D (radixin) are viewed from the opposite side. Yellow arrows in A and C indicate the areas showing different properties in two molecules. A red arrow in B indicates the IP$_3$-binding site found in the IP$_3$-bound radixin.
Table 1

<table>
<thead>
<tr>
<th><strong>Intensity data processing</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (Å)</td>
<td>30 - 2.9</td>
</tr>
<tr>
<td>$R_{\text{merge}}$ (%)$^a$</td>
<td>6.9 (31.6)$^b$</td>
</tr>
<tr>
<td>Number of measurements</td>
<td>32050</td>
</tr>
<tr>
<td>Number of independent reflections</td>
<td>8807</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>90.5 (75.8)$^b$</td>
</tr>
<tr>
<td>Mean $&lt;I/\sigma(I)&gt;$</td>
<td>8.4 (2.4)$^b$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Refinement statistics</strong></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Resolution (Å)</td>
<td>20 - 2.9</td>
</tr>
<tr>
<td>$R_{\text{cryst}}$ (%)$^c$</td>
<td>22.9 (37.3)$^b$</td>
</tr>
<tr>
<td>$R_{\text{free}}$ (%)$^c$</td>
<td>28.4 (40.7)$^b$</td>
</tr>
<tr>
<td>Mean B value (Å²)</td>
<td>50.8</td>
</tr>
<tr>
<td>Rms bond length (Å)</td>
<td>0.010</td>
</tr>
<tr>
<td>Rms bond angles (Ú)</td>
<td>1.5</td>
</tr>
<tr>
<td>Rms dihedral angles (Ú)</td>
<td>23.4</td>
</tr>
<tr>
<td>Rms improper angles (Ú)</td>
<td>1.24</td>
</tr>
</tbody>
</table>

---

$^a R_{\text{merge}} = \frac{100 \times \sum |I(h)| - <I(h)> | \sum I(h)}{\sum I(h)}$, where $<I(h)>$ is the mean intensity for reflection $h$.  

$^b$ Brackets are quantities calculated in the highest resolution bin at 2.9-3.03 Å. 

$^c R_{\text{cryst}} = \frac{100 \times \sum |F_o(h)| - F_c(h) | \sum F_o(h)}{\sum F_o(h)}$, where $F_o(h)$ and $F_c(h)$ are observed and calculated structure factors, respectively. $R_{\text{free}}$ is $R_{\text{cryst}}$, which was calculated using 10% of the data chosen randomly, and omitted from the subsequent structure refinement.
Figure 2 (Shimizu et al.)
Figure 3 (Shimizu et al.)
Figure 4 (Shimizu et al.)
Figure 5 (Shimizu et al.)
Structural basis for neurofibromatosis type 2: Crystal structure of the merlin FERM domain
Toshiyuki Shimizu, Azusa Seto, Nobuo Maita, Keisuke Hamada, Shoichiro Tsukita, Sachiko Tsukita and Toshio Hakoshima

J. Biol. Chem. published online December 27, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M109979200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2001/12/27/jbc.M109979200.citation.full.html#ref-list-1