

Structural characterization suggests models for monomeric and dimeric forms of full-length ezrin.

Juanita M. Phang¹, Stephen J. Harrop^{1†}, Anthony P. Duff², Anna V. Sokolova², Ben Crossett³, James C. Walsh¹, Simone A. Beckham⁴, Cuong D. Nguyen¹, Roberta B. Davies^{1§}, Carina Glöckner¹, Elizabeth H. C. Bromley⁵, Krystyna E. Wilk¹, Paul M. G. Curmi^{1,6*}

From the ¹School of Physics, The University of New South Wales, Sydney, NSW 2052, Australia.

²Australian Nuclear Science and Technology Organisation, Lucas Heights, NSW 2234, Australia.

³Mass Spectrometry Core Facility, University of Sydney, Sydney, NSW 2006, Australia,

⁴Department of Biochemistry and Molecular Biology, Monash University, Clayton, VIC 3800, Australia.

⁵Department of Physics, University of Durham, Durham, United Kingdom.

⁶Centre for Applied Medical Research, St Vincent's Hospital, Sydney NSW 2010, Australia

[†]Present address: MX Beamlines, Australian Synchrotron, 800 Blackburn Road, Clayton, Victoria 3168, Australia.

[§]Present address: The Victor Chang Cardiac Research Institute, Darlinghurst, NSW 2010, Australia.

*To whom correspondence should be addressed: Paul M. G. Curmi, School of Physics, The University of New South Wales, Sydney, NSW 2052, Australia, Tel.: 61-2-93854552; Fax: 61-2-93856060; E-mail: p.curmi@unsw.edu.au

ABSTRACT

Ezrin is member of the ERM (Ezrin-Radixin-Moesin) family of proteins that have been conserved through metazoan evolution. These proteins have dormant and active forms, where the latter links the actin cytoskeleton to membranes. ERM proteins have three domains: an N-terminal FERM (band Four-point-one ERM) domain comprising three subdomains (F1, F2 and F3); a helical domain; and a C-terminal actin-binding domain. In the dormant form, FERM and C-terminal domains form a stable complex. We have determined crystal structures of the active FERM domain and the dormant FERM:C-terminal domain complex of human ezrin. We observe a bistable array of phenylalanine residues in the core of subdomain F3 that is mobile in the active form and locked in the dormant form. As subdomain F3 is pivotal in binding membrane proteins and phospholipids, these transitions may facilitate activation and signaling. Full-length ezrin forms stable monomers and dimers. We used small-angle x-ray scattering to determine the solution structures of these species. As expected, the monomer shows a globular domain with a protruding helical coiled-coil. The dimer shows an elongated dumbbell structure that is twice as long as the monomer. By aligning ERM sequences spanning metazoan evolution, we show that the central helical region is conserved, preserving the heptad repeat. Using this, we have built a dimer model where each monomer forms half of an elongated anti-parallel coiled-coil with domain-swapped FERM:C-terminal domain complexes at each end. The model suggests that ERM dimers may bind to actin in a parallel fashion.

SUMMARY STATEMENT

Crystal structures of ezrin FERM and FERM:C-terminal domain complexes plus SAXS models for full-length ezrin monomer and dimer were elucidated. Crystal structures show changes in the protein core while SAXS shows that the ezrin dimer is an elongated dumbbell.

SHORT TITLE

Ezrin Structure via X-ray Crystallography and SAXS

KEYWORDS

Ezrin, x-ray-crystallography, small-angle x-ray scattering (SAXS), protein structure, FERM domain, protein dynamics

ABBREVIATION LIST

ERM, ezrin-radixin-moesin; EBP50, ERM-binding-phosphoprotein 50; FERM, Four-point-one (4.1)-ezrin-radixin-moesin; IP₃, inositol-(1,4,5)-triphosphate; ICAM, intercellular adhesion molecule; IPTG, isopropyl β-D-thiogalactopyranoside; LB, lysogeny broth media; NHERF, Na⁺/H⁺ exchanger regulatory factor; NEP, neutral endopeptidase; PEI, polyethyleneimine; PI(4,5)P₂, phosphatidylinositol 4,5-biphosphate; PSGL, P-selectin glycoprotein ligand; RMSD, root mean square deviation; SEC, size exclusion chromatography; SEC-MALS, SEC coupled with in line multi angle light scattering; SEC-SAXS, SEC coupled with small angle x-ray scattering; TEV, tobacco etch virus; MS, mass spectrometry.

INTRODUCTION

The ERM proteins: ezrin, radixin and moesin connect cell membranes and the underlying actin cortex [1, 2]. These three largely metazoan proteins are highly conserved paralogues (human ERM proteins have overall 73-81% sequence identity). ERM proteins are dynamically regulated by the Rho family of small GTPases, thus, the membrane-cell cortex interaction can be switched on and off locally. ERM proteins organize membrane domains by binding to the cytoplasmic tails of integral membrane proteins [1, 2], for example in the immune synapse and in vesicular trafficking, processing and phagocytosis. The ERM protein ezrin is directly involved in tumor metastasis, particularly in the pediatric cancers: rhabdomyosarcoma and osteosarcoma [3].

Each ERM protein contains three distinct domains [4]. The first is the highly conserved ~300-residue FERM domain (band **F**our-point-one, **E**zrin, **R**adixin, **M**oesin) at the N-terminus, which has the highest (~86%) sequence identity among the three human ERM proteins. The second is the intermediate or α -helical domain, which includes a proline-rich linker, where the latter is absent in moesin. The third domain is the C-terminal domain (sometimes referred to as the C-ERMAD) that shares ~71% sequence identity among human ERM proteins. The C-terminal domain contains a conserved phosphorylation site associated with activation (Thr567 in human ezrin) and a highly conserved F-actin binding sequence in the last 30 residues at its C-terminus [5].

The ERM proteins have at least two physiologically relevant states: the dormant or autoinhibited state and the active state [6]. In the dormant state, the molecules are biologically inert, as the functional binding sites on both the FERM domain and the C-terminal domain are masked by intramolecular interactions between these domains. In the active state, these interactions are disrupted, inducing the separation of the FERM and C-terminal domains. Subsequently, the FERM domain binds to the plasma membrane, while the C-terminal domain binds to F-actin. In the active state, the FERM domain is able to bind directly or indirectly (for example, through an adaptor protein EBP50/NHERF) to the cytoplasmic extensions of membrane proteins. It has been suggested that the last two residues of ezrin C-terminus are critical for interaction with both the FERM domain and the F-actin [6].

Studies have shown that ERM proteins are phosphorylated by Rho-kinase, protein kinase C θ and protein kinase C α . It has been proposed that the activation of ERM proteins *in vivo* occurs via a two-step mechanism [7], involving ERM proteins binding to phosphatidylinositol 4,5-biphosphate (PI(4,5)P₂) and phosphorylation of the conserved threonine residue in the C-terminal domain. It has been shown that the FERM domain of ezrin interacts with liposomes containing PI(4,5)P₂ [8]. Subsequent studies have shown that both phosphorylation and PI(4,5)P₂ binding regulate ERM activation and thus binding to F-actin [9, 10].

The first crystal structure of an ERM protein is that of moesin [11], comprising a complex between the FERM and the C-terminal domains. Subsequent crystal structures include: FERM domains of ezrin [12], radixin [13] and moesin [14]; a complex between radixin and IP₃, a mimetic of the head group of PI(4,5)P₂ [13]; several complexes between the radixin FERM domain and peptides mimicking the cytoplasmic tails of membrane proteins [15-19]; and complexes of radixin with the adaptor proteins NHERF-1 and NHERF-2 [20]. The only full-length crystal structure of an ERM protein is that of the monomeric

dormant state of *Sfmoesin* from the Fall armyworm (moth) *Spodoptera frugiperda* [21].

Ezrin exists in various states within cells including monomers, dimers, oligomers and heterodimers with other ERM proteins. The biological significance and function of these species is unknown. Gel filtration of Triton-X-100 solubilized microvilli extracts from human JEG-3 carcinoma cells show that ezrin is present as both a monomer and a larger homomeric species (possibly dimer or trimer) [22]. Two cytosolic species of ezrin have been purified from human placenta and have been characterized as a monomer and dimer that do not interconvert after isolation [23]. Solubilization of microvilli from human placenta show that ezrin occurs mainly as dimers and higher order oligomers [24]. These authors also show that stimulation of human A431 carcinoma cells with EGF induces the rapid formation of ezrin oligomers *in vivo*, which is correlated to tyrosine phosphorylation of ezrin [24]. In the LLC-PK1 kidney epithelial cell line, endogenous ezrin is partitioned into cytosolic and membrane-associated fractions where homo-oligomers (mainly dimers) make up 15 and 28%, respectively, of the total ezrin in each fraction [25]. In contrast, they find that endogenous (threonine) phosphorylated ERM proteins are largely monomeric. Expression of wild type and Thr567 mutant ezrins in these cell lines supports the observation that activation of ezrin by phosphorylating Thr567 produces mainly monomeric ezrin [25]. Micro-injection of bacterially expressed human ezrin into live gastric HGT-1 cells shows that it co-immunoprecipitates with endogenous ezrin, radixin and moesin, presumably in the form of homo and heterodimers [26]. Stable complexes between ezrin and moesin, presumed to be heterodimers, have been immunoprecipitated from human epidermoid carcinoma A431 cells [27]. Similarly, the ERM paralogue, merlin has been co-immunoprecipitated with ezrin from human U251 glioma cells, presumably as a heterodimer [28]. The above observations show that ezrin dimers exist in mammalian cells, however, the function of the dimer is currently unknown.

In this paper, we present the crystal structures of the active FERM domain and the inactive FERM:C-terminal domain complex of human ezrin. Comparison of these two high resolution structures shows changes in structure and mobility that may be important in ezrin activation. In particular, the hydrophobic core of subdomain F3 in the active FERM domain structure shows a bistable structure where an array of conserved phenylalanine residues can adopt two distinct arrangements. The formation of the dormant FERM:C-terminal domain complex stabilizes this structure by selecting one arrangement for the core phenylalanine residues. We also report solution SAXS structures of full-length ezrin monomer and dimer. The monomer structure is consistent with the full-length *Sfmoesin* crystal structure where the central α -helical domain protrudes from the FERM domain as an elongated α -helical coiled-coil. The SAXS analysis shows that the dimer is an extremely elongated dumbbell structure. This structure is modeled by domain-swapping, where the central α -helical domains from each protomer form a continuous anti-parallel coiled-coil linking the two domain-swapped FERM:C-terminal domain complexes. The domain-swapped nature of the dimer explains its stability and slow equilibration between monomer and dimer. Sequence analysis shows that the heptad repeat for the entire region of the central α -helical domain spanned by α -helices α B and α C including the linker (helices are labeled as per the *Sfmoesin* crystal structure) is conserved throughout metazoan evolution. Thus, evolution has preserved the ability of ERM proteins to form the continuous, elongated coiled-coil observed in the dimer model.

MATERIALS AND METHODS

Cloning

A plasmid encoding the FERM domain of human ezrin (residues 1-296, untagged, 35 kDa) was a gift from Mark Berryman (Ohio University, USA). Full-length human ezrin (residues 1-586, 69.5 kDa) was purchased from GeneArt AG, Germany, and subcloned into pET-15b vector (Novagen) for expression with an N-terminal His₈-tag that can be removed by TEV cleavage.

Protein Expression

The FERM domain of ezrin was expressed in *Escherichia coli* strain M15(pREP4) (Qiagen) grown in LB media containing 100 µg/ml ampicillin and 25 µg/ml kanamycin at 37 °C. Expression was induced at mid-log growth phase with 0.1 mM IPTG at 25 °C for 9-12 hours.

Full-length ezrin was expressed in *E. coli* BL21(DE3) (Invitrogen) grown in LB media containing 100 µg/ml ampicillin at 37 °C. Expression was induced at mid-log growth phase with 0.5 mM IPTG at 37 °C for three hours, producing a mixture of ezrin monomer and dimer.

Protein purification

Ezrin FERM domain: The FERM domain was purified using a modification of a previously published method [29]. The cell pellet was resuspended in lysis buffer (180 mM KH₂PO₄ pH 7.0) containing protease inhibitors (Complete, EDTA-free protease inhibitor, Roche). Following lysis, polyethyleneimine (PEI) was added to 0.15% (v/v) after lysis to precipitate nucleic acids. The lysate was centrifuged at 30,000 g and the supernatant applied to a pre-equilibrated (with lysis buffer) hydroxyapatite column at 4 °C. The protein was eluted with a gradient of 180 mM-900 mM KH₂PO₄ pH 7.0. Residual nucleic acid was removed using 0.15% (v/v) PEI and the protein was then subjected to size-exclusion chromatography (SEC) on a Superdex75 column (GE Healthcare) at 4 °C in 20 mM Tris pH 8.0, 300 mM NaCl. After SEC, fractions containing the FERM domain were concentrated to ~12 mg/ml and frozen using liquid nitrogen.

Full-length ezrin: The cell pellet was resuspended in 20 mM Tris pH 8.0, 500 mM NaCl, 10 mM imidazole, 0.5 mM DTT containing protease inhibitors. After lysis, PEI was added to 0.15% (v/v), and the lysate centrifuged at 30,000 g. Full-length ezrin was purified at room temperature. The clarified lysate was incubated with Ni-NTA resin (Qiagen) for 1 hour. The resin was washed with 20 mM Tris pH 8.0, 500 mM NaCl, 20 mM imidazole and 0.5 mM DTT, then protein was eluted with 20 mM Tris pH 8.0, 500 mM NaCl, 500 mM imidazole, 0.5 mM DTT. Eluted protein was dialyzed against 20 mM Tris pH 8.0, 500 mM NaCl, 10 mM imidazole, 1 mM DTT with a single change of dialysis solution. The His₈ tag was removed by incubation with His-tagged TEV protease for 16 hours. The incubated material was then passed through Ni-NTA resin. The tag-free ezrin was collected in the flow through and dialyzed against 20 mM Tris pH 8.5, 50 mM NaCl. Ezrin was applied to a Q-Sepharose column (GE Healthcare) and eluted with 20 mM Tris pH 8.5, 145-290 mM NaCl gradient. The protein was further purified using a Superdex200 26/60 SEC column (GE Healthcare) in 20 mM HEPES pH 7.0, 100 mM KCl, 1 mM DTT. This resulted in baseline separation of monomeric and dimeric ezrin (Fig. S1A). Purified ezrin was concentrated and frozen in liquid nitrogen.

Multi-Angle Light Scattering

Molecular masses were estimated by SEC coupled with in-line multi-angle light scattering (SEC-MALS). SEC was via a Superose 12 10/300 GL column (GE Healthcare) in 20 mM HEPES pH 7.0, 100 mM KCl, 1 mM DTT. The eluate was monitored by a miniDAWN TREOS light scattering instrument, consisting of a laser source at 690 nm and three discrete photodetectors (45°, 90° and 135°) coupled to an Optilab DSP differential refractometer (Wyatt Technology Corporation Pty Ltd). Data were collected and analyzed using ASTRA software, and molecular weight was calculated using a refractive index increment (dn/dc) value for protein of 0.19 mL/g.

Circular Dichroism

Circular dichroism (CD) experiments were performed using Jasco J-810 CD Spectropolarimeter fitted with Peltier temperature controller. Protein was diluted in 20 mM phosphate buffer pH 7.0 to a concentration of ~0.2 mg/ml. Spectra were collected in the far UV (190-260 nm) with 50 nm/minute scanning rate. The scans were repeated three times then averaged. CD responses were converted to mean residue ellipticity $[\theta]$ (deg cm² dmole⁻¹).

Thermal denaturation experiments were performed at 208 nm with a heating rate of 1 °C/minute between 20-80 °C. The mean residue ellipticity $[\theta]$ showed a linear dependence on temperature for both native/folded (N) and unfolded (U) states. Thus, the data were fit to two linear functions of temperature for N and U states with a sigmoidal function representing the unfolding transition:

$$[q] = \left(\frac{1}{1 + e^{-(T - T_{AM})/b}} \right) \cdot (Dm \cdot T + DC) + m \cdot T + C$$

where T_{AM} is the apparent melting temperature and b determines the steepness of the melting transition. The linear dependence of $[\theta]$ for N state is described by the slope m and the intercept C , while the linear dependence of $[\theta]$ for U state is described by the slope $m + \Delta m$ and the intercept $C + \Delta C$. The thermal unfolding of all proteins tested was irreversible.

Crystallization and x-ray diffraction data collection

Crystals of the ezrin FERM domain were grown at room temperature using hanging drop, vapor-diffusion by mixing 2 μ l protein (at ~12 mg/ml) and 2 μ l of reservoir solution containing 0.2 M ammonium sulphate, 0.1 M HEPES pH 7.5, 10% isopropanol, 16-18% (w/v) PEG4000. Crystals were transferred to a cryoprotectant solution consisting of reservoir solution supplemented with 15% PEG400 and flash-frozen in liquid nitrogen.

Crystals of the ezrin FERM:C-terminal domain complex resulted from attempts to crystallize monomeric and dimeric forms of full-length ezrin at room temperature using the hanging drop, vapor-diffusion. Crystals grew after 7-12 days in drops containing 2:1 ratio of protein (dimer: ~10-12 mg/ml, monomer: ~15-20 mg/ml) to reservoir solution consisting of 0.1 M Bis-Tris pH 5.5 or pH 5.9, 15-18% (w/v) PEG3350. Crystals were transferred to a cryoprotectant solution consisting of reservoir solution supplemented with either 15% PEG400 or 15% mixture of 1:1 glycerol to 50% (w/v) PEG3350. Crystals were flash-frozen in liquid nitrogen.

X-ray diffraction datasets for the ezrin FERM domain were collected at 100 K on an ADSC Quantum 210r detector installed at the Australian Synchrotron MX1 beamline. Data sets for crystals grown from full-length ezrin were collected at 100 K using an ADSC

Quantum 315r detector at the Australian Synchrotron MX2 beamline. Data collection was carried out using Blu-Ice [30]. Datasets were processed with MOSFLM [31] and scaled with SCALA [32].

Crystal structure determination and refinement

Crystal structures of both the ezrin FERM domain and the ezrin FERM:C-terminal domain complex were solved by molecular replacement using Phaser [33]. The N-terminal domain of ezrin (PDB ID: 1NI2 [12]) was used as a search model for the ezrin FERM domain while the *Sfmoesin* structure (PDB ID: 2I1J [21]) was used for crystals obtained from full-length ezrin. Model rebuilding was carried out using COOT [34]. The overall structures were refined using PHENIX [35]. Water molecules were initially built using PHENIX and then added manually during later stages of refinement. Statistics for data reduction and refinement are listed in Table 1.

The final model of the ezrin FERM domain contains two molecules in the asymmetric unit (molecule A – residues 2-296 and molecule B residues 2-137, 145-146, 150-296) with C_{α} RMSD of 0.39 Å. The main differences between the two molecules in the asymmetric unit are located at the flexible loop connecting α -helices $\alpha 2$ and $\alpha 2'$ in subdomain F2 (Fig. 1A). In molecule B, residues 138-149 are disordered with the exception of Gly145 and Tyr146. In both molecules, the side chain of Tyr146 forms a hydrogen bond with His176, however, the orientations of Gly145 and Tyr146 differ in molecules A and B.

The structure of the FERM:C-terminal domain complex was determined in two crystal forms: space group $C222_1$ (Ezrin-1) at 2.0 Å resolution and $P2_1$ (Ezrin-2) at 1.9 Å resolution. The former was an attempt to crystallize the full-length ezrin monomer while the latter, the dimer. After solving the structures, it was clear that the protein had inadvertently undergone limited proteolysis during crystallization. SDS-PAGE analysis of crystals resolved two bands at approximately 37 kDa and 10 kDa. Mass spectrometry revealed that the 37 kDa band contains the FERM domain, while the 10 kDa band contains the C-terminal domain, residues 510-586, with molecular weight 9336.2 Da (data not shown). Crystal structures in both space groups contain the FERM:C-terminal domain complex.

The structure of Ezrin-1 contains one molecule in the asymmetric unit (residues 1-297 and 516-586), while the structure of Ezrin-2 contains two molecules in the asymmetric unit (molecule A: residues 1-297, 516-586 and molecule B: residues 3-297, 515-586), with C_{α} RMSD of 0.73 Å over 365 residues. There are only small differences between Ezrin-1 and Ezrin-2 structures with RMSDs of 0.49 Å (368 residues) and 0.73 Å (365 residues) between Ezrin-1 and molecules A and B of Ezrin-2, respectively.

All coordinates and structure factors have been deposited with the Protein Data Bank with accession codes: 4RMA for the ezrin FERM domain, 4RM9 and 4RM8 for the FERM:C-terminal domain complexes derived from Ezrin-1 and Ezrin-2 crystals.

Sequence analysis

ERM protein sequences were obtained using the NCBI BLAST [36] server. Sequences were aligned using CLUSTAL W [37]. Phylogenetic trees were used to separate merlin protein sequences from other ERM sequences. BLAST searches restricted by taxa ensured that the final sequence set had the widest possible coverage of the ERM evolutionary tree.

Small-angle x-ray scattering

The solution structure of full-length monomeric and dimeric ezrin was determined by SEC coupled with SAXS (SEC-SAXS) using the SAXS/WAXS beamline at the Australian Synchrotron with an in-line Superose 6 PC 3.2/30 (GE Healthcare) column equilibrated with 50 mM HEPES pH 8, 200 mM Na₂SO₄, 50 mM K₂SO₄, 5 mM β-mercaptoethanol. 50 μl of sample (~24 mg/ml of monomer or 10.6 mg/ml of dimer) was loaded at a flow of 0.1 ml/min, with the eluent flowing directly into the SAXS sample chamber (SEC chromatograms shown in Fig. S1B). Data were collected with using a Dectris Pilatus 1M detector at 1.13 Å wavelength. SEC-SAXS data were collected at sample-detector distances of 7 m and 650 mm.

SAXS images were taken continuously as the protein eluted from the SEC column (SEC chromatograms shown in Fig. S1B). Data were averaged over 2-second periods, reduced to 1-dimensional scattering patterns and converted to an absolute scale using the program Scatterbrain (written and provided by the Australian Synchrotron and available at <http://www.synchrotron.org.au>). Useful data were obtained for Q values between $5.5 \times 10^{-3} \text{ \AA}^{-1}$ and 0.855 \AA^{-1} ($Q = 4\pi \sin q / \lambda$ where λ is the wavelength and $2q$ is the scattering angle). Data were binned in groups of five consecutive 2-second frames and averaged allowing the radius of gyration R_g and maximum chord D_{max} to be estimated. Analysis of these R_g and D_{max} values confirmed that the configuration of the protein was not changing over the SEC peak. For structure determination, only the ten brightest frames were analyzed to obtain maximum contrast between protein and background scattering. R_g , D_{max} and the pair distance distribution function $P(r)$ were calculated using Fourier methods implemented in the program GNOM [38]:

$$P(r) = \frac{r}{2\pi^2} \int_0^{\infty} Q \times I(Q) \times \sin(Q \times r) dQ$$

As an independent quality check, R_g was also estimated using the Guinier approximation:

$$I(Q) = I(0) e^{-\frac{Q^2 R_g^2}{3}}$$

Low and high Q data were merged using program PRIMUS [39] to produce final scattering patterns.

Ab initio methods based on a simulated annealing algorithm implemented in DAMMIF/DAMMIN [40] and GASBOR [41] were used to generate sets of models for both the ezrin monomer and dimer, using $P(r)$ functions with input D_{max} values of 172 Å and 335 Å, respectively. Models generated by GASBOR were favored as the program utilizes data over the entire Q range. A set of best models was selected for monomer and dimer based on C^2 with the proviso that the fit was good in the low Q region (monomer C^2 values: 0.70-1.7; dimer C^2 values: 0.7-2.2). These models were aligned by the DAMAVER program suite [42] to generate average and “filtered” (volume reduced to match expected value for molecule) shape functions. The program FoXS [43] was used to calculate SAXS profiles from atomic models. The complete experimental data sets ($Q_{max}=0.954 \text{ \AA}^{-1}$) were used to calculate C values.

Homology Modeling

Homology models for full-length ezrin monomer and dimer were generated by extending the crystal structure of the FERM:C-terminal domain complex. *Sfmoesin* (PDB: 2I1J and 2I1K [21]) was used as a template. A full model for human ezrin was initially generated by the I-TASSER server [44]. An ideal α -helical coiled-coil region was generated using the CCBuilder server [45]. These two models were used to supply the missing segments. Splicing was carried out manually using the program COOT [34].

RESULTS

Structural plasticity in the crystal structure of the ezrin FERM domain

The crystal structure of the ezrin FERM domain (residues 1-296) was determined at 1.75 Å resolution (Fig. 1A; Table 1) using a crystal form that is isomorphous to the previously reported 2.3 Å structure [12]. There are two copies of the FERM domain in the asymmetric unit and each molecule contains three subdomains, identified as F1 (ubiquitin fold), F2 (all α -helical acyl-CoA binding domain fold) and F3 (phosphotyrosine binding, PTB, or pleckstrin homology domain fold), respectively, forming a clover-shaped molecule as per all ERM FERM domains [11-14] (Fig. 1A).

A distinguishing feature of this higher resolution structure of the ezrin FERM domain is that in one copy of the FERM domain, subdomain F3 shows two distinct conformations in the packing of an array of phenylalanine side chains in the core of the β -sandwich (Fig. 2A). Four phenylalanine residues are arrayed across the core of the outer β -sheet (β -strands β 5- β 7) of subdomain F3: Phe250 (β 5), Phe255 (β 6), Phe267 (β 7) and Phe269 (β 7). Each of these phenylalanine residues can adopt one of two rotamer conformations (cyan and magenta in Fig. 2A). In the two conformations of the phenylalanine array in subdomain F3, each phenylalanine side chain points in the same direction like a stack of dominos. To alternate between configurations, each side chain must switch rotamer in concert with the others (Fig. 2A). The refinement of the structure indicates that these conformations are equally populated in one copy of the FERM domain (molecule A, Fig. 2A), while the other copy (molecule B, Fig. 2B) shows only one conformation which is similar to the one seen in most other ERM structures (see Discussion for exceptions).

The flexibility of the outer β -sheet region of subdomain F3, containing β -strands β 5, β 6 and β 7 (Figs 1A, 2A and 2B), is reflected in the elevated B-factors for this region in both copies of the structure in the asymmetric unit (Fig. 2D), as shown previously [12]. The main chain B-factors of β -strands β 5- β 7 are 20-30 Å² higher than either the overall FERM domain or subdomain F3 (Fig. 2E).

Crystal structures of the ezrin FERM:C-terminal domain complex obtained by crystallizing purified monomer and dimer samples

Crystals were obtained from both purified full-length ezrin monomer and dimer. The structures determined from these crystals were essentially identical, showing complexes between the FERM domain and the C-terminal domain with no density (or room) for the central helical domain. SDS-PAGE and mass spectrometry confirmed that limited proteolysis had occurred during crystallization (see Methods).

The crystal structure of the complex between the FERM domain and the C-terminal domain was determined in two different space groups at 1.9 Å (dimer sample) and 2.0 Å (monomer sample) resolutions, respectively (Figs 1B and 1C, Table 1). These structures show the C-terminal domain (residues 516-586) is intimately wrapped around the FERM domain (residues 1-297) as seen in the previously reported structures of dormant ERMs [11, 21]. The C-terminal domain is composed of four α -helices: α 1C- α 4C (Fig. 1B).

The only difference between structures of the ezrin FERM:C-terminal domain complex are the proximity of subdomain F1 to subdomain F3 (Fig. 1C). The largest difference can be seen on comparing the two molecules in the asymmetric unit of 4RM8 (Fig. 1C, green and purple), where the separation between the subdomains is closer by about 1-2 Å in molecule B, which has lower overall B-factors.

The fact that crystals obtained from both purified monomer and dimer samples resulted in the same structure after limited proteolysis suggests that this structure is present in both forms of the protein.

Structural changes in the FERM domain on binding the C-terminal domain

There are no major changes in the structure of the FERM domain of ezrin on forming a complex with the C-terminal domain (C_{α} RMSDs of 0.93-1.3 Å between FERM only and FERM:C-terminal domain complexes compared with C_{α} RMSDs of 0.49-0.73 Å among the complexes). However, there are specific regions in the FERM domain where the binding of the C-terminal domain alters the structure (Fig. 1D).

As noted from the radixin FERM domain structure [13], binding of the C-terminal domain alters the structure of the outer β -sheet (β -strands β_5 , β_6 and β_7) of subdomain F3 (Fig. 1D), the region showing plasticity in the FERM only structure. These three β -strands translate by approximately 3 Å towards subdomain F1 on binding C-terminal domain (Fig. 1E). Large changes are seen in the loop connecting β -strands β_6 and β_7 , which is directly adjacent to helix α_4C in the C-terminal domain (Fig. 1E).

Binding of the C-terminal domain to subdomain F3 stabilizes its plastic hydrophobic core structure. The binding of α -helix α_4C positions the side chain of Phe583 so that it selects only one possible conformation for the four phenylalanine residues (Phe267, Phe269, Phe255 and Phe250) that form the phenylalanine domino array in the core of subdomain F3 (Fig. 2C). This stabilization of the F3 core results in a reduction of the main chain B-factors for the outer β -sheet (β -strands β_5 - β_7) so that they are within ~ 10 Å² of the remainder of the structure (Fig. 2F).

Ezrin FERM subdomain F2 is an all α -helical structure (helices α_1 , α_2 , α_2' , α_3 and α_4 ; Fig. 1). The binding of the C-terminal domain results in the rotation of helices α_2' and α_3 and changes in the long loop connecting helices α_2 to α_2' (Fig. 1F), while helices α_1 , α_2 and α_4 make only small movements. These changes in FERM subdomain F2 accommodate the binding of helix α_1C and the loop connecting it to α_2C in the C-terminal domain. These changes are consistent with those seen on comparing the FERM domain structures of radixin [13] and moesin [14] to the moesin FERM:C-terminal domain complex [11].

Multi-angle light scattering shows bacterially expressed ezrin forms stable monomer and dimer fractions

The expression of full-length ezrin in *E. coli* results in two fractions that can be separated by SEC that are called monomer and dimer (Fig. S1A). To confirm that these SEC peaks actually correspond to ezrin monomer and dimer species, we performed multi-angle light scattering in-line with the SEC column (SEC-MALS). The estimated molecular masses of the monomer and dimer peaks were $70,800 \pm 900$ Da and $140,000 \pm 2,000$ Da, respectively. These are consistent with the calculated molecular masses of monomer and dimer (69,470 Da for 586 residue monomer plus N-terminal glycine). Recycling either monomer or dimer fraction through the SEC column did not result in re-equilibration of the protein into monomer and dimer (data not shown). Thus, both monomer and dimer are stable under the conditions of the experiment.

Thermal stability of ezrin as measured by circular dichroism

Figure 3A shows the far UV circular dichroism (CD) spectra of the ezrin FERM

domain and the monomeric and dimeric forms of full-length WT ezrin before (at 20 °C) and after thermal denaturation (at 80 °C). The thermal unfolding transition is not reversible. Figure 3B shows the mean residue ellipticity $[\theta]$ at 208 nm as a function of temperature for all proteins studied. They all show a clear unfolding transition. By fitting these curves with a sigmoidal function, we obtained the apparent melting temperatures, T_{AM} , for each protein (Fig. 3B). The data show that the full-length structures are significantly more stable than the isolated FERM domain (T_{AM} of $44.8 \pm 0.1^\circ\text{C}$ for the FERM only, compared to 65.0 ± 1.2 and $66.2 \pm 1.1^\circ\text{C}$ for full-length monomer and dimer, respectively). The apparent melting temperatures of monomer and dimer species are indistinguishable. We note that prior to the sigmoidal unfolding transition, the mean residue ellipticity shows a linear dependence on temperature for both full-length monomer and dimer samples. Such a linear dependence on temperature is typical for coiled-coil proteins [46] as well as stable single α -helices [47].

Solution structure of full-length ezrin monomer and dimer by SAXS

In order to gain insight into the structure of full-length ezrin, we collected small-angle x-ray scattering (SAXS) data in-line with a SEC column on previously purified monomer and dimer (Fig. 4; Table 2). The scattering curves for the monomer and dimer are almost indistinguishable except at very small angles (scattering vector magnitude $Q < 0.1 \text{ \AA}^{-1}$ corresponding to Bragg spacing of $D > 63 \text{ \AA}$). At very low angles ($Q < 0.02 \text{ \AA}^{-1}$, $D > 314 \text{ \AA}$), the dimer scattering increases more rapidly than the monomer, as would be expected for a larger molecule (Fig. 4A plus inset). For Q values between 0.02 and 0.06 \AA^{-1} ($D = 314 \text{ \AA}$ and 105 \AA , respectively), damped oscillations are observed in the difference between the monomer and dimer scattering (Fig. 4A inset). We note that a previous study failed to distinguish between the ezrin monomer and dimer via SAXS [48]. The reason for this appears to be the Q ranged used in the previous study, which had a lower limit of 0.014 \AA^{-1} and hence the key difference between monomer and dimer SAXS (the upswing below 0.02 \AA^{-1}) is effectively missing in their data.

Guinier (Fig. 4B inset) and $P(r)$ analysis (Fig. 4C) of the data are consistent with each other, giving the following structural parameters for the monomer: radius of gyration, $R_g = 40 \pm 1 \text{ \AA}$ and maximal chord, $D_{max} = 165 \pm 5 \text{ \AA}$; and dimer: $R_g = 95 \pm 5 \text{ \AA}$ and $D_{max} = 325 \pm 5 \text{ \AA}$ (Table 2). The D_{max} values suggest that the ezrin dimer is twice as long as the monomer.

The $P(r)$ plot for the monomer SAXS data shows a single peak around 35 \AA with an elongated tail at large pair distances (Fig. 4C). In contrast, the $P(r)$ plot for the dimer shows two peaks, one around a pair distance of 35 \AA that corresponds to the monomer peak and a second, at a pair distance around 210 \AA (Fig. 4C). The pair distribution for the dimer data is characteristic of a dumbbell-shaped protein.

Molecular models for the ezrin monomer

Ab initio models for the ezrin monomer using the SAXS data consistently resulted in comma shaped molecules (Fig. 5A). Given the similarity between these models and the crystal structure of full-length *Sfmoesin* [21], we used this structure combined with our crystal structure to build a complete homology model of human ezrin (Fig. 5B). Modeling the central helical domain as an extended coiled-coil is consistent with the experimentally determined maximum chord, D_{max} , of $165 \pm 5 \text{ \AA}$ for the ezrin monomer. The monomer model was docked into the envelope of the averaged (Fig. 5B, grey) and filtered (Fig. 5B, green) shapes generated from the family of *ab initio* SAXS models. Although the homology model fits the SAXS envelope, we stress that SAXS is a low-resolution technique, thus, the envelope equally fits a family of models where the globular domain is rotated with respect to

the coiled-coil.

Low resolution/SAXS models for the ezrin dimer

Ab initio models of the ezrin dimer using SAXS data produce very elongated structures (Fig. 5C). Each model has two globular regions at the extremities linked by a thinner connecting region. The panel of model structures shows that the ezrin dimer is extremely elongated with the separation between the globular structures conserved between models.

Analysis of the SAXS data shows that the maximum chord, D_{max} , for the dimer is twice that of the monomer. Additionally, the $P(r)$ plot shows that the ezrin dimer is a dumbbell shaped molecule. The simplest way to fulfill these two constraints is to dock two ezrin monomers so that their coiled-coil arms align tip to tip so as to resemble a single coiled-coil of double the monomer length. Such a model places the two FERM:C-terminal domain complexes at opposite ends of the central coiled-coil producing a dumbbell-shaped molecule. The only problem with this model is that such a dimer would be highly unstable if the only dimer contact was via the very tip of the coiled-coil (i.e. not a domain swapped dimer).

Sequence analysis of the central α -helical domain

To determine the role of the central α -helical domain in dimer formation, we analyzed a wide range of ERM protein sequences. A BLAST [36] search of the non-redundant protein sequence database (NCBI) showed that ERM proteins are present in nearly all metazoa. Alignment of representative ERM protein sequences spanning the metazoan evolutionary tree shows that ERM proteins are highly conserved with many residues in the FERM and C-terminal domains invariant.

We examined the conservation of the central α -helical domain (Fig. S2). The crystal structure of the full-length ERM monomer, *Sfmoesin*, was used to identify the three α -helices in the α -helical domain (α A, α B and α C, Fig. S2). The sequence of the N-terminus of the first helix, α A, is highly conserved. The length of the second α -helix, α B, appears to be conserved with the caveat that, in chordates, it appears to be longer by seven residues due to an insertion near its N-terminus. This insertion appears to have occurred during chordate evolution, as it is present in the ERM protein from the tunicate *Ciona intestinalis* (Fig. S2). The region between the C-terminus of α -helix α C and the C-terminal domain is poorly conserved in terms of residue identity and sequence length. Additionally, it contains low-complexity sequences.

Although the sequences of helices α B and α C are not conserved, the heptad repeat, which is a signature of coiled-coil structures [49], is conserved in these helices (Fig. S2). This can be seen in the pattern of hydrophobic residues at the 'a' and 'd' positions in the heptad repeat (highlighted in yellow in Fig. S2). More importantly, the length of the linker between α -helices α B and α C (as seen in the *Sfmoesin* crystal structure) is conserved [21]. The length of this linker is such that it preserves the register of the heptad repeat of the coiled-coil when the sequence is extended into one long α helix (Fig. S2). This has been noted previously, based on the alignment of the *Sfmoesin* sequence with the human ERM-merlin sequences [21]. By aligning ERM proteins spanning metazoa we see that the register of the heptad repeat is conserved throughout metazoan evolution (Fig. S2).

In order to accommodate both the monomer and dimer forms, the linker between helices α B and α C must be able to adopt two conformations, a loop and a helical coiled-coil,

respectively. In this regard, we note that the residues in the heptad 'd' and 'a' positions in this loop (residues numbered 112 and 116 in Fig. S2) are not hydrophobic, however, they are observed in these core positions in two-helix antiparallel coiled-coils. In the vertebrate sequences (Fig. S2), these two positions are occupied by glutamine residues (with one exception: arginine at position 'a' in *Callorhinchus milii* (elephant shark) sequence). In lower order metazoa, the 'd' site (residue 112) is occupied by: Glu (6), Thr (2), Asp or Ser while the 'a' site (residue 116) is occupied by: Lys (6), Arg (3) or Leu. Using the CC+ relational database of coiled-coil structures [50], we find that the Swiss-Prot normalized propensities in two-helix antiparallel coiled-coils for the 'd' site are: Gln 0.59, Glu 0.76, Thr 0.62, Asp 0.21 and Ser 0.52; while for the 'a' site they are: Gln 1.0, Lys 0.73, Arg 0.86 and Leu 2.66. Thus, these residues are not rare at these core positions in two-helix antiparallel coiled-coils. Therefore, the link between α B and α C is consistent with a metamorphic structure that can switch between a loop and a coiled-coil.

The conservation pattern of the central α -helical domain has several implications for ERM proteins. First, the preservation of the heptad repeat register through the sequence connecting α -helices α B and α C indicates that the monomer structure can be opened up to form a domain-swapped dimer where α -helices α B and α C form one, continuous α -helix that makes up half of a coiled-coil. Second, the conservation of the length of α -helices α B and α C (apart from a possible seven residue insertion in chordates) indicates that the coiled-coil domain serves a biological function.

Molecular models for the ezrin dimer

Given the sequence conservation in the central α -helical domain, we used the model of the ezrin monomer to build a model for the dimer where the coiled-coil is extended to form an elongated dumbbell (Fig. 5D, cartoon). This domain-swapped model satisfies the requirement of doubling the maximum chord, D_{max} , of the monomer. The model matches the averaged (Fig. 5D, grey) and filtered (Fig. 5D, green) shapes generated from the family of *ab initio* SAXS models. As per the monomer, SAXS is a low-resolution technique, hence, rotation of the globular domains relative to the coiled-coil will not affect the match between homology model and the SAXS shapes.

The fact that the monomer and dimer SAXS profiles are nearly identical except for very low angles supports the domain-swapped dimer model, as it indicates that the structural features of the dimer are identical to those of the monomer except at very long length scales. Given the crystallographic evidence for a coiled-coil in the monomer [21], the high correspondence between the monomer and dimer SAXS data argues that the same coiled-coil structure has to be present in the dimer.

Evaluation of the molecular models against the SAXS data

SAXS is a low-resolution technique, hence it is problematic to refine atomic models against SAXS data. However, SAX data should be able to select between competing putative atomic models. To evaluate atomic models, the program FoXS was used which generates a SAXS profile by evaluating the Debye equation [43]. The program outputs a χ value as a measure of quality of fit. The full homology models for the monomer and dimer resulted in χ values of 3.92 and 4.45 (Table 3) suggesting that portions of the model are incorrect. Given this, we created a series of partial models for both monomer and dimer, starting with the crystal structure of the FERM:C-terminal domain complex, which is a true atomic structure, and increasing the completeness of the models in a stepwise fashion by adding segments starting with those with the highest confidence (Fig. S3A).

For the monomer, the crystal structure alone gives a poor fit to the data (χ value 12.99, Table 3). Using the crystal structure of *Sfmoesin* as a guide, an ideal coiled-coil was added to the crystal structure, reducing the χ value to 4.12. Connecting the N-terminus of the coiled-coil to the FERM domain by adding helix αA and completing the N-terminus of helix αB further decreased the χ value to 2.26, which is a reasonable value for a single atomic model missing 7% of the structure [43]. The calculated SAXS profile for this model is shown in Fig. S3B along with the experimental profile. Further extension of the model reduced the quality of the fit (increased the χ value, Table 3). These segments of the model are likely to be poor as the sequence of human ezrin diverges from *Sfmoesin*, which is the template for building the model. They include the proline rich segment at the C-terminus of the helical domain.

The above set of partial ezrin monomer models were used to generate dimer models via domain swapping as described above. For the dimer models, the lowest χ value was observed for a model comprising only the two copies of the crystal structure plus the coiled-coil (χ value 3.19, Table 3). This is a reasonable χ value for model comprising 83% of the scattering mass [43]. Connecting the FERM domain to the coiled-coil domain increased the χ value to 3.58, in contrast to the reduction observed in the monomer model (Table 3). This suggests that either the orientations of the two globular domains of the dumbbell may be mis-oriented or there exists some degree of flexibility in the coiled-coil, preventing a single model fitting the data. The calculated SAXS profile for the dimer model where the FERM domain is connected to the coiled-coil via helix αA (equivalent to the best monomer model) is compared to the data in Fig. S3C.

DISCUSSION

The ERM proteins form a nexus between cell membranes and the underlying cortical actin cytoskeleton. As such, they must integrate signals from several sources and transmit information. To achieve this, the ERM proteins have several distinct structural states, dormant and activated, and appear to utilize higher order complexes including dimers. The results of our structural studies uncover potential mechanisms by which the ERM protein ezrin may transmit information between binding sites. Our SAXS-based structure for the ezrin dimer suggests a model for how ERM proteins couple membranes to F-actin so as to maintain the cortical actin cytoskeleton parallel to the membrane.

The FERM only crystal structure represents the activated FERM domain. Our key finding is that the hydrophobic core of subdomain F3 displays a bistable structure where a row of phenylalanine residues can adopt one of two mutually exclusive configurations (Fig. 2A). We are not aware of any other protein structure that displays such a dramatic bistability in its hydrophobic core. The core bistability is consistent with the thermal stability data (Fig. 3) and the elevated B-factors (noted previously [12]; Fig. 2E).

The crystal structure of the FERM:C-terminal domain complex represents the dormant form. The binding of α -helix α 4C from the C-terminal domain to the edge of subdomain F3 inserts the side chain of Phe583 into the hydrophobic core of subdomain F3. This selects a single arrangement of the core phenylalanine array (Fig. 2C) acting as a keystone [12] stabilizing the structure (Figs 2C, 2F). The binding of the adaptor proteins NHERF-1/EBP50 and NHERF-2 stabilizes subdomain F3 via the same mechanism [20].

The bistable structure of the core phenylalanine array may transmit information between functional sites on subdomain F3. This subdomain has three interaction sites: the opening between the two β -sheets adjacent to β -strand β 7 and β 4 where the C-terminal domain or NHERF adaptors bind; the opposite edge of the outer β -sheet (proximal to β -strand β 5) where cytoplasmic domains of membrane proteins bind [15-19]; and the outer surface of the subdomain F3 β -sandwich (β -strands β 5, β 6 and β 7) which is likely to be important in membrane binding. Thus, the core phenylalanine array in subdomain F3 may be involved in integrating information among these sites.

An examination of these core phenylalanine residues in all available ERM structures shows that, in general, the phenylalanine side chains adopt the same rotamers as seen in the dormant structure. The only exceptions are radixin FERM domain bound to the cytoplasmic domain of CD44 [16] and one molecule in the radixin FERM domain dimer [51]. In both of these structures, a peptide forms an additional anti-parallel β -strand that binds to β 5 of subdomain F3. However, in most structures of complexes between the radixin FERM domain and peptides representing membrane protein cytoplasmic domains, the rotamer conformation of the core phenylalanine array is identical to the dormant structure [15, 17-19].

The precise role of the central α -helical domain in ERM proteins is enigmatic. Models for the coupling of membranes to F-actin by ERM proteins show a membrane-bound FERM domain linked to an F-actin bound C-terminal domain via an extended central domain [2]. Studies on the isolated α -helical domain from radixin indicate that it may form a monomeric stable α -helix [52]. However, the crystal structure of *Sf*moesin shows that a large portion of the α -helical domain forms a coiled-coil with a conserved heptad repeat [21]. Thus, if in the activated state, the FERM and C-terminal domains were to be separated by either an extended stable α -helix or an unstructured chain, a significant hydrophobic surface

would be exposed, rendering the structure unstable.

A more likely scenario is that ERM proteins maintain a coiled-coil structure in the central α -helical domain. Our CD thermal denaturation experiments on full length ezrin show a linear dependence on temperature prior to unfolding (Fig. 3B), which is typical for coiled-coils [46], however, it does not exclude single, stable α -helices [47].

The SAXS data shows that the full-length, dormant ezrin monomer structure resembles that of *Sfmoesin* [21]. A similar result has been obtained via small-angle neutron scattering (SANS) using deuterated ezrin [48]. Thus, the central α -helical domain forms a coiled-coil that extends from the FERM domain which is consistent with SAXS and CD data.

Alignment of ERM protein sequences spanning metazoa shows that several features of the central α -helical domain are conserved. As noted previously, the heptad repeat is conserved which is consistent with the coiled-coil observed in the dormant state monomer structure [21]. Additionally, the length of the region linking the two α -helices (α B and α C) in the coiled-coil structure is conserved (Fig. S2). More importantly, the length of this linker preserves the phase of the heptad repeat between the two α -helices. This is a striking feature to conserve in the central α -helical domain over the whole of metazoan evolution. It is important for understanding the nature of the ERM dimer states.

In our model for the dimer, the monomer structure is opened up at the linker region connecting α -helices α B and α C of the coiled-coil. Because the linker preserves the phase of the heptad repeat, ezrin is able to form a continuous coiled-coil of twice the monomer length resulting in a domain-swapped dimer. This model is consistent with previous work and it explains previous results. The first detailed characterization of the ezrin dimer using SEC and analytical ultracentrifugation ascertained that the dimer was a stable, elongated species with little exchange between species [23]. The domain-swapped nature of our dimer model explains why monomer and dimer fractions do not exchange as this would require the complete unfolding of the elongated coiled-coil which is stabilized by a large hydrophobic interface.

Several studies have revealed structures for what is likely to be the dimeric state of ERM proteins. Low-angle rotary shadowing electron microscopy revealed two forms of radixin, one predominantly in samples of wild type and the phospho-resistant T564A, which shows a globular protein, consistent with the dormant monomer structure, while the phosphomimetic T564E showed a large fraction of elongated dumbbell molecules [53]. Although these were originally thought to be open monomers, their properties are consistent with our dimer model.

The dimer model is also consistent with images of ezrin obtained by atomic force microscopy (AFM) [54]. These authors collected images on bacterially-expressed wild type, T567A and T567D ezrin. The images show both globular structures resembling the dormant monomer and several types of dumbbell structures. The authors interpret the dumbbell structures as open and half-closed monomers, however, the images are consistent with our dimer model. We note that the AFM images also show higher order structures.

The conservation of the linker between the two coiled-coil helices in the monomer structure so as to maintain the phase of the heptad repeat indicates that evolution has preserved the ability of ERM proteins to form elongated domain-swapped dimer structures as

per our dimer model. This suggests that the ERM dimer plays a functional role in biology. The dimer structure places the two principal ERM functionalities, binding to membranes and binding to F-actin, at both ends of an elongated dumbbell. Thus, the dimer may bind along the axis of an actin filament. We speculate that such a structure may naturally link F-actin to a membrane so that the filaments lie parallel to the membrane surface.

ACKNOWLEDGEMENTS

This research was undertaken on the MX1, MX2 and SAXS/WAXS beamlines at the Australian Synchrotron, Victoria, Australia. SAXS experiments were carried out with the assistance of Nigel Kirby and Haydyn Mertens of the Australian Synchrotron. We thank Mark Berryman for the plasmid encoding the ezrin FERM domain, Robert Knott for assistance with preliminary SAXS measurements using the ANSTO Bruker NanoStar, Dek Woolfson for input regarding coiled-coils, Joel Mackay for the usage of MALS equipment and Sam Breit, Louise Brown, Katharine Michie, Sophia Goodchild, Till Böcking, Peter Gunning and Thomas Huber for fruitful discussions. MS analysis was supported by access to the University of Sydney's Mass Spectrometry Core Facility.

DECLARATION OF INTEREST

The authors declare that they have no conflicts of interest with the contents of this article.

FUNDING

This research was supported under the Australian National Health and Medical Research Council Project Grant scheme (APP1085718). J.M.P. was supported by an Australian Nuclear Science and Technology Organisation (ANSTO) Postgraduate Scholarship.

AUTHOR CONTRIBUTIONS

JMP, KEW, APD and PMGC conceived and coordinated the study and wrote the paper. JMP purified all proteins. RBD performed all molecular biology tasks. JMP and SJH determined the crystal structures. APD, AVS and SAB collected SAXS data. APD, AVS, JCW and PMGC analyzed SAXS data. BC performed MS. CDN performed MALS experiments. JMP performed and analyzed CD. CG initiated experiments on FERM. EHCB and PMGC analyzed the coiled-coil region. All authors reviewed the results and approved the final version of the manuscript.

REFERENCES

- 1 Neisch, A. L. and Fehon, R. G. (2011) Ezrin, Radixin and Moesin: key regulators of membrane-cortex interactions and signaling. *Curr. Opin. Cell Biol.* 23, 377-382
- 2 Fehon, R. G., McClatchey, A. I. and Bretscher, A. (2010) Organizing the cell cortex: the role of ERM proteins. *Nat. Rev. Mol. Cell Biol.* 11, 276-287
- 3 Clucas, J. and Valderrama, F. (2014) ERM proteins in cancer progression. *J Cell Sci.* 127, 267-275
- 4 Bretscher, A., Chambers, D., Nguyen, R. and Reczek, D. (2000) ERM-Merlin and EBP50 protein families in plasma membrane organization and function. *Annu. Rev. Cell Dev. Biol.* 16, 113-143
- 5 Turunen, O., Wahlstrom, T. and Vaheri, A. (1994) Ezrin has a COOH-terminal actin-binding site that is conserved in the ezrin protein family. *J. Cell Biol.* 126, 1445-1453
- 6 Gary, R. and Bretscher, A. (1995) Ezrin self-association involves binding of an N-terminal domain to a normally masked C-terminal domain that includes the F-actin binding site. *Mol. Biol. Cell.* 6, 1061-1075
- 7 Fievet, B. T., Gautreau, A., Roy, C., Del Maestro, L., Mangeat, P., Louvard, D. and Arpin, M. (2004) Phosphoinositide binding and phosphorylation act sequentially in the activation mechanism of ezrin. *J. Cell Biol.* 164, 653-659
- 8 Niggli, V., Andreoli, C., Roy, C. and Mangeat, P. (1995) Identification of a phosphatidylinositol-4,5-bisphosphate-binding domain in the N-terminal region of ezrin. *FEBS Lett.* 376, 172-176
- 9 Nakamura, F., Huang, L., Pestonjamas, K., Luna, E. J. and Furthmayr, H. (1999) Regulation of F-actin binding to platelet moesin in vitro by both phosphorylation of threonine 558 and polyphosphatidylinositides. *Mol. Biol. Cell.* 10, 2669-2685
- 10 Yonemura, S., Matsui, T. and Tsukita, S. (2002) Rho-dependent and -independent activation mechanisms of ezrin/radixin/moesin proteins: an essential role for polyphosphoinositides in vivo. *J. Cell Sci.* 115, 2569-2580
- 11 Pearson, M. A., Reczek, D., Bretscher, A. and Karplus, P. A. (2000) Structure of the ERM protein moesin reveals the FERM domain fold masked by an extended actin binding tail domain. *Cell.* 101, 259-270
- 12 Smith, W. J., Nassar, N., Bretscher, A., Cerione, R. A. and Karplus, P. A. (2003) Structure of the active N-terminal domain of Ezrin. Conformational and mobility changes identify keystone interactions. *J. Biol. Chem.* 278, 4949-4956
- 13 Hamada, K., Shimizu, T., Matsui, T., Tsukita, S. and Hakoshima, T. (2000) Structural basis of the membrane-targeting and unmasking mechanisms of the radixin FERM domain. *EMBO J.* 19, 4449-4462
- 14 Edwards, S. D. and Keep, N. H. (2001) The 2.7 Å crystal structure of the activated FERM domain of moesin: an analysis of structural changes on activation. *Biochemistry.* 40, 7061-7068
- 15 Takai, Y., Kitano, K., Terawaki, S., Maesaki, R. and Hakoshima, T. (2008) Structural basis of the cytoplasmic tail of adhesion molecule CD43 and its binding to ERM proteins. *J. Mol. Biol.* 381, 634-644
- 16 Mori, T., Kitano, K., Terawaki, S., Maesaki, R., Fukami, Y. and Hakoshima, T.

- (2008) Structural basis for CD44 recognition by ERM proteins. *J. Biol. Chem.* 283, 29602-29612
- 17 Terawaki, S., Kitano, K. and Hakoshima, T. (2007) Structural basis for type II membrane protein binding by ERM proteins revealed by the radixin-neutral endopeptidase 24.11 (NEP) complex. *J. Biol. Chem.* 282, 19854-19862
- 18 Takai, Y., Kitano, K., Terawaki, S., Maesaki, R. and Hakoshima, T. (2007) Structural basis of PSGL-1 binding to ERM proteins. *Genes Cells.* 12, 1329-1338
- 19 Hamada, K., Shimizu, T., Yonemura, S., Tsukita, S. and Hakoshima, T. (2003) Structural basis of adhesion-molecule recognition by ERM proteins revealed by the crystal structure of the radixin-ICAM-2 complex. *EMBO J.* 22, 502-514
- 20 Terawaki, S., Maesaki, R. and Hakoshima, T. (2006) Structural basis for NHERF recognition by ERM proteins. *Structure.* 14, 777-789
- 21 Li, Q., Nance, M. R., Kulikaukas, R., Nyberg, K., Fehon, R., Karplus, P. A., Bretscher, A. and Tesmer, J. J. (2007) Self-masking in an intact ERM-merlin protein: an active role for the central alpha-helical domain. *J. Mol. Biol.* 365, 1446-1459
- 22 Pakkanen, R. and Vaheri, A. (1989) Cytovillin and other microvillar proteins of human choriocarcinoma cells. *J. Cell. Biochem.* 41, 1-12
- 23 Bretscher, A., Gary, R. and Berryman, M. (1995) Soluble ezrin purified from placenta exists as stable monomers and elongated dimers with masked C-terminal ezrin-radixin-moesin association domains. *Biochemistry.* 34, 16830-16837
- 24 Berryman, M., Gary, R. and Bretscher, A. (1995) Ezrin oligomers are major cytoskeletal components of placental microvilli: a proposal for their involvement in cortical morphogenesis. *J. Cell Biol.* 131, 1231-1242
- 25 Gautreau, A., Louvard, D. and Arpin, M. (2000) Morphogenic effects of ezrin require a phosphorylation-induced transition from oligomers to monomers at the plasma membrane. *J. Cell Biol.* 150, 193-203
- 26 Andreoli, C., Martin, M., Le Borgne, R., Reggio, H. and Mangeat, P. (1994) Ezrin has properties to self-associate at the plasma membrane. *J. Cell Sci.* 107 (Pt 9), 2509-2521
- 27 Gary, R. and Bretscher, A. (1993) Heterotypic and homotypic associations between ezrin and moesin, two putative membrane-cytoskeletal linking proteins. *Proc. Natl Acad. Sci. U. S. A.* 90, 10846-10850
- 28 Gronholm, M., Sainio, M., Zhao, F., Heiska, L., Vaheri, A. and Carpen, O. (1999) Homotypic and heterotypic interaction of the neurofibromatosis 2 tumor suppressor protein merlin and the ERM protein ezrin. *J. Cell Sci.* 112 (Pt 6), 895-904
- 29 Smith, W. J. and Cerione, R. A. (2002) Crystallization and preliminary crystallographic analysis of the ezrin FERM domain. *Acta Crystallogr. D Biol. Crystallogr.* 58, 1359-1361
- 30 McPhillips, T. M., McPhillips, S. E., Chiu, H. J., Cohen, A. E., Deacon, A. M., Ellis, P. J., Garman, E., Gonzalez, A., Sauter, N. K., Phizackerley, R. P., Soltis, S. M. and Kuhn, P. (2002) Blu-Ice and the Distributed Control System: software for data acquisition and instrument control at macromolecular crystallography beamlines. *J. Synchrotron Radiat.* 9, 401-406
- 31 Leslie, A. G. W. (1992) Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB News-letter of Protein Crystallography,*

No. 26. CCP4, CCLRC Daresbury Laboratory, Cheshire, UK.

32 COLLABORATIVE COMPUTATIONAL PROJECT No. 4 (1994) The CCP4 Suite: Programs for Protein Crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 50, 760-763

33 McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. and Read, R. J. (2007) Phaser crystallographic software. *J. Appl. Crystallogr.* 40, 658-674

34 Emsley, P., Lohkamp, B., Scott, W. G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* 66, 486-501

35 Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* 66, 213-221

36 Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* 25, 3389-3402

37 Thompson, J. D., Higgins, D. G. and Gibson, T. J. (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673-4680

38 Svergun, D. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystallogr.* 25, 495-503

39 Konarev, P. V., Volkov, V. V., Sokolova, A. V., Koch, M. H. J. and Svergun, D. I. (2003) PRIMUS: a Windows PC-based system for small-angle scattering data analysis. *J. Appl. Crystallogr.* 36, 1277-1282

40 Franke, D. and Svergun, D. I. (2009) DAMMIF, a program for rapid ab-initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* 42, 342-346

41 Svergun, D. I., Petoukhov, M. V. and Koch, M. H. (2001) Determination of domain structure of proteins from X-ray solution scattering. *Biophys. J.* 80, 2946-2953

42 Volkov, V. V. and Svergun, D. I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystallogr.* 36, 860-864

43 Schneidman-Duhovny, D., Hammel, M., Tainer, J. A. and Sali, A. (2016) FoXS, FoXSDock and MultiFoXS: Single-state and multi-state structural modeling of proteins and their complexes based on SAXS profiles. *Nucleic Acids Res.*

44 Zhang, Y. (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics.* 9, 40

45 Wood, C. W., Bruning, M., Ibarra, A. A., Bartlett, G. J., Thomson, A. R., Sessions, R. B., Brady, R. L. and Woolfson, D. N. (2014) CCBUILDER: an interactive web-based tool for building, designing and assessing coiled-coil protein assemblies. *Bioinformatics.* 30, 3029-3035

46 Holtzer, M. E. and Holtzer, A. (1992) Alpha-helix to random coil transitions: interpretation of the CD in the region of linear temperature dependence. *Biopolymers.* 32, 1589-1591

47 Wang, E. and Wang, C. L. (1996) (i, i + 4) Ion pairs stabilize helical peptides derived from smooth muscle caldesmon. *Arch. Biochem. Biophys.* 329, 156-162

- 48 Jayasundar, J. J., Ju, J. H., He, L., Liu, D., Meilleur, F., Zhao, J., Callaway, D. J. and Bu, Z. (2012) Open conformation of ezrin bound to phosphatidylinositol 4,5-bisphosphate and to F-actin revealed by neutron scattering. *J. Biol. Chem.* 287, 37119-37133
- 49 Cohen, C. and Parry, D. A. D. (1986) Alpha-Helical Coiled Coils - a Widespread Motif in Proteins. *Trends Biochem. Sci.* 11, 245-248
- 50 Testa, O. D., Moutevelis, E. and Woolfson, D. N. (2009) CC+: a relational database of coiled-coil structures. *Nucleic Acids Res.* 37, D315-322
- 51 Kitano, K., Yusa, F. and Hakoshima, T. (2006) Structure of dimerized radixin FERM domain suggests a novel masking motif in C-terminal residues 295-304. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.* 62, 340-345
- 52 Hoeflich, K. P., Tsukita, S., Hicks, L., Kay, C. M. and Ikura, M. (2003) Insights into a single rod-like helix in activated radixin required for membrane-cytoskeletal cross-linking. *Biochemistry.* 42, 11634-11641
- 53 Ishikawa, H., Tamura, A., Matsui, T., Sasaki, H., Hakoshima, T. and Tsukita, S. (2001) Structural conversion between open and closed forms of radixin: low-angle shadowing electron microscopy. *J. Mol. Biol.* 310, 973-978
- 54 Liu, D., Ge, L., Wang, F., Takahashi, H., Wang, D., Guo, Z., Yoshimura, S. H., Ward, T., Ding, X., Takeyasu, K. and Yao, X. (2007) Single-molecule detection of phosphorylation-induced plasticity changes during ezrin activation. *FEBS Lett.* 581, 3563-3571
- 55 Chen, V. B., Arendall, W. B., 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* 66, 12-21

FIGURE LEGENDS

FIGURE 1. Crystal structure of the ezrin FERM domain and the FERM:C-terminal domain complex. **A.** Cartoon representation of the ezrin FERM domain showing subdomains: F1 (magenta, residues 2-82), F2 (orange, residues 96-198) and F3 (green, residues 204-296), with linker regions in blue. **B.** The crystal structure of the ezrin FERM:C-terminal domain complex colored as a rainbow from the N-terminus (blue) to the C-terminus (red). The orientation is similar to the FERM domain structure in panel **A**. **C.** Comparison of the three independent structures of the ezrin FERM:C-terminal domain complex (4RM8-A: green, 4RM8-B: purple, 4RM9: orange). **D.** A superposition of the FERM only structure (yellow) with the FERM:C-terminal domain complex (FERM domain: cyan, C-terminal domain: red). The panels show close up views of the structural changes in FERM subdomain F3 (**E**) and F2 (**F**). The arrow in panel **E** indicates movement of the upper β -sandwich/ β -sheet region in subdomain F3. Note the two close ups are rotated relative to the panel **D** for clarity.

FIGURE 2. Structural plasticity in the core of subdomain F3 in the ezrin FERM domain structure. **A-C.** The array of phenylalanine residues in the core of subdomain F3 for the FERM only structure (**A**: molecule A and **B**: molecule B both shown in green) and the FERM:C-terminal domain complex (**C**) where the FERM domain is shown in cyan and the C-terminal domain shown in pink. Molecule A in the FERM domain structure shows two alternative conformations of the phenylalanine array (cyan and magenta in **A**), while the other structures show a single conformation (shown in yellow in **B** and **C**). Electron density is shown as a grey mesh. **D.** A cartoon representation of the crystallographic B-factors for the FERM domain structure (molecule B). Blue represents 20 \AA^2 while red represents 50 \AA^2 . **E,F.** The distribution of main chain B-factors for the outer β -sheet of subdomain F3 (β -strands $\beta 5$ - $\beta 7$) compared to subdomain F3 and the remainder of the molecule for the ezrin FERM domain structure (**E**) and the FERM:C-terminal domain complex (**F**).

FIGURE 3. Circular dichroism spectroscopy and thermal denaturation of ezrin. **A.** CD spectra ($n=3$) of proteins before ($20 \text{ }^\circ\text{C}$ showing strong features) and after ($80 \text{ }^\circ\text{C}$ showing reduced features) thermal denaturation. **B.** Representative thermal denaturation curves of ezrin (FERM domain and full-length ezrin monomer and dimer) measured at 208 nm and $1 \text{ }^\circ\text{C}/\text{min}$ heating rate, with a sigmoidal function fitted to the data.

FIGURE 4. SAXS data analysis. **A.** Logarithm of the SAXS intensity plotted as a function of the magnitude of the scattering vector, Q , for ezrin monomer (grey) overlaid on the ezrin dimer (black). The inset plots the linear difference between the scattering intensity of the dimer and monomer samples as a function of Q (the region used to calculate the linear difference is shown by the dotted box in **A**). **B.** The scattering intensity versus Q plotted on a Log-Log scale. The inset shows the Guinier plots for the monomer (grey) and dimer (black). The shaded boxes indicate the data points used to calculate R_g . Note: the d values on the x-axis represents the Bragg spacings equivalent to the Q values. **C.** The pair distribution plot $P(r)$ for the monomer (grey) and dimer (black).

FIGURE 5. Modeling the solution structures of ezrin monomer and dimer. **A.** A panel of sixteen *ab initio* reconstructions derived from the SAXS data for the ezrin monomer. These reconstructions have been aligned using the program DAMAVER [42]. **B.** The homology model for the human ezrin monomer overlaid on surfaces representing the averaged (grey) and filtered (green) shapes obtained from the *ab initio* models as calculated by DAMAVER. **C.** A panel of nine *ab initio* reconstructions derived from the SAXS data for the ezrin dimer. These reconstructions have been aligned using the program DAMAVER [42]. **D.** The

homology model for the human ezrin dimer overlaid on surfaces representing the averaged (grey) and filtered (green) shapes obtained from the *ab initio* models as calculated by DAMAVER. Each panel shows the models in two orientations related by a 90° rotation about the vertical axis in the plane of the image.

Table 1. Crystallographic data and refinement statistics

PDB	FERM-Ezrin 4RMA	Ezrin-1 4RM9	Ezrin-2 4RM8
A. Data Collection			
Wavelength (Å)	0.954	0.954	0.954
Space Group	P2 ₁	C222 ₁	P2 ₁
Unit cell parameters (Å)	a=48.3 b=110.6 c=66.1 β=101.6°	a=67.4 b=113.5 c=111.8 α=β=γ=90°	a=65.0 b=111.8 c=68.4 β=113.0°
Resolution (Å) ^a	1.75 (1.84-1.75)	2.0 (2.11-2.00)	1.9 (2.00-1.90)
R _{merge} (%) ^a	7.0 (70.1)	11.8 (80.9)	11.2 (34.2)
Total reflections (Unique)	249580 (66327)	401954 (28759)	242283 (67737)
<I/(σI)> ^a	10.9 (1.9)	14.1 (3.7)	7.1 (2.0)
Completeness (%) ^a	97.0 (94.6)	98.3 (97.7)	95.9 (92.0)
Multiplicity ^a	3.8 (3.6)	14.0 (14.4)	3.6 (2.0)
Wilson B-factor (Å ²) ^b	21.2 (25.9, 17.9, 13.6)	24.9 (41.1, 27.5, 17.9)	24.4 (28.1, 26.7, 19.3)
B. Refinement			
Resolution Range (Å)	35.97-1.75	40.24-2.00	39.40-1.90
Total number of atoms (Protein atoms)	5437 (4942)	3233 (3086)	6776 (6198)
R-factor	20.0	20.2	17.4
R-free	24.3	25.3	21.2
Average B factor (Å ²)	32.3	45.8	33.3
Ramachandran plot (%) ^c			
Most favored	98.46%	96.2%	99.2%
Outliers	0.2% (Asp252-B)	0.3% (Lys63)	0.1% (Asp252-A)
RMS bond length (Å) ^d	0.013	0.003	0.005
RMS bond angles ^d	1.345°	0.736°	0.894°
MolProbity Clash Score ^c	4.42	6.29	3.46
MolProbity Overall Score ^c	1.49	1.80	1.14

^aPair values correspond to overall and outer shell (in parentheses) statistics

^bValues in parentheses correspond to maximum likelihood-based estimation of overall anisotropic Wilson B tensor obtained using *phenix.xtriage* tool from PHENIX suite [35]

^cValues calculated using MolProbity [55]

^dfrom PHENIX [35]

Table 2. SAXS analysis

	Ezrin Monomer	Ezrin Dimer
Data		
Q range	0.00531 – 0.954 \AA^{-1}	0.00531 – 0.954 \AA^{-1}
$D = 2\pi/Q$	1200 – 6.6 \AA	1200 – 6.6 \AA
Guinier analysis		
Points in Guinier region (point numbers)	31 (3-33)	8 (2-9)
$Q \times R_g$ range	0.209 – 0.997	0.537 – 0.933
R_g (Guinier best fit)	39.4 \pm 0.3 \AA	91 \pm 4 \AA
$P(r)$ analysis		
Q_{\max} used for $P(r)$	0.5 \AA^{-1} (points 3-375)	0.5 \AA^{-1} (points 2-370)
Regularization (smoothing) parameter "alpha" [†]	1.0	1.0
D_{\max} , best estimate, (range)	165 \AA (160-175 \AA)	325 \AA (320-335 \AA)
R_g ($P(r)$), using best D_{\max} ,	41 \pm 1 \AA	95 \pm 5 \AA

[†]Smoothing parameter used in GNOM [38]

Table 3. Fitting atomic models to SAXS data

Model	Crystal Structure only[†]	Crystal structure + Coiled-coil	Crystal structure + Coiled-coil + helix αA[‡]	Crystal structure + Coiled-coil + helix αA + $\beta 8$[§]	Full homology model
Residues	1-297 516-586	1-297 354-467 516-586	1-467 516-586	1-467 496-586	1-586
Completeness of model	63.1%	82.6%	92.7%	95.6%	100%
Monomer					
χ value [¶]	12.99	4.12	2.26	2.89	3.92
# atoms	3075	4027	4516	4659	4873
Dimer					
χ value [¶]	10.98	3.19	3.58	3.98	4.45
# atoms	6150	8054	9032	9318	9746

[†]Crystal structure of the ezrin FERM:C-terminal domain complex (4RM8).

[‡]FERM domain continuously linked to coiled-coil via helical domain helix αA plus full helix αB .

[§]C-terminal domain extended N-terminally so that it starts with the additional β -strand $\beta 8$ which is antiparallel to β -strand $\beta 5$ in subdomain F3.

[¶] χ value calculated by FoXS [43].

Figure 1

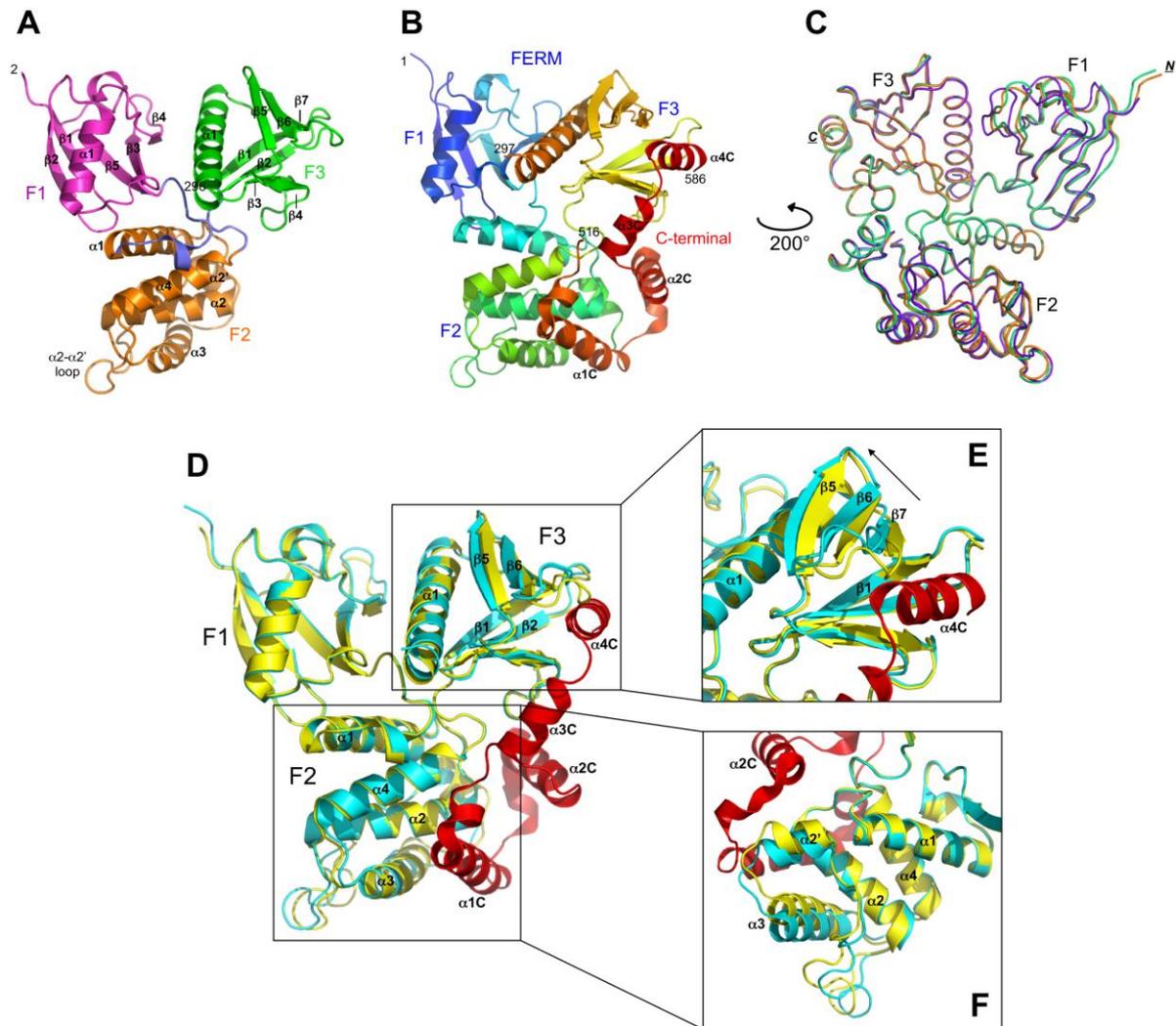


Figure 2

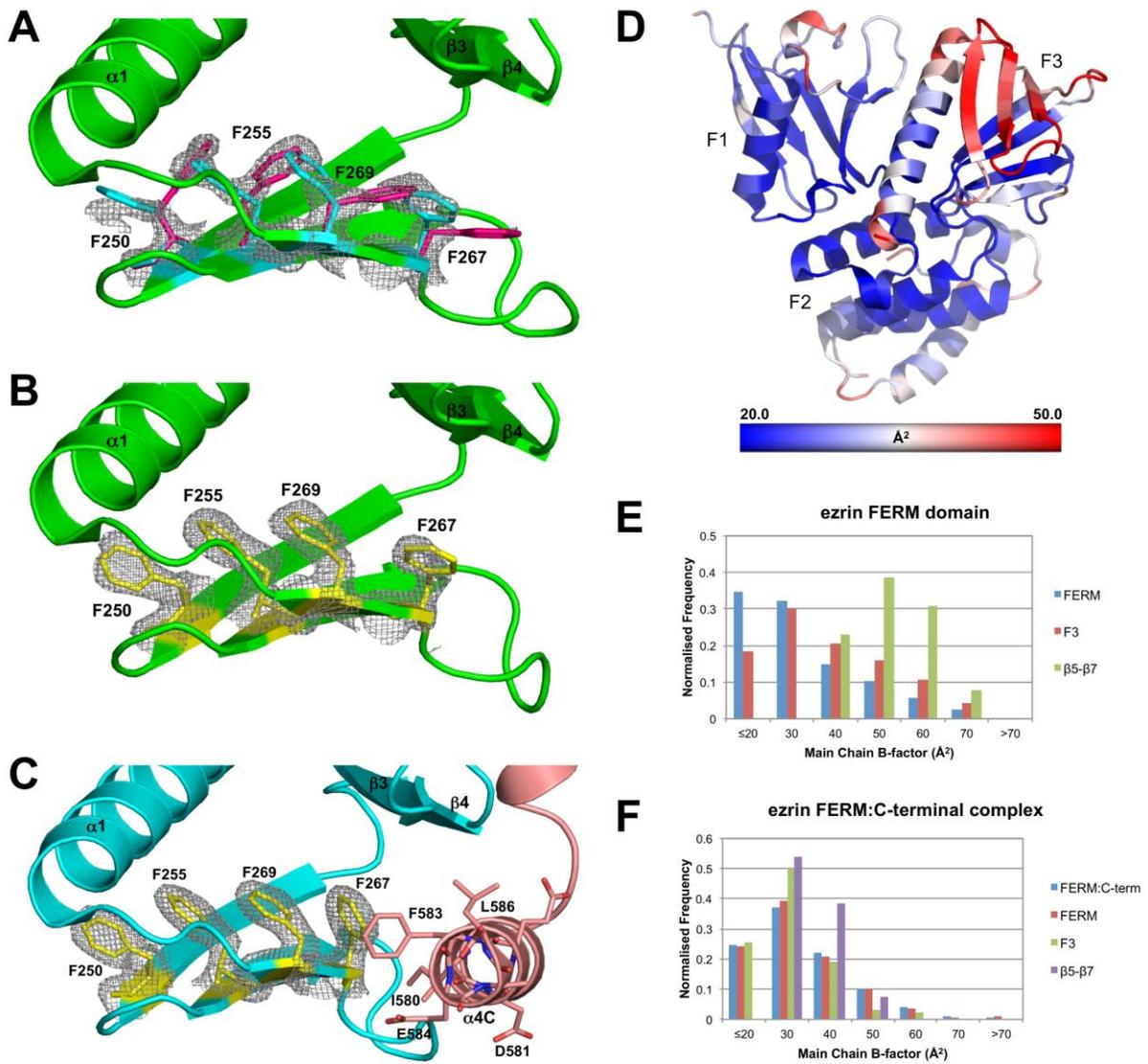
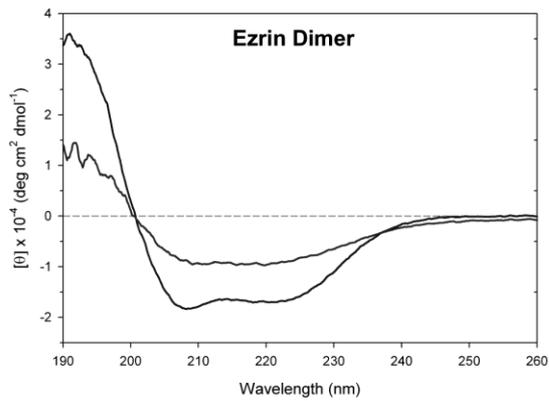
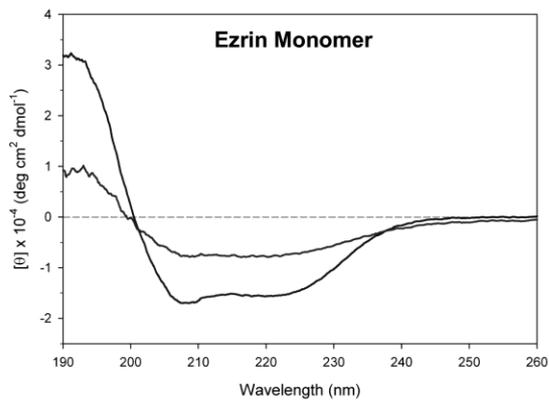
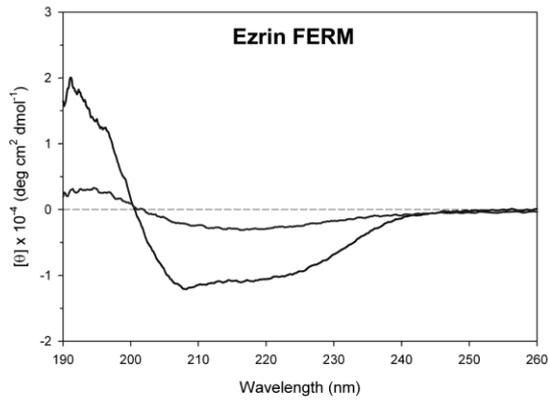


Figure 3

A



B

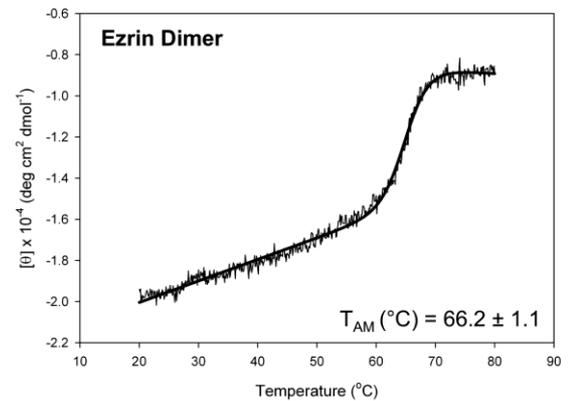
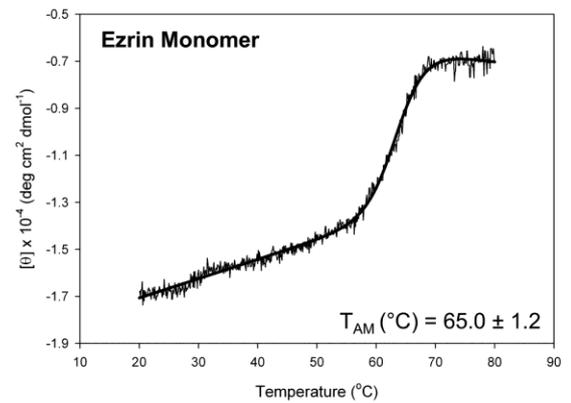
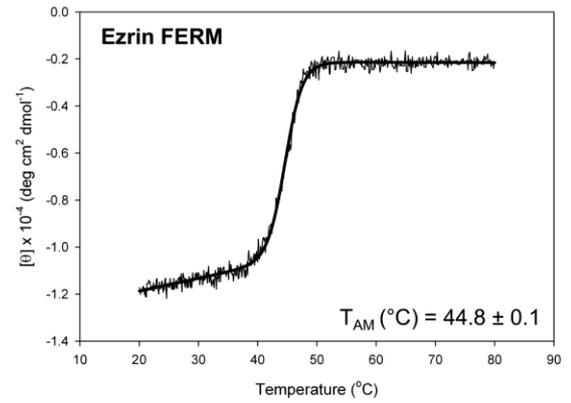


Figure 4

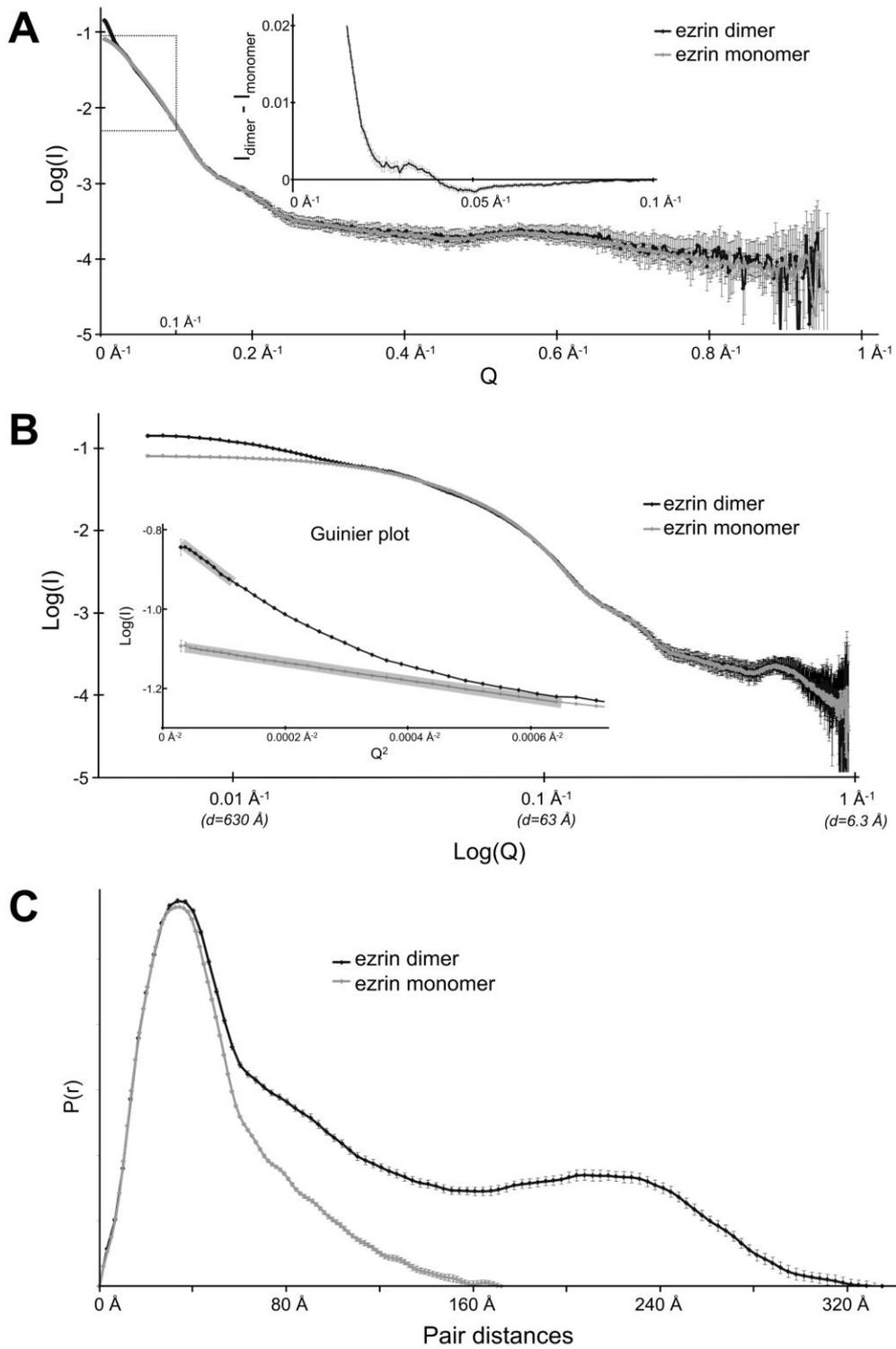


Figure 5

