The EH1 Domain of Eps15 Is Structurally Classified as a Member of the S100 Subclass of EF-Hand-Containing Proteins†

Brian Whitehead,‡,§ Marco Tessari,‡,§ Alfonso Carotenuto,‡,|| Paul M. P. van Bergen en Henegouwen,⊥ and Geerten W. Vuister*,‡

Nijmegen NSR Center for Molecular Structure, Design and Synthesis, Laboratory of Biophysical Chemistry, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands, and Department of Molecular Cell Biology, Institute Biomembranes, University of Utrecht, Padualaan 8, 3584 CH Utrecht, The Netherlands

Received April 23, 1999; Revised Manuscript Received June 14, 1999

ABSTRACT: The Eps15 homology (EH) domain is a protein−protein interaction module that binds to proteins containing the asparagine-proline-phenylalanine (NPF) or tryptophan/phenylalanine-tryptophan (W/FW) motif. EH domain-containing proteins serve important roles in signaling and processes connected to transport, protein sorting, and organization of subcellular structure. Here, we report the solution structure of the apo form of the EH1 domain of mouse Eps15, as determined by high-resolution multidimensional heteronuclear NMR spectroscopy. The polypeptide folds into six α-helices and a short antiparallel β-sheet. Additionally, it contains a long, structured, topologically unique C-terminal loop. Helices 2−5 form two EF-hand motifs. Structural similarity and Ca^{2+} binding properties lead to classification of the EH1 domain as a member of the S100 subclass of EF-hand-containing proteins, albeit with a unique set of interhelical angles. Binding studies using an eight-residue NPF-containing peptide derived from RAB, the cellular cofactor of the HIV Rev protein, show a hydrophobic peptide-binding pocket formed by conserved tryptophan and leucine residues.

Growth factors are involved in the regulation of various cellular functions such as cell growth, differentiation, and movement. Continued activation of growth factor-induced signaling results in aberrations in these functions, often leading to malignant transformation of cells and tissues. Receptor downregulation is an important mechanism in the negative control of this signaling. Upon stimulation of the growth factor receptor, a process is initiated leading to the internalization and subsequent degradation of the activated receptors in the lysosomes. Recent data have revealed that Eps15[1] and the closely related Eps15R are important factors in the receptor-mediated endocytosis of growth factor receptors. Eps15 was originally discovered as a substrate of the EGF receptor (*1*). It is constitutively bound to the ear of α-adaptin and has been localized in coated pits and coated vesicles (*2*−*4*). Direct evidence for a role of Eps15 in endocytosis has recently come from experiments where antibodies directed against Eps15 and Eps15R were microinjected into the cell and were found to inhibit internalization of EGF and transferrin. In addition, expression of fragments of Eps15 was also shown to inhibit EGF receptor internalization (*5*, *6*).

Analysis of the primary sequence indicated that Eps15 is composed of three different domains (*1*). The C-terminal part of the protein contains two proline-rich motifs that were shown to bind to the SH3 domains of either Grb2 or Crk. In addition, this part of the protein contains three binding sites for α-adaptin and multiple Asp-Pro-Phe repeats with unknown function. The central third of Eps15 is assumed to form a coiled-coil region, which induces oligomerization. In the N-terminal part of the molecule, three copies of a domain of approximately 100 amino acids are present which has been designated as the Eps15 homology (EH) domain (*7*). The EH domain is a protein−protein interaction module that binds to proteins containing the asparagine-proline-phenylalanine (NPF) or tryptophan/phenylalanine-tryptophan (W/FW) motif (*8*, *9*). Among the proteins that bind to the EH domain of Eps15 are Epsine, an AP-2 binding protein, RAB, the cellular cofactor of HIV REV protein, and synaptojanin 1, a major Grb2-binding protein in nerve terminals (*10*−*12*). This suggests a general role for EH domains in signaling, neuronal development, and synaptic transmission.

The EH domain has been found in different proteins originating from yeast, nematodes, and mammals. So far,

† This research was supported by the Netherlands Foundation for Chemical Sciences (CW) with financial assistance from the Netherlands Organization for Scientific Research (NWO).

* To whom correspondence and requests for material should be addressed. E-mail: vuister@nmr.kun.nl. Fax: +31-24-265-2112. Phone: +31-24-365-2321.

‡ University of Nijmegen.

§ Both authors contributed equally to this work.

|| On leave from the Department of Pharmaceutical Sciences, University of Salerno, Piazza Vittorio Emanuele 20, Penta di Fisciano (Salerno), Italy.

⊥ University of Utrecht.

[1] Abbreviations: EGF, epidermal growth factor; EH, Eps15 homology domain; Eps15, EGF receptor substrate clone 15; Eps15R, Eps15-related protein; hEH2, human Eps15 EH2 domain; HNHA, amide proton to α-proton quantitative *J* correlation experiment; HSQC, heteronuclear single-quantum correlation; mEH1, mouse Eps15 EH1 domain; mEH2, mouse Eps15 EH2 domain; NMR, nuclear magnetic resonance; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; NPF, SSTNPFL peptide; SH2, Src homology 2 domain; SH3, Src homology 3 domain.

10.1021/bi990922i CCC: $18.00 © 1999 American Chemical Society
Published on Web 08/10/1999

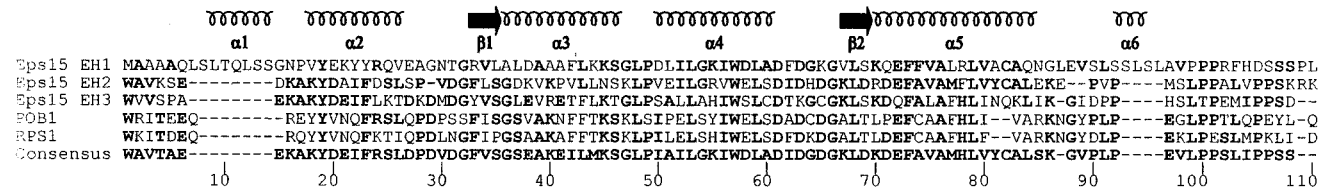

FIGURE 1: Secondary structure of the apo mEH1 domain and sequence alignment of various mammalian EH domains. Alignment was performed using the Clustal algorithm (Megalign, DNAstar Inc.). The consensus sequence is derived on the basis of both mammalian and yeast EH domains.

five mammalian EH-containing proteins have been described, Eps15, Eps15R, Reps1, POB1, and intersectin. While Eps15 and Eps15R contain three copies of the EH domain, two copies are found in intersectin. Only one EH domain is present in Reps1 and POB1. A sequence alignment of various mammalian EH domains is shown in Figure 1. Compared to the consensus EH domain, the EH1 domain of Eps15 is characterized by two inserts of eight and four residues, respectively, which are an integral part of the domain (vide infra).

Here, we report the solution structure of the apo form of the N-terminal EH domain of mouse Eps15 (mEH1). The mEH1 domain is classified as a member of the S100 subclass of EF-hand-containing proteins, albeit with a unique set of interhelical angles. The mEH1 peptide-binding pocket appears to be formed by a cluster of conserved hydrophobic residues.

MATERIALS AND METHODS

NMR Spectroscopy. mEH1 samples [1−1.5 mM in 100 mM KH_2PO_4/100 mM NaCl (pH 5.2)] were prepared as described by Whitehead et al. (*13*). Additionally, for Ca^{2+} titrations a 1.8 mM sample in 20 mM Tris/100 mM NaCl (pH 7.0) was used. A Ca^{2+} titration on mEH1 at pH 5.2 revealed no significant Ca^{2+} binding. Studies with calmodulin have shown that the Ca^{2+} affinity is reduced at pH 5.8 compared to those at neutral or basic pH values, presumably because of protonation of side chain carboxylate groups involved in Ca^{2+} ligation (*14*). At pH 7.0, large effects are observed in the $^1H-^{15}N$ HSQC spectrum of mEH1. NPF binding was monitored by acquiring two-dimensional $^1H-^{15}N$ HSQC spectra for EH1 at NPF concentrations of 0, 0.3, 0.6, 0.9, 1.2, 1.5, 1.8, and 2.1 equiv of the concentration of the EH1. NPF additions were made by adding aliquots of a 12.5 mM NPF solution in acetate buffer (pH 5.2) to a 1.5 mM apo EH1 sample, mixing thoroughly within an Eppendorf tube. A three-dimensional ^{15}N-separated NOESY−HSQC spectrum was acquired on the final pH 5.2 2.1:1 NPF/EH1 sample.

The assignment procedure for all ^1H, ^{15}N, and ^{13}C resonances was reported previously (*13*). ^{13}C NOESY−HSQC (*15*, *16*), ^{15}N NOESY−HSQC (*17*), and two-dimensional NOESY (*18*), all with a 100 ms mixing time, and HNHA experiments (*19*) were carried out at 25 °C on Varian Unity*Inova* 500, Bruker DRX 600, and Varian Unity*Inova* 750 spectrometers. The program NMRPipe (*20*) was used for transformation of all data. Analysis was performed using the program XEASY (*21*) and yielded a total of 1330 distance constraints (495 intraresidual, 394 sequential, 190 medium-range, 197 long-range, and 54 hydrogen bond) and 38 ϕ dihedral constraints. Distance constraints originating from the three different NOE spectra were calibrated in three separate sets using the routine Caliba of the program Dyana 1.5 (*22*), yielding a contiguous distribution of upper limits. After a first round of structure calculations and convergence, each of the three sets of NOEs was recalibrated with respect to the average distances observed in the initial ensemble of structures. Dihedral constraint ranges were derived on the basis of the Karplus curve for $^3J_{HNH\alpha}$ (*19*) assuming a 1 Hz error on the experimental value.

Structure Calculations. The data served as input for a torsion angle dynamics structure calculation as implemented in the program Dyana 1.5 (*22*). One hundred structures were calculated, 30 of which were selected on the basis of low overall target function score (0.24−2.03 Å^2). Quality assessment using the program Procheck NMR (*23*) shows 72.1% in the most favored regions, 24.3% in the additional allowed regions, 2.3% in the generously allowed regions, and 1.3% in the disallowed regions. In the ensemble of 30 structures, a grand total of only nine violations larger than 0.4 Å (maximum violation of 0.6 Å) and no dihedral violation larger than 5° were observed. rmsd values of experimental NOE and dihedral restraints were 0.017 ± 0.005 Å (all), 0.016 ± 0.006 Å (intraresidual), 0.019 ± 0.010 Å (sequential), 0.015 ± 0.005 Å (medium-range), 0.015 ± 0.004 Å (long-range), 0.006 ± 0.002 Å (hydrogen bond), and 0.5 ± 0.2° (dihedral). Pairwise rmsds were as follows: 0.91 ± 0.21 Å for the backbone (residues 7−105), 1.40 ± 0.21 Å for heavy atoms (residues 7−105), 0.59 ± 0.16 Å for backbone secondary structural elements (residues 9−14, 18−26, 33−46, 50−60, and 67−84), and 1.15 ± 0.15 Å for heavy atom secondary structural elements. Solvent accessibilities were calculated using a probe radius of 1.4 Å. Figures were generated using the program MOLMOL 2.6 (*24*). The ensemble of 30 NMR structures was submitted to the PDB (file name 1QJT).

RESULTS AND DISCUSSION

The solution structure of the apo mEH1 domain was determined using high-resolution multidimensional heteronuclear NMR spectroscopy following established protocols (*25*). The $^{15}N-\{^1H\}$ heteronuclear NOE data (not shown) and the deviations from random coil chemical shifts of $^1H^\alpha$, $^{13}C^\alpha$, and $^{13}C^\beta$ (*13*) indicated that the polypeptide was folded between residues Leu^7 and Asp^{105}. Indeed, no $^1H-^1H$ NOEs indicative of secondary or tertiary structure were found for the residues outside this region.

The superposition of the resulting ensemble of 30 NMR structures with the lowest torsion angle dynamics target function is shown in Figure 2B. Analysis of backbone ϕ and ψ angles, sequential- and medium-range NOEs, and H-bond

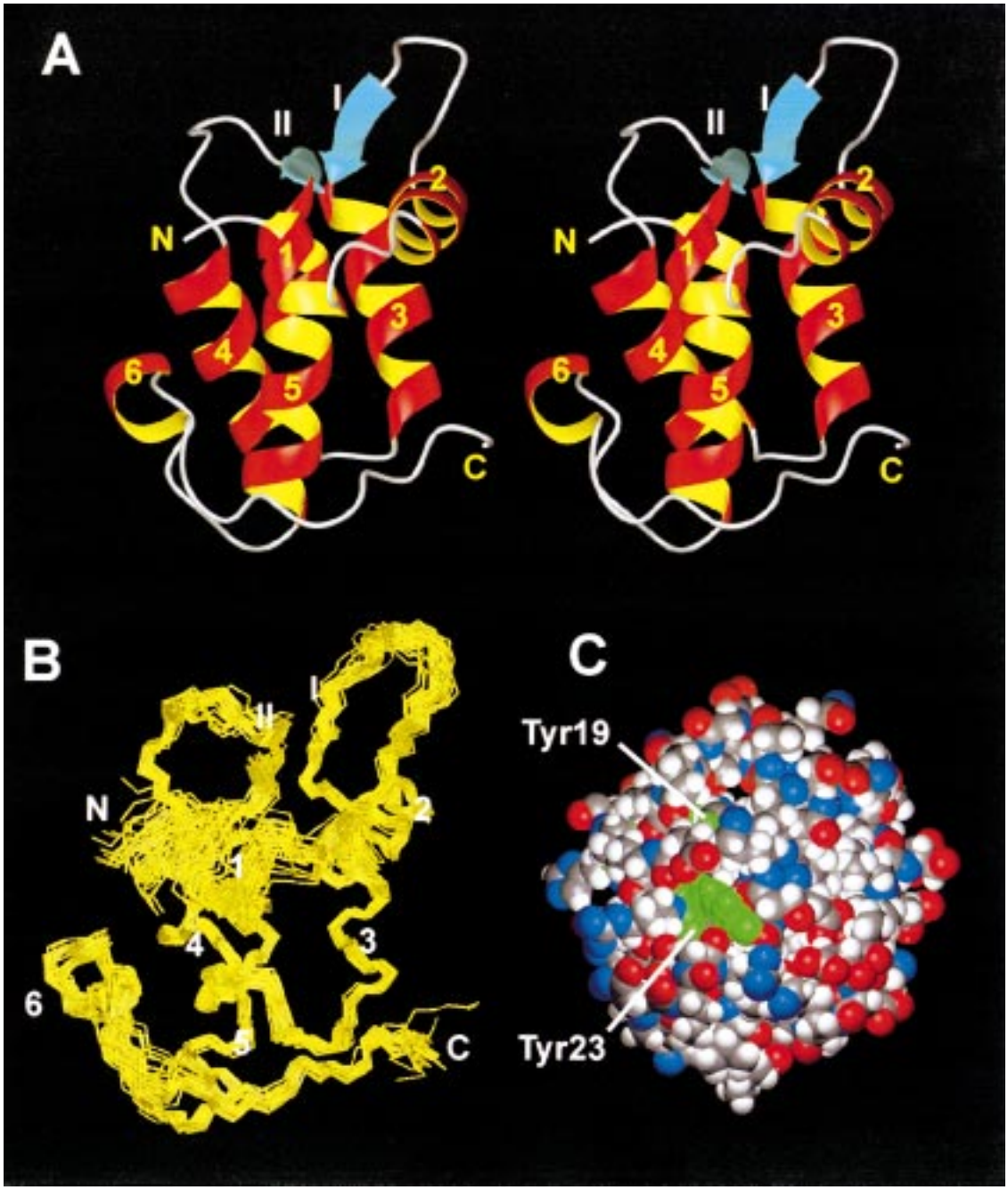

FIGURE 2: (A) Stereoview of the ribbon diagram of the apo mEH1 solution structure with the lowest overall target function score. The five helices (Leu^{9}−Ser^{14}, Val^{18}−Val^{26}, Ala^{36}−Ser^{46}, Asp^{50}−Ala^{60}, and Lys^{70}−Gln^{84}), the short antiparallel β-sheet (Arg^{33}−Leu^{35} and Val^{67}−Ser^{69}), and the helical turn (Ser^{92}−Leu^{94}) are denoted. A tightly packed hydrophobic core is formed by residues Val^{26}, Val^{34}, Ala^{36}, Ala^{39}, Ser^{46}, Ile^{56}, Ala^{60}, Val^{67}, Phe^{73}−Leu^{77}, and Leu^{79}−Ala^{83}. Of these, several are absolutely conserved within the EH domain family (cf. Figure 1). Residues Val^{89}, Ser^{90}, Leu^{94}, Leu^{96}, Val^{98}, Pro^{101}, and Phe^{103} of the long C-terminal loop also pack into the hydrophobic core, which is exemplified by the many long-range NOE contacts of these residues. (B) Superposition of the ensemble of 30 structures of apo mEH1. (C) Space-filling representation of the apo mEH1 solution structure. Tyr^{19}, Tyr^{22}, and Tyr^{23} are green. Solvent exposure of Tyr^{22} is more pronounced at the protein face that is invisible in this view.

patterns shows the presence of five α-helices (Leu^{9}−Ser^{14}, Val^{18}−Val^{26}, Ala^{36}−Ser^{46}, Asp^{50}−Ala^{60}, and Lys^{70}−Gln^{84}). Additionally, two extended stretches of three residues each (Arg^{33}−Leu^{35} and Val^{67}−Ser^{69}) form a short, strongly twisted, antiparallel β-sheet. The C-terminal part of the protein, Asn^{85}−Phe^{103}, forms a long structured loop that contains a three-residue helical turn (Ser^{92}−Leu^{94}). The secondary structural elements are well-defined, displaying an overall backbone rmsd of 0.45 Å with respect to the average. In several of the loop regions connecting the secondary structural elements, a somewhat larger spread is observed. In all, the rmsd for the backbone is 0.65 Å (residues 7−105).

Helices 2−5 and the small β-sheet adopt the characteristic fold of two helix−loop−helix EF-hand motifs (*26*). This motif was also identified in the structure of the Ca^{2+}-complexed form of the EH2 domain of the human Eps15 (hEH2) (*27*) and Ca^{2+}-complexed form of the EH domain of POB1 (*28*). We also identified this motif in the EH2 domain of mouse Eps15 (mEH2) (*29*). Helix 2 in apo mEH1 is two residues shorter and has shifted toward the C-terminus by three residues when compared to the equivalent helices

in both mEH2 and hEH2. These latter two differences allow for folding of the eight-residue insert of apo mEH1 (cf. Figure 1) into helix 1. This helix is followed by a short four-residue turn that orients helix 1 approximately perpendicular to helix 2.

The two EF-hands are connected by a short three-residue loop ($Gly^{47}-Pro^{49}$) and structurally connected by the small antiparallel β-sheet. Since helix 3 is extended at its C-terminus by two residues with respect to the equivalent helix in the hEH2 structure, the loop connecting helices 3 and 4 is concomitantly shorter. The complete C-terminal loop forms an integral part of the apo mEH1 domain. In particular, Phe^{103} is in close proximity to many residues in helix 3 and helix 5. Indeed, this result appears to be in agreement with EH domain interaction studies that revealed this C-terminus to be essential for interaction with its targets (*9*). The loop almost traverses one face of the protein. As a result, the first structured N-terminal residue and last structured C-terminal residue of apo mEH1 are located at opposite sites of the protein. In contrast, the N- and C-termini of the hEH2 domain are found in close proximity (*27*). The EH1 and EH2 domains in Eps15 are separated by a linker of 17 residues ($Ser^{106}-Trp^{122}$). Consequently, the relative orientation of the two domains will also depend on the conformation and flexibility of the linker residues. In the absence of the EH2 domain, this linker appears to be largely unstructured.

Structural comparison using the Dali server at the EMBL (http://www2.embl-ebi.ac.uk/dali) shows the class of EF-hand-containing proteins, in particular calbindin D_{9K} (PDB file name 5ICB), is structurally related. No structural similarities, however, are observed for the C-terminal part of the apo mEH1 domain.

Upon activation of the EGF receptor, Eps15 becomes phosphorylated on one or more tyrosine residues (*1*). The mEH1 domain can be phosphorylated in vitro by tkx bacteria which harbor an inducible lck kinase, and mutation of the three tyrosine residues of mEH1 (Tyr^{19}, Tyr^{22}, and Tyr^{23}) in Eps15 abolishes its monoubiquitination upon stimulation of the cells with EGF (data not shown). Tyr^{19}, Tyr^{22}, and Tyr^{23} in mEH1 are all part of helix 2. Of these, Tyr^{23} is most accessible, in particular for its aromatic side chain atoms (Figure 2C), which would allow for relatively easy access of a kinase. However, the aromatic side chains of Tyr^{19} and Tyr^{22} are also (partially) solvent-exposed, so their involvement cannot be absolutely ruled out. Interestingly, all other mammalian EH domains contain a phenylalanine residue at the position homologous to Tyr^{23}, indicating that the presence of this tyrosine residue in mEH1 may be specific and tied to the regulation of Eps15.

The EF-hand (*26*) is an important Ca^{2+} binding motif found in a large number of proteins that control a wide range of cellular activities, including cell division and metabolic processes. The EF-hand protein family includes calmodulin, calbindin D_{9K}, the S100 proteins, and troponin C. Structures of apo and Ca^{2+}-complexed EF-hand proteins have revealed that Ca^{2+} binding results in the exposure of a hydrophobic surface, allowing interaction with target proteins (*30*).

Two types of EF-hands have been identified so far (Figure 3). The first type, the canonical EF-hand, as found in calmodulin and calbindin D_{9K} site II, contains a calcium-binding loop of 12 residues (*26*). A second type is found in site I of the S100 family, examples of which include S100B, S100A6, and calbindin D_{9K}. This type, known as the pseudo EF-hand, contains a modified Ca^{2+} binding loop consisting of 14 residues (*31*). In the mEH1 sequence, residues $Gln^{25}-Asp^{38}$ and $Asp^{61}-Glu^{72}$ resemble a pseudo and canonical EF-hand motif, respectively.

Calcium binding of EH1 was tested by titrating an ^{15}N sample of apo mEH1 and by monitoring the ^{1}H and ^{15}N chemical shifts. mEH1 binds Ca^{2+} with low affinity as was evident from the fast exchange regime. The ^{15}N chemical shift changes upon titrating Ca^{2+} are mapped onto a ribbon diagram of the apo mEH1 structure in Figure 3C. The largest changes were observed for residues Val^{26}, Ala^{28}, Gly^{29}, Thr^{31}, and Val^{34} in site I and Leu^{68} in site II. In canonical EF-hands, the Ca^{2+} is coordinated by the backbone carbonyl in position 7 of the loop. ^{15}N chemical shift changes upon Ca^{2+} binding for the nucleus in position 8 were found to be in the range of 4−8 ppm (*32*). Leu^{68} is at the eighth position of site II, in perfect agreement with the observed shift ranges for other EF-hands. In pseudo EF-hands, the Ca^{2+} is coordinated by the backbone carbonyls in positions 1, 4, 6, and 9 of the loop (cf. Figure 3B), and ^{15}N chemical shift changes for positions 2, 5, and 7 are characteristic for this coordination scheme. For site I of mEH1, we observed the largest shift changes for positions 2, 4, 5, 7, and 10. Our titration indicated that this site is not yet saturated, even at a Ca^{2+}:EH1 ratio of 10:1. The partial occupancy of site I results in smaller chemical shift changes as compared to those of the S100 class of proteins, but our results still clearly identify the Ca^{2+}-binding sites. S100 proteins have affinities of 10−50 μM for the canonical EF-hand and 200−500 μM for the pseudo EF-hand (*33*), but the Ca^{2+} affinity of EH1 is lower. On closer examination, however, it is clear that neither Ca^{2+} binding loop in mEH1 contains the full EF-hand consensus sequence; the residue in position 11 normally provides a side chain oxygen ligand for metal binding in pseudo EF-hands. In the mEH1 domain, however, the residue at this position is a leucine (Leu^{35}). Similarly, the residue in position 5 of canonical EF-hands is normally an aspartate/asparagine, providing a side chain carbonyl for Ca^{2+} ligation. In mEH1, this residue is a lysine (Lys^{65}). Studies with EF-hands containing mutated residues in such important positions have shown dramatically reduced Ca^{2+} affinity in many cases (*34*, *35*). Interestingly, the Eps15 hEH2 domain, the POB1 EH domain, and the Eps15 mEH2 domain contain an aspartate at the homologous position, in agreement with the observed high affinity for Ca^{2+} for this site (*27*−*29*). It has recently been suggested that Ca^{2+} binding may not be important for the biological activity of the S100 family (*36*), as the affinities are too low to allow significant Ca^{2+} complexation at intracellular Ca^{2+} concentrations ($10^{-6}-10^{-5}$ M). The low affinity of EH1 for Ca^{2+} certainly precludes significant Ca^{2+} binding at such calcium concentrations, unless the formation of complexes with target proteins significantly raises the Ca^{2+} affinity.

The presence of a pseudo and canonical EF-hand classifies the mEH1 domain as a member of the S100 subclass of EF-hand-containing proteins. Proteins of this class serve important roles in regulatory processes such as endocytosis, exocytosis, intracellular signal transduction, and regulation of cell structure (*37*). The S100 proteins form homo- or heterodimers with a tissue-specific expression and show modest variation in structure upon binding Ca^{2+} (*36*, *38*).

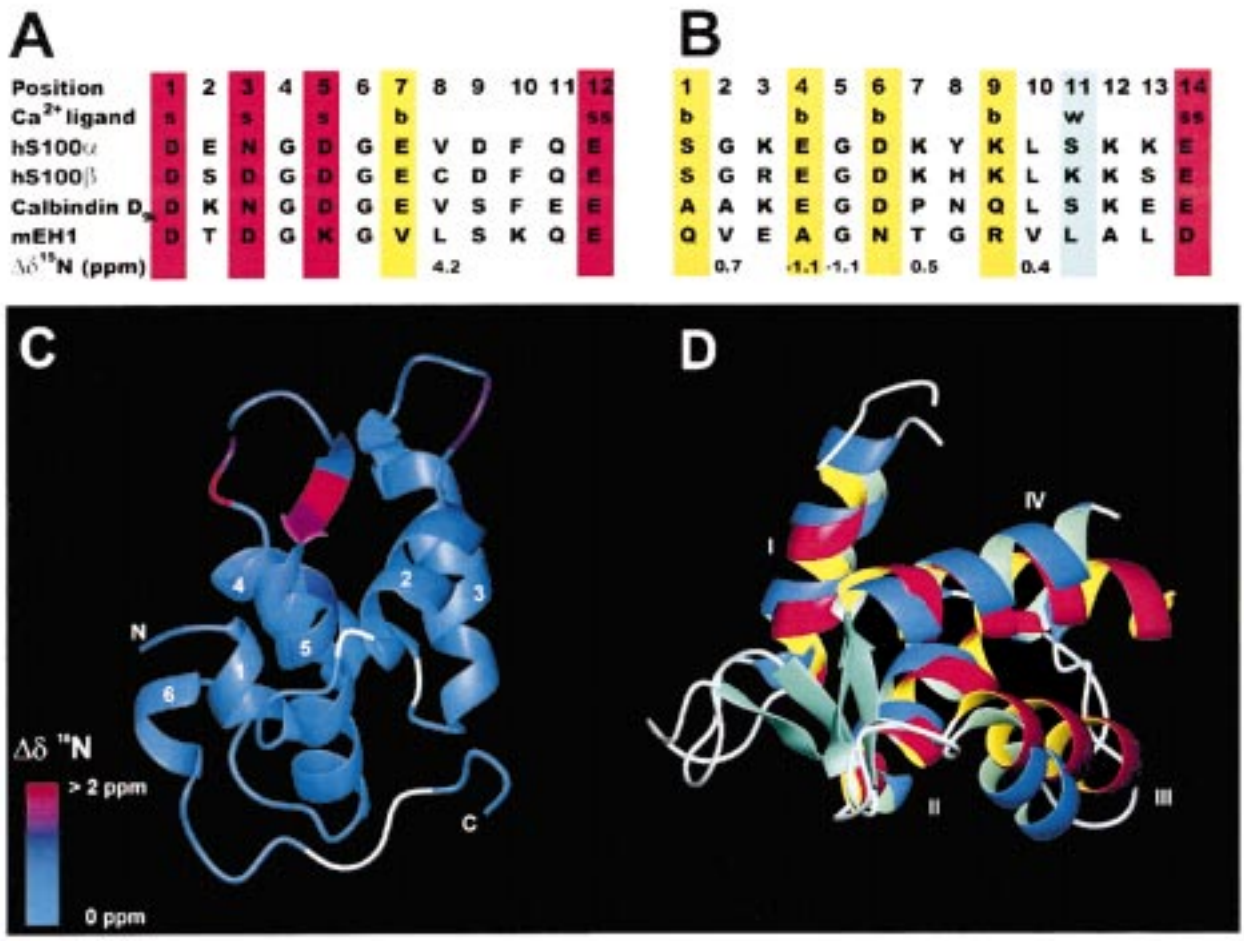

FIGURE 3: (A) Alignment of sequences containing a canonical EF-hand. Canonical EF-hands contain a calcium-binding loop of 12 residues, coordinating the Ca^{2+} through three side chain carbonyl oxygens and one bidentate side chain carboxylate (positions 1, 3, 5, and 12; boxed red) and one backbone carbonyl oxygen (position 7; boxed yellow) (*26*). (B) The pseudo EF-hand contains a modified Ca^{2+} binding loop consisting of 14 residues (*31*), coordinating the Ca^{2+} by the backbone carbonyl oxygens (positions 1, 4, 6, and 9; boxed yellow), one bidentate side chain carboxylate (position 14; boxed red), and a side chain oxygen ligand mediated through a water molecule (position 11; boxed blue). (C) Ribbon diagram of the apo mEH1 solution structure color-coded according to the observed ^{15}N shift changes ($\Delta\delta$ ^{15}N) upon binding Ca^{2+}. (D) Ribbon diagrams of the superposition of the apo mEH1 (red and yellow helices) and apocalbindin D_{9k} (blue and cyan helices; PDB code 1CLB). The backbone atoms of the two EF-hand motifs (residues 18−44 and 54−81 in mEH1 and 6−32 and 47−74 in calbindin D_{9k}) can be superimposed with an rmsd of 2.8 Å.

Table 1: Interhelical Angles

	I−II (deg)	I−III (deg)	I−IV (deg)	II−III (deg)	II−IV (deg)	III−IV (deg)
mEH1 (apo)[a]	115.9 ± 1.0	−111.6 ± 1.4	92.2 ± 1.0	133.1 ± 1.3	−34.7 ± 1.1	145.1 ± 1.6
calbindin D_{9k} (apo)[b]	123 ± 3	−109 ± 9	128 ± 4	124 ± 7	−34 ± 4	118 ± 8
calbindin D_{9k}[b]	130 ± 3	−111 ± 6	135 ± 4	113 ± 6	−28 ± 4	110 ± 6
S100A6 (apo)[c]	128 ± 11	−84 ± 18	118 ± 13	147 ± 12	−23 ± 8	147 ± 13
S100A6[c]	110 ± 8	−81 ± 16	126 ± 5	161 ± 10	−36 ± 13	135 ± 12

[a] Mouse EH1: helix I, residues 18−26; helix II, residues 36−46; helix III, residues 50−60; and helix IV, residues 70−84. [b] Data from Skelton et al. (*38*). [c] Data from Sastry et al. (*36*). S100A6 is also known as calcyclin (*42*).

Interhelical angles of the EF-hand motifs of apo mEH1 are compared in Table 1 with those of proteins containing the pseudo EF-hand motif. Of these, only calbindin D_{9k} is monomeric in solution. Significant differences exist for interhelical angles I−IV and III−IV of EH1 with respect to both apo and Ca^{2+}-bound forms of calbindin D_{9k}. The two EF-hand motifs in EH1 and apocalbindin can be superimposed with an rmsd of 2.8 Å. The resulting superposition is shown in Figure 3D. The canonical EF-hand motif in mEH1 is rotated with respect to the motif in calbindin, and the helices of this motif also exhibit a larger interhelical angle, which potentially may be related to its role in peptide binding (vide infra). Interhelical angle III−IV of apo S100A6, a homodimeric S100 class protein, is within experimental error equal of that of the mEH1 domain. However, in this case, interhelical angles I−III and I−IV appear to be different, again pointing to a rotation of the first EF-hand motif with respect to the second motif. A rotation of the canonical EF-hand motif was also observed for the Ca^{2+}-complexed POB1 EH domain (*28*).

In mEH1, chemical shift changes resulting from binding Ca^{2+} are not localized in the Ca^{2+}-binding EF-hand motifs, but also significantly affect more distant regions such as residues Ser^{8} and Leu^{9} in helix 1 and Ala^{41} and Leu^{43} in

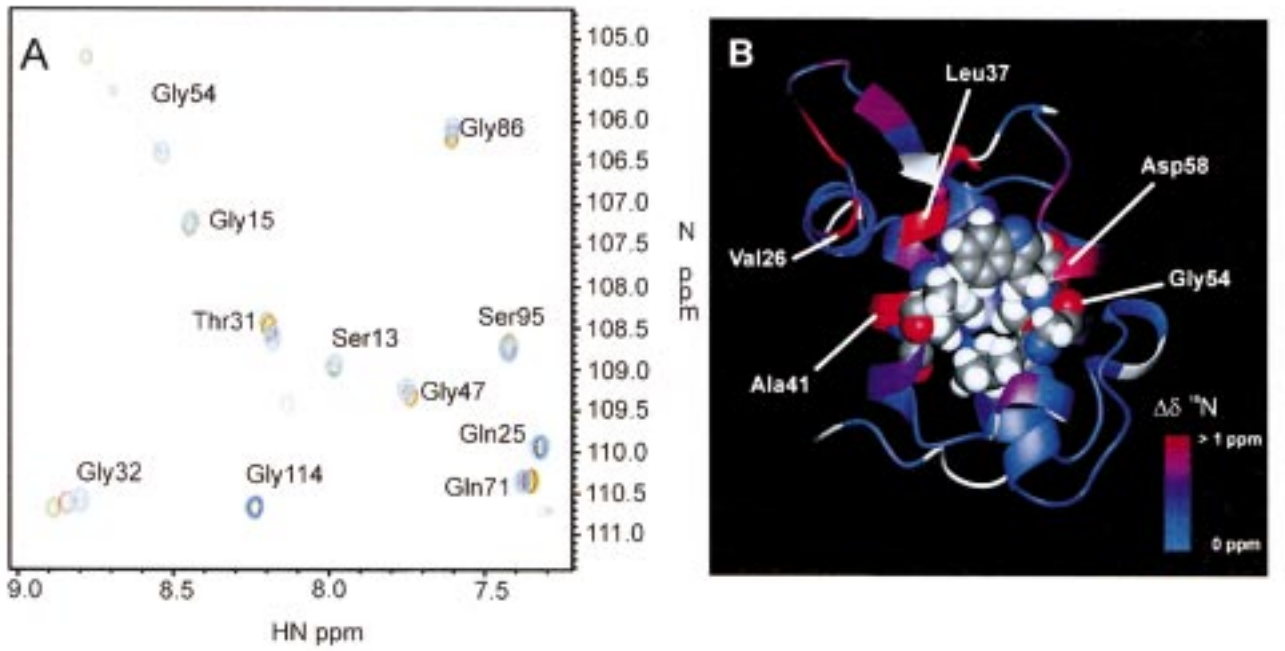

FIGURE 4: (A) Superposition of a small region of the $^{15}N-^{1}H$ HSQC spectrum of EH1 at 0 (blue), 0.9 (red), and 2.1 equiv (green) of the SSSTNPFL peptide. (B) Ribbon diagram of the apo mEH1 solution structure color-coded according to the observed ^{15}N shift changes ($\Delta\delta$ ^{15}N) upon binding the SSSTNPFL peptide. Residues in the large hydrophobic cleft (Ala[40], Leu[43], Leu[53], Gly[54], Ile[56], and Trp[57]) are shown in a space-filling representation.

helix 3. As ^{15}N chemical shifts in particular are sensitive to backbone conformation, this suggests conformational rearrangements in this part of the polypeptide. However, our data also indicate that the secondary structural elements remain conserved upon binding Ca^{2+}.

The EH domain binds to proteins containing the NPF motif or a W/FW motif (*8*, *9*). Mutation studies with an eight-residue peptide SSSTNPFL from RAB, the cellular cofactor of HIV REV protein, which strongly binds to full-length Eps15, revealed the specificity of the NPF binding (*8*). Mutation of any of the amino acids in this motif (e.g., F-Y) completely abolished the binding to Eps15. Alanine scanning revealed that residues −2, −1, 1, and 2 with respect to the NPF motif are also important, whereas mutations of residues −3 and −4 did not affect the binding.

To investigate NPF binding to mEH1, we titrated the SSSTNPFL peptide into an ^{15}N-labeled sample of apo mEH1 and acquired an $^{15}N-^{1}H$ HSQC spectrum after each titration point. A gradual shift in resonance positions was observed for several residues (Figure 4A) which establishes the fast exchange regime for complex formation. The data also allow for the determination of the binding constant, which was calculated to be 1700 ± 80 M^{-1}. This value agrees within experimental error with the value obtained for the Ca^{2+}-complexed form of the EH2 domain using surface plasmon resonance techniques (*27*).

Mapping of the chemical shift changes of mEH1 upon addition of NPF onto the tertiary structure of mEH1 (Figure 4B) reveals the interface of interaction. NPF binding predominantly affects residues in helix 3 and helix 4; in particular, Val[26], Leu[37], Ala[41], Phe[42], Gly[54], and Asp[58] experience the largest changes. A large hydrophobic pocket is formed between these helices, which provides a suitable binding site for the NPF peptide. This site differs from the target-binding sites observed for the S100 proteins. The latter site is located at the dimeric interface formed by residues in the linker between helices II and III (i.e., helices 3 and 4 in mEH1) and residues in the C-terminus of helix IV (i.e., helix 5 in mEH1) (*39*−*41*).

Studies with the hEH2 domain (*27*) also showed a hydrophobic pocket similar to that of EH1 as the primary interaction site. Indeed, several of the residues involved are conserved among the EH domain family (cf. Figure 1), indicating the general nature of the recognition of the NPF motif. It was previously shown for the hEH2 domain that the absolutely conserved tryptophan (Trp[57] in mEH1) is essential for binding (*27*). Several of the affected residues in mEH1 are part of the hydrophobic core, suggesting a significant rearrangement of the structure. Recent deletion studies showed the long structured C-terminal loop of the EH domain to be indispensable for peptide binding (*9*). However, in mEH1 only small changes are observed for residues Asn[85]−Asp[105], which points to a more structural role for this part of the protein. Alternatively, the interaction between the EH1 domain and peptide might be mediated by side chains only or involve parts of the EH1 target more distant from the NPF recognition motif.

In analogy to recognition by other protein interaction domains, such as SH2 and SH3, it is to be expected that specificity of interaction is conferred by residues other then the NPF motif, interacting with regions outside the hydrophobic pocket. Indeed, the discrimination between EH domains that preferentially bind type I (NPF) or type II (F/WW) peptides has been correlated with residue +3 with respect to the absolutely conserved Trp residue of the EH domains (Ala[60] in mEH1). Interestingly, this residue is located at the very end of helix 5 and is well buried in the mEH1 structure, rendering it somewhat puzzling how this residue would act as a specificity determinant. Studies designed to unravel these and related questions are currently underway.

The results presented here lead the first structure of an EH1 domain as well as the first structure of an apo EH-domain. The domain is classified as a member of the S100 subfamily of EF-hand-containing proteins. The structure of the apo mEH1 domain differs considerably from those of the Ca^{2+}-complexed hEH2 domain and the Ca^{2+}-complexed POB1 EH domain. Detailed comparisons of all these structures will reveal the structural basis of EH domain recognition and specificity, which forms an important first step in unraveling the role in signaling and processes connected to transport, protein sorting, and organization of the subcellular structure of this novel class of protein interaction domains.

ACKNOWLEDGMENT

We thank Prof. Cees Hilbers for stimulating discussions and support, Jan Aalen for sample preparation, and the SONNMR Large-Scale Facility for the use of the Varian 750 spectrometer.

REFERENCES

1. Fazioli, F., Minichiello, L., Matoskova, B., Wong, W. T., and Di Fiore, P. P. (1993) *Mol. Cell Biol. 13*, 5814−5828.
2. Benmerah, A., Begue, B., Dautry-Varsat, A., and Cerf-Bensussan, N. (1996) *J. Biol. Chem. 271*, 12111−12116.
3. Tebar, F., Sorkina, T., Sorkin, A., Ericsson, M., and Kirchhausen, T. (1996) *J. Biol. Chem. 271*, 28727−28730.
4. van Delft, S., Schumacher, C., Hage, W., Verkleij, A. J., and van Bergen en Henegouwen, P. M. (1997) *J. Cell. Biol. 136*, 811−821.
5. Carbone, R., Fre, S., Iannolo, G., Belleudi, F., Mancini, P., Pelicci, P. G., Torrisi, M. R., and Di Fiore, P. P. (1997) *Cancer Res. 57*, 5498−5504.
6. Benmerah, A., Lamaze, C., Begue, B., Schmid, S. L., Dautry-Varsat, A., and Cerf-Bensussan, N. (1998) *J. Cell. Biol. 140*, 1055−1062.
7. Wong, W. T., Schumacher, C., Salcini, A. E., Romano, A., Castagnino, P., Pelicci, P. G., and Di Fiore, P. (1995) *Proc. Natl. Acad. Sci. U.S.A. 92*, 9530−9534.
8. Salcini, A. E., Confalonieri, S., Doria, M., Santolini, E., Tassi, E., Minenkova, O., Cesareni, G., Pelicci, P. G., and Di Fiore, P. P. (1997) *Genes Dev. 11*, 2239−2249.
9. Paoluzi, S., Castagnoli, L., Lauro, I., Salcini, A. E., Coda, L., Fre', S., Confalonieri, S., Pelicci, P. G., Paolo, D. F., and Cesareni, G. (1998) *EMBO J. 17*, 6541−6550.
10. Di Fiore, P. P., Pelicci, P. G., and Sorkin, A. (1997) *Trends Biochem. Sci. 22*, 411−413.
11. Haffner, C., Takei, K., Chen, H., Ringstad, N., Hudson, A., Butler, M. H., Salcini, A. E., Di Fiore, P. P., and De Camilli, P. (1997) *FEBS Lett. 419*, 175−180.
12. Chen, H., Fre, S., Slepnev, V. I., Capua, M. R., Takei, K., Butler, M. H., Di Fiore, P. P., and De Camilli, P. (1998) *Nature 394*, 793−797.
13. Whitehead, B., Tessari, M., Versteeg, H. H., van Delft, S., van Bergen en Henegouwen, P. M. P., and Vuister, G. W. (1998) *J. Biomol. NMR 12*, 465−466.
14. Milos, M., Schaer, J. J., Comte, M., and Cox, J. A. (1986) *Biochemistry 25*, 6279−6287.
15. Zuiderweg, E. R., McIntosh, L. P., Dahlquist, F. W., and Fesik, S. W. (1990) *J. Magn. Reson. 86*, 210−216.
16. Ikura, M., Kay, L. E., Tschudin, R., and Bax, A. (1990) *J. Magn. Reson. 86*, 204−209.
17. Marion, D., Kay, L. E., Sparks, S. W., Torchia, D. A., and Bax, A. (1989) *J. Am. Chem. Soc. 111*, 1515−1517.
18. Ernst, R. R., Bodenhausen, G., and Wokaun, A. (1987) *Principles of nuclear magnetic resonance in one and two dimensions*, Clarendon Press, Oxford, U.K.
19. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
20. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
21. Bartels, C., Xia, T. H., Billeter, M., Guntert, P., and Wuthrich, K. (1995) *J. Biomol. NMR 6*, 1−10.
22. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
23. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
24. Koradi, R., Billeter, M., and Wuthrich, K. (1996) *J. Mol. Graphics 14*, 51−32.
25. Clore, G. M., and Gronenborn, A. M. (1994) *Methods Enzymol. 239*, 349−363.
26. Strynadka, N. C., and James, M. N. (1989) *Annu. Rev. Biochem. 58*, 951−998.
27. de Beer, T., Carter, R. E., Lobel-Rice, K. E., Sorkin, A., and Overduin, M. (1998) *Science 281*, 1357−1360.
28. Koshiba, S., Kigawa, T., Iwahara, J., Kikuchi, A., and Yokoyama, S. (1999) *FEBS Lett. 442*, 138−142.
29. Carotenuto, A., Tessari, M., Whitehead, B., Aelen, J. M. A., van Bergen en Henegouwen, P. M., and Vuister, G. W. (1999) *J. Biomol. NMR 14*, 97−98.
30. Ikura, M. (1996) *Trends Biochem. Sci. 21*, 14−17.
31. Heizmann, C. W., and Hunziker, W. (1991) *Trends Biochem. Sci. 16*, 98−103.
32. Biekofsky, R. R., Martin, S. R., Browne, J. P., Bayley, P. M., and Feeney, J. (1998) *Biochemistry 37*, 7617−7629.
33. Kligman, D., and Hilt, D. C. (1988) *Trends Biochem. Sci. 13*, 437−443.
34. Evenas, J., Thulin, E., Malmendal, A., Forsen, S., and Carlstrom, G. (1997) *Biochemistry 36*, 3448−3457.
35. Gagne, S. M., Li, M. X., and Sykes, B. D. (1997) *Biochemistry 36*, 4386−4392.
36. Sastry, M., Ketchem, R. R., Crescenzi, O., Weber, C., Lubienski, M. J., Hidaka, H., and Chazin, W. J. (1998) *Structure 6*, 223−231.
37. Schafer, B. W., and Heizmann, C. W. (1996) *Trends Biochem. Sci. 21*, 134−140.
38. Skelton, N. J., Kordel, J., and Chazin, W. J. (1995) *J. Mol. Biol. 249*, 441−462.
39. Rustandi, R. R., Drohat, A. C., Baldisseri, D. M., Wilder, P. T., and Weber, D. J. (1998) *Biochemistry 37*, 1951−1960.
40. Smith, S. P., and Shaw, G. S. (1998) *Structure 6*, 211−222.
41. Matsumura, H., Shiba, T., Inoue, T., Harada, S., and Kai, Y. (1998) *Structure 6*, 233−241.
42. Schafer, B. W., Wicki, R., Engelkamp, D., Mattei, M. G., and Heizmann, C. W. (1995) *Genomics 25*, 638−643.

BI990922I