Diversity of polyproline recognition by EVH1 domains

Francis C. Peterson, Brian F. Volkman

Department of Biochemistry, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, WI 53226

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Functional organization of the EVH1 family
4. Polyproline recognition by EVH1 domains
5. Peptide binding specificity of EVH1 domains
6. Verprolins contain an extended EVH1 binding motif
7. LIM domain recognition by the Mena EVH1 domain
8. Conclusions
9. References

1. ABSTRACT

Enabled/VASP Homology-1 (EVH1) domains function primarily as interaction modules that link signaling proteins by binding to proline-rich sequences. EVH1 domains are ~115 residues in length and adopt the pleckstrin homology (PH) fold. Four different protein families contain EVH1 domains: Ena/VASP, Homer, WASP and SPRED. Except for the SPRED domains, for which no binding partners are known, EVH1 domains use a conserved hydrophobic cleft to bind a four-residue motif containing 2-4 prolines. Conserved aromatic residues, including an invariant tryptophan, create a wedge-shaped groove on the EVH1 surface that matches the triangular profile of a polyproline type II helix. Hydrophobic residues adjacent to the polyproline motif dock into complementary sites on the EVH1 domain to enhance ligand binding specificity. Pseudosymmetry in the polyproline type II helix allows peptide ligands to bind in either of two N-to-C terminal orientations, depending on interactions between sequences flanking the prolines and the EVH1 domain. EVH1 domains also recognize non-proline motifs, as illustrated by the structure of an EVH1:LIM3 complex and the extended EVH1 ligands of the verprolin family.

2. INTRODUCTION

Protein-protein interactions regulate most cellular processes including gene expression, the cell cycle, protein trafficking, targeted proteolysis, and cytoskeletal reorganization. Cell surface receptors are often coupled to intracellular signaling pathways by recognition of modular protein interaction domains. Regulated signal transduction also depends on the ability to deactivate a pathway in response to changing conditions. Accordingly, protein interaction domains often bind with affinities in the low micromolar range to allow for rapid dissociation of active signaling complexes. Over 80 different protein interaction domain families have been characterized, many of which bind to short linear sequences containing specific amino acid patterns or modifications. Typically, these domains are self-contained modules consisting of 35-150 residues that can be expressed in isolation from their parent protein while retaining their ability to bind their physiological partners (1). This feature has allowed for the isolation and biochemical characterization of a number of protein interaction domains.

Protein interaction domains can be classified based on sequence homology, ligand-binding properties, or
Diversity of polyproline recognition by EVH1 domains

Figure 1. The EVH1 fold is a member of the Pleckstrin Homology superfamily. Structural comparison of the EVH1 domain from Mena (PDB code: 1EVH) with the PH domain from phospholipase C δ (PDB code: 1MAI) and the phosphotyrosine binding domain (PTB) from tensin-1 (PDB code: 1WVH). The PH fold consists of two perpendicular anti-parallel β-sheets followed by a C-terminal α-helix that assemble to form a ligand binding groove between β-strands 1, 2, 6, and 7. In the PH and PTB domains, this groove is occupied by an additional α-helix that moves the recognition of inositol-(3,4,5) triphosphate and phosphotyrosines peptides by PH and PTB domains, respectively, to distinct regions of the EVH1 domain. Rotation by 90° about the vertical axis (lower panel) shows a side view of the PH fold.

structural similarity (2). Thus, a typical class of ligand-binding proteins may contain a variety of protein interaction domains that recognize a common ligand whereas a family classified based on sequence homology contains a single fold that may recognize a variety of ligands. Some families function in a narrow cellular context, while others participate in a diverse range of processes. For example, PDZ domains typically help assemble protein complexes at the plasma membrane (3), but Src homology 2 (SH2) domains are found in a wide variety of signaling proteins (4). The paradigm originally established by the SH2 family, which binds phosphotyrosine-containing sequences, has expanded to include domains that recognize C-terminal sequences, acetyl-lysine residues, polyproline helices and other structural motifs. Four distinct families that recognize proline-rich sequences are now known: Src homology-3 (SH3), WW, glycine-tyrosine-phenylalanine (GYF), and Enabled/VASP homology-1 (EVH1) (5).

The EVH1 domain is a member of the Pleckstrin homology (PH) domain-like superfamily (Pfam ID – PF00169, CL0266). The PH fold consists of two perpendicular anti-parallel β-sheets followed by a C-terminal α-helix (Figure 1). Because sequence homology is low and loops regions are of variable length, sequence-based identification of PH domains can be difficult. Phosphotyrosine-interaction/phosphotyrosine-binding domains (PID/PTB), Ran-BP1, GRAM, FERM-C, and EVH1 domains all share the PH fold, and appear in a wide array of intracellular signaling proteins (Table 1) (6-8). These domains recognize a diverse set of ligands such as inositol lipids, phosphotyrosines, and proline-rich sequences, demonstrating that the PH fold is a stable scaffold that has been adapted to multiple functions. Uniquely in the PH domain superfamily, EVH1 domains interact with proline-rich sequences that bind a groove formed by β-strands 1, 2, 6 and 7. In contrast, PH and PTB domains contain an extra helix that occludes this binding groove shifting the binding sites for inositol lipids and phosphotyrosines to other regions of the PH superfold (9). Interestingly, a recent report shows that the EVH1 domain from Mena forms a stable complex with a Lim3 domain that lacks a polyproline motif (10). This interaction occludes the binding groove and competes with polyproline ligands. Lim binding reveals a new class of potential interactions for EVH1 domain containing proteins and
Diversity of polyproline recognition by EVH1 domains

<table>
<thead>
<tr>
<th>EVH1 domain</th>
<th>Target</th>
<th>PDB ID</th>
<th>Structure</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hs SPRED-2</td>
<td>Unknown</td>
<td>2ip2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Xi SPRED-1</td>
<td>Unknown</td>
<td>3xol</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rn Homer 2</td>
<td>mGlurR15, IP, receptor types 1 and 3</td>
<td>1dwy (1dwy)²</td>
<td>TPPSPF</td>
<td>(34, 69-78)</td>
</tr>
<tr>
<td>Rn Homer 1c</td>
<td>Shank 1 and 3, TRPC1, PIK-4-L</td>
<td>1t2h</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mm Homer 2c</td>
<td>DynaminII, Oligophrin-1, and DrebrinE</td>
<td>2ppk</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hs Homer 3</td>
<td>Aca, Zyxin, LPP, Vinculin, Fyb/SLAP</td>
<td>1q66</td>
<td>FEFPPPTDEE</td>
<td>(21-24, 26, 29, 30, 32, 79-93)</td>
</tr>
<tr>
<td>Mm EVL</td>
<td>Robo, Semaphorin 6A-1, Drk, F6E65, and Ena/VASP</td>
<td>1evh</td>
<td>FPPPP</td>
<td></td>
</tr>
<tr>
<td>Hs VASP</td>
<td>Profilin I and II</td>
<td>1evh</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hs N-WASP</td>
<td>WIP, WIG, WCH, and CR16</td>
<td>2if8 (1mke)³</td>
<td>ESRFVHPSDSLPPPVPQKTSSPSLARNERS</td>
<td>(43-49, 94)</td>
</tr>
</tbody>
</table>

Species codes are as follows: Hs – Homo sapiens, Xi – Xenopus tropicalis, Rn – Rattus norvegicus, and Mm – Mus musculus

²PDB code 1dwy is the structure of the rat Homer EVH1 domain in absence of ligand. ³PDB code 1mke is the NMR structure of the rat N-WASP with the 25 C-terminal residues of the peptide listed in Table 1.

suggests that EVH1 complexes may be dynamically regulated in the cell.

EVH1-containing proteins function in several cellular contexts including cytoskeletal dynamics, postsynaptic signal transduction, proliferation, and differentiation. Missense mutations in EVH1 domains that disrupt binding to their target proteins have been identified as the cause of inherited disease in humans (11-16). EVH1 domains recognize their target sequences in a different manner than other polyproline binding domains. This review aims to highlight the unique aspects of EVH1 complexes, including the enhancement of specificity and protein stability by extended binding motifs flanking the central proline motif.

3. FUNCTIONAL ORGANIZATION OF THE EVH1 FAMILY

EVH1 domains have been found in ~630 human genes and are classified into four distinct protein families based on amino acid sequence analysis: Wiskott-Aldrich-Syndrome protein (VASP); Enabled/vasodilator-stimulated phosphoprotein (Ena/VASP); Homer/Vesl; and Sprouty-related proteins with an EVH1 domain (SPRED) (Figure 2A). In all instances the EVH1 domain is ~115 residues in length and found solely at the N-terminus. Regions C-terminal to the EVH1 domain are highly divergent in sequence and domain composition (Figure 2B). Each EVH1 subclass recognizes a distinct pattern of amino acids, but all of them bind a proline-rich sequence in the left-handed polyproline type II conformation. Residues flanking the polyproline motif contribute to the binding specificity of EVH1 complexes as discussed below in more detail.

Enabled-VASP Homology 1 domains were first identified in proteins of the Ena/VASP family that includes Ena, the product of the Drosophila enabled gene (17, 18), the mammalian protein vasodilator-stimulated phosphoprotein (VASP) (19, 20), mammalian Enabled (Mena), and Enabled/VASP-like (EVL) protein (21). Ena/VASP proteins also share two other common structural elements: a central proline-rich region and a C-terminal EVH2 domain, each of which varies in size among the family members. The EVH1 domains within this family recognize protein ligands that contain a core (F/L)PPP motif (Figure 3) (22-24). A number of proteins such as zyxin (25), vinculin (25), roundabout (Robo) (26), Semaphorin 6A-1 (24), RIAM (27), lamellipodin (28), and the Fyn-binding and SLP-76 associated protein are ligands for Ena/VASP family members and play roles in focal adhesions (21, 29, 30), cell-cell adherens junctions (31), and cell motility. For example, zyxin and vinculin are focal adhesion proteins while Robo and Semaphorin 6A-1 participate in axon guidance. In contrast, RIAM and lamellipodin participate in cell motility and regulate the formation of lamellipodia. RIAM functions as an adaptor protein that provides a link between Rap1-GTP and profilin (27) while lamellipodin colocalizes with Ena/VASP proteins at the tips of lamellipodia and filopodia and plays a regulatory role (28). In addition, lamellipodin is recruited by Vaccinia virus and Enteropathogenic E. coli to facilitate their own motility (28). Similarly, the ActA protein from the intracellular bacterial pathogen Listeria monocytogenes utilizes both VASP and Mena to hijack the host’s actin nucleation and polymerization machinery to become mobile within the cell (32). ActA binding to the VASP or Mena EVH1 domain is mediated by the consensus sequence (D/E)FPPPPT (D/E) (D/E)EL, in which the core FPPP is essential for the interaction (22, 33).

The Homer/Vesl family of adapter proteins consists of three members, Homer1, Homer2, and Homer3, each of which is subject to alternative pre-mRNA splicing to yield long and short isoforms (34-36). All Homer splice variants retain the N-terminal EVH1 domain but the long isoforms contain a C-terminal coiled coil that is absent in the short form. EVH1 domains from the Homer family interact with ligands that contain a core PXxF motif, where x is any amino acid (Figure 3) (37). In the cell, members of the Homer family are primarily expressed in the nervous system and are localized to the post-synaptic density (PSD), an actin-rich structure found on the spines of neuronal dendrites. Homer proteins interact via their EVH1 domain with a number of receptors localized to the PSD including the metabotropic glutamate receptor (mGlur5), inositol 1,4,5-trisphosphate receptors, ryanodine receptors and Shank, an adaptor for the N-methyl-D-aspartate receptor complex. Within these complexes, Homer proteins can regulate the synaptic localization of target proteins or modulate cross talk between signaling proteins localized in the PSD. For example, overexpression...
### Figure 2

**Domain organization and phylogenetic tree of EVH1 domain proteins.**

A) Protein sequences corresponding to the EVH1 domain boundaries identified by Pfam (95) for human (*Homo sapiens*, Hs), fly (*Drosophila melanogaster*, Dm), fish (*Danio rerio*, Dr), worm (*Caenorhabditis elegans*, Ce), slime mold (*Dictyostelium discoideum*, Dd) and yeast (*Saccharomyces cerevisiae*, Sc) were aligned. NCBI accession numbers for the proteins are indicated in the brackets. Multiple sequence alignments and the phylogenetic analysis were performed using ClustalW version 1.83 (96). EVH1 domains within a particular branch are more related to each other than to those within a particular species. B) Domain organization of representative members from each branch of the four EVH1 protein families. In all families the EVH1 domain is located at the N-terminus and is followed by a variety of domains including proline rich domains (PRD), coiled-coil (CC), low complexity regions (LCR), Ena-VASP homology 2 domains (EVH2), GTPase binding domains (GBD), verprolin-cofilin-acidic motifs (VCA), Sprouty-like cysteine-rich domains (SPR), and c-Kit binding domains (KBD).

**Table 2.1**

<table>
<thead>
<tr>
<th>Domain Family</th>
<th>EVH1</th>
<th>PRD</th>
<th>EVH2</th>
<th>SPR</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPREDs</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homer 1a</td>
<td>EVH1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Homer 2a</td>
<td>EVH1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VASP/EVL</td>
<td>EVH1</td>
<td>PRD</td>
<td>EVH2</td>
<td></td>
</tr>
<tr>
<td>Ena</td>
<td>EVH1</td>
<td>LCR</td>
<td>PRD</td>
<td>EVH2</td>
</tr>
<tr>
<td>Mena</td>
<td>EVH1</td>
<td>LCR</td>
<td>PRD</td>
<td>EVH2</td>
</tr>
<tr>
<td>WASP</td>
<td>EVH1</td>
<td>GBD</td>
<td>PRD</td>
<td>VCA</td>
</tr>
<tr>
<td>N-WASP</td>
<td>EVH1</td>
<td>GBD</td>
<td>PRD</td>
<td>VCA</td>
</tr>
</tbody>
</table>

of Homer1b (long form) retained the mGluR5 receptor in the endoplasmic reticulum preventing it from being localized to the cell-surface (38, 39). In contrast, overexpression of Homer1a (short form) induced cell-
Diversity of polyproline recognition by EVH1 domains

Figure 3. Interaction of polyproline containing ligands with their EVH1 domains. A) Ribbon and surface diagrams depicting the interaction of Mena with ActA peptide (PDB code: 1EVH), of Homer-1 with the mGluR peptide (PDB code: 1DDV), of N-WASP with the WIP peptide (PDB code: 2IFS), and SPRED-1 (PDB code: 1XOD). No ligands for SPRED domains are known. The purple surface represents residues that are within 5Å of the bound peptide.

surface localization and clustering of the mGluR5 receptor (38). These results were confirmed in a number of cell types and suggest that the Homer proteins regulate cell-surface targeting and clustering of the mGluR5 receptor. Members of this family also participate in neuronal development and contribute to behavioral phenomena like drug addiction (40). While no clear role for has been established from Homer proteins in neurological disease, recent studies have suggested roles for certain family members in schizophrenia, X-linked mental retardation and Fragile X syndrome.

The WASP family consists of three members: WAS protein (WASP), expressed in cells of the hematopoietic lineage, its ubiquitously expressed homolog neuronal-WASP (N-WASP), and the yeast homolog Bee1p. WASP proteins are multidomain proteins that contain an N-terminal EVH1 domain, also referred to as a WASP homology 1 (WH1) domain, followed by a basic motif, a GTPase binding domain, a proline-rich region and verprolin-cofilin-acidic motif that binds directly to the actin-related protein (Arp)2/3 actin nucleating complex (41, 42). The functional importance of the EVH1 module is reflected in the high proportion of disease-causing missense mutations that occur within the domain (12-16).

In contrast to the Ena/VASP and Homer families, WASP proteins recognize an extended peptide motif that is far longer than the canonical EVH1 ligands (Figure 3) (43-45). Members of the mammalian verprolin family, including WASP interacting protein (WIP) (43-46), WIP-related protein (WIRE/WICH) (47, 48), and CR16 (49), contain a sequence near the carboxyl terminus that interacts with the WASP EVH1 domain. A polyproline sequence binds the conserved hydrophobic groove as in other EVH1 domains, but is flanked on both ends by additional sequences required for stable binding and biological WASP activity. The WASP protein is stabilized through its constitutive association with WIP (50, 51), and PKC θ-mediated phosphorylation of serine 488 near the WIP C-terminus reportedly weakens the interaction (52). To assess the structural impact of phosphorylation, we used 2D NMR to detect changes in an EVH1-WIP complex upon the introduction of mutations that mimic WIP phosphorylation (S488D and S488E). Our NMR analysis suggested that WIP phosphorylation does not disrupt the WIP-WASP complex (F. C. Peterson and B. F. Volkman, unpublished results). Consistent with our observations, a recent study by Dong et al. (53) showed that phosphorylation of serine 488 does not disrupt the WIP-WASP complex in cells. Thus, WIP-WASP interactions appear to be unaffected by phosphorylation at serine 488.

Disease causing mutations in the EVH1 domain are thought to destabilize the WASP/WIP complex leading to degradation of WASP (11, 51). Inherited mutations in the wasp gene lead to Wiskott-Aldrich Syndrome (WAS), an X-lined recessive disorder first identified in 1937 and characterized by immunodeficiency, eczema and thrombocytopenia (54). In lymphocytes, greater than 95% of WASP is normally bound to WIP (52). Most missense mutations identified in WAS families occur within the EVH1 domain and likely disrupt its interaction with WIP. This hypothesis is correlated with the observation that the most severe disease-causing WASP mutations result in the greatest reduction of WASP levels in circulating platelets, even though WASP mRNA levels remain unchanged (51). This observation lead Lutskiy et al. to conclude that WIP plays a WASP-protective role in leukocytes, and that disease-causing mutations lead to degradation of WASP protein and abnormal cytoskeletal regulation (51). These cytoskeletal defects result in impaired leukocyte chemotaxis and reduced T and natural killer cell engagement with targets cells leading to immune dysfunction in WAS patients (55).
Diversity of polyproline recognition by EVH1 domains

Member of the SPRED family of proteins contain an N-terminal EVH1 domain that is followed by a central c-Kit binding domain, and a C-terminal Sprouty-like cysteine-rich domain (56, 57). First identified in *Drosophila* as inhibitors of the RAS/MAPK signaling pathway (56-58), homologs have now been identified in *Xenopus*, mice and humans. At present, the binding partners recognized by the SPRED EVH1 remain unknown and direct interactions with other proteins through the EVH1 domain have not been shown. Structural studies with the SPRED1 EVH1 domain from *Xenopus tropicalis* suggest that the conserved polyproline binding groove is narrower than in other EVH1 domains (59). Harmer et al. concluded that SPRED domains are likely to bind polyproline peptides that are less proline-rich than other EVH1 ligands (59). Potential roles for SPRED-1 and SPRED-2 in disease have been identified through the use of knock-out mouse models (58, 60-62). These models showed that SPRED-1 and SPRED-2 were not essential for fertility or development, but showed a dwarf phenotype similar to hypochondroplasia in young adult mice. Roles for SPRED-1 in mature late phase hematopoiesis and for SPRED-2 as a negative regulator of embryonic hematopoiesis has also been suggested (58, 62). Presently, no functional role has been demonstrated for SPRED-3.

The lack of known binding partners suggests that SPRED domains may utilize an interaction surface distinct from the conserved EVH1 polyproline docking site. Like many proteins, the conserved PH domain fold has been adapted to support multiple interaction surfaces (Figure 1). Some SH3 domains also have binding sites distinct from the conserved proline-binding surface (51). Recent structural studies show that some EVH1 domains are multifunctional, recognizing binding partners that are not restricted to polyproline sequences. For example, the N-WASP EVH1 domain employs three distinct epitopes to form a WIP binding surface that extends well beyond the conserved polyproline binding site. Likewise, the Mena EVH1 domain binds to the third LIM domain of testin, which lacks a polyproline motif altogether (10). Further experimental work is required to establish whether SPREDs are functionally distinct from EVH1 domains that bind proline-rich sequences.

4. POLYPROLINE RECOGNITION BY EVH1 DOMAINS

Polyproline-rich regions are among the most numerous sequence motifs in the fly and nematode genomes (5). Proline is unique among the 20 naturally occurring amino acids for its cyclic side chain, which creates a substituted amide nitrogen and restricts the range of available conformations. As a consequence, proline-rich sequences can adopt a left-handed helical structure called the polyproline type II (PPII) helix. Proline recognition domains from the SH3, WW, GYF and EVH1 families target distinct proline-rich sequences, but structural studies have shown that the polyproline motif adopts the PPII helix upon binding to domains of each type. With a periodicity of three residues per turn, the PPII structure is triangular in cross-section, and displays both hydrophobic surface (aliphatic side chains) and hydrogen bond acceptors (backbone carbonyls). Strikingly, amide substitution in the PPII helix results in a structure that is nearly superimposable on itself in either the N- to C-terminal or C- to N-terminal orientation. This twofold pseudosymmetry enables the bidirectional ligand binding modes observed for members of the SH3, WW and EVH1 families (44, 63-66).

It has been noted previously that SH3, WW, and GYF domains use a relatively flat surface that mates with one face of the prism-shaped polyproline helix (Figure 4). In contrast, a wedge-shaped hydrophobic groove in the EVH1 family recognizes an apex of the triangular PPII helix which is rotated by ~60° relative to the other proline binding domains (9). Aromatic side chains and other conserved residues at five sequence positions in the β1, β2 and β6 strands of the EVH1 domain line the sides of the binding cleft (Figure 5A). An invariant tryptophan in the β2-strand creates a ridge across the center of the hydrophobic groove. Ligand peptides dock over the indole ring of the Trp sidechain, defining the register and orientation of the bound polyproline motif. Depending on the EVH1 subfamily, the first position is always Phe (Homer), Tyr (WASP) or Met (Ena/VASP and SPRED), while position 2 requires Ile (Homer), Tyr (Ena/VASP) or Arg (SPRED). The third position is absolutely conserved as a Phe residue in all EVH1 sequences, and position 4 contains a Gln (Homer and Ena/VASP) or His (SPRED). Positions 2 and 4 in the WASP family are less strongly conserved, but typically occupied by Ala and Thr, respectively.

The rotational difference of ~60° in how the PPII helix binds relative to other proline recognition domains may explain the less restrictive sequence requirements for EVH1 ligands. Proline is preferred but not absolutely conserved at each of five positions (-FPPP-), and few EVH1 ligands contain all five proline residues, as illustrated in Figure 5B. Thus it appears that while SH3, WW and GYF domains directly recognize the proline side chain at specific positions in the target sequence, EVH1 domains simply require that proline content be sufficient to stabilize a PPII helical conformation. However, since EVH1 domains recognize the general features of a proline-rich motif, but do not bind all proline-rich sequences, additional factors are required to achieve high peptide binding specificity as described in section 5.

Interestingly, while the polyproline motif is required for binding to most EVH1 domains, no peptide ligands have been identified for the SPRED family. Substitution of the positively charged Arg and His sidechains at positions 2 and 4 of the binding cleft may significantly alter the binding preferences for these domains, even though the invariant Phe and Trp residues are still present. A similar degree of sequence divergence within the aromatic binding cleft of WASP EVH1 domains probably explains why a proline-rich sequence is necessary but not sufficient for binding to WASP EVH1 domains (44, 45). However, low WASP EVH1 binding affinity to the PPII motif is compensated by extension of the peptide binding site relative to the Ena/VASP and Homer domains.
Diversity of polyproline recognition by EVH1 domains

Figure 4. Polyproline recognition by SH3, GYF, WW and EVH1 domains. The interaction of proline rich ligands is shown for the four distinct families that recognize proline rich sequences. A) The SH3 domain from Sem5 (PDB code: 1SEM). B) The GYF domain from CD2BP2 (PDB code: 1L27). C) The WW domain from dystrophin (PDB code: 1EG4). D) The EVH1 domain from N-WASP (PDB code: 2IFS).

5. PEPTIDE BINDING SPECIFICITY OF EVH1 DOMAINS

Recognition of a secondary structure configuration like the PPII helix is, by itself, insufficient for selective binding of a short peptide ligand. As with other polyproline binding domains, EVH1 binding specificity is enhanced by flanking residues that dock in complementary sites adjacent to the conserved aromatic binding cleft. For example, a phenylalanine (or other hydrophobic) residue adjacent to the proline motif in ligands for Ena/VASP (FPxφP) or Homer (PPxxF) family members binds a corresponding hydrophobic pocket in the EVH1 domain. Complementary electrostatic interactions between acidic residues of the peptide and basic residues on the EVH1 surface also contribute to ligand binding specificity in some cases (9).

Ena/VASP and Homer EVH1 domains recognize prototypical ligands that are analogous to the short sequence motifs recognized by SH3, WW, and GYF domains in that they combine a common proline recognition site with an adjacent binding pocket, typically for a hydrophobic side chain (5). Structures have been solved for relatively few EVH1-ligand complexes, so general rules for the binding preferences of all EVH1 domains are lacking. However, it is clear that EVH1 domains utilize interaction surfaces that extend well beyond the conserved polyproline-binding groove, as illustrated by novel complexes formed by N-WASP and Mena.

6. VERPROLINS CONTAIN AN EXTENDED EVH1 BINDING MOTIF

WASP interacting protein (WIP) forms a constitutive complex with an N-terminal EVH1 domain in WASP/WAVE proteins. WIP, WIP-related (WIRE/WICH), and CR16 comprise the mammalian verprolin family and function as regulators of the actin cytoskeleton (67). The verprolins are proline-rich proteins that bind profilin and various SH3-domain containing proteins, and contain an elongated EVH1 binding domain near the C-terminus. WIP residues 451–485 wrap more than halfway around the EVH1 domain, making specific contacts with three separate epitopes that correspond to regions of high sequence
Diversity of polyproline recognition by EVH1 domains

Figure 5. EVH1 family members use a conserved aromatic-rich binding site to recognize proline-rich ligands. A) Residues used in the recognition of the proline-rich ligands are conserved in the WASP, Homer, Ena/VASP and SPRED families. However, the orientation of the bound PPII helix is reversed when the Homer and Ena/VASP families are compared with the WASP proteins. B) Representative ligand peptide sequences for the Homer, Ena/VASP and WASP EVH1 domain families.
Diversity of polyproline recognition by EVH1 domains

Figure 6. Atypical EVH1 complexes. A) Three distinct WIP epitopes are required for N-WASP binding. Epitopes correspond to WIP amino acids 454-459 (magenta), 461-468 (gold), and 473-478 (blue). B) Hydrophobic contacts between the aromatic WIP residues Phe 454 and Phe 456 (magenta) and Ala 119 (cyan) from N-WASP. C) The WIP polyproline motif straddles the conserved Trp side chain of the EVH1 domain, but in the opposite orientation relative to peptide ligands for Mena and Homer. D) Lys 477 in WIP epitope 3 makes a conserved salt bridge to N-WASP residue Glu 90. E) Structures of the Mena EVH1 bound to the Tes LIM3 domain (left) and the FPPPP peptide (right) employ overlapping binding sites shown in green on the EVH1 domain surface (center).

conservation in the verprolin family (Figure 6A). A central polyproline motif occupies the canonical binding site, but in a reversed orientation relative to other EVH1 complexes. Specific interactions involving WIP residues on either side of the polyproline motif specify the direction of peptide binding.

Aromatic WIP residues in epitope 1 (454FYFHPIS460) bind a hydrophobic surface on the EVH1 domain (Figure 6B). Hydrophobic WIP residues Phe 454, Phe 456 and Ile 459 contribute ~20% of the total EVH1 contact surface. This N-terminal WIP epitope is linked to the polyproline sequence (epitope 2) by a well-defined structural element consisting of the His 457-Pro 458 cis peptide bond followed by a tight helical turn (residues 459–461).

WIP residues 461–468 contain the canonical polyproline motif and form the second binding epitope, which contributes just over 40% of the total buried surface of the WIP-EVH1 complex. WIP residues 461-LPPP465 form
Diversity of polyproline recognition by EVH1 domains

a PPII helical turn that surrounds the conserved Trp side chain of the EVH1 binding surface (Figure 6C), similar to other EVH1-peptide complexes, but with the polypeptide chain running in the opposite direction (44).

The third WIP epitope, \(473^{KSYPSK478}\), is separated from the polyproline motif by a flexible linker and contributes slightly less than 40% of the interface. While Lys 478 does not make extensive contact with the EVH1 surface, it is positioned by the WIP \(475^{–478}\) helical turn to form a salt bridge with Glu90 of N-WASp (Figure 6D), which corresponds to the site of a disease-causing mutation in the WASp sequence (11-16, 43).

These three WIP epitopes comprise a ~30 residue EVH1 binding domain that is conserved throughout the verprolin family of actin binding proteins (Figure 5B). Disruption of any of the three WIP epitopes reduces WASP or N-WASp binding in cells, demonstrating a functional requirement for a significantly longer sequence than the polyproline ligands recognized by other EVH1 domains. The WIP/N-WASp structure shows how semi-independent linear recognition motifs (68) can be used to make composite recognition motifs that are recognized in an extended conformation with enhanced specificity.

7. LIM DOMAIN RECOGNITION BY THE MENA EVH1 DOMAIN

While the verprolins contain unusually long EVH1 binding sequences, they still use the canonical PPII helix to recognize the conserved aromatic-rich binding groove in the same manner as all other known EVH1 ligands. Recently, however, the first non-polyproline binding partner was identified for an EVH1 domain. The EVH1 domain from Mena forms a specific complex with a small zinc-binding domain that lacks a proline-rich motif (10). The EVH1 domains of Mena, VASP and EVL all bind the FPPPP motif using the conserved proline recognition surface (9), but only Mena binds to the LIM3 domain of Tes, a putative tumor suppressor protein. The LIM3 and FPPPP binding surfaces overlap significantly (Figure 6D), consistent with the observation that they compete for binding to the EVH1 domain of Mena. Characterization of the Tes:Mena interaction raises the possibility of EVH1 complexes with other LIM domain proteins. By using a single site for multiple interaction partners, Mena also expands the potential range of EVH1 functions to include regulation of complex formation through competitive binding to overlapping recognition surfaces.

8. CONCLUSIONS

Structural studies of numerous EVH1-ligand complexes have defined a conserved aromatic-rich groove that recognizes a single turn of polyproline type II helix. Amino acids flanking the proline motif enhance ligand specificity and define the N-to-C orientation of peptide binding. Short (6–10 amino acids) peptide ligands of the Ena/VASP and Homer families bind in one orientation, but extended sequences (~30 amino acids) from the verprolin family wrap around EVH1 domains from the WASP family in the opposite orientation. The versatility of the EVH1 structure is further illustrated by the novel complex formed by the Mena EVH1 domain and the LIM3 domain from Tes, the first example of an EVH1 binding partner that lacks a polyproline motif. The unorthodox binding modes employed by N-WASp and Mena demonstrate that EVH1 interactions are more diverse than originally thought, and that competition between multiple EVH1 binding partners may be used to regulate the formation of signaling complexes in the cell.

9. REFERENCES

Diversity of polyproline recognition by EVH1 domains


Diversity of polyproline recognition by EVH1 domains


Diversity of proline recognition by EVH1 domains


79. Krause, M., A. S. Sechi, M. Konradt, D. Monner, F. B. Gertler & J. Wehland: Fyn-binding protein (Fyb)/SLP-76...
Diversity of polyproline recognition by EVH1 domains


Key Words : EVH1 domain, polyproline, Wiskott-Aldrich-Syndrome, Pleckstrin homology domain, Review

Send correspondence to: Brian Volkman, Medical College of Wisconsin, Department of Biochemistry, 8701 Watertown Plank Road, Milwaukee, WI 53226, Tel: 414-955-8400, Fax: 414-456-6510, E-mail: bvolkman@mch wis.edu