The solution structure of pGolemi, a high affinity Mena EVH1 binding miniature protein, suggests explanations for paralog-specific binding to Ena/VASP homology (EVH) 1 domains

Nina M. Link¹,*, Cornelia Hunke², Jonathan W. Mueller¹,³, Jutta Eichler¹ and Peter Bayer¹,⁺

¹ Medizinische und Strukturelle Biochemie, Zentrum für Medizinische Biotechnologie, Universität Duisburg-Essen, Universitätsstr. 2-5, D-45117 Essen, Germany
² Division of Structural and Computational Biology, School of Biological Sciences, Nanyang Technological University, 60 Nanyang Drive, Singapore 639798, Singapore
³ National Institute for Medical Research (MRC), The Ridgeway, London, NW7 1AA, UK
⁺ Department Chemie und Pharmazie, Pharmazeutische Chemie, Universität Erlangen-Nürnberg, Schuhstraße 19, D-91052 Erlangen, Germany

* Corresponding authors
e-mails: nina.link@uni-due.de; peter.bayer@uni-due.de

Abstract

Ena/VASP homology 1 (EVH1) domains are polyproline binding domains that are present in a wide range of adaptor proteins, among them Ena/VASP proteins involved in actin remodeling and axonal guidance. The interaction of ActA, a transmembrane protein from the food-borne pathogen Listeria monocytogenes, with EVH1 domains has been shown to be crucial for recruitment of the host's actin skeleton and, as a consequence, for the infectivity of this bacterium. We present the structure of a synthetic high-affinity Mena EVH1 ligand, pGolemi, capable of paralog-specific binding, solved by NMR spectroscopy. This peptide shares the common pancreatic peptide fold with its scaffold, avian pancreatic peptide, but shows pivotal differences in the amino-terminus. The interplay of spatial fixation and flexibility appears to be the reason for its high affinity towards Mena EVH1. Combined with earlier investigations, our structural data shed light on the specificity determinants of pGolemi and the importance of additional binding epitopes around the residues Thr74 and Phe32 on EVH1 domains regulating paralog specificity. Our results are expected to facilitate the design of other high-affinity, paralog-specific EVH1 domain ligands, and serve as a fundament for the investigation of the molecular mode of action of EVH1 domains.

Keywords: Ena/VASP homology 1 (EVH1); homology modeling; Listeria monocytogenes; NMR spectroscopy; peptides; structure elucidation.

Introduction

Adaptor proteins enable a smooth flow of signaling events by arranging spatial proximity of reaction partners through their substrate-specific protein interaction domains (Pawson and Nash, 2003). A set of such interaction domains that bind to proline-rich motifs, Ena/VASP homology 1 domains (EVH1) (Ball et al., 2002), are predominantly found in adaptor protein families engaged in cytoskeletal dynamics and remodeling. One of those is the Homer protein family (Shiraishi-Yamaguchi and Furuichi, 2007), which is located in the postsynaptic density and is part of the regulation of neuronal processes, such as behavioral plasticity and addiction; another one is the Wiskott-Aldrich syndrome protein family (Zigmond, 2000) nucleating actin polymerization in filopodia. EVH1 domains are also found in the newly discovered Sprouty-related protein family (Bundschu et al., 2007), which is not completely characterized in function yet and which regulates among others bone morphogenesis via inhibition of Ras/ERK pathways. Finally, EVH1 domains are present in the eponymous Enabled/vasodilator-stimulated phosphoprotein (Ena/VASP) family involved in lamellipodial protrusion and axonal guidance (Krause et al., 2003). The three mammalian members of the latter family, Mammalian enabled (Mena), Ena-VASP-like (Ev) and VASP have been shown to be enriched in regions of heavy actin remodeling, such as focal adhesions, stress fibers and lamellipodia (Reinhard et al., 1992; Gertler et al., 1996).

Ena/VASP proteins share a tripartite domain organization with a highly homologous amino terminal EVH1 domain responsible for interaction with focal adhesion proteins, such as lamellipodin, zyxin, and vinculin (Reinhard et al., 1995b, 1996; Brindle et al., 1996; Gertler et al., 1996), a polyproline stretch contacting profilin and SH3 domains (Reinhard et al., 1992a; Ahern-Djamali et al., 1998) and a carboxy terminal EVH2 domain coordinating multimerization and contact to F- and G-actin (Bachmann et al., 1999; Laurent et al., 1999; Walders-Harbeck et al., 2002; Kuhnelt et al., 2004). While their EVH2 domains mediate regulation of actin polymerization (Huttelmaier et al., 1999; Barzik et al., 2005), only the correct localization to the leading edge via their EVH1 domains confers on them the ability to influence actin geometry (Smith et al., 1996; Rottner et al., 1999; Bear et al., 2002).

A lot of information about the role of Ena/VASP proteins in processes at the leading edge was obtained by studying their interaction with actin nucleator protein A (ActA), a surface protein of the food-borne pathogen Listeria monocytogenes, with EVH1 domains regulating axonal guidance (Krause et al., 2003). The three mammalian members of the latter family, Mammalian enabled (Mena), Ena-VASP-like (Ev) and VASP have been shown to be enriched in regions of heavy actin remodeling, such as focal adhesions, stress fibers and lamellipodia (Reinhard et al., 1992; Gertler et al., 1996).

Ena/VASP proteins share a tripartite domain organization with a highly homologous amino terminal EVH1 domain responsible for interaction with focal adhesion proteins, such as lamellipodin, zyxin, and vinculin (Reinhard et al., 1995b, 1996; Brindle et al., 1996; Gertler et al., 1996), a polyproline stretch contacting profilin and SH3 domains (Reinhard et al., 1992a; Ahern-Djamali et al., 1998) and a carboxy terminal EVH2 domain coordinating multimerization and contact to F- and G-actin (Bachmann et al., 1999; Laurent et al., 1999; Walders-Harbeck et al., 2002; Kuhnelt et al., 2004). While their EVH2 domains mediate regulation of actin polymerization (Huttelmaier et al., 1999; Barzik et al., 2005), only the correct localization to the leading edge via their EVH1 domains confers on them the ability to influence actin geometry (Smith et al., 1996; Rottner et al., 1999; Bear et al., 2002).

A lot of information about the role of Ena/VASP proteins in processes at the leading edge was obtained by studying their interaction with actin nucleator protein A (ActA), a surface protein of the food-borne pathogen Listeria monocytogenes, with EVH1 domains regulating axonal guidance (Krause et al., 2003). The three mammalian members of the latter family, Mammalian enabled (Mena), Ena-VASP-like (Ev) and VASP have been shown to be enriched in regions of heavy actin remodeling, such as focal adhesions, stress fibers and lamellipodia (Reinhard et al., 1992; Gertler et al., 1996).

Ena/VASP proteins share a tripartite domain organization with a highly homologous amino terminal EVH1 domain responsible for interaction with focal adhesion proteins, such as lamellipodin, zyxin, and vinculin (Reinhard et al., 1995b, 1996; Brindle et al., 1996; Gertler et al., 1996), a polyproline stretch contacting profilin and SH3 domains (Reinhard et al., 1992a; Ahern-Djamali et al., 1998) and a carboxy terminal EVH2 domain coordinating multimerization and contact to F- and G-actin (Bachmann et al., 1999; Laurent et al., 1999; Walders-Harbeck et al., 2002; Kuhnelt et al., 2004). While their EVH2 domains mediate regulation of actin polymerization (Huttelmaier et al., 1999; Barzik et al., 2005), only the correct localization to the leading edge via their EVH1 domains confers on them the ability to influence actin geometry (Smith et al., 1996; Rottner et al., 1999; Bear et al., 2002).

A lot of information about the role of Ena/VASP proteins in processes at the leading edge was obtained by studying their interaction with actin nucleator protein A (ActA), a surface protein of the food-borne pathogen Listeria monocytogenes, with EVH1 domains regulating axonal guidance (Krause et al., 2003). The three mammalian members of the latter family, Mammalian enabled (Mena), Ena-VASP-like (Ev) and VASP have been shown to be enriched in regions of heavy actin remodeling, such as focal adhesions, stress fibers and lamellipodia (Reinhard et al., 1992; Gertler et al., 1996).

Ena/VASP proteins share a tripartite domain organization with a highly homologous amino terminal EVH1 domain responsible for interaction with focal adhesion proteins, such as lamellipodin, zyxin, and vinculin (Reinhard et al., 1995b, 1996; Brindle et al., 1996; Gertler et al., 1996), a polyproline stretch contacting profilin and SH3 domains (Reinhard et al., 1992a; Ahern-Djamali et al., 1998) and a carboxy terminal EVH2 domain coordinating multimerization and contact to F- and G-actin (Bachmann et al., 1999; Laurent et al., 1999; Walders-Harbeck et al., 2002; Kuhnelt et al., 2004). While their EVH2 domains mediate regulation of actin polymerization (Huttelmaier et al., 1999; Barzik et al., 2005), only the correct localization to the leading edge via their EVH1 domains confers on them the ability to influence actin geometry (Smith et al., 1996; Rottner et al., 1999; Bear et al., 2002).
Listeria monocytogenes (Kocks et al., 1992). With several proline-rich stretches for Ena/VASP interaction (Chakraborty et al., 1995; Niebuhr et al., 1997) and specific protein interaction sites for Arp2/3 and actin (Lasa et al., 1997; Welch et al., 1998), this transmembrane protein has proven to be absolutely crucial and also sufficient for Listeria monocytogenes to hijack the host’s actin machinery for rapid propulsion in the cytoplasm and invasion of neighboring cells (Pistor et al., 1995; Smith et al., 1996; Welch et al., 1998; Laurent et al., 1999; Skoble et al., 2000). ActA has been shown to interact with the EVH1 domains of all three mammalian members of the Ena/VASP family, Evi, VASP and Mena, via its proline-rich sequence stretches in the central part of the protein (Niebuhr et al., 1997). Depletion of these EVH1 domains in the host cells severely inhibits Listeria motility and actin nucleation (Niebuhr et al., 1997; Laurent et al., 1999).

In order to investigate the paralog-specific actions of the Ena/VASP family with ActA, Golemi-Kotra and co-workers designed a miniature protein with a 10-times enhanced affinity to Mena EVH1 compared to ActA, a peptide ligand derived from the first proline-rich motif (PR) of ActA with a high affinity to Mena (Niebuhr et al., 1997; Golemi-Kotra et al., 2004). This 30 amino acid peptide, termed pGolemi, was designed by grafting the polyproline motif onto the N-terminus of avian pancreatic polypeptide (aPP), a stably folded hormone (Blundell et al., 1981). pGolemi was used to study the binding kinetics for all three Ena/VASP paralogs with alanine mutants by fluorescence perturbation measurements (Holtzman et al., 2006). Nevertheless, no experimentally determined structure of pGolemi is available yet.

We now succeeded in determining the solution structure of pGolemi. Additionally, homology models of pGolemi and the second proline-rich repeat (PRII) of ActA bound to Mena EVH1 were built. With these data, the enhanced binding affinity of pGolemi can be explained, and light is shed on the spatial arrangement of specificity determinants for paralog-specific binding to EVH1 domains. In addition, characteristic binding determinants formed by conserved residues on the PRII of ActA are identified in our models.

Results and discussion

Structure determination of pGolemi – a monomeric, back-folded miniature protein

The solution structure of synthetic pGolemi was determined by two-dimensional NMR spectroscopy. Structure calculations based on a total of 373 constraints were carried out. The final ensemble was selected for acceptance criteria of minimal energies, NOE violations less than 0.25 Å and angle violations less than 5°. An ensemble of the 10 lowest energy structures is depicted in Figure 1A, an alignment for the polyproline part is shown in Figure 1B. Despite the high proline content in the amino terminal part of pGolemi, the number of constraints was sufficient to obtain low energy and well-defined structures (Table 1, Figure 1). For each residue except the first, a sufficient number of cross-peaks could be assigned (Figure 2A). Throughout the whole sequence, strong $d_{\alpha}(i,i+1)$ connectivities could be found for the majority of non-proline residues.

The existence of $d_{\alpha}(i,i+3)$ NOEs from residue 14 through to 27 and $d_{\alpha}(i,i+4)$ NOEs from the same residues, clearly indicate an $\alpha$-helical conformation for this part of pGolemi. Due to spectral overlap in this region of

![Figure 1](image)

**Figure 1** Representation of the final structure ensemble of pGolemi. (A) Solution structure of pGolemi as represented by the bundle of lowest energy structures. The traced backbone atoms have been superimposed to an overall RMSD of 0.525 Å by the program YASARA (Krieger et al., 2002, 2004). (B) Alignment of residues 2–8 of pGolemi. The helix residues were omitted for clarity.

![Table 1](image)

**Table 1** Structural statistics for the energy-minimized ensemble of pGolemi structure calculated with Xplor-NIH.
Structure of pGolemi, a specific Mena EVH1 binder

Figure 2  Graphical depiction of the unambiguous distance constraints used for structure calculation of pGolemi. (A) Short- and medium-range connectivities. (B) $^1$H$_{L}$-Chemical shift index versus residue position.

the spectra, only a limited number of $d_{ab}(i,i+3)$ type NOEs could be resolved. In accordance to the secondary chemical shift analysis (Figure 2B) using the method of Wishart et al. (1992), initial structure calculations based on the resolved unambiguous constraints have extended the helical conformation further from V13 to V30.

During refinement, hydrogen bond restraints for helical conformation were therefore introduced to the corresponding region for a better definition of the helix.

In the designing process of pGolemi several critical residues from aPP have been mutated (Figure 3). However, the overall backbone RMSD of the minimized average structure to the parental scaffold amounts to only 1.12 Å deviation for the whole sequence and 0.61 Å deviation for the helix from V13 to V30, respectively. In agreement with circular dichroism experiments previously performed by the Schepartz group with helix contents between 43% and 49% (calculated in the helix-coil-transition model; Golemi-Kotra et al., 2004; Holtzman et al., 2007), the secondary structure of pGolemi comprises a carboxy-terminal $\alpha$-helix from residues 13 to 30 and a classical $\beta$-turn (residues 9–12) confirmed by the existence of strong $d_{ab}(i,i+1)$ NOEs and the absence of $d_{nn}$ connectivities for these residues. An extended amino terminal polyproline-rich part ranges from residues 1 to 8 (Table 2). Contrary to dimeric aPP and other designed miniature proteins derived from this peptide (Rutledge et al., 2003; Jurt et al., 2006), pGolemi remains monomeric even at millimolar concentrations, substantiating previous fluorescence polarization results (Golemi-Kotra et al., 2004). In all recorded spectra, residues corresponding to the aPP dimer interface showed no intermolecular NOEs. By and large, amino acid mutations for the design of pGolemi are divided into two types: (a) residues responsible for spatial arrangement and structure preservation, and (b) residues for paralog specificity and ligand binding in general (Figure 3). First, the structural residues will be considered.

pGolemi has a typical pancreatic peptide (PP) fold with a modified hydrophobic core

Similar to the well-folded and stable aPP (Blundell et al., 1981; Glover et al., 1983) (Figure 4A), pGolemi adopts a pancreatic peptide (PP) fold (Figure 4B). The stability of these hairpin structures is provided by the interaction of a long amphipathic helix with several hydrophobic residues in the first part of the peptides, namely of the pairs P1-Y26, P4-L23, Y6-F19, P7-L16 (for clarity, residues are referred to in pGolemi nomenclature) in pancreatic peptides. For aPP an additional hydrophobic contact between P1 and V29 contributes to this core (denoted with triangles in Figure 3). All except one of the hydrophobic core residues remained unchanged for pGolemi construction. Intriguingly, this substitution of a large aromatic tyrosine with a proline at position 6 does not seem to weaken the fold notably. Rather, the existence of a long range NOE from P6 (Table 3) and several medium range NOEs indicates a relatively fixed position of this residue in the hydrophobic core of pGolemi. Although the perpendicular stacking of the two aromatic rings is eliminated by the changed geometry, profound contacts between both residues P6-F19, now nearly in parallel arrangement, still exist (Figure 4B).

The altered hydrophobic core leads to increased flexibility

The structural ensemble of pGolemi (Figure 1A; PDB entry 2k76) can be grouped into two populations with equal energy and visible changes extending up to the turn region. One population shares a perpendicular orientation of F2 relative to the hydrophobic core leading
Table 2 Overview of the dihedral angles of the PP-rich regions of different EVH1 ligands and aPP.

<table>
<thead>
<tr>
<th>Pos.</th>
<th>pGolemi</th>
<th>FP, in Mena</th>
<th>PRIII in Evl</th>
<th>aPP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>αα</td>
<td>F</td>
<td>αα</td>
<td>αα</td>
</tr>
<tr>
<td>0</td>
<td>-119.45</td>
<td>-99.79</td>
<td>E</td>
<td>G</td>
</tr>
<tr>
<td>1</td>
<td>-54.33</td>
<td>88.99</td>
<td>F</td>
<td>P</td>
</tr>
<tr>
<td>2</td>
<td>-60.17</td>
<td>-46.72</td>
<td>P</td>
<td>-76.08</td>
</tr>
<tr>
<td>3</td>
<td>-119.05</td>
<td>138.76</td>
<td>T</td>
<td>P</td>
</tr>
<tr>
<td>4</td>
<td>-75.10</td>
<td>152.34</td>
<td>P</td>
<td>-75.10</td>
</tr>
<tr>
<td>5</td>
<td>-120.31</td>
<td>-144.84</td>
<td>T</td>
<td>-75.10</td>
</tr>
<tr>
<td>6</td>
<td>64.20</td>
<td>169.82</td>
<td>G</td>
<td>G</td>
</tr>
</tbody>
</table>

Ideal PPII-helix: -78° 146°

Figure 4 Hydrophobic cores and electrostatic surface potentials of aPP and pGolemi.
(A) and (B) Hydrophobic cores of aPP and pGolemi. The molecular surface of involved residues is depicted; residues mentioned in the text are in pGolemi nomenclature. (C) and (D) Side-on view on the PPII part of the electrostatic surface potential of aPP and pGolemi. Both potentials were calculated in the YAMBER2 force-field using YASARA (Krieger et al., 2002, 2004). Residues important for binding specificity are labeled. Note the acidic cluster on top of both miniature proteins.

Table 3 List of the long-range NOEs used for structure calculation.

<table>
<thead>
<tr>
<th>F</th>
<th>2</th>
<th>H_y</th>
<th>↔</th>
<th>Y</th>
<th>26</th>
<th>H_y</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>2</td>
<td>H_y</td>
<td>↔</td>
<td>Y</td>
<td>26</td>
<td>H_y</td>
</tr>
<tr>
<td>F</td>
<td>2</td>
<td>H_y</td>
<td>↔</td>
<td>Y</td>
<td>26</td>
<td>H_y</td>
</tr>
<tr>
<td>P</td>
<td>3</td>
<td>H_y</td>
<td>↔</td>
<td>Y</td>
<td>26</td>
<td>H_y</td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>H_y</td>
<td>↔</td>
<td>F</td>
<td>19</td>
<td>H_y</td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>H_y</td>
<td>↔</td>
<td>L</td>
<td>23</td>
<td>H_y</td>
</tr>
<tr>
<td>P</td>
<td>4</td>
<td>H_y</td>
<td>↔</td>
<td>L</td>
<td>23</td>
<td>H_y</td>
</tr>
<tr>
<td>T</td>
<td>5</td>
<td>H_y</td>
<td>↔</td>
<td>F</td>
<td>19</td>
<td>H_y</td>
</tr>
<tr>
<td>P</td>
<td>6</td>
<td>H_y</td>
<td>↔</td>
<td>F</td>
<td>19</td>
<td>H_y</td>
</tr>
<tr>
<td>P</td>
<td>7</td>
<td>H_y</td>
<td>↔</td>
<td>D</td>
<td>15</td>
<td>H_y</td>
</tr>
<tr>
<td>P</td>
<td>7</td>
<td>H_y</td>
<td>↔</td>
<td>D</td>
<td>15</td>
<td>H_y</td>
</tr>
</tbody>
</table>

to the P7 pyrrolidine ring being oriented more to the solvent. The other population has a parallel alignment of F2 relative to the hydrophobic core region and P7 turned 40° compared to the other population (Figure 1A). The heterogeneity of the ensemble mirrors in the NOE spectra: line broadening can be seen for residues 2 and 3. In Figure 1B, this inhomogeneity can be found as well: in an alignment of the polyproline part only the central residues 5 and 6 are very well defined, the other residues show deviations from the mean. Given the complete assignment of our NOE spectra, these findings indicate either dynamics or conformational changes in this part of the peptide. Similar to other structures of shortened PP variants (Nygaard et al., 2006), the abolishment of the P1-V29 interaction by truncating the sequence could leave an altered hydrophobic core but at the same time increase the dynamics or flexibility of the polyproline part of pGolemi. Since four hydrogen bonds of the C-terminal α-helix of aPP are eliminated by shortening the last turn (Figure 4A), a higher mobility caused by a lack of helical stabilization is likely. Due to the conformational changes in the carboxy-terminal residues, P1 cannot contact the helix and is turned towards the solvent. Even more dramatic changes for the amino terminus might be caused by substitution of the SQ residues for FP at positions 2 and 3, what can be seen by comparing the dihedrals of the respective residues in pGolemi and aPP (Table 2). This alteration comes quite surprising as earlier studies have shown that phenylalanine possesses a comparable PPII helix propensity to serine (Rucker et al., 2003). Presumably, the substitution of Q3 to a conformationally highly restricted proline in pGolemi leads to a dihedral angle combination (F_s=-63.8°, C_s=126.8°) usually found in right handed helices, whereas in the refined structure of aPP (Glover et al., 1983) (PDB code 2bf9) the respective glutamate roughly adopts a PPII helix dihedral combination. This turn-like change in the peptide might be driven by entropic gain. For a loss of higher dynamics in the amino terminus and consequently considerably reduced degrees of freedom hydrophobic interactions can be formed which usually are thermodynamically more favorable. As a result, the first two amino terminal residues of pGolemi seem flexible and point outwards to the solvent. With standard values of F_s=-78° and C_s=146° (Cowan and McGavin, 1955) for a polyproline helix type
II (PPII) only one of the other residues in the amino terminal part of pGolemi falls within this range: P6. Interestingly, this residue is the only residue besides F2 contacting Mena EVH1 via its side chain in a published complex structure (Prehoda et al., 1999) (PDB code 1evh), and also mutation of this proline to alanine diminishes binding significantly (Holtzman et al., 2007). The central proline P6 might induce a PPII helix for the otherwise flexible amino terminal residues after initial binding of this correctly oriented residue to the narrow EVH1 binding groove has occurred. The sum of mutations shown in Figure 3 therefore makes pGolemi a classic PP fold miniature protein with a functional hydrophobic core providing spatial fixation for the grafted specificity residues and a flexible amino terminus and turn region.

The binding determinants Phe2 and Thr5 are spatially arranged for highly specific binding

By inspection of the electronic surface potential of pGolemi and aPP significant differences in shape and the overall charge distribution can be seen. Although the acidic loop region of pGolemi is similarly pronounced like in aPP (Figure 4C), the amino terminal part in pGolemi (Figure 4D) carries less charge; possibly benefitting the hydrophobic binding to EVH1. The proline-rich region of pGolemi is very narrow relative to aPP due to spatial fixation of its specificity determinants by the amphipathic helix (Figure 4C and D) and favors binding to the constricted EVH1 binding groove. In contrast to aPP, the amino terminal of pGolemi up to the acidic region around residues 9 and 10 forms a narrow wedge and protrudes into the solvent with a fixed proline part and a flexible phenylalanine as binding switch.

In fluorescence measurements, the $K_a$ value of pGolemi for one of the natural ligands of ActA, the EVH1 domain of Mena, was found to be 0.6 μM, 10 times smaller than that of PRI of ActA binding to this domain (Golemi-Kotra et al., 2004). The essential binding residues in the proline-rich repeats of ActA were identified as xFPPΦPTxEL (Niebuhr et al., 1997) matching perfectly the motif for class 1 EVH1 domains with a consensus sequence of (D/E)(F/L/W/Y)Px1–3 (Ball et al., 1997). The specificity residues to be discussed now were pre-arranged for highly specific binding

...
conserved acidic residues contributing to high-affinity binding in vivo, as they are involved in domain contacts.

The model for EVH:pGolemi (Figure 5A) clearly shows that due to sterical restrictions, only E9 can contact Thr30 on the domain via its side chain while E10 remains oriented to the solvent. In addition to ligand contacts with the conserved hydrophobic triad, the domain surface is contacted by pGolemi between sheets 1 and 2 directly adjacent to the peptide binding site. F19 in the ligand contacts K22 and W23 on Mena, and an additional interaction can be seen between L23 on pGolemi and the residues K21 and Y16 on Mena, the latter being like W23 a part of the hydrophobic triad (yellow ball representation in Figure 5A). The helix as a second interaction partner obviously strengthens binding and can explain the higher binding free energy for pGolemi of approximately 1.5 kcal/mol compared to PRIII (Holtzman et al., 2007). Our model also clarifies the massive lowering of binding affinity with the mutants F19A and L23A (Holtzman et al., 2007), as both residues are directly involved in binding and provide inter- and intramolecular stabilization. As T5 contributes a hydrogen bond to Thr74 on Mena, the low affinity of the T5A mutant mentioned above can also be explained.

While pGolemi cannot reach further in Mena's binding groove, the extended conformation of PRIII in our second model (Figure 5B) allows additional contacts of the acidic stretch. First, a group of threonines further up the binding groove (T30, T74, T76) is affected, and also polar contacts of the peptide backbone and Mena's Q54 arise. Our MD simulations underline previous studies (Ball et al., 2000) in which a second, hydrophobic epsilon-determinant on VASP EVH1 is contacted by a conserved EL epitope of PRIII (positions 10 and 11, Figure 3). But while in this study contacts are mediated by both E and L, our model shows contacts of E10, but not L11 to the hydrophobic patch (F33, M54, A32 in VASP). Instead, L13 makes multiple contacts to these residues. In particular, the extensive contacts to a phenylalanine present in all Ena/VASP members, F32 (Mena nomenclature), contribute to binding. In our model F32 in Mena's hydrophobic patch interacts with the residues R12 and L13 of PRII. Intriguingly, this phenylalanine 32 is also well oriented for interactions with E8 in the acidic stretch on ActA. This hydrophobic region therefore seems to be important for binding of the carboxy terminal part of the repeats in all EVH1 paralogs.

Our modeled complexes show that pGolemi and PRIII exploit a different set of additional contacts in spite of their common core binding residues.

**Paralog specificity arises from amino acid insertions and different loop conformations**

Earlier fluorescence measurements underline the importance of so-called epsilon-determinants on binding of Ena/VASP domains (Ahern-Djamali et al., 1998; Ball et al., 2000; Aasland et al., 2002). Those residues confer, in
addition to the primary binding site, the characteristic substrate specificity of single domain family members. Mutation of the acidic residues in pGolemi results in a halved affinity to Mena. Intriguingly, the same mutation leads to a significantly higher affinity to its paralog domains Evl and VASP (Holtzman et al., 2007). It has also been shown that removing the residues \textit{aDEL1} from the listerial PRIII repeat (Figure 3) reduces binding to Mena 5-fold and binding to VASP even 10-fold (Ball et al., 2000). As proposed by Prehoda et al. (1999), another epsilon-determinant for Mena EVH1 domains seems to be a group of basic residues located 16 Å apart from the hydrophobic triad. One of the conserved residues of this cluster, R10, is contacted by the side chain of L13. Nevertheless, the chemical shift perturbation data for VASP indicate a binding of this residue also (Ball et al., 2000). Moreover, the basic cluster present in all paralogs could readily interact with further acidic residues located in the PRs. Our study therefore rather points at a conserved binding mode of the PR repeats through both the basic and the hydrophobic cluster proposed by previous structural investigations (Prehoda et al., 1999; Ball et al., 2000) and paralog specificity arising from an acid insertion in the loop between sheets 2 and 3 and varying side chain conformations in the three paralogs. Due to this steric constriction, the cavity provided for a ligand is very small for both Evl and VASP EVH1 (Figure 5C). Especially for the EVH1 domain of VASP proteins, the increase in affinity of pGolemi by exchanging the glutamates at positions 9 and 10 for alanines (Holtzman et al., 2007), which can also contribute to interaction with the juxtaposed hydrophobic residues on the domain, can hence be explained.

**Structure preservation is pivotal for high-affinity binding**

Another critical residue for paralog specificity of pGolemi is A11. Mutation to leucine decreases affinity for Mena 20-fold, while affinity for VASP doubles and for Evl a quadruplication was observed (Holtzman et al., 2007). By its position in the turn it has a crucial role in stabilizing the overall shape of pGolemi (Figure 4B), which seems to be important for high-affinity binding to the domain surface. Several strong medium-range NOEs to P7, V13, and D15 prove the contacts to the PPII part and to the helix. Mutation of this structurally important residue to a leucine with its high spatial demands can cause large-scale changes of the turn conformation. A distortion or even disruption of the turn could lead to a more stretched conformation of the PPII part and to the helix. The affinity of different paralogs to conformationally different substrates underlines this hypothesis: while pGolemi has by far the highest affinity to Mena, followed by VASP (20-times decreased) and Evl (1% of Mena’s affinity), the affinities for the more stretched PRI of ActA are reversed in order (Holtzman et al., 2007). Assuming that pGolemi forces a very tense and bent conformation of the PPII part by its helical frame (Table 2), a preference of Evl to bind stretched PPII motifs while Mena prefers a bent conformation is conceivable.

L23 as an integral part of the hydrophobic core stabilizes the overall shape of pGolemi (Figure 4B) and also contacts the EVH1 surface when bound (Figure 5A). A mutation to alanine in this position destabilizes the core, which is also reflected by a severely reduced helicity of this mutant (Holtzman et al., 2007). As a consequence, the mutant’s affinity to Mena is 130-times smaller, whereas Evl affinity increases 4-fold. Therefore, the binding affinities of the L23A support the hypothesis of an inclination of Mena EVH1 towards compact substrates. The spatial differences in the loop region between \( \beta_1 \) and \( \beta_2 \) could readily explain these differences in affinity for pGolemi (Figure 5C, arrow). Due to the different steric demands in their binding sites, both Evl and VASP cannot form the important secondary contacts by the amphipathic helix of pGolemi and consequently prefer extended substrates, where other epsilon-determinants on the domains can be reached by the ligand.

**Biological implications for the processes of actin polymerization**

The differences in binding profiles of the three paralog domains raise the question about the physiological relevance in the cell. While the affinities of the paralog members for different PR differ significantly (Niebuhr et al., 1997; Ball et al., 2000; Holtzman et al., 2007), also the affinity of the four proline-rich stretches is strongly dependent on the sequence context as proven by motif swapping (Niebuhr et al., 1997). It is not elucidated yet whether each of these motifs has a special EVH1 target or if the ActA PR motif is the best consensus sequence for allocation of a multitude of EVH1 domains to the nucleating actin tail in order to potentiate actin polymerization. Due to the fact that all four PRs in ActA are occupied by an EVH1 domain \textit{in vitro} (Machner et al., 2001), the preferential binding mode of the tetrameric Ena/VASP proteins (Machner et al., 2001; Zimmermann et al., 2002; Kuhnle et al., 2004) to the tandem repeats is neither understood. Apart from one single protein binding to all four PR motifs, two or more Ena/VASP members could also cross-link other effectors to the actin tails (Zimmermann et al., 2002). The exact role of each one of the Ena/VASP family members in actin nucleation as well as their molecular actions in the cell and during recruitment of the actin polymerization machinery by the intracellular pathogen \textit{Listeria monocytogenes} therefore are still not clear (Auerbuch et al., 2003). As was found recently, Mena EVH1 domains are also regulated differentially by specific binding of Tes, a potential tumor suppressor protein, which by its third LIM domain competitively occludes the PPII binding groove of Mena, but not the other Ena/VASP members (Boeda et al., 2007). This opens up totally unexplored possibilities for the regulation of EVH1 domains. Based on the solution structure of pGolemi in combination with our modeling results, further optimized, paralog-specific, high affinity binders for all EVH1 family members can be created. Therewith, a powerful tool will be gained in order to study structural aspects of paralog-specific binding.

**Materials and methods**

Peptide synthesis was carried out using standard solid phase chemistry. The crude peptide was purified by high-pressure
liquid chromatography and verified by high-pressure liquid chromatography coupled to mass spectrometry (HPLC-MS).

**Nuclear magnetic resonance spectroscopy**

pGolemi was dissolved at a concentration of 1 mM in 45 mM potassium phosphate buffer, 10% (v/v) D$_2$O (pH 7.0), NMR experiments were performed at 300 K on a 700 MHz Ultra-shield NMR spectrometer (Bruker Biospin, Rheinstetten, Germany) equipped with a cryoprobe. Two-dimensional TOCSY (Braunschweiler and Ernst, 1983), COSY (Aue et al., 1976) and NOESY (Jeener et al., 1979) spectra were recorded using water suppression by a WATERGATE pulse sequence (Piotto et al., 1992) with 2048 data points in $t_2$ and 512 increments in $t_1$ (48 scans each). Mixing times were 200 ms for the NOESY and 80 ms for the TOCSY spectra. Processing was carried out with the Bruker TopSpin program (Bruker). The data were apodized in both dimensions by a shifted sine-bell window function and zero-filled to 1024 data points for the TOCSY and 2048 data points for the NOESY and COSY in $F_1$, respectively, prior to Fourier transformation.

**Structure calculation**

After assignment of the proton signals via standard procedures (Wüthrich, 1986), NOE peak volume integration was carried out using the program AUREMOL (Gronwald and Kalbitzer, 2004). The volumes were converted into distance constraints by calibrating the nuclear Overhauser intensity against known standard bond lengths to obtain a total of 373 distance restraints (for details see Table 1). Additionally, TALOS (torsion angle likelihood obtained from shift and sequence similarity), a prediction method for dihedral angles, was used for generation of 50 backbone torsional angle restraints based on the $\chi_1$ shifts (Cornilescu et al., 1999). Secondary structure elements were identified using the TALOS predictions and chemical shift indices (Wishart et al., 1992). A set of 200 three-dimensional structures were calculated from a random coil start conformation in the National Institutes of Health version of Xplor (Schwieters and Clore, 2002; Schwieters et al., 2003, 2006) using the above distance and angle restraint data in a standard simulated annealing protocol based on torsion angle dynamics. For refinement, the RAMA potential for torsional conformations was included (Kuszewski et al., 1996). Ten structures out of all conformers were selected for lowest total energy. The representative structure ensemble complies with the following criteria: no NOE violations greater than 0.25 Å, RMSD for bond deviations less than 0.01 Å, RMSD for angle deviations less than 5°. Quality of the structures was checked with WHATIF (Vriend, 1990). Visual inspection of the models and graphic representations were carried out using the program YASARA (Vriend, 1990; Krieger et al., 2002). Structure coordinates have been deposited in the Protein Data Bank (PDB) under the accession code 2k76, chemical shifts in the BMRB database under the entry number BMRB-15946.

**Homology model of pGolemi and PRII in complex with Mena EVH1**

Homology models of the EVH1 complexes were performed on the crystal coordinates of Mena EVH1 in complex with an FP$_1$ peptide (Prevoda et al., 1999; PDB code 1evh), the sequence of PRII (aa 290–311 of SwissProt accession number P33379) and the minimized geometrical average structure of pGolemi derived from the minimized ensemble by a standard XPLOR-NIH protocol. All modeling steps were carried out with Yasara Model in the Yasara interface. After 1 ns run of 1 ns at physiological salt concentration, a final simulated annealing minimization step of a representative simulation snapshot in explicit solvent on the lowest energy complex was performed.

**Acknowledgments**

We thank Professor Dr. Daniel Hoffmann (Bioinformatics Department, University of Duisburg-Essen, Germany) for fruitful discussions and valuable advice regarding the modeled complexes, and the Deutsche Forschungsgemeinschaft (PB 1624/7-1) for financial support.

**References**


Received December 3, 2008; accepted February 10, 2009