Structure-optimized ProM scaffolds address Ena/VASP as a possible antimetastatic target

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and Pharmacy
of Freien Universität Berlin

by

Msc. Matthias Barone
from Wolhusen, Switzerland

2017

The research for this PhD thesis was conducted from 04.2013 until 03.2017 under the supervision of Dr. R. Kühne at the Leibniz Forschungsinstitut für molekulare Pharmakologie, Berlin.

1. Reviewer: Dr. R. Kühne
2. Reviewer: Prof. Dr. U. Heinemann

Date of the oral examination: June 26, 2017

Dedicated to my parents and my wife

Zusammenfassung

Metastasen sind die häufigste Todesursache der Krebserkrankung. Neue Arzneimittel erreichen jedoch häufig nicht ausreichende Steigerung der Überlebensrate von Patienten und die medikamentösen Tumortherapien zeigen nur geringen Erfolg. Es gibt daher grossen Bedarf an neuen Ansätzen. Obwohl die Aktin-Regulatoren der Ena/VASP Proteinfamilie als Marker für Karzinogenese der Brust und als Signatur invasiver Krebsentwicklung gelten, gab es bislang keine zellgängigen Inhibitoren gegen diese Proteine. Wir haben kürzlich gezeigt, dass eine Interferenz mit Ena/VASP die Invasivität von Brustkrebszellen signifikant reduziert. Dazu haben wir di-Prolin-Mimetika entwickelt, genannt ProM, welche die prolinvermittelte Protein-Protein Interaktion der Ena/VASP EVH1-Domäne nachahmen. Aufbauend auf der Verknüpfung zweier ProM Bausteine konnten wir einen nicht toxischen, selektiven und zellgängingen Inhibitor entwickeln, der Ena/VASP erfolgreich von seinen natürlichen Interaktionspartnern verdrängt. Die moderate Bindungsstärke dieses Inhibitors wird jedoch die Validierung in zukünftigen Tierexperimenten erschweren. In dieser Arbeit wird die Affinität des Inhibitors mit Hilfe struktureller Informationen aus 16 Ena/VASP-EVH1 Kristallstrukturen optimiert. Dabei entschlüsseln wir den Bindungsmodus der C-terminal verlängerten Bindesequenz des Proteins ActA von *Listeria monocytogenes*. Wir ahmen die gefundenen Interaktionen von ActA mit neuen ProM Bausteinen nach und entwickeln einen 734 Da Inhibitor, der als neuer Ausgangspunkt für *in vivo* Studien verwendet werden kann. Wir konservierten dabei die strukturelle Einfachheit, das kleine Molekulargewicht und pharmakologischen Eigenschaften des Inhibitors und erhöhten die Bindungsstärke der ProM Scaffolds gegen ein flaches Epitop drastisch. Die hier präsentierte Struktur-Affinitäts-Beziehung demonstriert die Vorteile der modular aufgebauten ProM Moleküle und ebnet den Weg für zukünftige *in vivo* Studien. Die strukturoptimierten ProMs könnten eine neue Klasse von Medikamenten in der Tumortherapie darstellen, die ihre Wirkung am Ende der Rezeptor Tyrosinkinasen und Integrin Signalkaskaden entfalten.

Summary

Metastasis is the major lethal attribute of cancer. However, the lack of antimetastatic drugs and the limited progress of metastasis-directed drug development efforts make new approaches essential in drug design. Despite being a marker of breast carcinogenesis and signature of invasive cancer progression, potent inhibitors against the actin regulatory protein family Ena/VASP remained elusive. Recently, we showed that interfering with Ena/VASP causes significant reduction in breast cancer cell invasion. We designed and evaluated diproline mimicking scaffolds, coined ProM, to inhibit the proline-mediated protein-protein interaction of the Ena/VASP EVH1 domains with a non-toxic, selective and cell-membrane-permeable inhibitor. However, the moderate affinity of the initial inhibitor will restrict the validation in future animal studies. In this work, we optimize the affinity of this inhibitor by using structural informations of 16 Ena/VASP EVH1 crystal structures. We solve the binding mode of the elongated peptide of the actin assembly-inducing protein (ActA) of *Listeria monocytogenes*. We mimic the found interactions of ActA with new scaffolds and develop a 734 Da compound with nanomolar affinity, which serves as novel starting point for *in vivo* studies. We conserved the structural simplicity, low-molecular weight and pharmacological properties of the inhibitor while drastically increasing the affinity of the scaffolds against the flat protein surface. The herein presented structure-affinity relationship study demonstrates the powerful modular architecture of our scaffold toolbox and paves the way for future *in vivo* studies. Structure-optimized ProM scaffolds might represent a novel class of antimetastatic drugs acting at the very end of converging receptor kinase signaling and integrin pathways.

Contents

List of Figures

List of Tables

Part I.

Introduction and Literature Survey

1. Searching for cancer metastasis-related targets

1.1. Metastatic disease indicates terminal illness

Tumor metastasis and the consequences of their treatment are the leading cause of mortality among cancer patients.[1,2] The pathogenesis of cancer metastasis consists of sequential steps that select for enhanced metastatic capacity and enable cells of the primary tumor to progressively colonize specific metastatic sites within distant organs.[1,3,4] A requirement for the metastatic process involves the invasion of the tissue surrounding the cancer site and the migration towards the bloodstream or lymphatic vessels.[3,4] Hereby the formation of adhesions and invadopodia is fundamental for the development of metastasis.[5–7] Importantly, these compartments allow cells of the primary lesion to reenter the systemic circulation to form secondary distant metastases. If diagnosed with such a distant metastatic disease assessed to the four most common types of cancer[1], fewer than 20% of the patients survive the next 5 years.[2,8] Strikingly, the last 10 years (2005–2015) of cancer research improved the overall 5-year survival rate for patients with diagnosed metastatic disease not even 3%.[1,8,9] How comes that metastasis-directed cancer therapy is nowhere further than 10 years ago?

[1] prostata, breast, lung, colorectal

1.2. Targeting the traction control of invasive cells

Integrins provide the necessary traction for focal adhesions and invadopodia to affect angiogenesis, viability, invasion and colonization of foreign tissue.[7,10–13] As integrins cooperate with oncogenes to initiate tumorigenesis and invasion,[14] the few approved drugs, notably cilengitide and dasatinib,[1,15,16] both intervene with the integrin-mediated pathways. While cilengitide specifically antagonizes $\alpha v\beta 3$ and $\alpha v\beta 5$ integrins[1,16] involved in angiogenesis and tumor growth,[15,16] dasatinib inhibits the downstream Src kinase important for regulation of actin dynamics[17,18] (Figure 1.1). Both antimetastatic drugs showed promising preclinical results but failed in clinical trials.[1,18,19] Inhibition of other antimetastatic targets like Src/Abl[19,20] or matrix metalloproteinases[21,22] showed no survival benefit, while FAK[23,24] kinase inhibitors are under ongoing clinical study trials.[25]

Downstream of these targets, receptor kinase signaling and integrin pathways ultimately recruit actin regulatory complexes such as Lpd, Ena/VASP and Scar/Wave (Fig. 1.1).[7,26,27] These non-enzymatic proteins connect the actin cytoskeleton with the extracellular matrix via dynamic protein clusters composed of integrins and signaling proteins.[5,6,10,11,28] Even though Gertler and Krause have been speculating that these proteins are part of a pro-metastatic pathway for over 10 years,[26,29] none of the actin regulatory complexes has been validated as druggable target so far. And this is not due to very recent discovery of the complexes – the major components of actin self-organization were discovered nearly two decades ago.[30] The apparent undruggability of the actin interactome lays in the fact that its players are unstructured to a large extend and exhibit weak affinities.

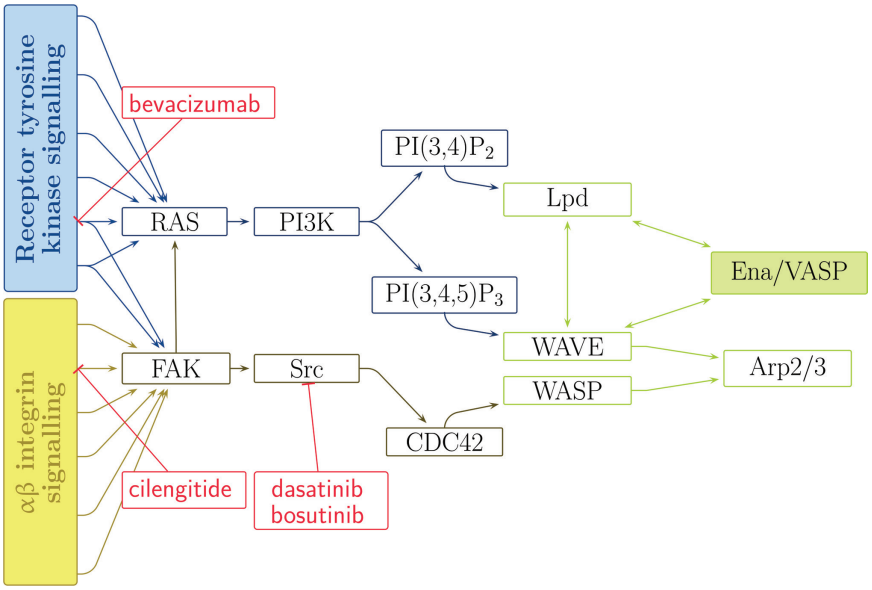


Fig. 1.1.: **Actin cytoskeleton at the very end of converging receptor kinase signaling**
Ena/VASP proteins are elongation factors of F-actin at the very end of converging receptor kinase signaling (blue) and integrin (yellow) pathways. Antimetastatic drugs (red) inhibit pro-metastatic signaling towards the actin interactome (green) that connects the actin cytoskeleton to the extracellular matrix via integrin-mediated adhesions and invadopodia.

2. Fibrous polymers as driving force for membrane protrusions

Fibrous networks traversing the inner cell lumen are composed of asymmetric monomers. Ultimately, the fiber serves as routing polymer and defines in an unambiguous bottom-up approach what is front and what is rear. This property is referred to as cell polarity.[28] Even lowest eukaryotic cells take advantage of such fibrous polymers to temporally restrict polarity during the cell cycle.[31] Controlled polymerization at the front of the fibrous lining and de-polymerization at the rear enables cells to fulfill complex tasks like the locomotion along chemotactic gradients. This activity requires the rearrangement of fibers according to extracellular signals.

2.1. Biochemistry of fast growing actin filaments

Directional motility needed for embryogenesis, wound healing, immune response, chemotaxis or metastasis is based on the rearrangement of one cellular fibrous network, the actin cytoskeleton. The motility mechanism by polymerizing fibrous actin (F-actin) is ancient and functionally conserved between protozoa and metazoa.[32]

However, spontaneous assembly of actin monomers is unfavorable as actin dimers and trimers are unstable. Cells therefore use nucleation promoting factors (NPFs) to overcome the instability of small actin oligomers. As the actin monomer is the most abundant protein in many eukaryotic cells, activity of NPFs is kept under tight control.[33,34] Once the filament is formed, diffusion limits the addition of monomers to the elongating tip of the fiber (barbed end). Consequently, the rate of elongation is proportional to the concentration of monomers. Hence, within a bounded compartment like a cell, growth at steady state is limited by the dissociation from the pointed end. Under physiological ion conditions the dissociation rate of around

$0.2\,s^{-1}$ corresponds to a growth speed of only 0.04 µm/min, which is much slower than the speed of locomoting keratocytes moving with up to 10 µm/min. Given their size, this speed requires the turnover of the entire actin network within a single minute.[32]

Migrating cells like keratocytes run actin polymerization not in steady state, but feed the growing fibers from a vast pool of unpolymerized actin monomers. A concentrated monomer pool sustains high rates of filament elongation and is ready to respond explosively when new barbed ends are created.[32,34] To inhibit spontaneous nucleation followed by an uncontrollable polymerization chain reaction, the actin monomers are bound to either profilin or thymosin-β4. Profilin binds the barbed end of actin monomers, thereby inhibiting polymerization and allowing the same time actin-profilin to bind to the barbed end of a fiber close to the rates similar to free actin.[32] Profilin furthermore binds to actin elongation factors.[34] This interaction captures actin monomers close to the growing fibers and enables actin assembly even above the rates of diffusion-limited growth.[35]

At steady state, rapid growth at barbed ends would deplete the pool of monomers substantially and eventually slow down the polymerization rate. To overcome this bottleneck, cells employ two mechanisms. Either disassembly of the pointed ends is sped up to refill the monomer pool, or the rate of drawdown is decreased by selectively terminating growth of barbed ends (capping). The latter mechanism is more efficient as global capping of barbed ends avoids nonproductive consumption of actin subunits elsewhere in the cell and gains local control over the length of fibers and their physical properties.[28]

2.2. An interactome regulates the leading edge

F-actin architecture is based on four activities: nucleation, branching, elongation and capping. Different mixtures of these activities causes the actin-polymer to exhibit different physical properties. Predominant elongation without capping leads to thin, finger-like membrane protrusions, while on the other hand, enhanced branching and capping activity creates a densely woven actin network. If the nucleation reaction is triggered at the surface of intracellular pathogens, the comet tail of actin filaments provides enough rocketing thrust by which the pathogen aims to infect neighboring

cells. Since the early 1990s, studies of the actin assembly-inducing protein (ActA) of the intracellular pathogen *Listeria monocytogenes* generated leaps in our understanding of how actin-based polymerization exerts pressure needed for directional cellular motility.[36–38]

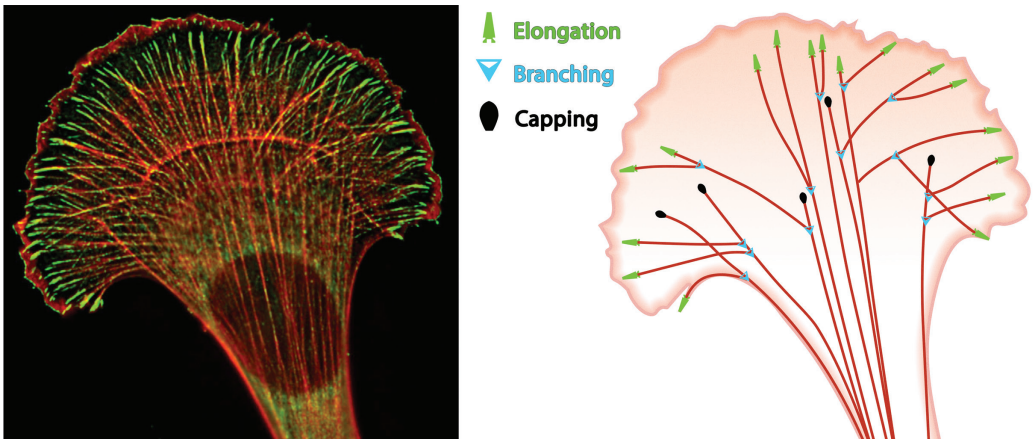


Fig. 2.1.: **Simplified cartoon of the leading edge**
Immunofluorescence staining for the elongation factor VASP (green) and F-actin (red) in a fibroblast.[39] The simplified cartoon on the right shows the main actors of the actin interactome. The protrusive force in the leading edge is formed by branching (blue triangles) of actin filaments (red). F-actin in close contact to the membrane is elongated (green), thereby acting against capping activity (black).

2.2.1. Branching actin fibers by Arp2/3

At the leading edge, the actin structure is composed of a highly branched meshwork. The branching activity of actin-related protein (Arp) 2/3 complex in cooperation with high capping activity produces short fibers that are able to sustain the pressure of a protruding membrane without buckling. Only growing filaments that are in close contact to the membrane remain effective generators of the propulsive force, as the elongation factors protect them from capping proteins and favor persistent fibers to push against the inner surface (Fig. 2.1). Branching by the Arp2/3 complex seems to be the more important activity than elongation, as it provides the stress in the growing network to protrude the membrane.[40,41] However, after intrinsically inactive Arp2/3 complexes formed new filaments, actin itself is co-activator and the reaction proceeds in an auto-catalytic manner.[32] Hence, the Arp2/3 complex is more effector

than a central orchestrator of the interactome. Consequently, inhibition of the Arp2/3 complex by selective small molecules has shown to not stop the gliding motility of keratocytes and revealed that the lamellar motility is less sensitive to inhibition of Arp2/3 compared to the *Listeria* comet tails.[42]

No direct interactions between the Arp2/3 complex and the elongation factors have been reported. However, such a coupling is strongly proposed by mathematical models that explain how a polymerizing filament rod conjoined to the membrane could produce unidirectional force. In the most refined model,[36] thermal fluctuations bend the filament away from the membrane. This ratcheting motion creates a gap sufficient to permit intercalation of a monomer. When the filament straightens again, it works against the load force of the membrane. Interesting consequence of such models is that the distance from the tip to the last branching point must be quiet short, as longer fibers take up the thermal motion to vibrate with inner modes and pushing becomes ineffective. To ensure optimal propulsive forces against the membrane, cells must have found a way to synchronize the rates of branching, capping and elongation within the meshwork of the leading edge.

The cross-talk between the elongation factors and the Arp2/3 complex happens not directly, but via the NPFs that activate the Arp2/3 complex. The Arp2/3-activating segment within these proteins is a homologous sequence called VCA domain and usually includes three short functional motifs composed of verprolin homology, connecting, and a segment of acidic regions. The already mentioned rate-limiting step of forming trimeric actin nuclei is catalyzed by the Arp2/3 complex and the V domain, wherein Arp2 and Arp3 mimic an actin-pseudodimer and the V domain captures actin monomers in close proximity for an efficient nucleation reaction.[43] The VCA domain is present in not less than 13 proteins forming the WASP and WAVE family.[43] Homologs of this family are found throughout all eukaryotic species from fungi to vertebrates and even plants, suggesting that the evolutionary history of the family extends further back than the divergence of eukaryotic lineages. Presumably, evolution of actin cytoskeleton itself and WASP/WAVE family are intimately connected as cells must have gained tight spatiotemporal control over actin dynamics by establishing this highly conserved WASP/WAVE-Arp2/3 axis.[43]

The nucleation-promoting activity of the C-terminal VCA domain is in both subfamilies strongly autoinhibited. The mechanisms differ as WASP exists indepen-

dently, while WAVE is sequestered within the pentameric WAVE regulatory complex.[44] WASP/WAVE activation is under the control of receptor kinases and integrins that signal via the Rho-family GTPases (RAS,[45] Cdc42,[46] Rac[47]) and cytosolic kinases (FAK and Src,[48] Abl,[49] Cdk5[50,51]) as well as phosphoinositides (Fig. 1.1). WASP/WAVE are therefore activated by membrane-bound signal pathways that convert shallow external gradients of attractants into steep intracellular signals. Downstream signaling then provokes the Arp2/3 complex to nucleate branching points.

2.2.2. Elongation by Ena/VASP and formins

The interactome consists of two sets of elongation factors that perform seemingly the same function: the *Drosophila* enabled/vasodilator stimulated phosphoprotein family (Ena/VASP)[29,52] and formin proteins.[53–55] While Ena/VASP is a pure polymerization machinery that shows no nucleation activity,[56] formins are efficient *de novo* nucleation factors.[35] Consequently, only Ena/VASP has shown to enhance Arp2/3 complex-based actin assembly by recruitment to activated Arp2/3 via members of the WAVE complex (Abi[57] or WAVE1[58,59]). F-actin structures based on the nucleation activity of either Arp2/3 or formins exhibit diametric mechanical properties.[28] While the Arp2/3 complex and Ena/VASP produce laminar membrane protrusions, formins and Ena/VASP, but not formins alone,[60] create long persistent plasma membrane extensions.

2.3. Ena/VASP orchestrates the interactome at the leading edge

Both mentioned actin structures relay on Ena/VASP and are needed for directed movement. The formin-based bundles extend as filopodia[61,62] into the environment as exploratory compartments, while the Arp2/3-based meshwork forms the leading edge (lamellipodia).[63–66] Cancer cells use a unique form of the lamellipodia,[67] termed invadopodia, to couple directional movement to the degradation of extracellular matrix.[5] This raises two immediate questions. How does the rearrangement of actin mediate between exploration and locomotion/invasion? Does a

central node of the interactome trigger the rearrangement of filopodia to lamellipodia/invadopodia?

Two recent studies show that WAVE and the Arp2/3 complex jointly capture formins to inhibit filopodia formation[68] and that Ena/VASP can accumulate and pause active formins to temporally block formin-dependent nucleation.[69] Hence, the interactome is organized in a hierarchical way so that formin-driven actin assembly is suppressed by the WAVE-activated Arp2/3 complex and Ena/VASP. Of the latter two regulators, Ena/VASP can enhance Arp2/3-based actin assembly[59] but shows no nucleation activity itself.[56] This restriction appears to be of crucial importance as the interactome only needs to control the explosive polymerization activity of Arp2/3[70–72] via its activators. Once the WAVE complex signals into the interactome,[73] Ena/VASP orchestrates the rearrangement of the actin structure[57–60,68,69,74] and connects the barbed ends of actin filaments to the extracellular matrix via adhesion sites.[75,76] Otherwise the established load force of the protruding membrane would drive the filaments rearward into the lumen and not the plasma membrane forward. The adhesion sites are dynamic integrin-mediated protein clusters that are recognized by Ena/VASP directly.[6,10,11,63] Consequently, the absence of Ena/VASP has dramatic effects on the outside-in signaling to anchor the cell in the proper tissue,[77,78] as well as the ability to move along gradients.[79,80]

Metastasis is a hallmark of cancer[81] and is based on the actin cytoskeleton that provides the pressure needed for directional cellular motility through the extracellular matrix. Ena/VASP thereby provides the physical linkage of the actin cytoskeleton to the extracellular matrix via invadopodia[82,83] and has recently shown to be a requirement in the inside-out signaling of tumor cells to remodel the extracellular matrix and the haptotactic response on fibronectin in the periphery of the tumor tissue.[84] Indeed, Ena/VASP is discussed as part of the invasive signature and as a marker of breast carcinogenesis[85] and plays a significant role in the transformation of benign lesions into malignant invasive and metastatic cancer.[80,86] The protein family is further over-expressed in the advanced tumor stage[87–89] and disruption has shown to drastically effect *in vitro* and *in vivo* migration.[90] Yet, potent drugs to interfere with Ena/VASP *in vivo* remain elusive.

3. Structure and recognition motif of Ena/VASP EVH1

3.1. Targeting the localization of Ena/VASP by interfering with EVH1

The three mammalian Ena/VASP family paralogs, enabled homolog (ENAH, also called Mena), VASP, Ena-VASP-like (EVL) share a tripartite structural organization in which two homology domains (Ena/VASP homology (EVH) 1 and 2) are interspaced by a more divergent proline-rich central part. Elongation of barbed end filaments, protection from capping proteins and tetramerization is mediated by the C-terminal EVH2 domain,[52,56,91] which harbors binding sites for monomeric and filamentous actin. The subcellular distribution of Ena/VASP is regulated by the N-terminal Ena/VASP EVH1 domain. It interacts with adherens-type cell-matrix or cell-cell junction proteins at focal adhesions,[63,64] the leading edge[65,66] and invadopodia[5,82] to ensure attachment to the substrate.[92] In the same way the interactome recruits Ena/VASP to interaction partners like Lpd[66] (also known as RAPH1), vinculin[63,75] or zyxin,[76,93] *L. monocytogenes* captures Ena/VASP by presenting the recognition motif of EVH1 in four repeats by the surface protein ActA.[37,94]

The central proline-rich part and the EVH2 domain are intrinsically disordered segments and hence display undruggable targets for small molecules. In sharp contrast, EVH1 folds into a structured globular domain of the Pleckstrin homology domain-like superfamily (Pfam ID – PF00169, CL0266) and recognizes short linear proline-rich sequences (PRS) in poly-*L*-proline type II helix (PPII) conformation (Fig. 3.1).

The Ena/VASP EVH1 represents a promising stable target as the invasion-specific isoforms of Ena/VASP show no alternation in the splicing of the EVH1 domain. [95–98]

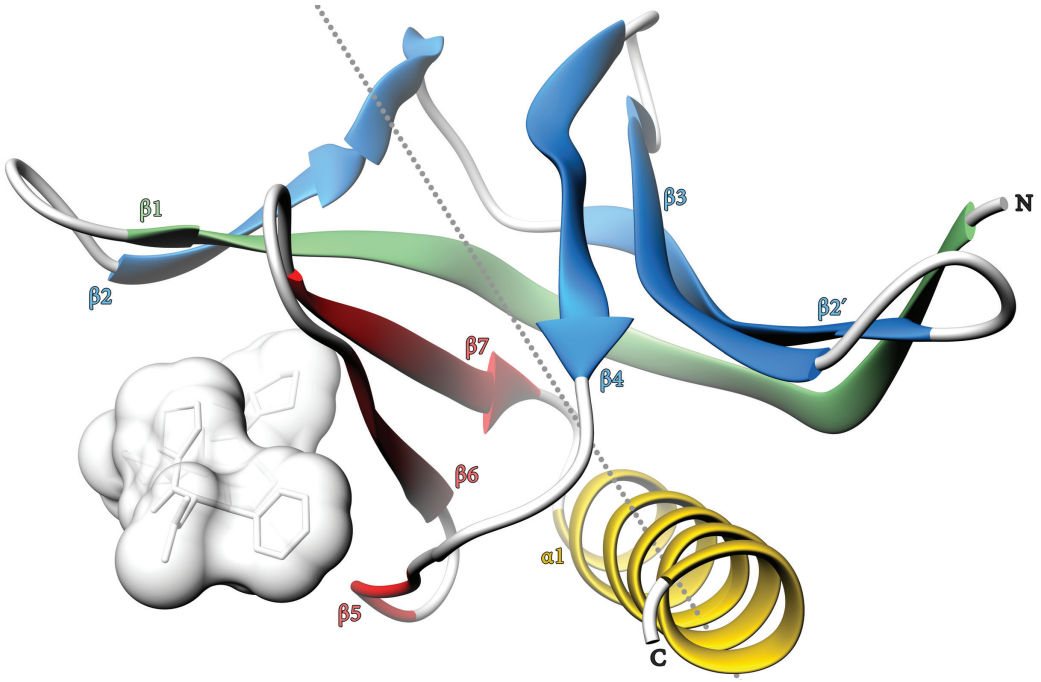


Fig. 3.1.: **Fold of the Ena/VASP EVH1 domain**
The EVH1 domains fold into two β-sheets (red/green and blue/green) that form an incomplete β-barrel (barrel axis as dotted gray line), which is C-terminally capped by a α-helix (yellow). Strands β1, 2, 5-7 provide a shallow binding groove for the ligand. Depicted is Mena EVH1 with Ac–FPPPPT–NH_2 as white sticks and solvent accessible surface (PDB code 1EVH). Image view inspired by Volkman *et al.*. [99]

3.2. Structural description and recognition pattern of the EVH1 binding epitope

The roughly 115 amino acid long EVH1 adapter domains are classified according to their recognition motifs [100] [F/W/L/Y]PxϕP (ϕ hydrophobic, x any: class I Ena/VASP), PPxx(F/Y) (class II Homer-Vesl) and class III found in the WASP/N-WASP protein family with untypical far-spread consensus motif. [99] The different binding mode of class I and II ligands originates from mutations on the protein surface,

which are discussed later when comparing the Homer1 EVH1 interaction site with ENAH EVH1 binding C-terminally elongated ActA-derived peptides (Sec. 12.2.3).

Selectivity towards these proline-rich motifs originates primary from recognizing prolines over any other amino acid and not by establishing complementarity with deep binding pockets.[101,102] Target proteins are bound weakly via a short, continuous, shallow, non-polar binding groove which is surrounded by several polar residues, such as Gln79, Arg81, Asn90 (ENAH EVH1 numbering if not mentioned otherwise, Fig. 3.2).

Binding affinity towards Ena/VASP EVH1 is mainly achieved by a target sequence commonly found as as ^{1}FPPPP.[103] EVH1 binding is successfully abolished with single mutations within this pentamer,[104] which is accordingly referred to as the core recognition motif. In both[1] available crystal structures of EVH1 domains, only the core recognition motif is resolved as it is structurally bound by two hydrogen bonds (H-bonds) via the carbonyl oxygens of ^{2}Pro and ^{3}Pro to Gln79 and Trp23 (Fig. 3.2b). The N-terminal non-proline amino acid ^{1}Phe (blue in Fig. 3.2) resides in the deepest pocket that EVH1 domains provide for the ligand. As shown by single residue peptide substitution experiments,[103] the pocket of VASP EVH1 is flexible enough to host bulky aromatic residues like ^{1}Trp. This is, as the N-terminal side of the main binding groove is shaped by two flexible, positively charged Lys69 and Arg81 that encircle the non-proline amino acid of the ligand and cause a strong preference for N-terminally acetylated ligands.[105]

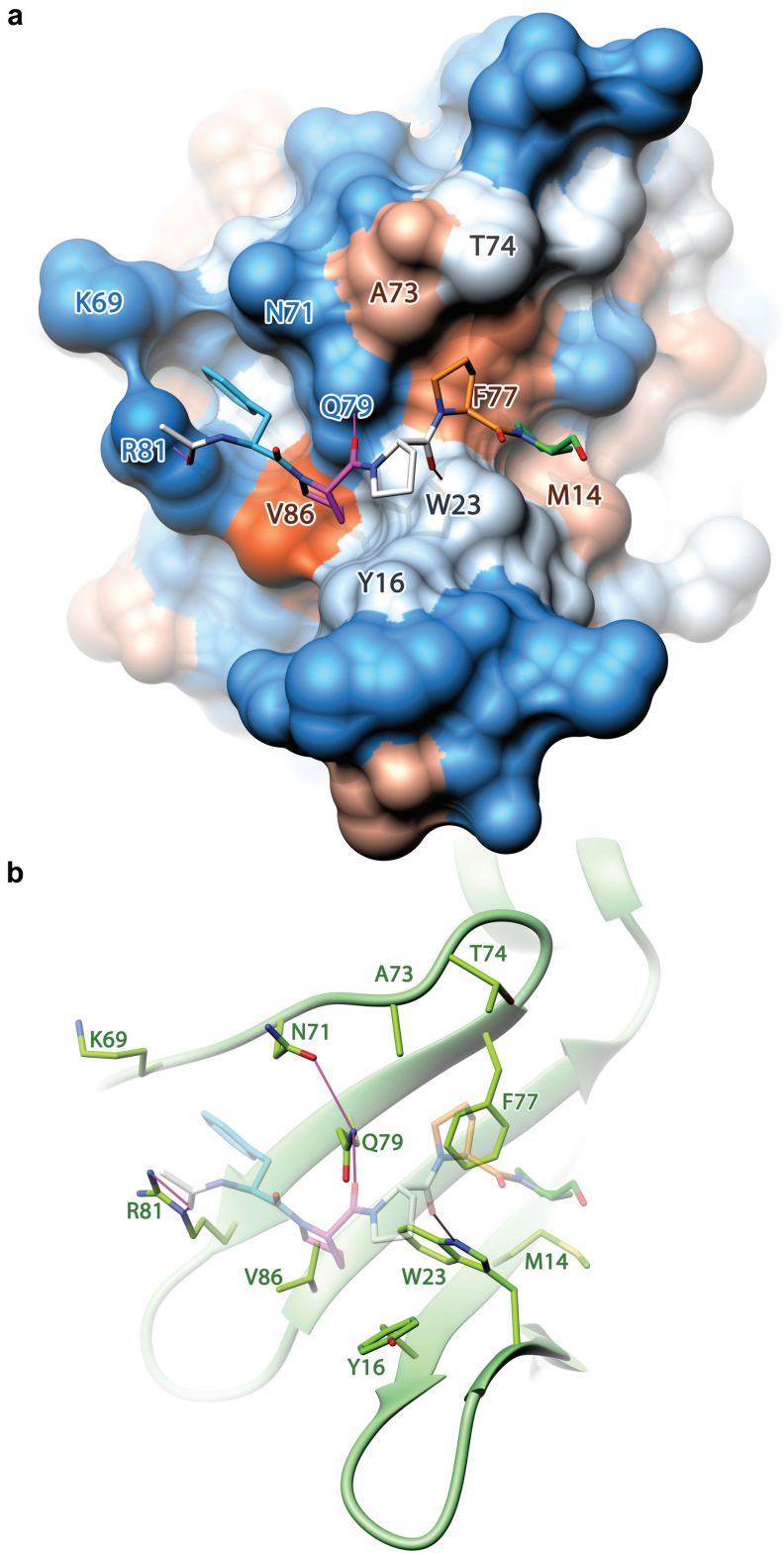
Opposite of Arg81, the pocket is pre-structured by the H-bond providing doublet Asn71-Gln79 and Val86 at the bottom (Fig. 3.2b). Val86 is part of the main binding groove that is shaped by three π stacked aromatic residues Tyr16, Trp23 and Phe77. Trp23 is sandwiched within the aromatic triad and provides the central second H-bond (Fig. 3.2b). As a consequence of the directed interactions by the carbonyl groups of the core recognition motif, the prolines do not interact equally intimate with the epitope of EVH1. The pyrrolidine rings facing the protein surface the closest (^{2}Pro magenta and ^{5}Pro green in Fig. 3.2a) can not be exchanged by any naturally occurring amino acid,[106] while the solvent-exposed ^{3}Pro (white) is exchangeable with any amino acid. Note that the discrimination mechanism of EVH1 domains is not by establishing complementary with the entire proline sidechain, but by fixing the ligand

[1] Mena EVH1 and Ac–FPPPPT–NH_2; EVL EVH1 and Ac–FEFPPPPTDEE–NH_2

with hydrogen bonds in an orientation in which two C_α-C_β exit vectors point straight on solvent-exposed aromatic rings Tyr16 and Phe77. The hydrogen-bonding network selects for PPII backbone conformation and induces heavy restraints for the side chains of amino acids 2 and 5. At these positions, the unique amide N-substitution of proline is not compensated by any mutation, as C_δ is needed at a position only reachable by proline. The pyrrolidine ring of ^{4}Pro (orange in Fig. 3.2a) however is positioned flat over Phe77 in an orientation where N-amide substitution is not needed to maintain contact to the protein. The patch underneath position 4 of the core recognition motif is pharmacologically interesting. Phe77 is surrounded by more hydrophobic residues Ala73 and Met14 that open the binding groove to a 10 Å-long hydrophobic basin that accepts aliphatic residues beside proline.

Fig. 3.2.: **The binding epitope of ENAH EVH1**

(**a**) Detailed depiction of amino acids shaping the main binding groove of ENAH EVH1. Solvent accessible surface is color coded by hydrophobicity (blue-white-red for hydrophilic-hydrophobic). Proline-rich ligands like Ac–FPPPPT–NH_2 are bound with three H-bonds (magenta) onto a shallow binding groove. (**b**) For better visibility of the H-bond network and the solvent-exposed amino acids, the ligand is transparent. Both panels are aligned and have the same viewing angle and scale (PDB code 1EVH).


a
T74
A73
K69
N71
F77
Q79
R81
M14
V86
W23
Y16
b
A73
T74
K69
N71
F77
Q79
R81
M14
V86
W23
Y16

4. Proline-rich motifs

Transduction of external signals to rearrange the actin cytoskeleton requires a large amount of site-specific assemblies of multi-protein complexes that are transient by necessity.[102,107] To establish the protein-protein interactions, nature strikingly often chose solvent-exposed proline-rich segments to interact with adapter domains.[108] As we saw above, this holds true also for the EVH1 domain, which we aim to target with a small molecule. To successfully mimic proline-mediated interactions, it is therefore needed to understand the behavior of polyproline peptides.

4.1. Proline-rich segments

Solvent exposed PRS exhibit three outstanding properties that arise from the cyclic nature of proline. Unlike the other 19 naturally occurring amino acids, the side chain is cyclized back on to the backbone amide position. The severe backbone angle restriction of proline limits the available backbone dihedral angle to around $\phi = -65°$. In a consecutive proline sequence, the mobility restrictions become so dominant that poly-*L*-proline peptides tend to adopt an extended structure known as the left-handed polyproline II (PPII) helix.[109] The secondary structure element looses little conformational entropy upon binding and thus binds more favorably than other exposed peptide sequences.[108] The dihedral angles of $\phi = -78°$ and $\psi = +146°$[110,111] align the carbonyl groups of a polyproline peptide nearly perpendicular to the helical axis. As polyproline can not form any intramolecular interactions, the carbonyl groups become fully available for intermolecular H-bond networks with the protein epitope. The replacement of the amide proton by the methylene group with higher electron-donating potential[112] causes the amide carbonyl to be electron-rich and a strong hydrogen bond acceptor. This property assures PRS to be highly soluble in water.[113,114]

These three chemical properties of high solubility, full accessibility of the backbone carbonyl groups, and a well defined secondary structure even in the unbound state explain the fact that proline-rich segments belong to the most abundant sequence motifs of the eukaryotic proteome.[115] However, developing small potent inhibitors that mimic the secondary structure element of PRS *and* meet necessary pharmacological properties remains highly challenging.[116–119]

4.2. Recognition modules of proline-rich segments

PRS-recognizing protein domains (PRDs) take advantage of the 3_2 screw symmetry along the PPII helical axis. Compared to an α-helix, PPII helices possess three residues per turn. The trigonal profile of PRS conserved only two ligand recognition motifs.[100,103] Depending on the interaction site with the helix translation, the six distinct families of PRDs recognize either PxxP[1] or xPPx[2] as visible in figure 4.1. In all recognition motifs, proline predominates in PPII helices and polar amino acids are favored over nonpolar, especially over aromatic residues.[114]

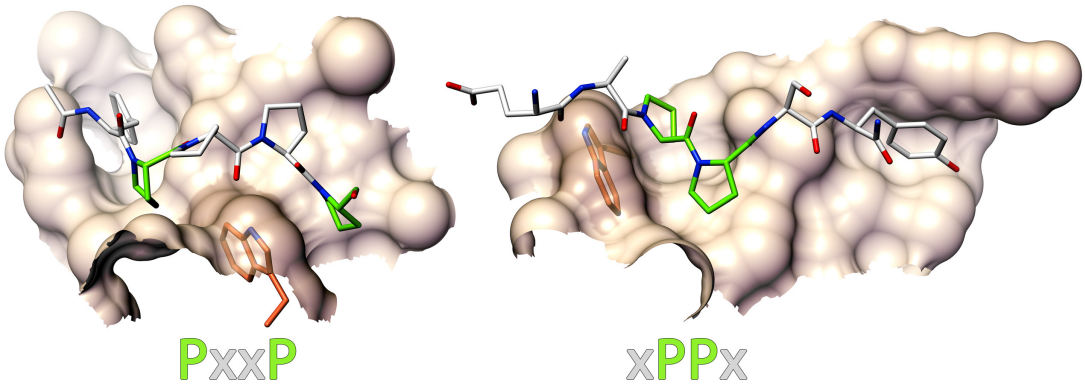


Fig. 4.1.: **Polyproline helix II as structural recognition motif**
Polyproline helices display a trigonal profile. The threefold axial symmetry conserved two different recognition patterns PxxP or xPPx, depending on which helical face is bound by a central aromatic residue (Trp, orange). Left panel: Core recognition motif FPxxP bound by Mena EVH1 (PDB code 1EVH). Right panel: Core recognition motif PPxY bound by Nedd4 WW (PDB code 4N7H).

[1] SH3,[120,121] EVH1,[75,93,122] and the single-domain profilin[123,124]
[2] WW,[125,126] GYF,[127,128] and UEV[129,130] domains

Beside the pseudo threefold rotational symmetry along the helical axis, PPII helices possess a twofold pseudo-symmetry perpendicular to the helix axis. The rotational symmetry transforms the backbone atoms approximately to the same position and allows formation of very similar hydrogen-bonding networks in both orientations. Consequently, peptides could bind in two directions relative to the domain (often termed forward and reverse). Both ligand binding modes have been reported for SH3, WW domains, and profilin. For example, the Src SH3 domain binds two near-palindromic peptides RPLPPLP and the reversed version PPVPPR in accordance to the underlined class I and II recognition patterns RxLPxxP and PxLPxR.[131]

5. Previous attempts to interfere with Ena/VASP EVH1

To the best of our knowledge, all previous attempts to design small molecules against the binding epitope of Ena/VASP EVH1 have failed so far. Ligands with non-peptidyl drug-like features either kept only moderate binding affinity for VASP EVH1, or biased miniproteins never went past *in vitro* experiments for obvious reasons of membrane impermeability or degradation by cytosolic proteases.

5.1. Mimicking the interface of EVH1 domains

Using the structural insights of the ActA-derived peptide bound to Mena EVH1, the group of Alanna Schepartz used the well-folded miniature protein avian pancreatic polypeptide (aPP) to mimic the core recognition motif and flanking residues of ActA (Fig. 5.1). Composed of a PPII-helix–turn–α-helix motif, pGolemi[132] was optimized during the next few years by coworkers. pGolemi and its successor, pGol-2,[133,134] have shown to compete against ActA for Ena/VASP EVH1 domains, both, *in vitro* as well as in *Xenopus laevis* egg cytoplasmic extracts. pGolemi and pGol-2 reached paralog specificity[1] and were able to maintain the binding strength of the ActA-derived 11-mer Ac–^{1}DFPPPPTDEEL–NH_2 used that lab.[135]

It is important to note that the measured affinities hold only in the context of the entire miniprotein, as removing the α-helical epitope abolishes binding to Mena EVH1 entirely. To explain this circumstance, it was argued that pGolemi pays a lower entropic cost upon binding, as the α-helix prestructures the crafted binding epitope already in PPII helix conformation.[134] However, the correlation between the level

[1] Mena EVH1: 600 nM with pGolemi, VASP EVH1: 13 µM with pGolemi, EVL EVH1: 10 µM with pGol-2

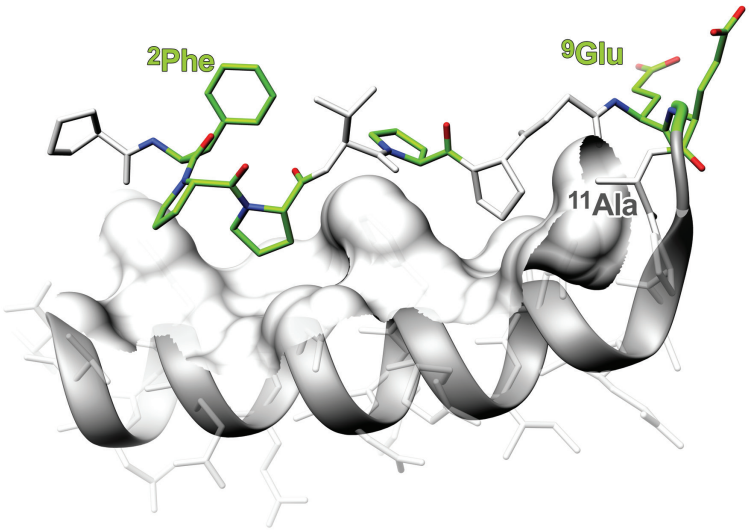


Fig. 5.1.: **FPPxP pasted into miniprotein pGolemi**
pGolemi is a stable miniprotein derived from avian pancreatic polypeptide. Crucial amino acids from the ActA 11-mer were pasted into the PPII portion (FPPxPxxEE, green) of pGolemi. For better visibility, the unimportant amino acids of the α-helix are faded out. Image view inspired by Holtzman *et al.*[134]

of α-helicity[2] of the unbound pGolemi and ΔG is slightly visible only for the binding to Mena EVH1.[135] One mutation (A11L) clearly not following this trend is situated in the β-turn (Fig. 5.1) and was concluded to contact the protein surface, too.[135] Interestingly, ^{11}Leu and the preceding ^{10}Glu are conserved in all four ActA repeats[3] and seem to be important for binding.[103,136] However, in pGolemi only A11L is able to modulate the affinity, while the alanine scan of adjacent E9A and E10A do not.[134] The discrepancy within this β-turn could indicate that pGolemi is not binding canonically to the EVH1 domains. An altered binding mode of pGolemi would explain why the α-helix contributes –3.5 kJ/mol to the free energy upon binding.[132] Probably the α-helix does not solely stabilize the PPII-fold as claimed, but interacts with the Ena/VASP EVH1 domains as well.

Given the shallow main binding groove of EVH1, the miniproteins might have never interacted solely with their PPII-helix portion, but instead took advantage of the additional surface provided by the β-turn and the α-helix. In a follow-up study, the sequence of pGol-2 was screened in phage display assays for even higher affinities.[135]

[2] mean residue ellipticity Θ at 222 nm
[3] $^{265/300}$FPPPPTDEEL, ^{335}FPPPPTEDEL, ^{380}FPPIPTEEEL

Doing so, mutations of pGol-2 were prone to establish alternative interaction sites to bypass the demanding proline-based EVH1 interaction. The endeavor to keep pGol-2 within the canonical binding groove was not successful, even though crucial prolines[4] within the class I consensus motif were conserved. The miniproteins obtained from the phage display, called EVH1-binding peptides (Ebp), were not competing with the ActA 11-mer anymore. Furthermore, truncated Ebp 13-mers containing all randomized amino acids did not bind to VASP EVH1.[135] With 3.46 kDa weight, such a peptide-based construct was unsuitable as a drug, anyway.

Based on the work of the Schepartz lab, another attempt was reported in 2006 in which the interaction was addressed by mimicking the peptide binding groove of EVH1.[137] Via thioether-linked synthetic peptides, the epitope was partly reconstructed and was able to compete against immobilized Mena EVH1 for the ActA 11-mer and pGolemi. The approach was not investigated further as the binding affinity to natural ligands was lower compared to the full-length EVH1 domain.

5.2. Mimicking proline-rich recognition motifs

Single glycine N-substituted residues were successfully incorporated into the ligands of SH3 domains, yielding in so called peptoids that completely replaced the PxxP core by non-proline residues. Thereby either single[138] or double[101] substitutions of the critical ^{5}Pro and ^{8}Pro of ^{1}YEVPPPVPPRRR increased the affinity for N-Grb2 SH3 up to 50-fold compared to the parent peptide. Beside replacements of prolines, a combination of leucine substitutions and decorated termini was successful.[102] Here, using the starting sequence ^{1}RALPPLP yielded in a ligand whose affinity was boosted nearly 1000 times to 25 nM for Fyn SH3. Based on such encouraging results, amide N-substituted peptoid building blocks were also incorporated into the ActA-derived 13-mer SFE^{1}FPPPPTEDEL to replace ^{2}Pro and/or ^{5}Pro.[136] Using a combination of sarcosine-derived residues, the obtained hybrid ligands lost at least half of their affinity for VASP EVH1. This work showed though that VASP EVH1 accepts exchange of critical and conserved prolines of the recognition motif.

The astonishing selectivity and affinity boosts of ligands for SH3 domains emphasize that the requirement for successful proline mimicry is the conservation of

[4] PxxP or PxPP in two different phage libraries

the amide nitrogen substitution rather than the cyclic nature of proline.[138] However, it is entropically advantageous to restrict the ligand in a cyclized conformation that resembles the bound state. Three groups reported such compounds, although not to target Ena/VASP. Morken and coworkes screened a chemical library of conformally restricted hydrophobic non-natural amino acids to combine three building blocks that mimic ^{10}PLP in Ac–^{1}VSLARRPLPPLP–OH. The best compounds exhibited nearly 6-8 times weaker binding to Src SH3 compared to the parent 12-mer.[139]

Recently, Aillard *et al.* reported a Xaa-*trans*-Pro mimicking scaffold with high similarity to our approach (Chapter 6). The azabicycloalkane-based scaffold was selected from a set of *in silico* minimized candidates to replace Xaa-Pro in PPII conformation (Fig. 5.2). The scaffold was evaluated in the peptide recognition motif Ac–^{1}RPLPVAPG–NH_2 of Fyn SH3 as replacement of ^{3}LP and/or ^{6}AP. While the parent peptide exhibited a K_d of 41 µM, the single-substitution of ^{6}AP as best ligand composition lost over 7 times affinity and bound with 300 µM.[117]

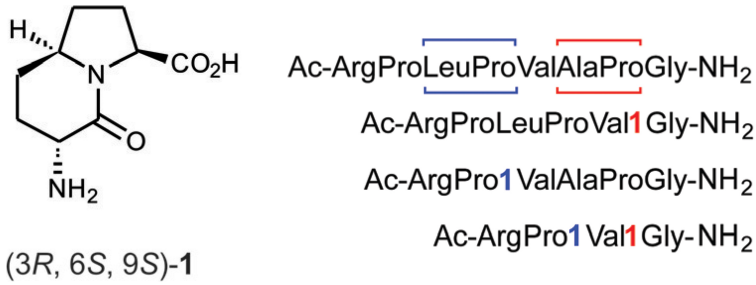


Fig. 5.2.: **Azabicycloalkane-based peptidomimetic replacing Xaa-Pro**
Aillard and coworkers reported 2015 a chiral compound replacing ^{3}LP and/or ^{6}AP in ^{1}RPLPVAPG. While the parent peptide bound with a K_d of 41 µM, the single-substitution of ^{6}AP bound with 300 µM to Fyn SH3. Binding was not detectable for compositions with ^{2}LP replacement. Modified reprint of the original figures in Aillard *et al.*[117]

Only one group succeeded with a conformationally constrained scaffold that maintained the binding strength when incorporated into the starting peptide. Witter *et al.* designed a chiral spirolactam-based scaffold mimicking with ^{5}PV a critical proline of the consensus motif RxxPxxP within ^{1}PRPLPVAPG (Fig. 5.3). Ligands containing scaffolds in the proper configuration displayed IC_{50} values for Lyn SH3 similar to the parent peptides.[118]

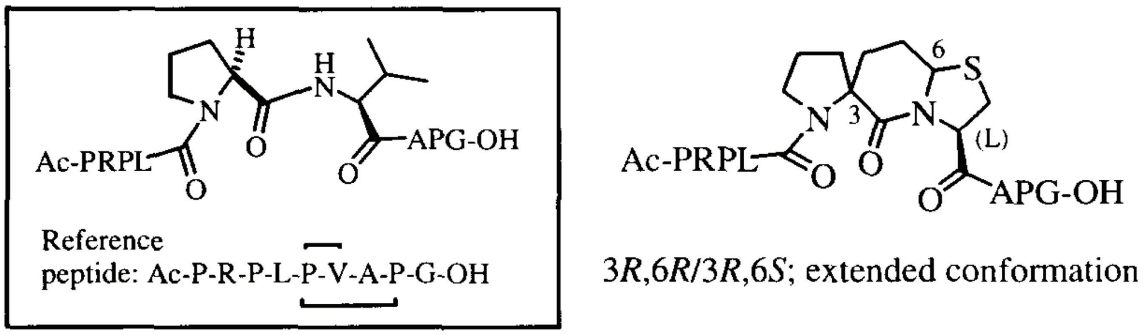


Fig. 5.3.: **Spirolactam-based peptidomimetic replacing Pro-Val**
Witter and coworkers reported 1998 a chiral compound replacing ^{5}PV in ^{1}PRPLPVAPG (left). The diastereoisomers with the proper conformation of the pyrrolidine ring ($3R$) bind with IC_{50} similar to the canonical peptide. Modified reprint of the original figure in Witter *et al.*[118]

6. A novel approach in drug design renders Ena/VASP a druggable target

6.1. The architecture of ProM scaffolds

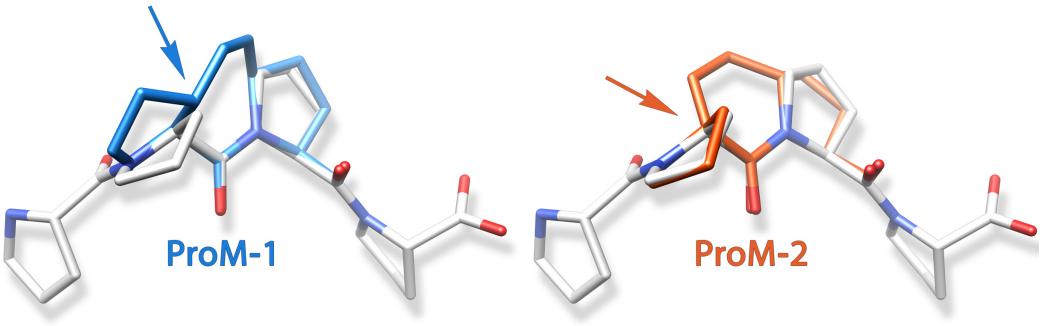


Fig. 6.1.: **Fixation of two subsequent prolines by a vinylidene bridge**
PRS (white) in PPII helix conformation are stiffened by stereoselective fixation. The N-terminally proceeding proline is either connected by the C_β (ProM-1) or the C_α (ProM-2), forming two different proline mimetic scaffolds.

Compared to the aforementioned tricyclic spirolactam scaffold (Fig. 5.3), the proline mimetic scaffolds (ProMs) used in this work exhibit twice the amount of chiral centers and a highly challenging synthesis strategy. The synthesis route produces a distinct vinylidene bridge as part of a six to eight membered stiff heteroaromatic ring that is flanked by pyrrolidine rings. However, the steric demands of the binding grooves are too diverse to address a broad panel of PRDs with only one scaffold. Therefore we chose a different approach in which the scaffold is composed of two vinyl building blocks that contribute to a central ring. The modular concept allows

independent adaptations of either building block and produced nearly a dozen scaffolds. While the re-design of building blocks was guided by *in silico* docking studies and crystal structures produced at the Leibniz-Institut für Molekulare Pharmakologie in Berlin, the new synthesis strategies were found by our collaborators at the University in Cologne in the group of Prof. Dr. H.-G. Schmalz. With this multidisciplinary approach, ProM scaffolds can be tuned to accurately fit onto any PRD. For obvious reasons the design is not restricted to natural amino acids: beside connecting two (*L*)-proline precursors mimicking Pro-Pro in PPII conformation [140–142] (ProM-1, ProM-2), the pyrrolidine rings on either side are extend- or shrinkable (ProM-4, ProM-8[1], ProM-16[1]), stereoselectively decorated (ProM-9, ProM-13), can contain heteroatoms (ProM-6[1]), or display inversed C_α chirality (ProM-12). Similar to the scaffold of Aillard and coworkers, [117] the proline ring can be opened to mimic Leu or even smaller moieties attached to the C_β (ProM-3, ProM-10[1], ProM-15[1]). The common structural element remains the recognizable vinylidene bridge-containing ring at which either the N- or the C-terminal prolines were modified. As a result, the current ProM suite contains scaffolds not even mimicking PPII conformation but serving as α-helix inducer (ProM-5). [143]

[1] unpatented scaffold

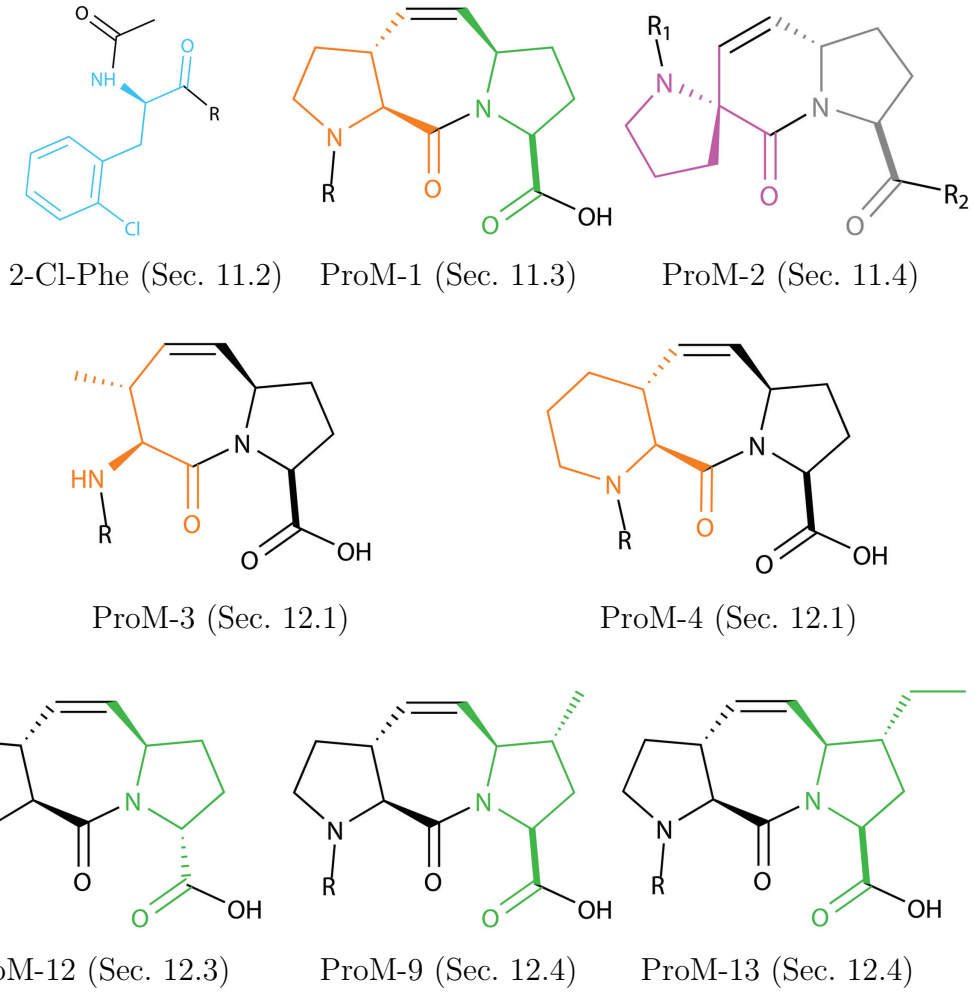


Fig. 6.2.: **The ProM scaffold suite and the corresponding sections in the discussion**
Ena/VASP is addressed by replacing the core recognition motif FPPPP with non-natural moieties. Each replacement is discussed in a separate section. *In silico* designed di-proline mimicking (ProM) scaffolds successfully replaced prolines of PRS. The enantioselective synthesis strategy via Ru-catalyzed ring closing metathesis of two vinyl building blocks yielded in sp^3-rich chiral scaffolds that fulfill specific structural requirements of the addressed PRS recognizing domain.[140–143]

6.2. Can ProM scaffolds be tuned for a high-affinity inhibitor against a pro-metastatic target?

Protein-protein interactions are viewed as challenging targets for small molecules.[116] As reviewed in section 5, the previous attempts to mimic the proline-mediated interaction of Ena/VASP have failed so far, even though the protein family represent a promising target to address metastasis downstream of EGF-dependent protrusive and migratory activity of tumor cells.[96]

This work aims to design a high-affine cell-membrane-permeable inhibitor against Ena/VASP. Previous work[105] succeeded in mimicking the proline-mediated interaction of Ena/VASP EVH1 domains but failed to provide structural informations of the bound inhibitor. In light of a possible drug, high-precision protein structures were a crucial prerequisite to improve the binding strength of the initial inhibitor even further.[116] Above all, the thesis investigated whether the protein purification could be optimized to crystallize Ena/VASP EVH1. Production of high-resolution structures could then serve as templates for *in silico* docking studies to optimize the initial inhibitor. Using such a novel structure-based approach, the thesis should furthermore validate whether docking studies and biophysical characterizations can be implemented in new scaffold synthesis routes. Such a work-flow could feed itself with new inhibitor generations that are iteratively crystallized, characterized, *in silico* optimized and synthesized.

High affinity is required but not sufficient for the design of an antimetastatic drug which must achieve additional pharmaceutical properties such as cell-membrane-permeability, selectivity, oral bioavailability, non-toxicity, as well as metabolic, microsomal and plasma stability. This thesis should pave the way for further *in vivo* evaluation of a possible novel antimetastatic target by providing precise crystal structures that allow selective modifications of the drug lead without inflating the molecular weight nor loosing beneficial pharmacological properties, while confirming a conserved binding mode.

Part II.

Materials and Methods

7. Materials

7.1. Chemicals

Chemicals were purchased from the following providers and were sufficiently pure to be used without further purification:

Roth	
Hepes	pufferan ≥ 99.5%
2-mercaptoethanol	78.13 g/mol, 99%, 4227.3
L-glutathion reduced (GSH)	307.33 g/mol, ≥ 98%, 6382.2
ethylendiaminetetraacetic acid disodium salt dihydrate (EDTA)	372.24 g/mol, ≥ 99%, 8043.2
isopropyl-β-D-thiogalactoside (IPTG)	238.30 g/mol, ≥ 99%, CN08.3
TRIS-hydrochloride	$C_4H_{11}NO_3$·HCl ≥ 99%, 157.60 g/mol, 9090.3
ethanol	≥ 99.5%, 46.07 g/mol, 5054.3
milk powder	blot grade
2×YT Broth	X966.3: 16 g/l Trypton, 10 g/l yeast, 5 g/l NaCl, pH 7
Merck	
1,4-Dioxane	$C_4H_8O_2$ ≥ 99.8%, 88.11 g/mol, 31.05.06
di-ammonium hydrogen phosphate	$(NH_4)_2HPO_4$, ≥ 99%, 132.05 g/mol, 1.01207.0500
ammonium sulphate	$(NH_4)_2SO_4$ ≥ 99%, 132.14 g/mol, 1.01216.5000
potassium chloride	KCl ≥ 99.5%, 74.56 g/mol, 1.04936.1000
citric acid monohydrate	$C_6H_8O_7$·H_2O for analysis, 210.14 g/mol, 1.00244.0500
poly(ethylene glycol) dimethyl ether (PEG1000)	950-1050 g/mol, 8.07488.1000
Sigma	
Ampicillin sodium salt	371.39 g/mol, A9518-25G

continued on next page

sodium phosphate monobasic	NaH_2PO_4, 119.98 g/mol, ≥ 99%, S0751-1KG
sodium phosphate tribasic dodecahydrate	$Na_3PO_4 \cdot 12H_2O$, 380.12 g/mol, ≥ 98%, 222003-500G
Fluka	
ammonium nitrate	NH_4NO_3 ≥ 99%, 80.04 g/mol, 09891
lithium chloride anhydrous	LiCl ≥ 99%, 42.39 g/mol, 62478
lithium acetate dihydrate	$C_2H_3LiO_2 \cdot 2H_2O$ ≥ 97%, 102.01 g/mol, 62395
lithium nitrate	$LiNO_3$ ≥ 98%, 68.94 g/mol, 62575
potassium bromide	KBr ≥ 99.5%, 119.01 g/mol, 60093
potassium nitrate	KNO_3 ≥ 99%, 101.11 g/mol, 60415
sodium thiocyanate	NaCNS ≥ 98%, 81.07 g/mol, 71938
MES monohydrate	$C_6H_{13}NO_4 \cdot H_2O$ ≥ 99%, 213.25 g/mol, 69892
Poly(ethylene glycol) dimethyl ether (PEG400)	ultra, 380-420 g/mol, 91893
ammonium formate	CH_5NO_2 ≥ 99%, 63.06 g/mol, 09735
dimethylsulfoxide (DMSO)	78.13 g/mol, ≥ 99.9%, 41640
Fischer Scientific	
hydrochloric acid 36%	HCl, ≈36%, H/1200/PB17
Invitrogen	
SDS Mark12 (unstained)	100006637
Roche	
COmplete Tables, Mini EDTA-free, EASYpack	Protease inhibitor cocktail tablets, 04 693 159 001
Gold biotechnology	
Tris (2-carboxyethyl) phosphine HCl (TCEP-HCl)	286.65 g/mol, 51805-45-9
Stratagene	
E. coli BL21(DE3)	One Shot, F- ompT gal [dcm] hsd SB (r-B m-B) λ-Prophage DE3 T7 RNA-polymerase
AppliChem	
Tris (2-carboxyethyl) phosphine HCl (TCEP-HCl)	286.70 g/mol, ≥ 98%, A2233.0001
sodium choride	NaCl ≥ 99.9%, 58.44 g/mol, A4661,5000
Biochrom	
PBS Dulbecco (PBS)	9.55 g/mol, L 182-10

continued on next page

Calbiochem	
Thrombin (Bovine)	37000, Activity: 2611.0 U/ml, 605157-1KU
TH Geyer	
Sodium chloride	NaCl ≥ 99%, 1367.5000
Sodium hydroxide	NaOH, ≥ 98.8%, 40.00 g/mol, 1375.1000
Quiagen Sciences Suites	
AmSO4	130705
MPDs	130706
Classics II	130723
pHClear	130709
Anions	130707
Jena Bioscience	
JBS classic (composed of JBS classic 1-4)	CS-121 to CS-124
JBS JCSG (composed of JBS-JBSG++ 1-4)	CS151 to CS154)

7.2. Equipment

Protein expression	
Electroporator Gene Pulser X cell	Bio-Rad
Incubation shakers	Infors HT Multitron
Shaker DRS-12	neoLab
Thermoshaker CHB-202	BioEr
Chromatography	
ÄKTApurifier 10	GE
Glutathione Sepharose 4 Fast Flow	GE
Glutathione Sepharose 4B	GE
Q Sepharose HP	Pharmacia Biotech
HiLOad 16/60 Superdex 75	GE
Gel electrophoresis	
PowerPac basic	Bio-Rad

continued on next page

Horizontal DNA electrophoresis chamber	Peqlab
SDS-PAGE mini Protean Tetra Cell	Bio-Rad
Thermoblock heater	Stuart Scientific
Centrifuges	
Refrigerated bench-top	Eppendorf 5415R, 5810R
Non-refrigerated bench-top	Eppendorf 5418, 5424
Refrigerated stand-alone	Beckman Avanti J-25
SpeedVac	Thermo Servant SPD 1010
Measuring	
Laboratory balance	Denver Instr. TP-214
Micro scales	Mettler Toledo AT21 comparator
pH meter (ligands)	Mettler Toledo server easy, InLab Ultra-Micro
pH meter (solutions)	Schott CG-840, Mettler Toledo InLab Expert
UV/Vis spectrophotometer	Shimadzu UV mini 1240
UV/Vis spectrophotometer	Thermo NanoDrop ND-1000
Affinity measurement	
Fluorescence spectrometer	Jasco FP-6500, Spectra Manager v1.54.03
Isothermal calorimetry	MicroCal GE ITC200
600 MHz spectrometer	Bruker AV3, TCI CryoProbe/TXI probe heads 1 axis shielded gradients
Crystallization	
Screen set-up	Formulatrix Formulator 16, Rock Maker Experimental Designer v2.6.6.5
Crystallization plate	Greiner CrystalQuick plate 96 well, low base (609171-067)
Plate sealing	Greiner bio-one Platesealer viewseal, transparent (676070)
Screen sealing	Greiner silverseal, aluminium (676090)
Pipet robot	Thermo Scientific Matrix Hydra II
Pipet robot	Art Robotics Instruments Crystal Gryphon (Software 1.4.1.0)
Plate hotel	Formulatrix RockImager 1000, v1.11.0.862

8. Methods

8.1. Protein preparation

8.1.1. Vector construction

Glutathione-(S)-Transferase(GST)-fused human VASP EVH1 vector was a gift from Dr. Linda Ball. The human lung cDNA library encoding ENAH EVH1 and EVL EVH1 was a gift from Dr. Anne Diehl (Quick Screen Human cDNA library panel; Clontech).

ENAH EVH1

ENAH EVH1 was amplified by PCR using the primer in table 8.2 and insert was verified by sequencing. This construct crystallized without further optimization on protein sequence level.

EVL EVH1

The initial construct of EVL EVH1 (using primer termed EVL EVH1 $_{\text{ini}}$) did not crystallize. Comparing the construct with the sequence stored in the deposited [144] EVL EVH1 structure (PDB code 1QC6) revealed differences on both termini. N-terminal $^{-2}$GS was probably left out in the PDB remark as it originates from the thrombin digestion site. Importantly, the crystallized construct is missing ^{2}AT and exhibits an elongated C-terminus of 18 amino acids. Already in 1999, mRNA sequences of two EVL EVH1 isoforms were reported. While one splicing product is missing ^{2}AT, the usage of an alternative exon 1 includes ^{2}AT. Fedorov and coworkers, among which also the Gertler lab [63] used the EVH1 domain of first isoform (GenBank accession number AAF17197.1) while our initial construct contained ^{2}AT (GenBank accession number

NM_016337.2, updated July 28th 2004).[145] We chose this isoform due to its expression in more diverse tissues like hypothalamus, hippocampus, B cells, T-lymphocytes, thymus, breast, placenta, spleen, liver, stomach, tumor tissue. In contrast, the first splicing variant of EVL is expressed in hippocampus, brain, frontal lobe and lung.[146] Whether Fedorov and coworkers ever tried to crystallize EVL EVH1 without the unstructured C-terminus could not be traced back[1].

As our initial construct differed in two positions from the successfully crystallized sequence, three constructs were cloned: the exact copy of the construct deposited (termed EVL EVH1 $_{\Delta AT,tail}$), one with shortened C-terminus and missing ^{2}AT (termed EVL EVH1 $_{\Delta AT,-}$) and a full length with ^{2}AT and the C-terminal extra 18 amino acids (termed EVL EVH1 $_{AT,tail}$). Combining EVL EVH1 $_{ini}$ and EVL EVH1 $_{\Delta AT,tail}$ forward and reverse primer, all three inserts were produced with hotstart KOD-polymerase and the EVL EVH1 $_{ini}$ vector. Inserts were purified on gel, extracted and eluted in water. Inserts were subsequently truncated with BamHI FD and NotI (Biolabs) and cloned into pGEX-4T1 vector using T4 DNA ligase (Biolabs) HF, amplified using the existing EVL EVH1 vector as template and verified by sequencing.

VASP EVH1

The initial VASP EVH1 construct did not crystallize. Based on the packing of ENAH EVH1 complex structures, three VASP EVH1 constructs were designed. The mutations introduced should allow VASP EVH1 to establish ENAH EVH1-typical crystal contacts found in the different crystal systems. The design of these constructs and the sequence alignment are found in the supplementary informations (SI) section 19.2 (Tab. 19.1). However, none of these constructs crystallized.

[1] personal communication with E. Fedorov

Domain	Sequence	
VASP EVH1	MSETVICSSRATVMLYDDGNKRWLPAGTGPQAFSRVQIYHNPTANSFRVVGRKMQPDQQV	60
EVL EVH1	MSEQSICQARASVMVYDDTSKKWVPIKPGQQGFSRINIYHNTASNTFRVVGVKLQ-DQQV	59
ENAH EVH1	MSEQSICQARAAVMVYDDANKKWVPA-GGSTGFSRVHIYHHTGNNTFRVVGRKIQ-DHQV	58
	*** **.:**:**:*** .*:*:* * .***:.***. *:***** *:* *:**	
VASP EVH1	VINCAIVRGVKYNQATPNFHQWRDARQVWGLNFGSKEDAAQFAAGMASALEALEG	115
EVL EVH1	VINYSIVKGLKYNQATPTFHQWRDARQVYGLNFASKEEATTFSNAMLFALNIM--	112
ENAH EVH1	VINCAIPKGLKYNQATQTFHQWRDARQVYGLNFGSKEDANVFASAMMHALEVL-	111
	*** :* :*:****** .**********:****.***:* *: .* **: :	

Tab. 8.1.: **Sequence alignment of Ena/VASP EVH1 domains**
Sequence alignement the EVH1 paralogs reveals three highly conserved domains that differ mostly in loop regions.

Primer	Sequence
ENAH EVH1	5′-TTTTTGGATCCATGAGTGAACAGAGTATCTGTCAG-3′
	5′-TTTTTTTTTTGCGGCCGCTCATAACACTTCTAAGGCATGCATCATGG-3′
VASP EVH1 $_{\text{S8Q, fwd}}$	5′-CGCGTGGATCCATGAGCGAGACGGTCATCTGTCAGAGCCGGGCC-3′
EVL EVH1 $_{\text{ini}}$	5′-GAAAGGGATCCATGGCCACAAGTGAACAGAGTATC-3′
	5′-TTTTTGCGGCCGCTTACATGATGTTCAGGGCAAACAGC-3′
EVL EVH1 $_{\Delta\text{AT,tail}}$	5′-GCTGTAGGATCCATGAGTGAACAGAGTATCTGCCAAGC-3′
	5′-ACGTATATGCGGCCGCTTAAGACGGACCGTTCTGAACCTGACGCTGGG
	TAGACGGGCCACCTTCCTGAGAGTTCATGATGTTCAGGGCAAACAGC-3′

Tab. 8.2.: **Primer list to generate all EVH1 domains**
Primer sequences used to create the inserts of different EVH1 domains. First line shows the forward, second line the reverse primer

Construct	Sequence
EVL EVH1 $_{\text{ini}}$	GSMATSEQSICQARASVMVYDDTSKKWVPIKPGQQGFSRINIYHN
EVL EVH1 $_{\text{1QC6}}$	--M--SEQSICQARASVMVYDDTSKKWVPIKPGQQGFSRINIYHN
EVL EVH1 $_{\text{ini}}$	TASNTFRVVGVKLQDQQVVINYSIVKGLKYNQATPTFHQWRDARQ
EVL EVH1 $_{\text{1QC6}}$	TASSTFRVVGVKLQDQQVVINYSIVKGLKYNQATPTFHQWRDARQ
EVL EVH1 $_{\text{ini}}$	VYGLNFASKEEATTFSNAMLFALNIM----------
EVL EVH1 $_{\text{1QC6}}$	VYGLNFASKEEATTFSNAMLFALNIMNSQEGGPSTQRQVQNGPS

Tab. 8.3.: **Sequence alignement of initial and crystallized EVL EVH1 construct**
Comparison of the primary sequences of the initially used EVL EVH1 $_{\text{ini}}$ and the extract of the deposited structure (PDB code 1QC6). EVL EVH1 $_{\text{1QC6}}$ was cleaved with thrombin.[144] Therefore, the construct most probably contained $^{-2}$GS as well, even though it is not annotated as expression tag.

8.1.2. Protein expression and purification

The detailed purification protocol can be found in SI section 17. In short, GST-fusion constructs were expressed in *E. Coli* BL21(DE3) electrocompetent cells in 2× YT medium supplemented with 0.3 mg/ml ampicillin. Cell cultures were induced with 0.1 mM IPTG at OD_{600} 0.4-0.5 and expression took place over night at 18°C. Cell pellets were slightly re-suspended in concentrated PBS buffer and stored at -20°C.

The entire purification protocol for EVH1 domains was designed to achieve maximal solubility in reductive environment and took advantage of the circumstance that EVH1 domains display a rather high isoelectric point (pI). This holds for both pI's of the GST-constructs (around 8) and the single EVH1 domains (over 9). Cells were lysed by sonification and the soluble fraction loaded on a fast-flow glutathion column at pH 7 already in reductive environment. Subsequently, the GST tag was cleaved off and a strong cation exchange column run at pH 7.5 efficiently separated the EVH1 domain from the remaining impurities. As final step, the buffer was exchanged by running a size exclusion column equilibrated with crystallization buffer, in which the EVH1 domains were concentrated, flash frozen in liquid nitrogen and stored at -80°C.

8.2. Scaffold and peptide preparation

Scaffold synthesis was subject to the collaborating group of Prof. Dr. Hans-Günther Schmalz at the Institute of Organic Chemistry, University of Cologne.[140,142,143,147,148] Except for the ProM scaffolds, Fmoc-protected amino acids were purchased from IRIS Biotech. Peptides were synthesized automatically on an ABI 433A peptide synthesizer (Applied Biosystems) following the standard N^{α}-9-fluorenylmethoxycarbonyl (Fmoc) solid-phase protocol[149] and purified on a preparative reversed-phase high-performance liquid chromatography system (LC-10AD, Shimadzu). Final ligand purity was monitored at 220 nm and usually exceeded 95 %, while the expected molar masses were confirmed by MALDI-TOF mass spectrometry (Voyager-DE STR, Applied Biosystems). Ligand assembly was carried out in-house at the FMP Berlin by Dr. R. Volkmer, I. Kretzschmar and K. Franke.

8.3. Macromolecular crystallography

8.3.1. Generalized overall workflow for EVH1 crystallization

Purified protein and ligands were mixed and diluted to concentrations between 8-15 mg/ml. Incubation over night was not necessary. Ligand and protein were mixed on the same day and kept on ice until pipetting the crystallization plates.

Crystallization experiments on VASP EVH1 and later on with ENAH EVH1 were performed by extensive screening trials against a range of diverse chemical environments provided by the Quiagen Sciences Suites. This screening aimed for an optimal combination of precipitating agents, buffers, organic additives, inorganic salts at two different temperatures (20 and 4°C). The 96-well sitting drop plates were prepared by pipetting robots and stored in plate hotels. Obtained crystals from the initial screens often grew in sub-optimal conditions that severely limited the diffraction power. Only very rarely, the resolution limit of initial crystals fulfilled the criterion of high-resolution structures (<1.5Å). This made the preparation of well diffracting crystals the most demanding and time consuming bottle-neck in the whole workflow.[150] Three methods were used to increase the crystal quality in this work. For most complexes, an extensive fine-screening of the initial condition in combination with micro seeding technique was sufficient. To minimize bias, the crushed crystals always contained a ligand with weaker affinity than the ligand of interest. Towards the end of the thesis, a whole panel of crystal seeds grown in different space groups and cell dimensions was available. This method made material-consuming initial screens obsolete and allowed to grow crystals with high scattering power in only 3 days. Only in some cases, micro seeding nevertheless produced crystals that diffracted less than 1.8 Å. Optimization was then successful by rising the temperature to 27°C. The mentioned techniques are explained in the following section.

8.3.2. Optimization of crystal and diffraction quality

Fine-screening

Fine-screening turned out to be the most efficient of all techniques used throughout the thesis. ENAH EVH1 strongly preferred diammonium sulfate over any other precipitation agents at concentration of 1.8-2.4 M $(NH_4)_2SO_4$. Luckily, the AmSO4

screen covered 48 salt additives at 200 mM concentration with 1.8 M $(NH_4)_2SO_4$. These conditions were usually sufficient to obtain initial hits, upon which fine-screen plates were produced as follow-up experiments. Thereby NH_4NO_3 was the most promising salt additive, ranging from as low as 50 mM up to 800 mM. To cover these concentrations, several fine-screens were pipetted and stored at -20°C.

The protein-ligand samples were prepared to final volumes of around 120 µl in low-bind eppendorf tubes. After initial screening experiments the tube was flushed with nitrogen gas, flash frozen in liquid nitrogen and stored at -80°C. After finding initial hits and preparing fine-screens accordingly, the sample was thawed and re-used. Pipetting 100-300 nl protein per screening conditions, the preparations lasted out for at least three additional fine-screens, which was usually sufficient to achieve high quality crystals.

Microseeding

Microseeding technique was used towards the end of thesis to further improve crystal quality of important complexes. For an initial seed-stock, a fine-screen was prepared to re-crystallize ENAH EVH1 in complex with Ac–[2-Cl-Phe]PPPP–OH. This inhibitor displayed the weakest affinity and was assumed to be displaced in the crystal by any ligand of interest. A selection of crystals was measured and the whole crystallization plate harvested in the mother liquor condition of the best diffracting crystal. For the mentioned seed-stock, ENAH EVH1 crystals were crushed in filtered, ice-cold 1.8 M $(NH_4)_2SO_4$, 350 mM NH_4NO_3 using Seed Bead kit (Hampton Research) kindly gifted by Alain Chambovey (Actelion Ltd., Basel). Cross-seeding with this seed-stock improved the resolutions of two complexes up to 0.22Å. Cross-seeding became obsolete during cocrystallization assays with crystals of not satisfactory quality grown in fine-screen plates. These plates usually produced enough micro crystals and tangly grown material, fresh enough to be harvested and crushed in the mother liquor condition of the so-far best diffracting crystal.

For pipetting, seeds were diluted 2-20× with mother liquor to 20 µl and kept on ice until pipetted by the robot. The crystallization droplets were prepared in the following order and volume ratio: 200/300/100 nl protein/reservoir/seed. In combination with adapted fine-screens, this protocol resulted in well diffracting crystals within 2-5 days.

Increasing the temperature

All C-terminally elongated ligands produced excessive amorphous aggregation, independent of $(NH_4)_2SO_4$ concentrations. These semicrystalline structures might have formed due to strong protein-protein interactions. Facing overnucleation or amorphous growth, the general approach is to introduce more repulsive forces between proteins. Doing so, one changes the solution condition by adding for example 1,4-Dioxane, an additive that enhances electrostatic interactions.

As the fine-screens did not suffer from over-nucleation but mainly produced spherolytes, the conditions were presumably still close to the solubility line and a simple temperature raise would be enough to drive the system out of the amorphous regime. [151] 27℃ thereby reflected the highest temperature ENAH EVH1 came in contact with during purification (thrombin digestion over night). Higher temperatures have never been explored, as especially the combination with seeds yielded in dozens of well diffracting crystals (Fig. 8.1).

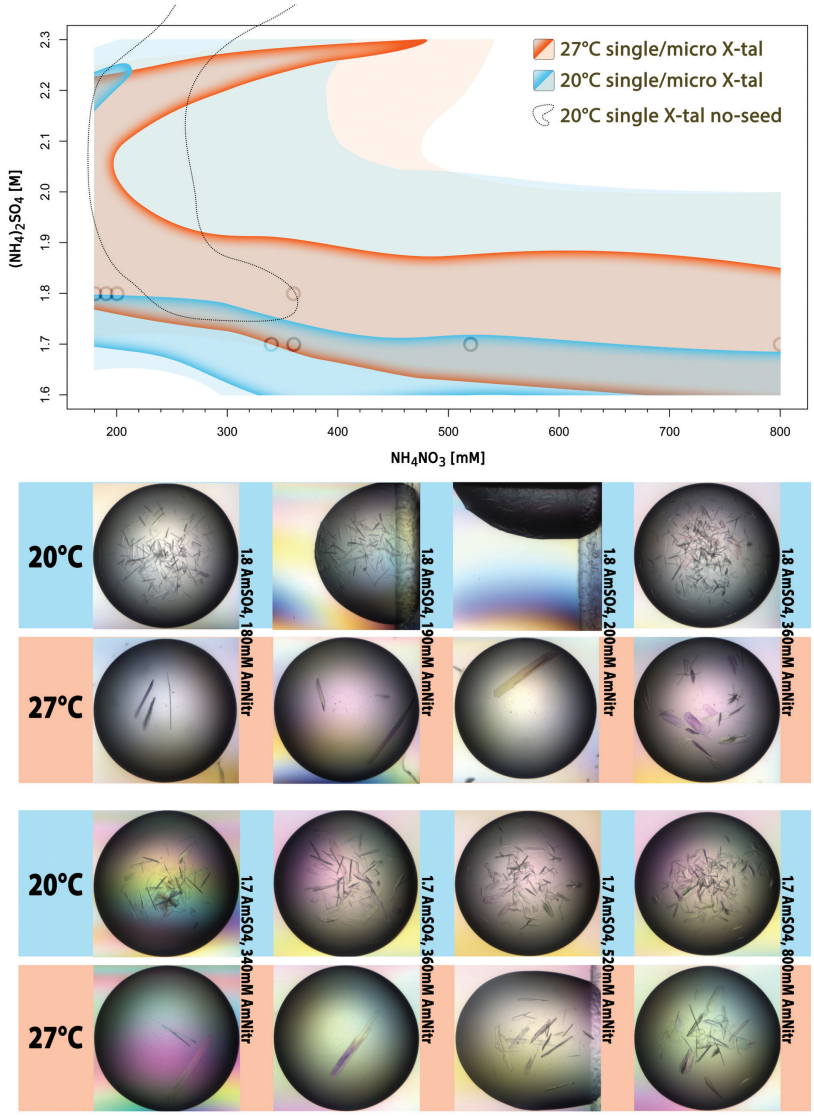


Fig. 8.1.: **Cocrystallization experiments at 20°C and 27°C of ENAH EVH1 in complex with Ac–[2-Cl-Phe]PP[ProM-1]TEDEL–NH_2**
Upper panel: graphic illustration of the phase boundaries producing single crystals (darker color) and regimes with micro crystals (brighter color) at 20°C (blue) and 27°C (red) using the micro seeding technique. Crystals of modest diffraction quality (1.8Å) grew also without seeds (dotted line), but needed 0.2-0.4 M higher $(NH_4)_2SO_4$ concentration to reach nucleation zones and grew only between 200-240 mM NH_4NO_3. The opaque circles belong to pictures in the lower panel: At 20°C, the conditions produced massive over-nucleation. Increasing the temperature to 27°C tilted the system towards the solubility line and produced crystals diffracting to 1.45Å.

8.3.3. Data acquisition, processing and graphic works

X-ray data collection and processing

Fully grown crystals were mounted on nylon loops and rarely reacted sensible to cryo-protectants. Most crystals were flash-frozen with mother liquor containing 20% (v/v) glycerol. Data sets were recorded at the BESSY-II radiation light source (MX-beamlines 14.1 and 14.2 in Adlershof, Berlin). Diffraction data was indexed and processed manually using XDS.[152] Optimization of the data sets was guided by asymptotic ISa,[153] standard deviations (rmsd) of spot and spindle positions and the number of misfits.[154] High-resolution limit was determined by $CC_{1/2}$[155] of 0.6, high resolution bin completeness and $\langle I/\sigma \rangle_{mrgd}$. $CC_{1/2}$ was used in a much more conservative way than proposed.[156,157] Not even the merged datasets displayed a multiplicity high enough to process data out to resolution limits with $CC_{1/2}$ of 0.2-0.1. Only in one case re-processing of a dataset to $CC_{1/2}$ 0.17 improved R_{free} marginally.

Structure determination

Molecular replacement started with unliganded EVH1 structures as search model (highest resolution available, starting with PDB code 1EVH) using Phaser,[158] whose solution had rarely to be cross-checked by MOLREP.[159] Ligands were visible in the experimental map right after molecular replacement but loaded into the structure in later rounds of model building. Non-natural ligands were built, minimized and docked with MOLOC,[160] and restraint files were produced using either JLigand[161] or the PRODRG2 server.[162] The MAB force field[163] of MOLOC was further used for real space refinement of weakly defined loops and crystal contacts using the H-bond pattern evaluation. As MOLOC can not read mtz files, $2mF_0$-DF_c composite, mF_0-DF_c difference and feature enhanced maps (FEM) were generated with the according programs (CCP4 fft,[164] phenix.fem[165]) and maps extended 5Å beyond the asymmetric unit (CCP4 mapmask). Some structures were refined in early stages by REFMAC5,[166] but were subsequently transferred to the PHENIX system[167] and rebuilt with COOT.[168] Guided by R_{free}, $rmsd_{bond}$, $rmsd_{angle}$, MolProbity,[169,170] and the PDB_REDO server,[171] water molecules, ions and double conformations were

added in later refinement cycles. Final complex structures were deposited in protein data bank (rcsb.org).

Structure depiction and graphic works

Molecular structure images were rendered by Chimera (v1.10.1). Chains within asymmetric units were superposed in MOLOC using the atomic positions W23NE, W23N, F77CG, Y16CG, and Q79CD to calculate a rigid body match. Ligands were therefore positioned over the epitope-shaping amino acids as riding atoms. For better visibility the surveying superpositions contain only one protein complex each, while in detailed views all chains of the asymmetric unit are shown.

8.4. Affinity measurements

8.4.1. Fluorescence titration

Fluorescence titration (FT) experiments were carried out at 25℃ using a 10×4 mm precision cell made from quarz suprasil with conical bore for a stirring magnet (Hellma, 109.004F-QS), filled to 800 µl. Titration experiments were made in 40 mM NaPi pH 7.3, 100 mM NaCl and 2 mM TCEP. Changes in the tryptophane fluorescence were detected at 340 nm, using slits opened to 1-3 nm (excitation) and 3-5 nm (emission) and a photomultiplier voltage to load the detector with around 30% arbitrary starting signal. Ligand was titrated with 20 injections to roughly five times nominal excess. Protein concentration was set to around 5 times K_d, and ligand stock solution prepared to inject 22×5 µl. Depending on the protein concentration, excitation wavelength was set between 298-305 nm to keep starting protein absorbance below 0.04 cm^{-1}. Data points were averaged during one second and spacing of one minute ensured homogeneous mixing of the cell content. In total three to four replicas were made, combined with two background measurements titrating ligand into buffer. Parameters were optimized iteratively until standard errors reached less than 20% for the K_d and less than 0.5 kJ/mol for ΔG.

A one-to-one model was assumed and the nominal ligand concentration was multiplied by a factor $0 < M \leq 1$ to determine the amount of active ligand. The fitting

procedure was carried out by an script written and kindly gifted by Dr. Robert Opitz in R language.

8.4.2. Isothermal calorimetry

Isothermal calorimetry (ITC) experiments were performed at 25°C using 40 mM NaPi pH7.3, 100 mM NaCl and 1 mM DTT. Protein concentration was set to around 5 times K_d, and ligand stock solution prepared to inject $11\times2\,\mu l$. Reference power was set to 18 µcal/sec with a stirring speed of $1000\,s^{-1}$ and 2 s filter period. In total three to four replicas were made, combined with two background measurements titrating ligand into buffer.

A one-to-one model was assumed and the nominal ligand concentration in equation 16.9 (SI Sec. 16.1.1) was multiplied by a factor $0 < M \leq 1$ to determine the amount of active ligand. The fitting procedure was carried out by an script written and kindly gifted by Dr. Robert Opitz in R language.

8.5. ^{1}H-^{15}N-HSQC-measurements

Fluorescence titration experiments were performed prior to the ^{1}H-^{15}N-HSQC experiments to assess the affinity and the purity of the ligand. The final molar ligand excess was calculated with equation 16.9 (SI Sec. 16.1.1) to saturate EVH1 domains to around 98% $[PL]/P_{tot}$. An unbound and a fully ligand-bound protein sample was prepared with 50-200 µM EVH1 domain. NMR measurements consisted of at least 5 titration experiments in which the ligand concentration was increased by subsequently adding ligand-bound EVH1 on top of the previous experiment. During the measurements, the remaining samples were kept on ice. Titration experiments were performed at 300 K.

Perturbation spectra of ENAH EVH1 and EVL EVH1 were assigned by M. Beerbaum, VASP EVH1 was assigned by Dr. L. Ball. 2D ^{1}H-^{15}N-HSQC spectra were processed with TopSpin. bruk2ucsf was used to convert 2rr raw files directly to ucsf format for import into Sparky (v3.114).[172] Following the suggestion of Dr. P. Schmieder, Euclidean distances were calculated from the peak tables by:

$$\Delta\delta_{\mathrm{TOTAL}} = \sqrt{(\Delta\delta^{1}\mathrm{H})^2 + (0.1\,\Delta\delta^{15}\mathrm{N})^2} \tag{8.1}$$

To calculate the affinites of amino acids involved in a putative second binding site, the titration experiment in Tab. 20.1 consisted of 26 ^{1}H-^{15}N-SOFAST-HMQC measurements. Changes in chemical shift $\Delta\delta$ of each assigned resonance peak were calculated according to equation 16.9 with a script written in R language. Shifts were fitted with and without M-value and evaluated with the analysis of variance (ANOVA).

8.6. SPOT array substitution

SPOT array was prepared by an automated spot synthesis protocol on whatman-50-Cellulose.[173] Membrane and bound peptides were washed with ethanol and washing buffer (3% (v/v) tween 20, 1×TBS pH 8.6) and subsequently blocked (5% (w/v) milk powder, 3% (v/v) tween 20, 1×TBS pH 8.6) for 3 min at room temperature. Each membrane was covered with GST-fused EVH1 domains (10µl/ml GST-VASP EVH1, GST-ENAH EVH1, GST-EVL EVH1, and GST) in sample buffer (3% (w/v) milk powder, 3% (v/v) tween 20, 1×TBS pH 8.6) and incubated at 7°C over night. After extensive washing, the anti GST antibody (Z-5 polyclonal rabbit IgG, Santa Cruz Biotech 200 µg/ml) was added as 1:1000 (v/v) dilution in sample buffer and incubated over night at 4°C. Membranes were then washed and covered with anti-rabbit antibody (IRDye800, Sant Cruz Biotech), diluted 1:20000 (v/v) in sample buffer. After binding took place over 45 min at room temperature, excess antibody was washed away and membranes transferred into 1×TBS. Binding was detected with fluorescence signal on Odessey infrared imaging system (Li-Cor, v2.0.40) at a resolution of 84 µm.

Part III.

Results

Abstract of the results

The first section 9.1 serves as an overview that lists the tables of diffraction and refinement statistics as well as the accession codes of the atom coordinates found in the Protein Data Bank (rcsb.org). For detailed description of each crystal structure, see SI section 18.

The second section 9.2 lists binding studies that guided the scaffold modifications and the crystallization of ligand compositions. Section 9.2 takes into account that initial ligand optimization was performed with little structural insights except for the deposited crystal structure of Mena EVH1 in complex with the core recognition motif Ac–^{1}FPPPPT–NH_2 (PDB code 1EVH,[174] Fig. 3.2). As starting point served the proline-rich 13-mer of the surface protein ActA from the intracellular pathogen *Listeria monocytogenes*. Optimization of ^{1}Phe to fit into the binding pocket gained at least –6 kJ/mol of free energy (Sec. 9.2.1). Replacement of ^{4}PP by ProM-1 added at least –3.5 kJ/mol of free energy (Sec. 9.2.2). Redesigning ProM-1 yielded in ProM-2 which successfully replaced ^{2}PP without any loss of affinity (Sec. 9.2.5). Subsequent optimization of the N-terminal building block of ProM-1 was not successful, as disrupting the backbone amide NH substitution decreased the affinity up to an amount comparable to di-proline (ProM-3, Sec. 9.2.6), while inducing more flexibility at the same position had no beneficial effect (ProM-4, Sec. 9.2.7). Further improvement of the binding strength of the initial inhibitor was achieved by *in silico* designed modifications of the C-terminal building block of ProM-1. Competition with a bound water molecule was not successful (ProM-12, Secs. 9.2.11 and 9.2.12), while mimicking the hydrophobic interaction of TEDEL by scaffolds ProM-9 (Sec. 9.2.14) and ProM-13 (Sec. 9.2.15) boosted nearly –8 kJ/mol of free energy and restored the affinity of the C- and N-terminally elongated initial inhibitor. Optimization was guided by the increase in binding affinity. Small tables list the release of free energy ΔG upon binding and $\Delta\Delta G$ relative to a given reference (ref) ligand.

Solution NMR studies were used to to gain insights related to possible crystal packing artifacts. The additional binding of ligands involved weak interactions that we sought to identify with ^{1}H-^{15}N-HSQC titration experiments. The shift perturbation tables are printed in section 10.2 and include Euclidean distances of roughly 0.1 ppm and further.

9. Deposited crystal structures and binding affinity studies

9.1. Diffraction and refinement statistics of deposited complex structures

Of most complex structures existed multiple crystal data sets that were recorded during the process of crystal quality optimization. Isomorphous data sets were merged together by XSCALE and paired refinement was carried out with the unmerged and merged data sets. For deposition, the most accurate model was chosen according to the free R value calculated at the same resolution limit. [155]

ENAH EVH1	Ac–[2-Cl-Phe] PPPP–OH	Ac–[2-Cl-Phe] PP[ProM-1]–OH	Ac–[2-Cl-Phe] PP[ProM-1]–NH_2
Diffraction statistics[a]			
Spacegroup	P1	P1	$P3_121$
Cell dimension [Å, °]	a = 34.91 b = 43.37 c = 43.65 α = 61.15 β = 84.25 γ = 84.33	a = 34.76 b = 43.39 c = 43.59 α = 61.40 β = 84.03 γ = 84.10	a = 47.27 b = 47.27 c = 202.72
Resolution [Å]	38-1.49 (1.58-1.49)	38-1.16 (1.23-1.16)	41-1.80 (1.91-1.80)
Unique reflections	33987 (5476)	71797 (11195)	25337 (4010)
Completeness [%]	93.0 (92.6)	93.1 (90.0)	99.4 (99.7)
Redundany	2.7 (2.7)	2.0 (1.9)	5.4 (5.6)
$\langle I/\sigma \rangle_{mrgd}$	12.58 (1.36)	11.24 (1.86)	14.61 (1.49)
$CC_{1/2}$	99.9 (59.5)	99.9 (82.6)	99.9 (58.1)
R_{meas}[b] [%]	5.5 (96.5)	4.8 (51.7)	8.6 (113)
ISa	26.18	36.57	45.41
Refinement statistics			
Solvent content [%]	45.5	45.9	52.3
V_M [$Å^3/Da$]	2.26	2.25	2.58
Molecules per AU	2	2	2
R_{fact}[c] [%]	17.9	13.9	20.21
R_{free}[d] [%]	20.1	14.1	21.66
$rmsd_{bond}$ [Å]	0.06	0.01	0.007
$rmsd_{angle}$ [°]	1.2	1.7	1.4
PDB code	5N91	5N9C	5N9P

[a] Numbers in parentheses correspond to outer shell
[b] R_{meas}: redundancy-independent R factor
[c] $R_{fact} = \sum |F_0 - F_c| / \sum |F_0|$ [175]
the observed and calculated structure factors
[d] R_{free} set contains 5% of total reflections [176]

Tab. 9.1.: Diffraction and refinement statistics of published crystal structures with ligands that successfully replaced the core motif of ActA-derived peptide FPPPP.

ENAH EVH1	Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH	Ac–[2-Cl-Phe][ProM-2][ProM-3]–OH	Ac–[2-Cl-Phe][ProM-2][ProM-4]–OH
Diffraction statistics[a]			
Spacegroup	C222	P2	P1
Cell dimension [Å, °]	a = 89.83 b = 131.42 c = 35.56	a = 34.61 b = 38.25 c = 44.00 β = 90.59	a = 34.72 b = 43.36 c = 44.18 α = 60.97 β = 84.19 γ = 84.16
Resolution [Å]	32-1.7 (1.74-1.70)	44-1.15 (1.22-1.15)	50-1.40 (1.48-1.40)
Unique reflections	23659 (1704)	38436 (6060)	41864 (6668)
Completeness [%]	99.8 (99.9)	93.5 (92.0)	94.6 (93.3)
Redundany	4 (4.1)	2.2 (2.1)	2.3 (2.3)
$\langle I/\sigma \rangle_{mrgd}$	12.86 (2.35)	13.15 (1.44)	10.39 (1.12)
$CC_{1/2}$	99.8 (76.8)	99.9 (59.1)	99.9 (60.2)
R_{meas}[b] [%]	8.2 (64.6)	4.5 (78.2)	7.3 (99.5)
ISa	19.92	34.82	36.01
Refinement statistics			
Solvent content [%]	40.9	47.2	46.3
V_M [$Å^3/Da$]	2.08	2.33	2.29
Molecules per AU	2	1	2
R_{fact}[c] [%]	19.0	15.7	22.0
R_{free}[d] [%]	23.6	18.6	26.0
$rmsd_{bond}$ [Å]	0.01	0.01	0.009
$rmsd_{angle}$ [°]	1.6	1.60	1.16
PDB code [Refcyc]	4MY6	5NBF	5NCF

[a] Numbers in parentheses correspond to outer shell
[b] R_{meas}: redundancy-independent R factor
[c] $R_{fact} = \sum |F_0 - F_c| / \sum |F_0|$ [175]
the observed and calculated structure factors
[d] R_{free} set contains 5% of total reflections [176]

Tab. 9.2.: Diffraction and refinement statistics of published crystal structures with ligands composed of scaffolds that were unsuccessful for replacing the core motif of ActA-derived peptide FPPPP.

ENAH EVH1	Ac–WPPPP TEDEL–NH_2	Ac–[2-Cl-Phe] PPPP TEDEL–NH_2	Ac–[2-Cl-Phe] PP[ProM-1] TEDEL–NH_2
Diffraction statistics[a]			
Spacegroup	C2	C2	C2
Cell dimension [Å, °]	a = 187.92 b = 34.42 c = 110.96 β = 91.53	a = 147.78 b = 44.02 c = 34.67 β = 102.13	a = 149.74 b = 44.24 c = 34.83 β = 101.47
Resolution [Å]	48-2.70 (2.80-2.70)	42-1.58 (1.68-1.58)	42-1.45 (1.54-1.45)
Unique reflections	20211 (2039)	29131 (4618)	39592 (6309)
Completeness [%]	98.8 (100)	96.8 (95.7)	99.4 (99.2)
Redundany	5.3 (5.3)	2.3 (2.3)	3.3 (3.1)
$\langle I/\sigma \rangle_{mrgd}$	8.60 (1.53)	10.23 (1.47)	7.69 (1.32)
$CC_{1/2}$	98.9 (58.7)	99.8 (60.5)	98.9 (58.7)
R_{meas}[b] [%]	23.0 (153)	8.5 (97.5)	11.3 (104)
ISa	18.40	29.34	13.37
Refinement statistics			
Solvent content [%]	65.2	44.2	44.7
V_M [$Å^3/Da$]	3.53	2.20	2.22
Molecules per AU	4	2	2
R_{fact}[c] [%]	20.1	24.6	18.2
R_{free}[d] [%]	27.0	19.9	20.4
$rmsd_{bond}$ [Å]	0.009	0.01	0.004
$rmsd_{angle}$ [°]	1.10	1.17	0.977
PDB code	5NC7	5NC2	5ND0

[a] Numbers in parentheses correspond to outer shell
[b] R_{meas}: redundancy-independent R factor
[c] $R_{fact} = \sum |F_0 - F_c| / \sum |F_0|$ [175]
[d] R_{free} set contains 5% of total reflections [176]

Tab. 9.3.: Diffraction and refinement statistics of published crystal structures with C-terminally elongated ActA-derived peptides.

ENAH EVH1	Ac–[2-Cl-Phe] PP[ProM-9]–OH	Ac–[2-Cl-Phe] [ProM-2][ProM-9]–OH	Ac–[2-Cl-Phe] [ProM-2][ProM-13]–OEt
Diffraction statistics[a]			
Spacegroup	C222	$P2_12_12_1$	P2
Cell dimension [Å, °]	a = 90.15 b = 131.52 c = 35.68	a = 35.02 b = 61.27 c = 89.26	a = 34.81 b = 44.36 c = 72.47 β = 90.56
Resolution [Å]	45-1.65 (1.75-1.65)	36-1.02 (1.08-1.02)	44-1.29 (1.39-1.29)
Unique reflections	25469 (3536)	172743 (26288)	53114 (9979)
Completeness [%]	97.3 (85.3)	91.4 (86)	99.4 (69.4)
Redundany	4.1 (2.8)	3.0 (2.9)	4.7 (2.7)
$<I/\sigma(I)>$	12.7 (1.7)	12.4 (2.49)	6.76 (1.94)
$CC_{1/2}$	99.8 (60.7)	99.9 (80.1)	99.4 (76.4)
R_{meas}[b] [%]	8.6 (74.6)	6.2 (52.6)	15.6 (65.6)
ISa	21.54	25.51	19.34
Refinement statistics			
Solvent content [%]	41.0	34.7	43.8
V_M [$Å^3/Da$]	2.08	1.88	2.19
Molecules per AU	2	2	2
R_{fact}[c] [%]	16.5	12.4	15.1
R_{free}[d] [%]	19.4	15.0	19.0
$rmsd_{bond}$ [Å]	0.01	0.01	0.008
$rmsd_{angle}$ [°]	1.26	1.72	1.21
PDB code [Refcyc]	5NBX	5NCG	5NEG

[a] Numbers in parentheses correspond to outer shell
[b] R_{meas}: redundancy-independent R factor
[c] $R_{fact} = \sum |F_0 - F_c| / \sum |F_0|$ [175]
[d] R_{free} set contains 5% of total reflections [176]

Tab. 9.4.: Diffraction and refinement statistics of published crystal structures with scaffolds yielded through structure-based drug design. These scaffolds allow replacement of the core motif of ActA-derived peptide Ac–FPPPPTEDEL–NH_2 and mimic it hydrophobic interaction.

ENAH EVH1	Ac–[2-Cl-Phe][ProM-1][ProM-1]-OH	Ac–[2-Cl-Phe][ProM-1][ProM-12]-OH	Ac–[2-Cl-Phe][ProM-1][ProM-12]-OMe
Diffraction statistics[a]			
Spacegroup	$P2_12_12$	C222	C2
Cell dimension [Å, °]	a = 73.71 b = 78.02 c = 80.85	a = 44.14 b = 141.37 c = 34.71	a = 136.15 b = 34.57 c = 44.24 β = 96.93
Resolution [Å]	45-1.46 (1.55-1.46)	42-1.65 (1.75-1.65)	44-1.42 (1.50-1.42)
Unique reflections	81310 (12951)	13548 (2144)	38981 (5889)
Completeness [%]	99.7 (99.6)	99.9 (100)	99.9 (99.9)
Redundany	5.4 (5.5)	8.5 (8.4)	6.4 (5.8)
$\langle I/\sigma \rangle_{mrgd}$	11.79 (1.66)	9.14 (0.97)	11.36 (1.25)
$CC_{1/2}$	99.8 (62.1)	99.7 (43.3)	99.9 (59.3)
R_{meas}[b] [%]	9.6 (101)	20 (210.6)	12 (171)
ISa	21.62	19.13	17.26
Refinement statistics			
Solvent content [%]	46.2	42.3	39.5
V_M [$Å^3/Da$]	2.28	2.13	2.03
Molecules per AU	4	1	2
R_{fact}[c] [%]	17.8	18.3	17.0
R_{free}[d] [%]	19.9	20.8	19.8
$rmsd_{bond}$ [Å]	0.008	0.009	0.02
$rmsd_{angle}$ [°]	1.28	1.27	1.86
PDB code	5NAJ	5NCP	5NDU

[a] Numbers in parentheses correspond to outer shell
[b] R_{meas}: redundancy-independent R factor
[c] $R_{fact} = \sum |F_0 - F_c| / \sum |F_0|$ [175]
[d] R_{free} set contains 5% of total reflections [176]

Tab. 9.5.: Diffraction and refinement statistics of published crystal structures with sub-optimal scaffold compositions or unsuccessful structure-based design.

9.2. ENAH EVH1 crystal structures and affinity measurements that guided inhibitor optimization

9.2.1. Ac–[2-Cl-Phe]PPPP–OH in complex with ENAH EVH1

The recognition motif of Mena EVH1 binding Ac–^{1}FPPPPT–NH_2 is based on a deep binding pocket for ^{1}Phe which is not filled entirely by the ligand. Single-point Phe-substitutions (SI Tab. 16.1) were used to determine alternate non-natural moieties. 2-chloro-*L*-Phe (2-Cl-Phe, **0b**) improved the affinity substantially.

Ac–SFE-(?)-PPPPTEDEL–NH_2		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
wt	Phe	20 (3)	–26.9 (0.3)	(ref)
0b	[2-Cl-Phe]	1.3 (0.2)	–33.6 (0.2)	–6.7 (0.4)
0d	[2-CF_3-Phe]	1.8 (0.1)	–32.9 (0.1)	–6.2 (0.3)
0e	[2,6-di-Cl-Phe]	1000 (100)	–17.0 (0.2)	+9.9 (0.4)

Tab. 9.6.: **Chlorination of ^{1}Phe gains at least –6 kJ/mol**
The wild type 13-mer **wt** is derived from the repeat three of the ActA protein from *L. monocytogenes*. Substituting the *ortho* position of Phe with chloride (**0b**) or tri-fluoro methyl (**0d**) increased ligand affinity. Symmetrical di-substitution of chloride (**0e**) disrupted binding. Table lists affinities as K_d and ΔG of ligands measured by fluorescence titration using VASP EVH1. Values in parentheses are standard errors.

Depending on the ligand composition, boosts in ΔG of up to –6.5 kJ/mol were measured. The modified core recognition motif of ligand **0b** was crystallized to investigate whether 2-Cl-Phe-containing compositions optimally filled the void. The structure showed no altered binding mode of Ac–[2-Cl-Phe]PPPP–OH compared to the deposited complex structure with Ac–FPPPPT–NH_2.

9.2.2. Ac–[2-Cl-Phe]PP[ProM-1]–OH in complex with ENAH EVH1

On the search to replace prolines, ProM-1 was incorporated into the ligand. Previous extensive affinity measurements[105] showed that the binding strength increases by –3 kJ/mol if the second proline pair is replaced by ProM-1 (SI Tab. 16.2).

Ac–SFE[2-Cl-Phe]PP-(?)-TEDEL–NH_2		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
0b	PP	0.79 (0.05)	–34.8 (0.2)	(ref)
1b	[ProM-1]	0.21 (0.05)	–38.1 (0.6)	–3.3 (0.6)

Tab. 9.7.: **Substitution of ^{4}PP by ProM-1 gains –3 kJ/mol**
ProM-1 can not be shuffled within ^{1}FPPPP to replace any Pro-Pro tandem within the ActA-derived 13-mer **0b**.[105] Only replacement of ^{4}PP by ProM-1 leads to a significant boost in affinity, yielding in a ligand with nanomolar affinity. Table lists affinities as K_d and ΔG of ligands measured by fluorescence titration using ENAH EVH1. Values in parentheses are standard errors.

The core motif of **1b**, Ac–[2-Cl-Phe]PP[ProM-1]–OH, was crystallized to confirm that the scaffold conserved the canonical binding mode. The complex structure showed that ProM-1 was not interfering with the overall positioning of the ligand compared to core motif of **0b**.

9.2.3. Ac–[2-Cl-Phe]PP[ProM-1]–NH_2 in complex with ENAH EVH1

However, both mentioned ligand compositions so far displayed a free C-terminus and were therefore not directly comparable with the deposited complex structure of Ac–FPPPPT–NH_2. To investigate the effect of the C-terminal negative charge of the ProM scaffold on the binding mode, the ligand Ac–[2-Cl-Phe]PP[ProM-1]–OH was masked with an amide function and Ac–[2-Cl-Phe]PP[ProM-1]–NH_2 was cocrystallized with ENAH EVH1.

Masking the core motif of **1b** with an amide function showed no visible alternation in the binding mode. Therefore, the charge of the C-terminus was not masked as these ligands crystallized faster and showed improved scattering power. For comparison,

the ENAH EVH1 complex with Ac–[2-Cl-Phe]PP[ProM-1]–NH_2 was solved at a high-resolution cut-off of 1.8 Å, while the same composition with free C-terminus diffracted to 1.1 Å.

9.2.4. Ac–[2-Cl-Phe][ProM-1][ProM-1]–OH in complex with ENAH EVH1

Especially for VASP EVH1, ProM-1 could not be used as replacement of ^{2}PP, where ProM-1 double-substituted 13-mers (**1f**) bound as weak as the **wt** ligand (SI Tab. 16.2). Previous work thereby suggests an altered binding mode for ^{1}F[ProM-1]-containing ligands.[141,142]

Ac–SFE-(?)-TEDEL–NH_2		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
wt	FPPPP	20 (3)	–26.9 (0.3)	(ref)
1e	FPP[ProM-1]	5.7 (0.3)	–29.9 (0.1)	–3.0 (0.3)
1f	F[ProM-1][ProM-1]	19 (6)	–26.9 (0.8)	–0.0 (0.9)
0b	[2-Cl-Phe]PPPP	1.5 (0.3)	–33.3 (0.5)	(ref)
1b	[2-Cl-Phe]PP[ProM-1]	0.31 (0.05)	–37.2 (0.4)	–3.9 (0.6)
1d	[2-Cl-Phe][ProM-1][ProM-1]	0.3 (0.1)	–37.1 (0.8)	–3.8 (0.9)

Tab. 9.8.: **ProM-1 changes the binding mode of VASP EVH1 when replacing ^{2}PP**
Table lists affinities as K_d and ΔG of ligands measured by fluorescence titration using VASP EVH1. Values in parentheses are standard errors.

To investigate whether ProM-1 double-substituted compositions indeed show a different binding mode, Ac–[2-Cl-Phe][ProM-1][ProM-1]–OH was cocrystallized with ENAH EVH1. This complex structure could not explain the altered binding dynamics measured for VASP EVH1. Other ^{2}ProM-1-containing compositions did not crystallize.

9.2.5. Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH in complex with ENAH EVH1

As we aimed for an inhibitor that binds canonically to all three paralogs, *in silico* design yielded in a new scaffold. ProM-2 fulfilled the requirements of the underlying epitope and allowed replacement of ^{2}PP without affinity loss (SI Tab. 16.2).

Ac–SFE-(?)-TEDEL–NH_2		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
1e	FPP[ProM-1]	5.7 (0.3)	–29.9 (0.1)	(ref)
2e	F[ProM-2][ProM-1]	4 (0.3)	–30.8 (0.2)	–0.9 (0.2)
1b	[2-Cl-Phe]PP[ProM-1]	0.31 (0.05)	–37.2 (0.4)	(ref)
2b	[2-Cl-Phe][ProM-2][ProM-1]	0.3 (0.1)	–37.7 (0.7)	–0.5 (0.8)

Tab. 9.9.: **ProM-2 allows replacement of ^{2}PP without affinity loss**
Table lists affinities as K_d and ΔG of ligands measured by fluorescence titration using VASP EVH1. Values in parentheses are standard errors.

The incorporation of 2-Cl-Phe, ProM-1 and ProM-2 into the 13-mer **2b** boosted the affinity drastically. Omitting the peptidic termini SFE and TEDEL yielded in the initial inhibitor Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH (**2a**). The complex structure of **2a** and ^{1}H-^{15}N-HSQC measurements (Fig. 10.2) confirmed the conservation of the canonical binding mode for all Ena/VASP EVH1 domains.

9.2.6. Ac–[2-Cl-Phe][ProM-2][ProM-3]–OH in complex with ENAH EVH1

To boost the affinity of the initial inhibitor **2a**, we searched for interaction sites that allowed a certain degree of ligand variability. We performed SPOT array substitution experiments with all three Ena/VASP paralogs and found that ^{4}Pro can be exchanged by aliphatic amino acids (Results Sec. 10.1). ProM-1 was modified on the N-terminal side to explore the variability of ^{4}Pro. The pyrrolidine ring was opened to form ProM-3, a conformationally restricted Val-Pro mimetica. Inhibitor compositions with this scaffold lost nearly fourfold affinity (SI Tab. 16.2).[141]

The complex structure of Ac–[2-Cl-Phe][ProM-2][ProM-3]–OH (**3a**) was solved to get insights into the binding mode of ProM-3. The structure revealed a unique ligand-mediated crystal contact caused by the unsubstituted backbone amide of the N-terminal building block of ProM-3. Both, the backbone NH proton and the carbonyl

Ac–[2-Cl-Phe][ProM-2]-(?)–OH		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
2a	[ProM-1]	2.3 (0.2)	–32.2 (0.2)	(ref)
3a	[ProM-3]	9.7 (0.8)	–28.6 (0.2)	+3.6 (0.3)

Tab. 9.10.: **Unsuccessful optimization of the N-terminal building block of ProM-1 by ProM-3**
Mimicking Xaa-*trans*-Pro with ProM-3 lost significant affinity. Table lists affinities as K_d and ΔG of ligands measured by fluorescence titration using ENAH EVH1. Values in parentheses are standard errors.

group of ProM-3 were contacted by the Asn43 side chain of the neighboring chain. The crystal contact biased the water shell around the ligand and aggravated the interpretation of the loss in affinity. Attempts to crystallize ENAH EVH1 in a different crystal packing were not successful (SI Sec. 18.6).

9.2.7. Ac–[2-Cl-Phe][ProM-2][ProM-4]–OH in complex with ENAH EVH1

We therefore decided to get further insights into the binding mode of ^{4}Pro by synthesizing a second scaffold, ProM-4.[141] This scaffold harbored a piperidine ring as N-terminal building block, which kept the N-amide substitution intact. Given the accuracy of the affinity measurement, the inhibitor Ac–[2-Cl-Phe][ProM-2][ProM-4]–OH (**4a**) lost insignificant affinity compared to the ProM-1-containing reference **2a**.

Ac–[2-Cl-Phe][ProM-2]-(?)–OH		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
2a	[ProM-1]	2.3 (0.2)	–32.2 (0.2)	(ref)
4a	[ProM-4]	3.0 (0.3)	–31.5 (0.2)	+0.7 (0.3)

Tab. 9.11.: **Unsuccessful optimization of the N-terminal building block of ProM-1 by ProM-4**
The amide N-substituted piperidine moiety of ProM-4 bound not significantly weaker than the reference inhibitor. Two-tailed t-test with α=5%, df=152: t_{crit}=1.98, t_{value}=1.32. 95% confidence intervals: **2a** 1.9–2.7 µM, **4a** 2.6–3.7 µM. Table lists affinities as K_d and ΔG of ligands measured by fluorescence titration using ENAH EVH1. Values in parentheses are standard errors.

We solved the crystal structure to ensure that the binding mode of ProM-4 was not altered and the negative effect of ProM-3 was solely based on the broken N-amide substitution.

9.2.8. Ac–WPPPPTEDEL–NH_2 in complex with ENAH EVH1

The modifications on the N-terminal side of ProM-1 were not successful. In a different approach we searched for unknown interaction sites, accessible by modifications of the C-terminal building block. Affinity measurements suggested that the C-terminal pentamer TEDEL contributes the major part of the affinity boost of 13-mers like the chimeric ligand **0b**.

Ligand composition		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
0b	Ac–SFE[2-Cl-Phe]PPPPTEDEL–NH_2	0.79 (0.05)	−34.8 (0.2)	(ref)
0g	Ac–SFE[2-Cl-Phe]PPPP–OH	4.6 (0.2)	−30.5 (0.1)	+4.3 (0.2)
0c	Ac–[2-Cl-Phe]PPPPTEDEL–NH_2	1.25 (0.04)	−33.7 (0.07)	+1.1 (0.2)
0a	Ac–[2-Cl-Phe]PPPP–OH	9.9 (0.8)	−28.6 (0.2)	+6.2 (0.3)

Tab. 9.12.: **The flanking epitope TEDEL affects ActA-derived peptides stronger than SFE**
Table lists affinities as K_d and ΔG of ligands measured by fluorescence titration using ENAH EVH1. Values in parentheses are standard errors.

We therefore solved crystal structures of the C-terminally elongated natural ligands. However, the N-terminally truncated **wt** peptide did not crystallize. To increase affinity without affecting the C-terminal binding mode, ^{1}Phe was substituted and Ac–WPPPPTEDEL–NH_2 was cocrystallized.

9.2.9. Ac–[2-Cl-Phe]PPPPTEDEL–NH_2 in complex with ENAH EVH1

To exclude that the adopted conformation of TEDEL–NH_2 was not biased by a packing artifact, the ligand composition was slightly modified and a non-natural moiety was incorporated. We synthesized Ac–[2-Cl-Phe]PPPPTEDEL–NH_2 (**0c**) and repeated the crystallization with ENAH EVH1.

The superposition of both complex structures revealed that TEDEL–NH_2 bound canonically to ENAH EVH1 with a mixture of polar and hydrophobic interactions. Single-point mutation affinity measurements further corroborated the conformation of TEDEL.

Ligand composition		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
0c	Ac–[2-Cl-Phe]PPPPTEDEL–NH_2	1.25 (0.04)	–33.7 (0.07)	(ref)
0h	Ac–[2-Cl-Phe]PPPPTEDDL–NH_2	3.8 (0.2)	–30.9 (0.1)	+2.8 (0.1)
0i	Ac–[2-Cl-Phe]PPPPTEDEA–NH_2	3.5 (0.1)	–31.4 (0.1)	+2.3 (0.1)

Tab. 9.13.: **Terminal ^{9}EL contribute substantially to binding affinity**
Single point mutations from ^{9}Glu to ^{9}Asp, as well as ^{10}Leu to ^{10}Ala lost significant affinity. Table lists affinities as K_d and ΔG of ligands measured by fluorescence titration using ENAH EVH1. Values in parentheses are standard errors

9.2.10. Ac–[2-Cl-Phe]PP[ProM-1]TEDEL–NH_2 in complex with ENAH EVH1

To facilitate the *in silico* re-design of ProM-1, we solved a third crystal structure with the hybrid ligand Ac–[2-Cl-Phe]PP[ProM-1]TEDEL–NH_2 (**1c**). The structure revealed that ProM-1 did not affect the binding mode of TEDEL. The finding opened the possibility to perform docking studies of new scaffolds with the chimeric composition. The docking studies yielded in three new scaffolds[1] ProM-9, ProM-12, and ProM-13.

9.2.11. Ac–[2-Cl-Phe][ProM-2][ProM-12]–OH in complex with ENAH EVH1

In silico docking studies suggested to use the C-terminus of ProM-1 to compete with a bound water while mimicking the polar interaction of TEDEL–NH_2. To reach the polar interaction site, the chirality of the C-terminal building block was inverted. This configuration used the exit vector of the C-terminal carboxyl group to point towards the polar site. However, ligand compositions with the new scaffold ProM-12, Ac–[2-Cl-Phe][ProM-2][ProM-12]–OH (**5a**) lost dramatic affinity. To investigate why the docking studies did not agree with the affinity measurements, inhibitor **5a** was cocrystallized.

[1] The numbering of ProM scaffolds is discontinuous as several other scaffolds were synthesized prior to the design of ProM-9

Ac–[2-Cl-Phe][ProM-2]-(?)		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
2a	[ProM-1]–OH	2.3 (0.2)	–32.2 (0.2)	(ref)
5a	[ProM-12]–OH	13.5 (0.4)	–27.8 (0.1)	+4.4 (0.2)

Tab. 9.14.: **ProM-12 fails to mimic the polar interaction of ^{9}Glu**
The scaffold ProM-1 was modified to mimic the H-bond of the negatively charged ^{9}Glu. The new scaffold, ProM-12 was incorporated into the parent inhibitor with free C-terminus but failed to boost the affinity.

9.2.12. Ac–[2-Cl-Phe][ProM-2][ProM-12]–OMe in complex with ENAH EVH1

The inhibitor **5a** was masked with a non-hydrolyzable methyl ester function that would render it cell-membrane-permeable. This aimed to investigate the behavior of a possible uncharged pro-drug composition. However, also the methyl ester composition **5c** lost as much affinity as the charged inhibitor.

Ac–[2-Cl-Phe][ProM-2]-(?)		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
2c	[ProM-1]–OMe	4.4 (0.7)	–30.6 (0.4)	(ref)
5c	[ProM-12]–OMe	15 (1)	–27.6 (0.2)	+3.0 (0.4)

Tab. 9.15.: **ProM-12 fails to mimic the polar interaction of ^{9}Glu**
The scaffold ProM-1 was modified to mimic the H-bond of water or ^{9}Glu. The new scaffold, ProM-12 was incorporated into the parent inhibitor but failed to boost the affinity.

Both crystal structures of ENAH EVH1 in complex with inhibitors **5a** and **5c** revealed a significant rearrangement of ProM-12 to minimize steric clashes with the underlying epitope. As different approaches to modify ProM-12 were out of reach, we switched to the second interaction of TEDEL–NH_2, namely the hydrophobic interaction of the terminal leucine.

9.2.13. Ac–[2-Cl-Phe]PP[ProM-9]–OH in complex with ENAH EVH1

The docking studies for a new scaffold addressing the hydrophobic interaction were partly based on the high-resolution structure of Ac–[2-Cl-Phe]PP[ProM-1]–OH (core motif of **1b**). The *in silico* docking studies yielded in the methylated derivate ProM-9,

which outperformed ProM-1 substantially. Given the reference model on which the new scaffold was docked, ProM-9 was crystallized first as Ac–[2-Cl-Phe]PP[ProM-9]–OH composition (**6b**).

Ac–[2-Cl-Phe]PP-(?)		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
	[ProM-1]–OH	2.4 (0.4)	–32.0 (0.4)	(ref)
6b	[ProM-9]–OH	0.23 (0.05)	–37.9 (0.5)	–5.9 (0.6)

Tab. 9.16.: **ProM-9 mimics the hydrophobic interaction of ^{10}Leu**
The alkylated derivate of ProM-1 successfully addressed the hydrophobic patch as predicted by docking studies.

The structure revealed a very similar ligand orientation and confirmed that the *in silico* docking worked satisfactorily to predict the hydrophobic interaction between ProM-1 derivatives and ENAH EVH1.

9.2.14. Ac–[2-Cl-Phe][ProM-2][ProM-9]–OH in complex with ENAH EVH1

The composition of the initial inhibitor **2a** was exchanged by ProM-9. This crystal structure of Ac–[2-Cl-Phe][ProM-2][ProM-9]–OH (**6a**) revealed that the hydrophobic patch is successfully addressed by ProM-9 and that the methylation did not cause steric clashes with the hydrophobic patch.

Ac–[2-Cl-Phe][ProM-2]-(?)		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
2a	[ProM-1]–OH	2.3 (0.2)	–32.2 (0.2)	(ref)
6a	[ProM-9]–OH	0.12 (0.02)	–39.5 (0.3)	–7.3 (0.4)

Tab. 9.17.: **ProM-9 mimics the hydrophobic interaction of ^{10}Leu**
Compared to the parent ligand **2a**, alkylation by a single methyl group boosted the affinity nearly 20-fold.

Among other ligands, difference electron density of **6a** was visible in a putative second binding site of ENAH EVH1 (Discussion Sec. 13.2). Isothermal calorimetry was performed to test whether a one-to-one model is sufficient to describe the binding process to ENAH EVH1.

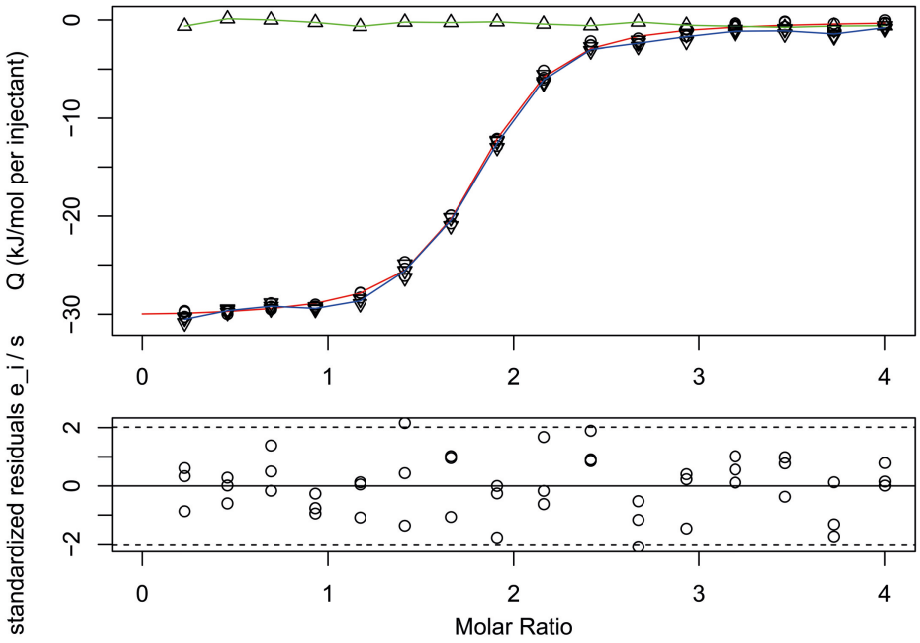


Fig. 9.1.: **ITC data of a ligand found in second binding site**
Raw ITC titration data of Ac–[2-Cl-Phe][ProM-2][ProM-9]–OH binding to 50 µM ENAH EVH1 (blue, triangle). Subtraction of the background titration (green, triangle) yielded in the ITC binding isotherms (circles) that were fitted by a one-to-one model (red curve). Standardized residuals from this fit are printed in the lower panel and showed normal distribution and no measurable competition for a second binding site. K_d 680 (40) nM, M=57.9% (Sec. 16.1.1), lack-of-Fit: F=1.999, P=0.059.

9.2.15. Ac–[2-Cl-Phe][ProM-2][ProM-13]–OEt in complex with ENAH EVH1

The ProM-9-containing structures suggested that the hydrophobic interaction can be accentuated by enlarging the coverage capacity of ProM-9. ProM-13 was synthesized with ethylated decoration and incorporated directly into the cell-membrane-permeable inhibitor Ac–[2-Cl-Phe][ProM-2][ProM-13]–OEt. Notably, this ligand fully recovered the affinity of the high affine peptidic chimera **2b**.

Ligand composition		$K_{d,FT}$ [µM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
2b	Ac–SFE[2-Cl-Phe][ProM-2][ProM-1]TEDEL–NH_2	0.15 (0.06)	–39 (1)	(ref)
2	Ac–[2-Cl-Phe][ProM-2][ProM-1]–OEt	4.1 (0.3)	–30.8 (0.2)	+8 (1)
6	Ac–[2-Cl-Phe][ProM-2][ProM-9]–OEt	0.38 (0.05)	–36.6 (0.3)	+2 (1)
7	Ac–[2-Cl-Phe][ProM-2][ProM-13]–OEt	0.18 (0.03)	–38.5 (0.4)	+1 (1)

Ligand composition		HA	LE_{FT} [J (mol HA)$^{-1}$]
2b	Ac–SFE[2-Cl-Phe][ProM-2][ProM-1]TEDEL–NH_2	115	–340 (9)
2	Ac–[2-Cl-Phe][ProM-2][ProM-1]–OEt	50	–620 (4)
6	Ac–[2-Cl-Phe][ProM-2][ProM-9]–OEt	51	–720 (6)
7	Ac–[2-Cl-Phe][ProM-2][ProM-13]–OEt	52	–740 (8)

Tab. 9.18.: **ProM-9 & ProM-13 successfully mimic the hydrophobic interaction of ^{10}Leu**
Upper table: The alkylated derivate of ProM-1, ProM-13, restored the entire binding energy of the flanking epitopes present in the peptidic chimera **2b**. Lower table: The introduced heavy atoms (HA) by methylation and ethylation lead to an over-proportional boost in affinity. The improved ligand efficiency (LE: ΔG per HA) for both inhibitors **6** and **7** justified the increase of the molecular weight. Table lists affinities as K_d and ΔG of ligands measured by fluorescence titration using ENAH EVH1. Values in parentheses are standard errors.

9.2.16. Ac–[2-Cl-Phe]PP[ProM-9]–OH in complex with EVL EVH1

Most of the above mentioned inhibitor compositions were used for cocrystallization experiments with the EVL EVH1 constructs as well. However, crystallization of EVL EVH1 turned out to be difficult and was successful only with the inhibitor Ac–[2-Cl-Phe]PP[ProM-9]–OH.

The binding mode of EVL EVH1 showed no differences to ENAH EVH1, since the the amino acids of the main binding groove are conserved. Even though the structure is not mentioned further in the discussion, it represents the highest resolution structure of EVL EVH1 which showed for the first time the loop region of Thr30 resolved.

10. Additional biochemical and biophysical studies

10.1. SPOT array substitution

Single residue substitution experiments are reported for VASP EVH1.[106] Here, mutational SPOT analysis was completed with all three paralogs, using an infrared dye anti-rabbit antibody for read-out, compared to the chemiluminescent approach used by Ball and coworkers.

In agreement with the reported mutational analysis, the SPOT array experiments showed that prolines of the ^{1}FPPPP core motif do not interact equally intimate with Ena/VASP EVH1. The prolines at position 2 and 5 (magenta, green) facing the epitope were not exchangeable by any naturally occurring amino acids. The second (white) proline, is fully solvent accessible and exchangeable by any amino acid. The hydrophobic interaction site with ^{4}Pro (orange) allows exchange by aliphatic amino acids.

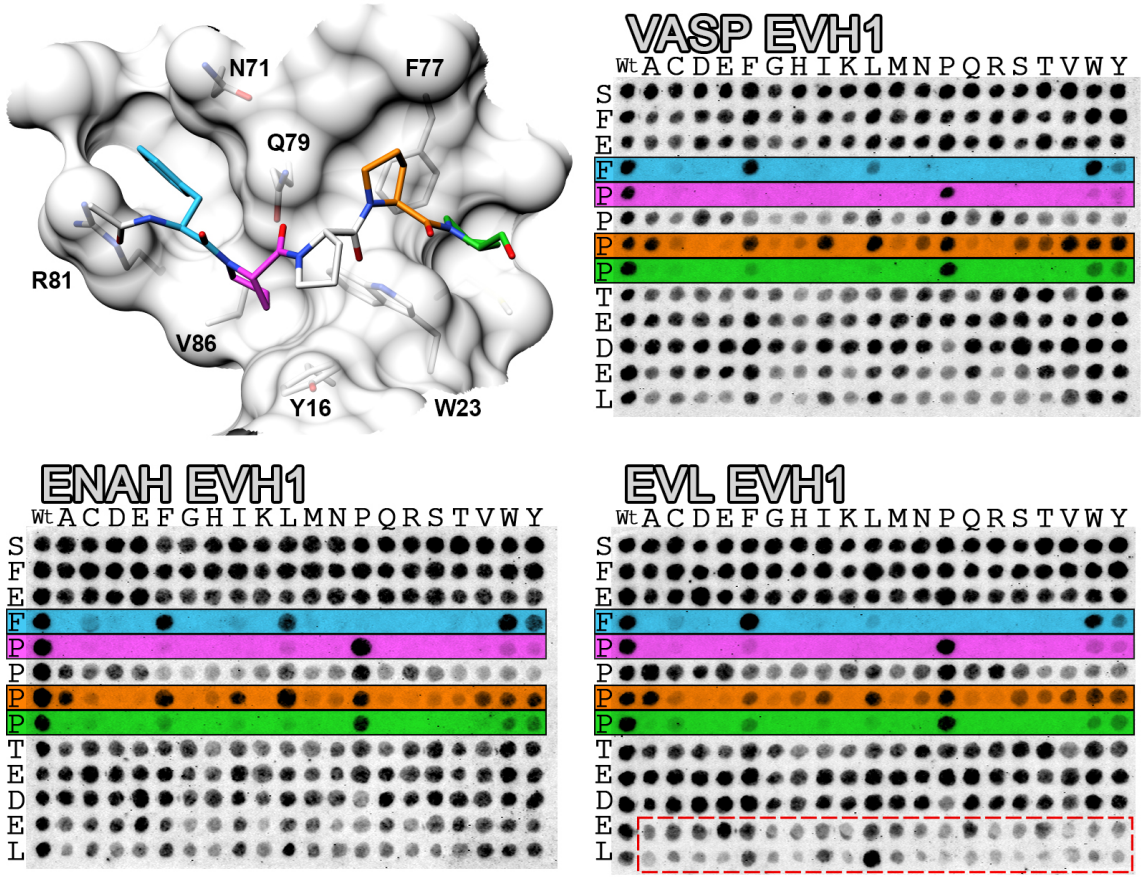


Fig. 10.1.: Mutational analysis of all three Ena/VASP EVH1 domains binding the proline-rich segment SFEFPPPPTEDEL, derived from ActA. Spot intensities correlate to the EVH1 binding affinities. For better understanding the complex structure of Mena EVH1 bound to Ac–FPPPPT–NH_2 is shown in the top left panel (PDB code 1EVH). The resolved parts of the ligand are color coded according to the amino acids driving selectivity in the SPOT experiments. Of the C-terminal epitope EL (dashed red box), no structural information is available.

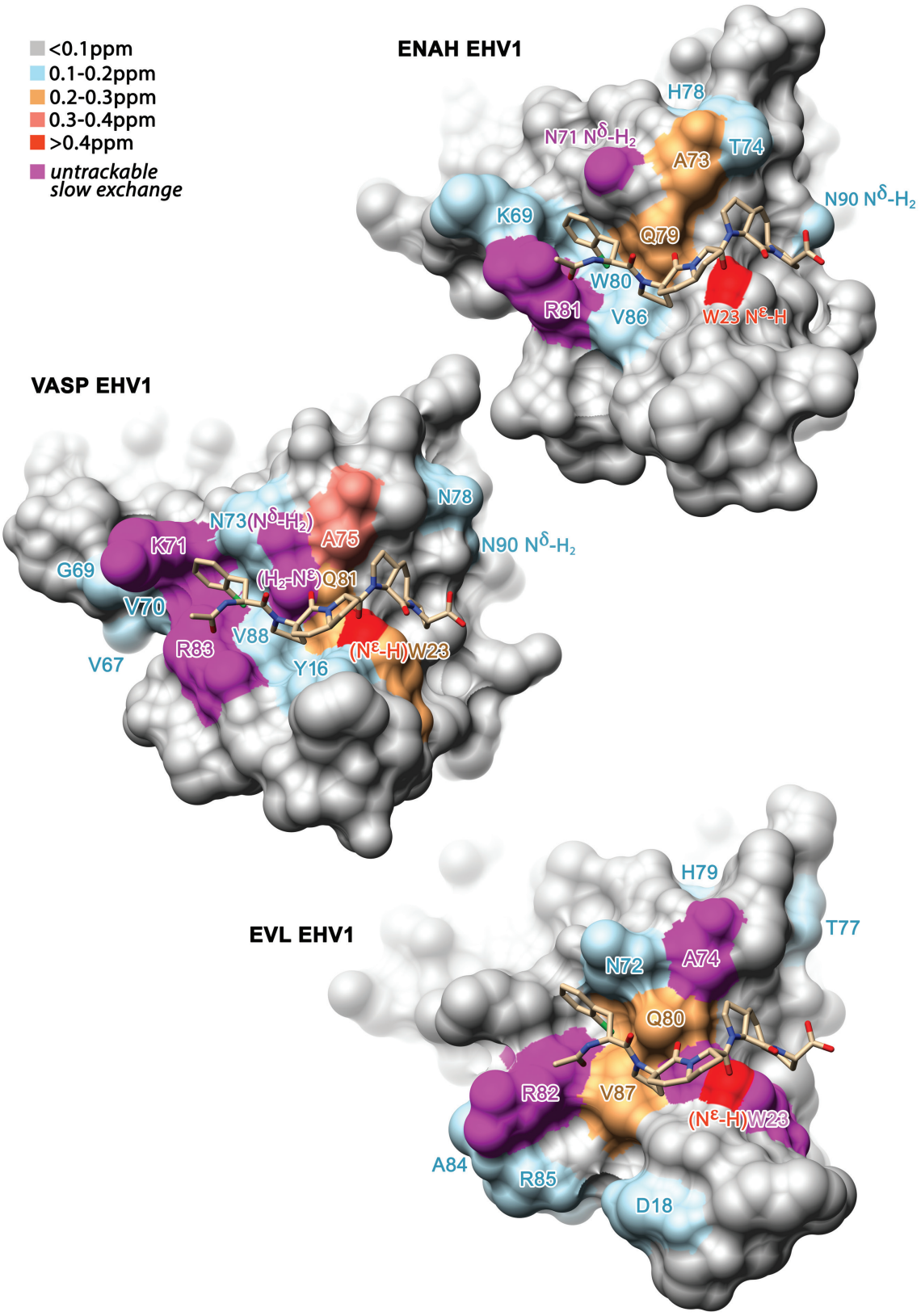
10.2. ^{1}H-^{15}N-HSQC experiments

The peak tables of the ^{1}H-^{15}N-HSQC measurements are listed in Sec. 20 of the SI. Mapping the perturbations on the solvent accessible surface of EVH1 domains confirmed that the initial inhibitor **2a** bound canonically to all three Ena/VASP paralogs (Fig. 10.2).

Additionally, ^{1}H-^{15}N-SOFAST-HMQC experiments with C-terminally elongated ligands indicated that TEDEL–NH_2 bound in the vicinity to the main binding groove. These experiments further suggested that ActA-derived 13-mers bound to a putative second binding site of ENAH EVH1. The perturbations of this ^{1}H-^{15}N-SOFAST-HMQC experiment are printed graphically only in the discussion (Fig. 13.5b).

Fig. 10.2.: **^{1}H-^{15}N-HSQC perturbations of all Ena/VASP EVH1 domains binding the inhibitor 2a**

Changes in chemical shifts of the ^{1}H and ^{15}N resonances larger than 0.1 ppm were color coded on the solvent accessible surfaces of the according EVH1 domain (PDB codes 4MY6, 1QC6, 1EGX). Except for ENAH EVH1, inhibitor **2a** was docked.


<0.1ppm
0.1-0.2ppm
0.2-0.3ppm
0.3-0.4ppm
>0.4ppm
untrackable
slow exchange
ENAH EHV1
H78
N71 Nδ-H2
T74
A73
N90 Nδ-H2
K69
Q79
W80
R81
V86
W23 Nε-H
VASP EHV1
N78
N73(Nδ-H2)
A75
K71
N90 Nδ-H2
G69
(H2-Nε)Q81
V70
V88
R83
(Nε-H)W23
V67
Y16
EVL EHV1
H79
T77
N72
A74
Q80
R82
V87
(Nε-H)W23
A84
R85
D18

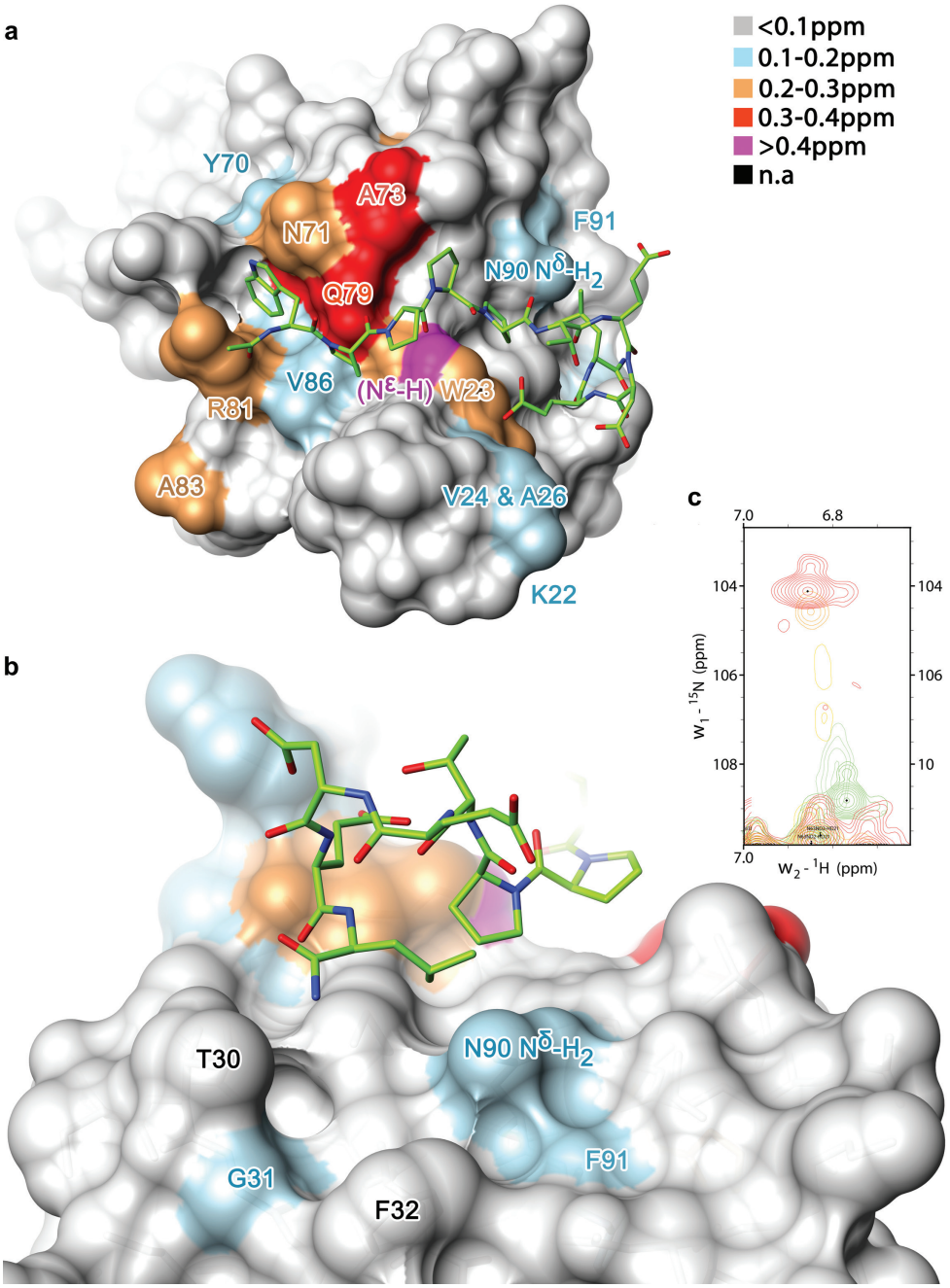


Fig. 10.3.: **^{1}H-^{15}N-HSQC perturbations of ENAH EVH1 binding Ac–WPPPPTEDEL–NH_2**
(**a**) Changes in chemical shifts of the backbone ^{1}H and ^{15}N larger than 0.1 ppm were color coded on the solvent accessible surface of ENAH EVH1. (**b**) Thr30 and Phe32 in vicinity to the main binding groove of ENAH EVH1 could not be assigned (n.a). (**c**) EVH1 domains show a dominant slow exchange binding mode that made tracking of peaks difficult. Shown are two different, unassigned peaks with fast line broadening (green unbound, red 6× excess)

Part IV.

Discussion

Abstract of the discussion

Despite of the recent advances in NMR-based drug development,[177] protein crystallography remains the suitable method for detailed structural elucidation of protein-ligand interactions.[151] I provide herein a set of 15 deposited crystal structures of ENAH EVH1 in complex with inhibitors representing all relevant scaffold compositions. Many of these models were refined well below a resolution limit of 1.7 Å and allow examination of the ProM scaffold suite on an atomic level. Much effort was invested to provide structural information also for ligand compositions that reflect early stages of the inhibitor design.

Sections 11.2 to 12.4 cover the drug development and optimization process that starts from the wild type ActA-derived 13-mer Ac–SFE^{1}FPPPPTEDEL–NH_2 peptide and yields in a high-affine cell-membrane-permeable inhibitor against the pro-metastatic target Ena/VASP. For better accessibility, ligand compositions and abbreviations are listed in the backside cover inlay. The quantitative structure-affinity relationship study involves the scaffolds presented in figure 6.2 and provides relevant crystal structures and according changes in the binding affinity. Within the mentioned parent peptide, we exchanged residues of the core recognition motif ^{1}FPPPP with 2-chloro-*L*-Phe (2-Cl-Phe, Sec. 11.2) and the scaffolds ProM-1 (Sec. 11.3) and ProM-2 (Sec. 11.4). This chimeric 13-mer was not suitable for *in cell* studies due to its peptidic nature, massive negative charge and large molecular weight. However, the affinity gain by the three modifications was high enough to shorten the 13-mer to the core motif pentamer. Masking the only negative charge of the C-terminus rendered the initial inhibitor cell-membrane-permeable.[141]

Previous work[105,141] already optimized 2-Cl-Phe and showed that ProM-2 did not improve the affinity significantly. The subsequent optimization of the initial inhibitor therefore focused on ProM-1. The modular synthesis strategy of ProM scaffolds allowed separate modifications of both building blocks to explore the epitope underneath prolines 4 and 5. Optimization of the N-terminal building block by the scaffolds ProM-3 and ProM-4 is discussed in section 12.1. These two alternations for ^{4}Pro were not successful.

On the search to optimize the C-terminal buiding block, three C-terminally elongated ActA-derived 10-mers were cocrystallized. For the first time, these structures gave insights in the adopted conformation of -TEDEL–NH_2. This interaction as well

as the additional binding epitope provided by ENAH EVH1 are discussed in section 12.2. The two found interaction sites were located in close vicinity to the initial inhibitor and were addressed by newly synthesized ProM-1 derivatives. Unsuccessful mimicry of the polar interaction by ProM-12 (Sec. 12.3) is discussed separately from the successful addressing of the hydrophobic patch by ProM-9 and ProM-13 (Sec. 12.4).

Incorporation of the latter scaffold boosted the affinity of the initial inhibitor up to 20-fold. The successful *in silico* docking studies and scaffold synthesis demonstrate the powerful modular architecture of our scaffold toolbox that allows adaptation to a very specific epitope, as claimed earlier in Opitz *et al.*.[141] In the context of Ena/VASP, the optimized, cell-membrane-permeable inhibitor displays an interesting lead substance and a potential novel class of antimetastatic drugs acting at the very end of converging receptor kinase signaling and integrin pathways.

Even though many of the crystal structures served as templates for the successful structure-guided drug design, they did not provide the measure which guided the optimization process. Throughout the discussion, ligand optimization is guided by the binding affinity, namely the decrease of ΔG. The derivation of ΔG as suitable measure can be found in the SI chapter 16, which also highlights that different parts of the interaction site of Ena/VASP EVH1 domains contributed additively to ΔG. This circumstance allowed comparison of scaffold modifications independent of the over-all ligand composition.

11. Non-natural moieties allow shortening of the wt ligand to an uncharged pentamer

11.1. The parent wild type ligand

As starting point of the ligand development served the third proline-rich repeat of ActA from the intracellular pathogen *L. monocytogenes*. The repeats were of particular interest as the evolutionary pressure selected these peptide sequences for maximal affinity and not for integrity within the actin-associated interactome presented in the introducing section 2.2. Indeed, sequences of natural ligands of comparable length show affinities in the range of hundreds of micromolar,[178] while the wild type ActA-derived 13-mer Ac–SFE^{1}FPPPPTEDEL–NH_2 (**wt**, Tab. 9.6) binds with roughly 20 µM to the EVH1 domains. However, initial attempts to crystallize 13-meric ligands like **wt** or **0b** all failed. Crystallization was initially only successful for the pentameric core recognition motifs with free carboxyl termini (–OH). Later complex structures of pentameric ligands with masked C-termini confirmed that the negative charge had no influence on the orientation of the ligand. Even C-terminally elongated ligands bound with the same mode as negatively charged pentameric compositions (Fig. 11.1a).

11.2. Replacement of ^{1}Phe by 2-chloro-*L*-phenylalanine

The spot array analysis of all three Ena/VASP EVH1 domains revealed a strong preference for ^{1}Phe, ^{1}Tyr, ^{1}Leu and ^{1}Trp within the core motif ^{1}FPPPP (Fig. 10.1). The substitution assay showed unreported selectivity among the three paralogs for

^{1}LPPPP and suggested that ENAH EVH1 bound ^{1}Leu with higher affinity than the other two paralogs. Such a selectivity within the core recognition motif is reported only for the miniature protein pGolemi. Studies are ongoing to investigate whether the alternate motif ^{1}LPPPP found in natural binding partners like Lpd,[66] RIAM,[65] or dAbi[57] shows indeed selectivity for Ena/VASP EVH1.

The preference for aromatic amino acids is explained by Val86 and Arg81 that provide hydrophobic and cation-π[179] interactions while shaping the bottom and the outer side of the pocket. Arg81 adopts a mixture of two rotamer conformations in many solved structures (Fig. 13.1b) and provides enough flexibility to host bigger amino acids than Phe. Due to steric restrictions caused by ^{2}Pro and Tyr16, the ^{1}Phe side chain can not approach the underlying Val86. This opened the possibility to fill the void by *ortho* substituted phenyl rings with few hydrophobic heavy atoms. Selected non-natural amino acids were incorporated into the parent ActA-derived 13-mer to replace ^{1}Phe. Binding studies on VASP EVH1 suggest that *ortho* substituted chloride provides optimal complementarity to the pocket. Neither smaller nor more hydrophobic decoration increase the binding affinity significantly (Tab. 9.6).[141]

The complex structure of Ac–[2-Cl-Phe]PPPP–OH (inhibitor **0a**) bound to ENAH EVH1 confirmed a hydrophobic contact to Val86 (Fig. 11.1b) that added at least –6 kJ/mol of free energy or more than 20-fold decrease in K_{d}. Chlorination of the ActA 13-mer released the same amount of free energy on ENAH EVH1 as structuring the binding motif by pGolemi. Holtzman and coworkers report a gain of –6.3 kJ/mol[134] compared to the solvent exposed ActA 11-mer (Sec. 5.1). Incorporation of 2,6-di-Cl-*L*-phenylalanine to remove the asymmetry and positively influence the on-rate abolished binding (ligand **0e**). The crystal structure of ENAH EVH1 in complex with Ac–WPPPPTEDEL–NH_2 confirmed that ligands with bulkier residues show no altered binding mode. Due to the dominant *cis-trans* isomerization of the Trp-Pro peptide bond and the fact that only the *trans* isomer is binding to the EVH1 domain, ^{1}Trp decoration has not been investigated.

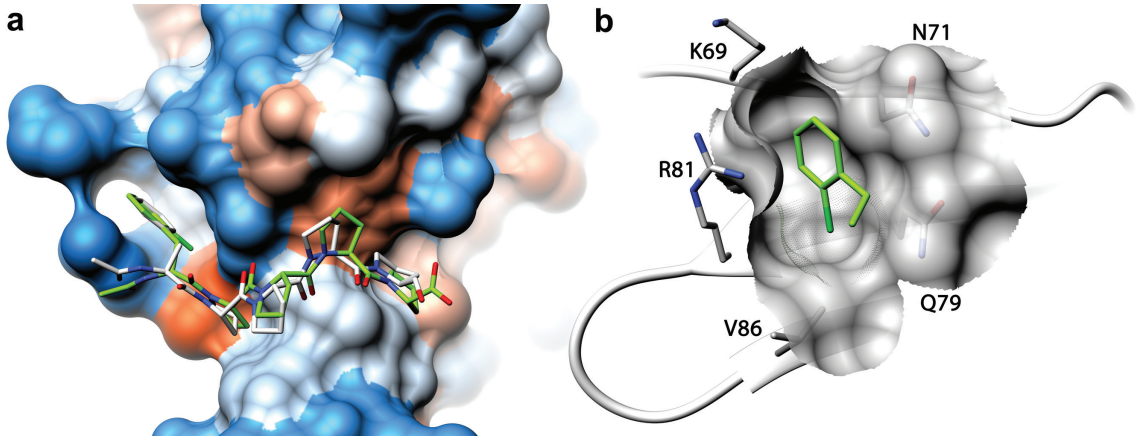


Fig. 11.1.: **Superposition of Ac–[2-Cl-Phe]PPPP–OH (0a) and Ac–FPPPPT–NH_2 in complex with ENAH EVH1**
(a) Chlorination of ^{1}Phe (**0a**, green) had no effect on the binding mode of the proline-rich motif. ^{6}Thr of the wild type ligand (white, PDB code 1EVH[174]) is not resolved. Solvent accessible surface of ENAH EVH1 is color coded by hydrophobicity (blue-white-red for hydrophilic-hydrophobic). **(b)** The N-terminal non-proline amino acid is hosted on a hydrophobic patch established by Val86. ^{1}Phe was *ortho* substituted with chloride to contract Val86. For better visibility, the rest of the ligand is transparent.

11.3. Replacement of ^{4}PP by ProM-1

The peptide binding groove establishes a flat apolar basin contacting the C-terminal ^{4}PP. Especially Phe77 in the center of the patch accepts substitutions of ^{4}Pro by any aliphatic residue, but exclusively selects for prolines at position 5 (Fig. 10.1). Replacement of ^{4}PP by ProM-1 was advantageous (Tab. 9.7). Isothermal titration calorimetry (ITC) data analysis suggests that the gain of affinity is solely enthalpy driven.[105] Superposition of the core motifs **0b** and **1b** confirmed that ProM-1 efficiently shielded Phe77 from the solvent without affecting the C_{α} chain trace of the ligand (Fig. 11.2). Incorporation of ProM-1 into the mentioned 13-mer **0b** gained at least –3.2 kJ/mol of free energy (SI Tab. 16.4) and boosted the affinity fourfold. The superposition in Fig. 11.2b suggests that the scaffold snugly fit over the hydrophobic patch.

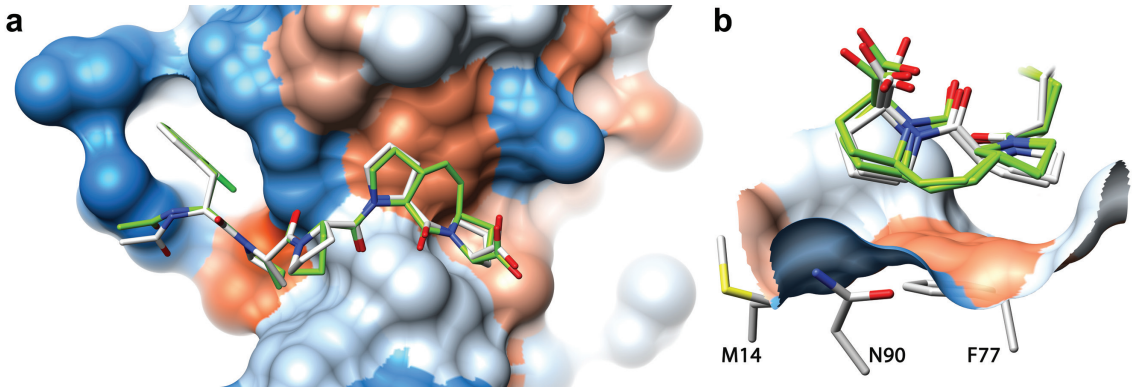
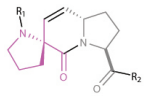


Fig. 11.2.: **Superposition of Ac–[2-Cl-Phe]PPPP–OH (0a) and Ac–[2-Cl-Phe]PP[ProM-1]–OH (1a) in complex with ENAH EVH1**
(**a**) Replacement of the second pair of prolines by ProM-1 (**1a**, green) did not affect the ligand C_α chain trace relative to ^{4}PP (**0a**, white). (**b**) The vinylidene bridge snugly fit over the aromatic ring of Phe77 and shielded the hydrophobic patch from the solvent. Complex structures of **0a** and **1a** each contain two chains in the asymmetric unit, all are shown.



11.4. Replacement of ²PP by ProM-2

Extensive binding studies reveal that ProM-1 can not be shuffled within the core recognition motif to replace any Pro-Pro tandem while maintaining the binding strength of the **wt** ligand.[105] Especially on VASP EVH1, the replacement of ²PP by ProM-1 is disadvantageous, and ProM-1 double-substituted 13-mers (**1f**) bind as weak as the parent **wt** ligand (Tab. 9.8).[141,142] Strikingly, when equipped with 2-Cl-Phe, VASP EVH1 binds 13-mers with the core motifs ²PP[ProM-1] and ²[ProM-1][ProM-1] with the same affinity (ligands **1b** and **1d**, Tab. 9.8). These measurements indicate that anchoring the ligand with 2-Cl-Phe on VASP EVH1 is sufficient to compensate the affinity loss of ProM-1 replacing ²PP. In sharp contrast, the single substituted **wt** composition with the core motif ¹F[ProM-1]PP lost only +1.2 (0.3) kJ/mol compared to **wt** on ENAH EVH1 (SI Tab. 16.2).

The VASP EVH1 domain could not be cocrystallized with any ¹F[ProM-1]-containing hybrid ligand. Only ENAH EVH1 crystallized in complex with the core recognition motif of **1d**, Ac–[2-Cl-Phe][ProM-1][ProM-1]–OH. Superposing ligands **1a** and **1d** revealed no rearrangements of the underlying epitope. In the complex structure of the ProM-1 double-substituted inhibitor **1d**, Tyr16 was contacted by the vinylidene bridge and not by the pyrrolidine ring. This unnatural contact caused the inhibitor to readjust in order to minimize the steric clashes (Fig. 11.3a). As we sought to replace the core motif ¹FPP[ProM-1] by scaffolds that contact the EVH1 domains canonically, the configuration of ProM-1 was modified to meet the structural requirements underneath ²PP.

In silico docking studies suggested to alter the connection point of the vinylidene bridge to run it further away from Tyr16 (Fig. 11.3b). ^{1}H-^{15}N-HSQC measurements of single substituted **wt** ligands confirmed that such a spirocyclic scaffold ProM-2 (Fig. 6.2) successfully mimics the natural binding mode of two prolines.[142]

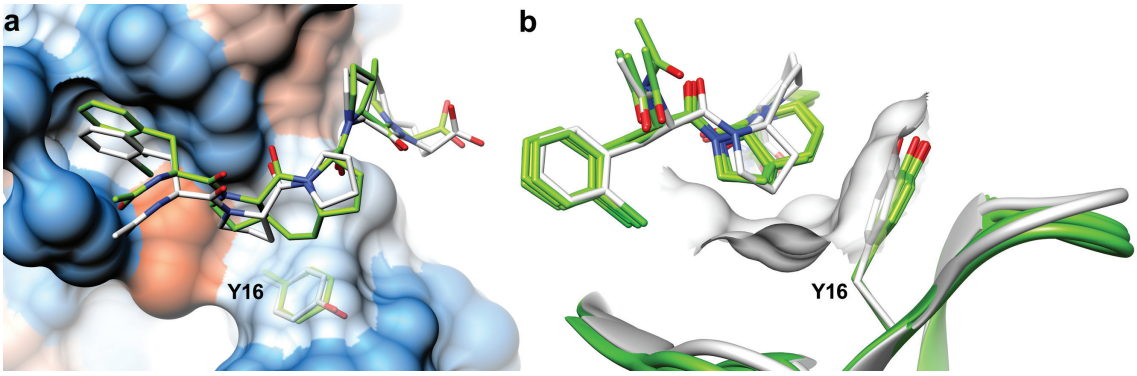


Fig. 11.3.: **Superposition of ENAH EVH1 crystal structures in complex with ligand compositions ^{2}PP, ^{2}[ProM-1] and ^{2}[ProM-2]**
(a) Superposition of Ac–[2-Cl-Phe][ProM-1][ProM-1]–OH (green, **1d**) and Ac–[2-Cl-Phe]PP[ProM-1]–OH (white, **1a**) revealed no readjustment of the underlying Tyr16. Instead, replacement of ^{2}PP by ProM-1 significantly elevated the ligand N-terminus from the epitope. **(b)** Superposition of **1d** (green) and parent inhibitor Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH (white, **2a**). ProM-2 mimicks the natural binding mode, as Tyr16 is contacted by the pyrrolidine ring and not the vinylidene bridge of ProM-1.[142] The different conformation of adjacent loop is due to the crystal contact. Complex structure of **1d** contains 4 chains, all are shown.

ProM-2 was incorporated into **1b** to form the high affine peptidic chimera **2b** (SI Tab. 16.2). The crystal structure of its core recognition motif Ac–[2-Cl-Phe][ProM-2]-[ProM-1]–OH (**2a**) confirmed that the vinylidene bridge of ProM-2 was not contacting Tyr16 (Figs. 11.3b and 11.4b). ^{1}H-^{15}N-HSQC perturbations showed that **2a** bound canonically to all EVH1 paralogs (Fig. 10.2). The lack of interaction between ProM-2 and Tyr16 was notable in the affinity. Replacement of ^{2}PP by ProM-2 generally showed no significant[1] boost in affinity (Tab. 9.9). Compared to an earlier attempt,[136] we were able to replace ^{2}Pro without any loss of affinity.

[1]meaning the error propagation of $\Delta\Delta G$ includes 0 kJ/mol

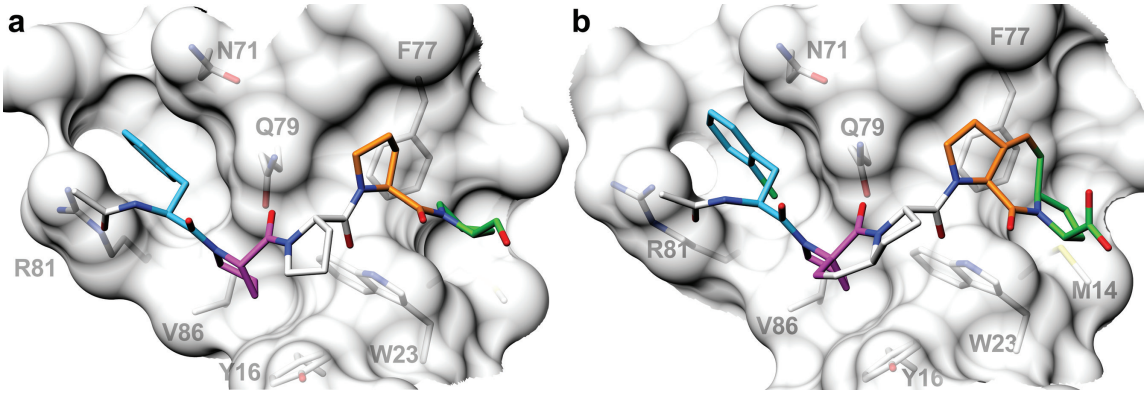


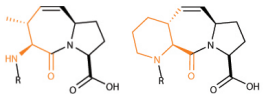
Fig. 11.4.: **The initial inhibitor Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH (2a) mimics the natural binding mode of Ena/VASP EVH1**
(a) Cocrystal structure of Mena EVH1 in complex with Ac–FPPPPT–NH_2. Only the core recognition motif FPPPP is resolved (PDB code 1EVH).[174] **(b)** Inhibitor **2a** mimics the binding mode of FPPPP.[141] Ligands are colored according to the pattern found in the SPOT array Fig. 10.1.

Taken together, ^{1}H-^{15}N-HSQC and affinity measurements suggest that VASP EVH1 reacts sensibly to ^{1}F[ProM-1] compositions.[105] However, the inhibitors are equipped with ^{1}2-Cl-Phe to boost affinity. The increased binding strength is presumably an effect of tightly anchoring the N-terminal side of the ligand over Val86. Consequently, ^{1}[2-Cl-Phe][ProM-1] compositions lost no affinity compared to ^{1}[2-Cl-Phe]PP (Tab. 9.8). Strikingly, ENAH EVH1 seemed to prefer ^{2}ProM-1 rather than ^{2}ProM-2 in single substituted 13-mers (Tab. 9.8). Future work must therefore elucidate if the affinity of the optimized inhibitor relies on ProM-2 or whether ^{1}[2-Cl-Phe][ProM-1]-containing compositions bind equally strong. All the more, as the synthesis of ProM-2 is highly challenging and optimization of ^{4}ProM-1 was successful only for the C-terminal building block. The usage of ProM-1 as replacement for ^{2}PP would therefore facilitate the supply of inhibitor, as both scaffolds would be composed of the same N-terminal building block.

12. Increasing the affinity of the initial inhibitor

The peptidic nature, massive negative charge and large molecular weight disqualifies the peptidic chimera **2b** from any *in cell* studies due to low membrane penetration and poor metabolic stability. Truncating the flanking residues of the chimera **2b** and masking the negative charge of the C-terminus was necessary to render Ac–[2-Cl-Phe][ProM-2][ProM-1]–OEt (**2**) cell-membrane-permeable.[141,178]

However, the moderate affinity of the parent inhibitor **2** made a concentration of up to 100 µM necessary to observe the biological effects. Such concentrations are unsuitable for *in vivo* experiments. For the optimization process, the configuration of ProM-2 remained unmodified as the scaffold showed no apparent boost in affinity (SI Tabs. 16.3 and 16.4). Instead, both building blocks of ProM-1 were modified. Derivatives of the N-terminal building block, 3-*trans*-vinylproline, are discussed in section 12.1 and yielded in two scaffold modifications (ProM-3 and ProM-4) which did not increase the affinity.[141] In a highly demanding approach,[116] we went to search for formerly unknown interaction sites beyond the main binding groove. We solved high-resolution crystal structures that were suitable for *in silico* docking studies to incorporate the newly found interaction sites into three new ProM scaffolds (ProM-9, ProM-12 and ProM-13). This structure-based approach yielded in the most potent, cell-membrane-permeable inhibitor known nowadays against Ena/VASP (Secs. 12.3 and 12.4).



12.1. Structural changes of N-terminal pyrrolidine ring in ProM-1

The results of the spot array analysis showed that the third proline was exchangeable by aliphatic residues. ProM-1-based Xaa-*trans*-Pro mimetica are interesting for other PRDs apart from Ena/VASP EVH1, and could serve as templates for WW domains that recognize PLPPLP motifs. Investigations for a ProM-2-based Leu-*trans*-Pro scaffold are ongoing. Like other Xaa-*trans*-Pro mimetica,[117,119] ProM-1 derivatives lack a freely rotating Xaa side chain as the vinylidene bridge fixates the C_β in one conformation. Hence, these scaffolds can not be seen as true Val-Pro (ProM-3), Ile-Pro (ProM-7) or Ala-Pro (ProM-10) mimetica and display an important alteration compared to Pro-Pro, namely the unsubstituted, protonated backbone amide of the Xaa amino acid (Fig. 6.2).

The secondary amide group of Ac–[2-Cl-Phe][ProM-2][ProM-3]–OH (**3a**) drastically altered the crystal packing. ProM-3 drove ENAH EVH1 into establishing a never-seen ligand-mediated crystal contact by binding Asn43 with both the amide proton and the carbonyl group. The unique crystal contact suggested that ProM-3 introduced more polar interactions on the solvent-exposed side of the bound ligand. Compositions with ProM-3 were therefore expected to loose affinity compared to ProM-1-containing ligands. Indeed, ^{4}ProM-3-containing inhibitors bound to ENAH- and VASP EVH1 with the same affinity as ^{4}PP, while parent compositions with ^{4}ProM-1 bound 4-5 times stronger (Tab. 9.10). In other words, the affinity boost caused by the vinylidene bridge of ProM-1 was canceled by the opening of its N-terminal pyrrolidine ring. The loss in binding strength for ProM-3-containing inhibitors held for inhibitor **3a** as well as for chimeric 13-mers (SI Tab. 16.2).

The presence of possible intrachain hydrogen bonds by ProM-3 presumably weakened the stabilization by water[180] and increased the solvation of the inhibitor, which in turn tended to interact less strongly with other solutes.[112] The hydration by more water molecules did not provide favorable conditions for PPII conformation[181] and the increased mobility of the scaffold finally raised the entropic penalty upon binding.[112] The resulting weakened affinity confirms the hypothesis that backbone solvation is the major determinant of PRS to adopt pre-structured, extended helical conformations.[181,182] In this context, the missing amide substitution and not the

conformational freedom of ProM-3 determined the PPII conformation and decreased affinity.[181]

This hypothesis was supported by the crystal structure and affinity measurements of compositions containing ProM-4 (Fig. 6.2). The enlarged ring size ensured full amide N substitution but loosened the rigidity of Ac–[2-Cl-Phe][ProM-2][ProM-4]–OH (**4a**). The piperidine moiety as N-terminal buiding block led to a different exit vector of ProM-4 and caused a slight rearrangement of **4a** relative to the parent inhibitor **2a** in the bound state (Fig. 12.2). However, the ProM-4 composition bound not significantly worse to ENAH EVH1 than the ProM-1-containing reference ligand (Tab. 9.11). As we aimed to maximize the ligand efficiency[183] (ΔG per heavy atom), ProM-4-containing ligands were not investigated further. The notion that even bulkier substitutions for prolines are valid as long as the backbone amide remains substituted is in agreement with reported ligands for SH3 domains, where replacement of single prolines by amide N-substituted peptoids yields in high-affinity inhibitors.[101,102,138]

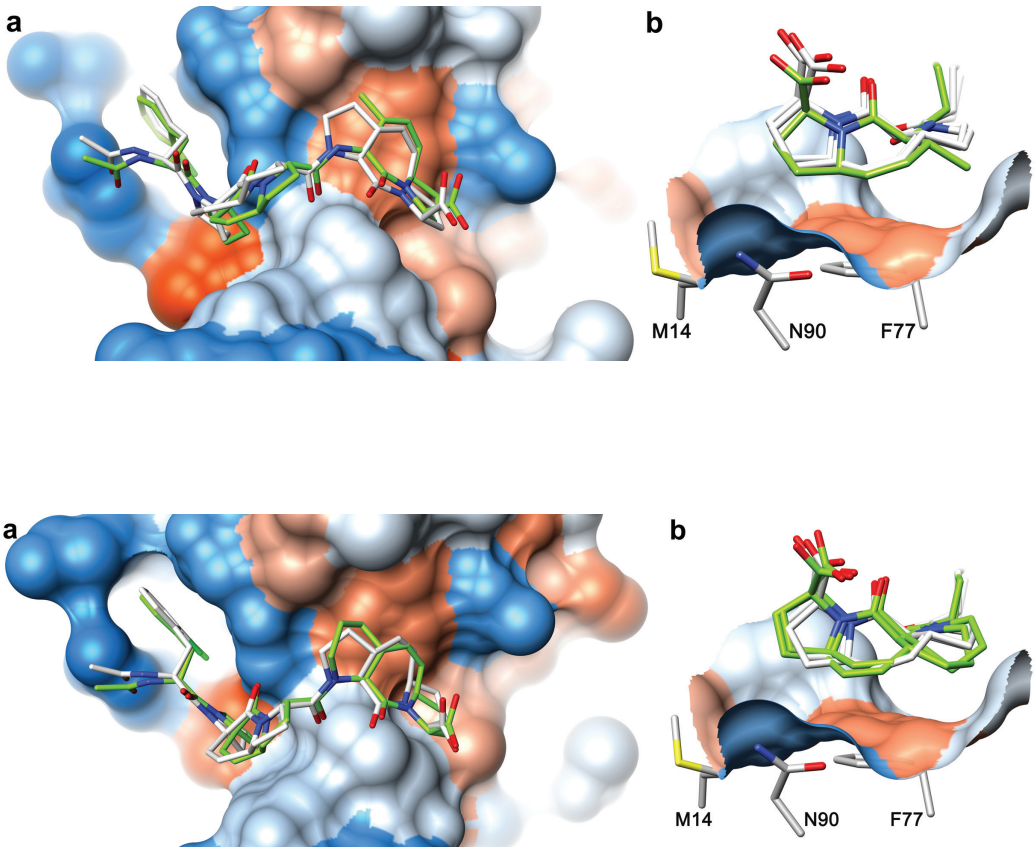


Fig. 12.2.: **Superposition of Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH (2a) and Ac–[2-Cl-Phe][ProM-2][ProM-4]–OH (4a) in complex with ENAH EVH1** Scaffold ProM-4 was designed to mimic the amide N-substitutions of Pro-*trans*-Pro interactions of ^{4}Pro (**a**) Enlarging the N-terminal ring of ProM-1 (**2a**, white) yielded in ProM-4 (**4a**, green) with slightly different exit vector visible at the C-terminal side. (**b**) The rearrangement was only local and did not affect the binding affinity significantly, as fluorescence measurements suggested. Solvent accessible surface of ENAH EVH1 is color coded by hydrophobicity (blue-white-red for hydrophilic-hydrophobic). Complex structures of **2a** and **4a** each contain two chains in the asymmetric unit, all are shown.

12.2. Finding interaction sites beyond the main binding groove: The C-terminal pentamer TEDEL

12.2.1. The crystal structure of TEDEL-elongated ligands

A comparison of the affinities of the chimera **2b** and its cell-membrane-permeable core motif Ac–[2-Cl-Phe][ProM-2][ProM-1]–OEt (**2**) showed that truncation of the flanking residues lost 8 kJ/mol of free energy upon binding. We hypothesized that the eight amino acids established specific interactions not addressed by the inhibitor **2** and tested whether the binding energy was concentrated on the N- or C-terminal flanking epitope. We found that the residues TEDEL–NH_2 affected the affinity of ActA-derived 13-mers significantly stronger than Ac–SFE, which is prominently visible in unrestricted peptides missing any ProM scaffolds (Tab. 9.12).

In the context of the Ena/VASP EVH1 domains, the protein surface on the C-terminal side of the binding groove is nearly featureless. To boost the affinity of the parent inhibitor we had to examine the interaction of TEDEL-containing ligands with EVH1 in close detail. I solved for the first time crystal structures of three TEDEL–NH_2-containing ligands (Fig. 12.3a) in complex with ENAH EVH1 with resolution limits up to 1.45 Å. While two of these ligands, Ac–WPPPPTEDEL–NH_2 and Ac–[2-Cl-Phe]PPPPTEDEL–NH_2 (**2d**, Tab. 9.13) resembled the proline-rich core of the **wt2** peptide, the third ligand **1c** contained the scaffold designated to mimic the additional interactions of ligand **2b**. The N-terminally truncated **wt** 10-mer Ac–^{1}FPPPPTEDEL–NH_2 could not be crystallized, similar to a previous failed attempt by Fedorov and coworkers.[144] Instead, ^{1}Phe had to be replaced by Trp or 2-Cl-Phe to increase the affinity.[37,141]

The crystal structures consistently revealed that the TEDEL pentamer adopted a short α-helical loop conformation exposing acidic residues to the solvent and returning with the terminal ^{9}EL on the protein surface in close vicinity to the binding groove (Fig. 12.3b). Consistent with the helical conformation found in the crystal structures, the secondary structure prediction server PSIPRED[184,185] calculates helix propensity for amino acids succeeding Thr for all four[1] proline-rich repeats[186,187] of ActA.

[1] ActA repeats $^{265/300}$FPPPPTDEEL, ^{335}FPPPPTEDEL, and ^{380}FPPIPTEEEL

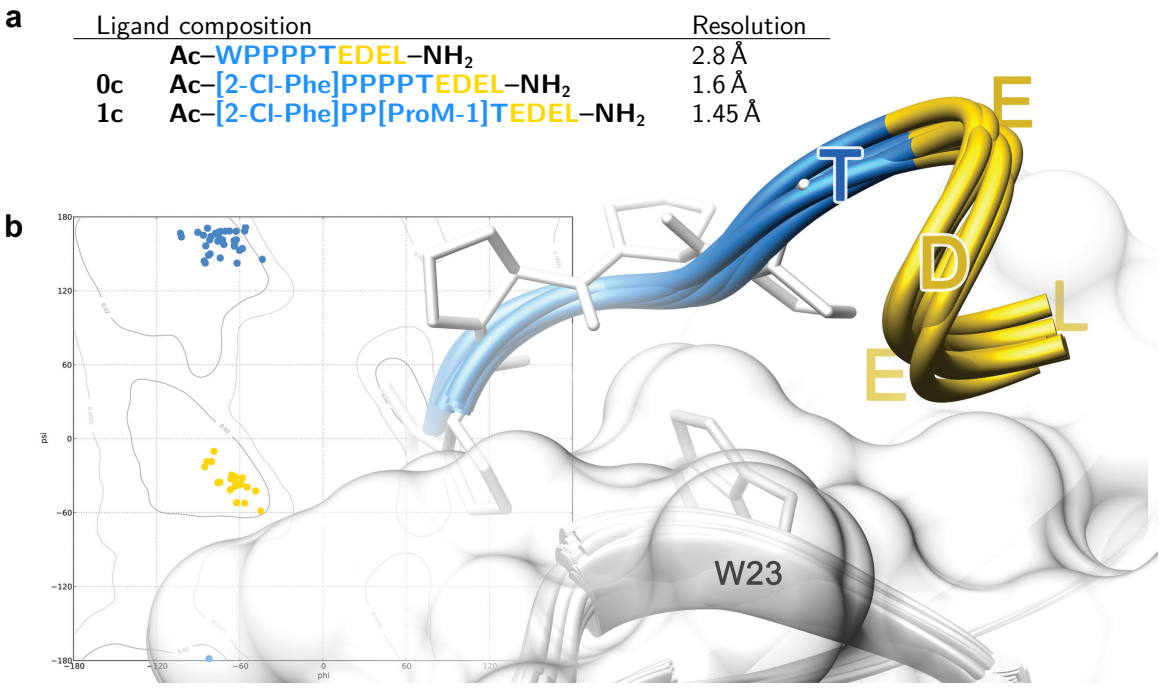
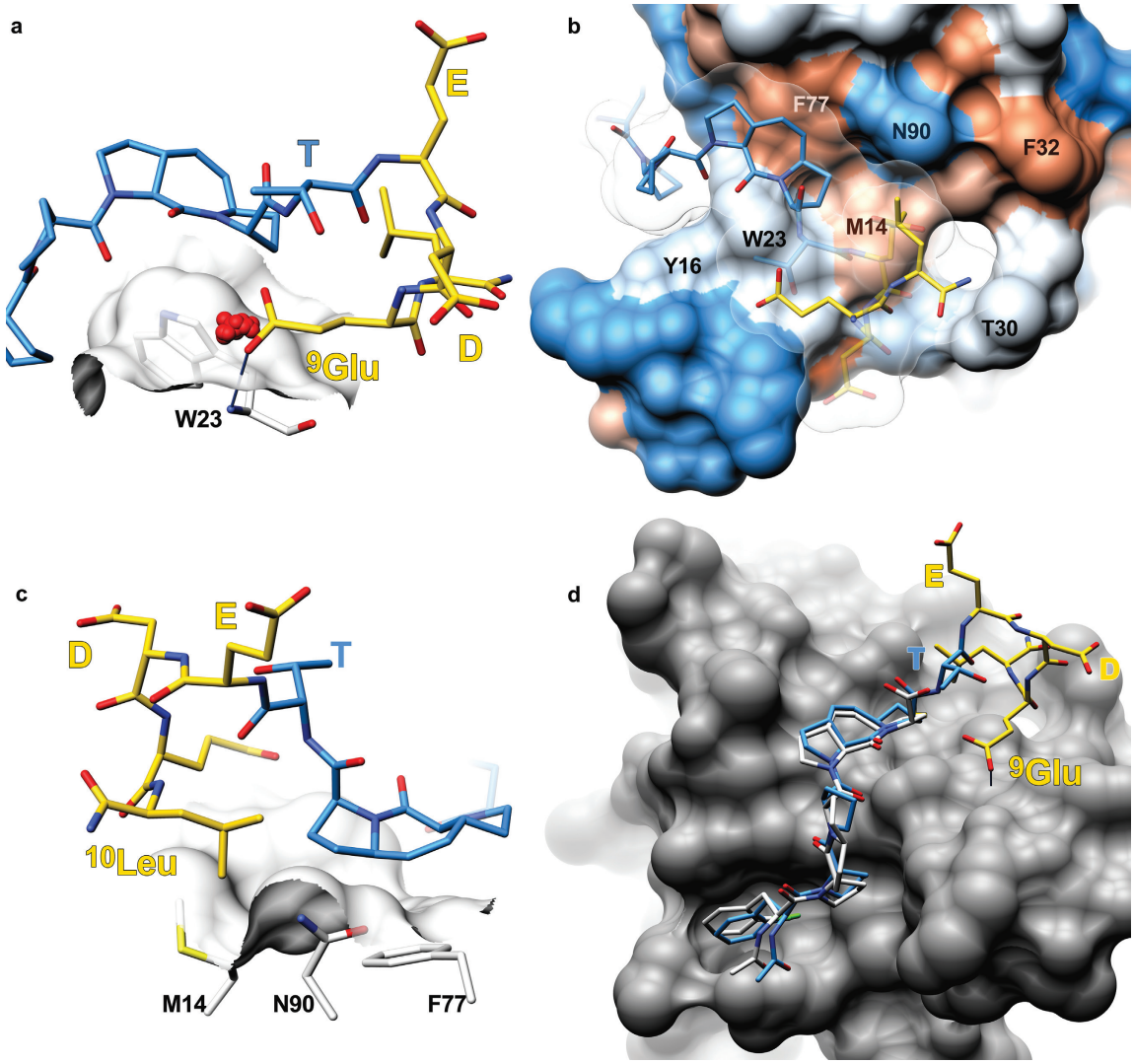


Fig. 12.3.: **Structural determination of the second epitope of *L. monocytogenes'* surface protein ActA**
(**a**) Three TEDEL-containing 10-mers cocrystallized with ENAH EVH1. (**b**) The ^{2}PPPPT and ^{2}PP[ProM-1]T sequence of the 10-mer adopts PPII conformation (blue), whereas the ^{7}EDEL forms an α-helical loop (yellow). Only the two terminal ^{9}EL residues contacted the protein while the rest of the structured C-terminal loop was solvent-exposed. Superposition of 27 ENAH EVH1 chains of 12 asymmetric units reveals a stable and conserved protein epitope.

^{6}Thr itself is solvent exposed but conserved, emphasizing its function as helix capping motif.[188] However, N-terminal capping of the helix seems important only in the context of the full length ActA protein, as common side chain acceptors, especially Thr and Asp,[188] were not preferred in the SPOT array pattern by any Ena/VASP EVH1 domain (Fig. 10.1). We hypothesized that relatively short 13-mers display a mobility high enough so that ^{9}EL found the interaction sites independent of a helix-stabilizing amino acid at position 6. The helical conformation provided structural basis to explain previously reported affinity measurements. Ball and coworkers propose ^{9}EL as an isolated second binding epitope, whose truncation results in 5.5-fold reduced affinity for VASP EVH1.[106,136,189]

Fig. 12.4.: **TEDEL interacts close to the main binding groove of ENAH EVH1 by a mixture of polar and hydrophobic interactions**
(a) Superposition of inhibitor **1c** and 19 H_2O positions (red spheres) of the ENAH EVH1 chains visible in Fig. 12.3b. In the absence of TEDEL, a highly conserved water molecule bound to the backbone amide of Trp23. The carboxylate group of ^{9}Glu bound Trp23 and displaced the bound water. **(b)** TEDEL–NH_2 contacted an apolar patch (Met14) that extends from the rigidly shaped binding groove (Trp23, Phe77). ProM-1 and TEDEL–NH_2 fully covered this hydrophobic patch. Solvent accessible surface of ENAH EVH1 is color coded by hydrophobicity (blue-white-red for hydrophilic-hydrophobic). For better visibility, ^{6}TED and the solvent accessible surface of inhibitor **1c** (white) are transparent. **(c)** Detailed view of the apolar patch covered by ^{10}Leu. To reach back into the main binding groove (Phe77), a hydrophobic amino acid longer than alanine is needed. **(d)** Superposition of the parent inhibitor **2a** (white) and the peptidic chimera **1c** (blue-yellow) bound to ENAH EVH1. View along the main binding groove reveals no steric clashes caused by ProM-2 (N-terminal scaffold in **2a**) and only a minor rearrangement of ProM-1 in the presence of TEDEL–NH_2. These rearrangements are insignificant as the relative contributions to ΔG are additive (Example 3 in section 16.3).


a
E
T
D
9Glu
W23
b
F77
N90
F32
W23
M14
Y16
T30
c
E
D
T
10Leu
M14
N90
F77
d
E
T
D
9Glu

The helical loop is stabilized by the the carboxylate group of ^{9}Glu that formed a hydrogen bond to the backbone amide of Trp23 (W23NH) and displaced a highly conserved water molecule found in the absence of TEDEL (Fig. 12.4a). That conformation suggested that the side chain length of ^{9}Glu was needed to reach W23NH. Consequently, mutation from ^{9}Glu to ^{9}Asp decreased the binding affinity three times (Tab. 9.13 ligand **0h**). The polar interaction site of ^{9}Glu is separated by 8 Å from a non-polar patch that contacts the terminal ^{10}Leu. The patch is caused by Met14 and extends towards the hydrophobic peptide-binding groove, where Phe77 contacts ^{5}Pro with high specificity (Figs. 12.4b and 10.1). Due to the helix pitch of TEDEL, a long hydrophobic amino acid like leucine is needed to reach back into the shallow binding groove (Fig. 12.4c). Consequently, a mutation from ^{10}Leu to ^{10}Ala lost up to three times binding affinity (Tab. 9.13 ligand **0i**). The ligand side chains ^{5}Pro and ^{10}Leu shielded the hydrophobic hot-spot of the protein surface entirely from the solvent (Fig. 12.4b).

12.2.2. Previously reported interaction site on VASP EVH1

Based on ^{1}H-^{15}N-HSQC perturbations in the range of 0.1-0.15 ppm for two subsequent amino acids Gln31 and Ala32 (VASP EVH1 numbering), Ball and coworkers modeled a different ligand conformation onto the solution NMR structure of VASP EVH1 (PDB code 1JNG).[106] The two mentioned amino acids are part of a loop rather far from the binding groove (Fig. 12.5) and shifted only with ligands containing the terminal ^{9}EL.[106] To explain the ^{1}H-^{15}N-HSQC perturbations of this loop region, the ligand was modeled in an entirely stretched conformation in order to reach Ala32 with the terminal ^{9}EL. This model has many weaknesses, and I personally do not believe that TEDEL can adopt a conformation so close to the epitope without being bound by strong, directed interactions. Most disturbingly, the model neither explains why truncation of ^{9}EL reduces the affinity, nor reflects the outcome of the SPOT array experiments, where ^{6}TED, but mostly ^{7}ED were exchangeable by any amino acid. Instead, the ligand conformation predicts polar interactions with ^{7}Glu. The model seems to be docked without water molecules, as the crucial solvation of the ligand backbone was ignored, and several carbonyl and amide groups point towards hydrophobic patches of VASP EVH1.

Interestingly, the TEDEL conformation found in the crystal structures interacted with the mentioned loop region Gln31 and Ala32 as well. Compared to VASP-, ENAH EVH1 is mutated in this loop. Of the according amino acids in ENAH EVH1, Thr30 could not be assigned, but Gly31 shifted 0.1-0.13 ppm as well[2] (Fig. 10.3). Compared to the theoretical model with stretched conformation, not ^{9}Glu interacted with this epitope, but the amidated C-terminus of ^{10}Leu (Fig. 12.5). The interaction of the Leu–NH_2 group with the secondary hydroxyl group of Thr30 is not present every complex structure, indicating that the measured perturbations of the loop might be an artifact of the C-terminal amidation. Cocrystallization of an even longer peptide would shed light on the binding mode at this position.

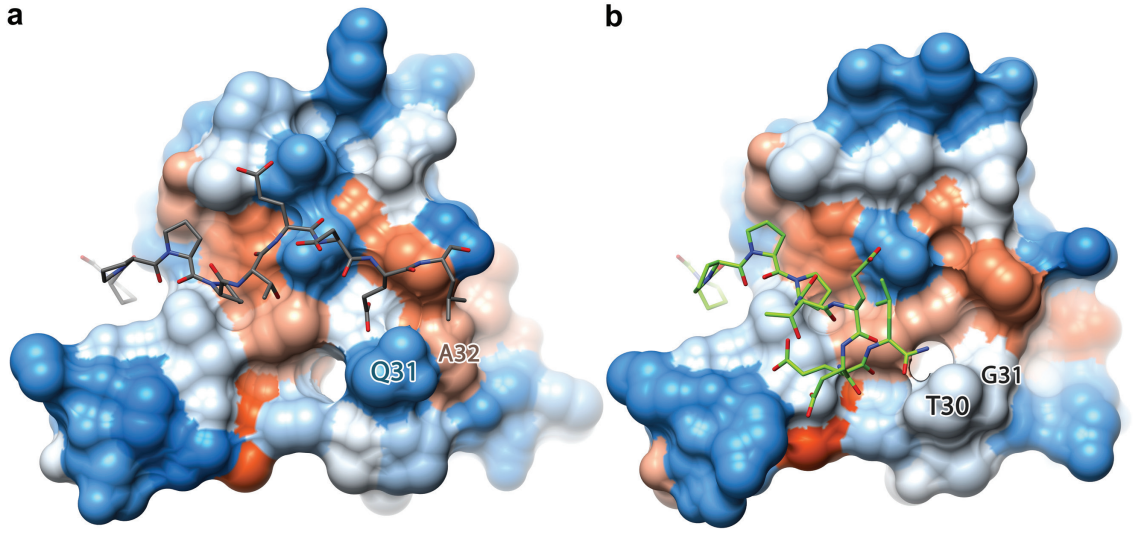


Fig. 12.5.: **Comparison with a TEDEL model in stretched conformation**
(**a**) Theoretical model of Ac–SFE^{1}FPPPPTEDEL–NH_2 docked on the solution NMR structure of VASP EVH1 (PDB code 1JNG). Amino acids ^{31}QA showed shifts only if the ligand contained ^{9}EL.[106] (**b**) The according amino acids are mutated in ENAH EVH1. In the crystal structures of ENAH EVH1 (Ac–^{1}WPPPPTEDEL–NH_2, green), not ^{9}Glu of the ligand interacted with this area, but the amidated C-terminus of ^{10}Leu. Solvent accessible surfaces are color coded by hydrophobicity (blue-white-red for hydrophilic-hydrophobic).

[2] A direct comparison is not possible as Ball *et al.* scaled the ^{15}N shifts with 0.2, while the perturbations reported here use 0.1 as scale factor (Sec. 8.5).

12.2.3. Previously reported interaction site on Homer EVH1

mGluR-derived peptide ^{1}TPPSPF is the only known ligand that interacts close to the hydrophobic patch that binds TEDEL. This ligand is bound by Homer EVH1, a class II EVH1 domain that displays point mutations in both epitopes that bind ^{1}Phe as well as ^{9}EL. The distinct class I core recognition motif ^{1}FPxϕP for Ena/VASP EVH1 is caused by a steep hydrophobic cleft at the N-terminal side of the ligand. In the context of the class II Homer1 EVH1, this cleft is nonexistent due to two mutations (R81A and Y16I), which flatten the epitope and allow TPPSPF to bind C-terminally shifted within the binding groove (Fig. 12.6a).

The second mutated epitope around Phe77 and Met14 is flattened due to an Asn90 to Gly89 mutation (Fig. 12.6b), which allows ^{1}TPPSPF to access the hydrophobic patch with an unique conformation. Thereby ^{3}Pro adopts *cis*-conformation that twists the succeeding ^{4}SP away from the protein surface and induces a type VIa β-turn. The tight turn is needed to bring ^{6}Phe straight down over Gly89 and bury it under ^{4}SP (Fig. 12.6b). The tight turn within the ligand and the conserved Gly89 are crucial to provide enough space for ^{6}Phe to access the patch. [190]

Even though Homer1 EVH1 displays additional mutations, like M14F, superposition of Homer1 and ENAH EVH1 reveals that the Asn to Gly mutation alters the epitope of the hydrophobic patch most severely. Like Gly89 in Homer EVH1, the solvent-exposed Asn90 is conserved among the Ena/VASP EVH1 domains and sterically blocks the hydrophobic patch for ligands with a binding mode seen in ^{1}TPPSPF (Fig. 12.6b). As the volume occupied by ^{6}Phe is blocked on Ena/VASP EVH1 by Asn90, the hydrophobic patch is reachable only from the other side via W23NH. Figure 12.6b highlights that ^{10}Leu of TEDEL and ^{6}Phe of TPPSPF bind to the same hydrophobic patch in close vicinity of the main binding groove and not to the loop region as claimed by Ball *et al.*.

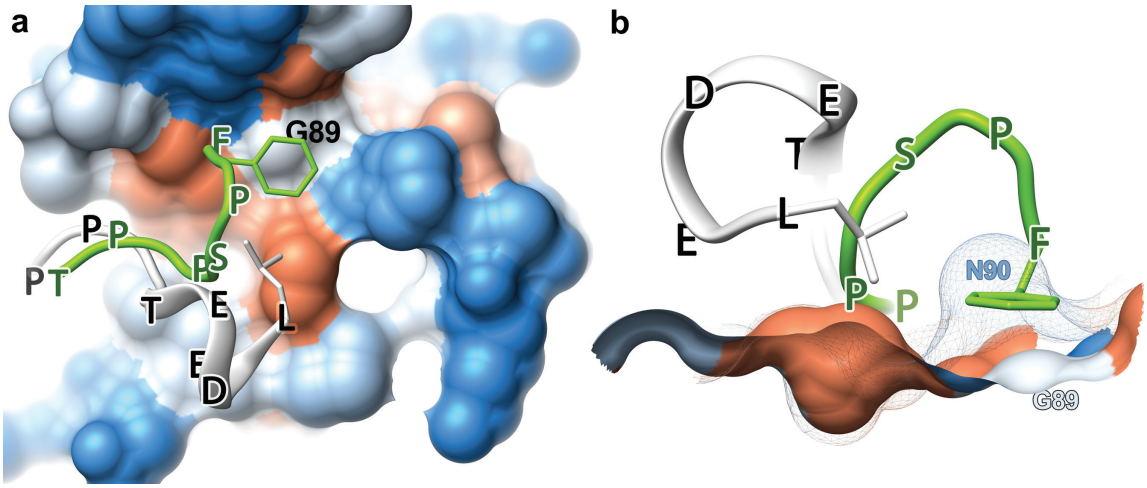
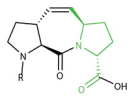


Fig. 12.6.: **Class I ENAH EVH1 and class II Homer1 EVH1 bind their ligands with the same hydrophobic patch**
(**a**) Superposition of ENAH EVH1 in complex with inhibitor **0c** (white) and Homer1 EVH1 bound to TPPSPF (green, PDB code 1DDV). Solvent accessible surface of Homer1 EVH1 is color coded by hydrophobicity (blue-white-red for hydrophilic-hydrophobic). The ligand TPPSPF binds Homer1 EVH1 C-terminally shifted and interacts at the end of the main binding groove with a conserved Gly89. (**b**) The same superposition and coloring but the solvent accessible surface of ENAH EVH1 added as mesh. Like Gly89 in Homer-, Asn90 is conserved in Ena/VASP EVH1. The Gly to Asn mutation sterically blocks the hydrophobic patch (red) for ligands with a binding mode seen in ^{1}TPPSPF. Instead, TEDEL–NH_2 adopted a helical conformation to access the hydrophobic patch form the other side.



12.3. ProM-12 mimics the polar interaction of glutamate

Our first attempt was to mimic the polar interaction between ^{9}Glu and W23NH. Docking studies with crystal structures of ProM-1-containing ligands suggested that the most simplistic shortcut towards W23NH and the bound water was by a new scaffold ProM-12 with inverted chirality at the C_α of the C-terminal pyrrolidine ring (Fig. 6.2). This modification used the carboxyl group itself to either contact or compete with the bound H_2O (Fig. 12.7a). ProM-1 in inhibitor **2** was replaced by ProM-12 and the new inhibitor analyzed as two derivatives, (i) with free C-terminus (**5a**) mimicking the negative charge of ^{9}Glu and (ii) in a cell-membrane-permeable version with non-hydrolyzable methyl ester (OMe, **5c**).

Both inhibitors **5a** and **5c** showed a significant loss in affinity relative to the parent inhibitors **2a** and **2c** (Tabs. 9.14 and 9.15). Superposition with the peptidic chimera **1c** revealed that the exit vector of ProM-12 pointed towards the interaction site of ^{9}Glu as intended, however the inhibitor did not rearrange within the binding groove to interact with the bound water (Fig. 12.7a). Superposition of three ENAH EVH1 crystal structures with bound inhibitors **5a**, **5c** and **2a** showed that ProM-12 displayed a prominently puckered *trans*-5-vinylproline that was lifted away from the protein surface to minimize steric clashes between the ProM-12 carboxylate oxygen and the underlying Trp23 side chain (Fig. 12.7b).

Figure 12.7b shows that the carboxylate group of ProM-12 might be too rigid to mold the scaffold over Trp23. Docking studies suggested that a reduction of the C-terminus of ProM-12 to a hydroxyl group would still not replace the bound water. However, the thereby generated methanol moiety would generate a non-hydrolyzable, cell-membrane-permeable compound. The idea should be investigated from a pharmacological point of view anyway. The inhibitors, even with masked C-terminus, display an incredible solubility up to several hundred millimolar in normal salt conditions. Increasing the hydrophobicity would increase plasma binding and decrease the excretion via the kidneys.

In the context of ProM-12 though, the methanol group is too short and would not even be able to contact the bound water. Another chemically feasible approach[3]

[3]personal communication with Prof. Dr. H.-G. Schmalz

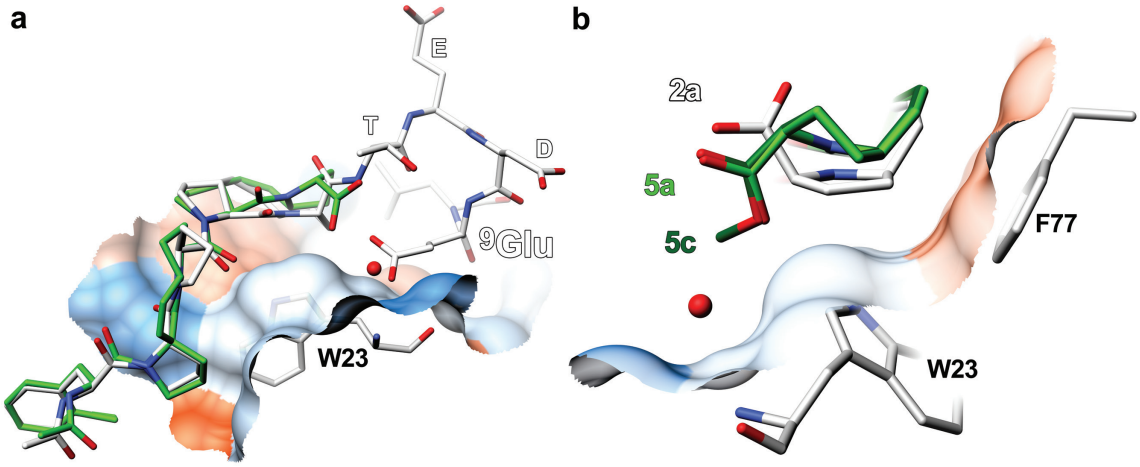
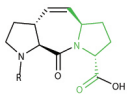


Fig. 12.7.: ***In silico* designed scaffold ProM-12 to mimic the polar interaction of ^{9}Glu**
(a) Superposition of the peptidic chimera **1c** (white) and inhibitor **3a** (green). ProM-12 did not interact with the the bound water molecule (red sphere) as intended. **(b)** Superposition of ligands **5a** and **5c** (greens) with the parent ligand **2a** (white) reveals the distorted orientation of ProM-12 to minimize the clash of the carboxylate oxygen with Trp23, lifting the scaffold away from the epitope.

would be the attachment of a diazole ring. Docking studies showed that especially 1,3-diazole (imidazole) would provide the distance needed to contact W23NH via a 2.9 Å long H-bond (Fig. 12.8). Masking the C-terminus of ProM-12 with imidazole has not been investigated either as the modification relays on a sp^2-hybridized C-terminus that appears to be too rigid. Hence, we concluded that the interaction site of ^{9}Glu was unreachable for the C-terminus of ProM-12 and accessible only with an elongated polar group. However, a synthesis strategy based on educts other than (*L*)-proline is currently out of reach. We decided therefore to postpone optimization of ProM-12 and to seek instead for modifications that address the hydrophobic patch contacted by the terminal ^{10}Leu.



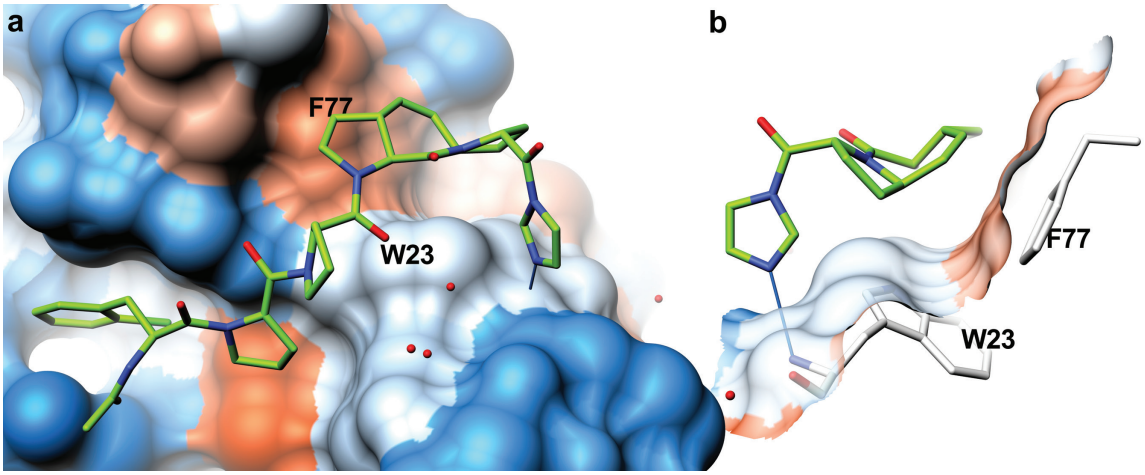
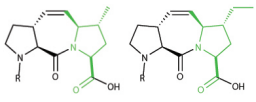


Fig. 12.8.: **Docking study of imidazole masked ProM-12 to reach W23NH with a 2.9 Å long H-bond**
(**a**) Docking of a virtual ProM-12 composition based on the 1.1 Å resolution crystal structure of ENAH EVH1 in complex with Ac–[2-Cl-Phe]PP[ProM-1]–OH. Of different moieties tested to reach W23NH, imidazole seemed the most simplistic and chemically feasible masking. (**b**) As in the *in silico* design of ProM-12, the docking studies do not predict the massive puckering of the pyrrolidine ring visible in both ProM-12 crystal structures (Fig. 12.7b). Solvent accessible surface of ENAH EVH1 is color coded by hydrophobicity (blue-white-red for hydrophilic-hydrophobic).



12.4. ProM-9 and ProM-13 mimic the hydrophobic interaction of leucine

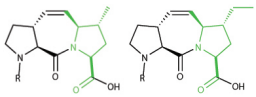
For the attempt to increase the hydrophobic interaction of ProM-1 with Phe77 and Met14, we used docking studies and retrosynthetic analysis, which revealed that enantioselective alkylation of the C_γ of the *cis*-5-vinylproline (Fig. 6.2) would be the most promising and synthetically feasible strategy. Methylated and ethylated derivatives of ProM-1 yielded in two new inhibitors **6** (Ac–[2-Cl-Phe][ProM-2][ProM-9]–OEt) and **7** (Ac–[2-Cl-Phe][ProM-2][ProM-13]–OEt), respectively. FT and ITC measurements showed that both scaffolds outperformed ProM-1 substantially and lowered the K_d 10 times (inhibitor **6**) or even 20 times (inhibitor **7**) to all three Ena/VASP EVH1 domains (Tab. 9.18). Noteworthy, inhibitor **7** regained the affinity of the peptidic chimera **2b**. Incorporation of a single ethyl group with 30 Da restored the entire binding energy of eight amino acids and increased the molecular weight of the inhibitor **2** by only 1%, yielding in a potent 734 Da compound. The alkylated derivatives of inhibitor **2** lead to a continuous increase in the ligand efficiency[183] (LE, ΔG per heavy atom) for **6** and **7**, which justified the increased molecular weight (Tab. 9.18). The apparent saturation in LE from inhibitor **6** to **7** indicated that ProM-13 was nearly fully optimized over the apolar patch and further enlargement by propyl was not investigated.

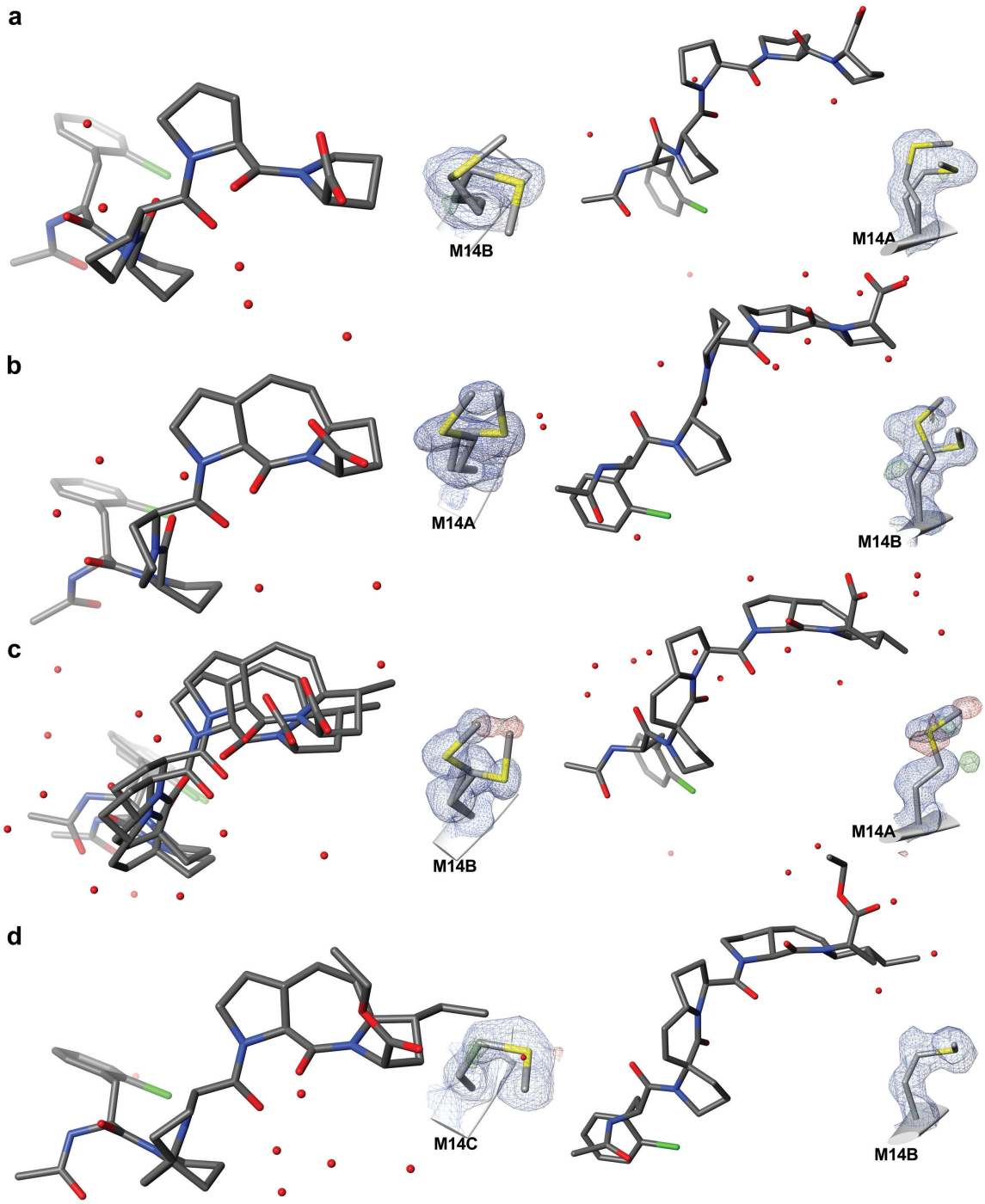
Inhibitor **6** crystallized with free carboxyl group in two variations (**6a** and **6b**, Tab. 9.18), while inhibitor **7** crystallized with the ethyl mask (OEt). The crystal structures of ENAH EVH1 confirmed that the scaffolds shielded the hydrophobic patch as intended by the rational design (Fig. 12.10b). Both ligands **6a** and **7** showed little conformational bias and marginal displacements relative to inhibitors **1c** or **2a** (Figs. 12.10a and 12.10b). The incorporation of ^{4}ProM-9 created no conformational restrictions for the underlying Met14, which adapted two rotamer conformations as in structures with bound ligands containing ^{4}PP or ^{4}ProM-1 (Fig. 12.9). These double conformations were significantly less dominant in the complex structure of **7**, suggesting that the ethyl moiety of ProM-13 contacted Met14.

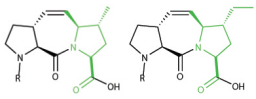
In agreement with the crystal structure, ITC data show a lower entropic change for inhibitor **7** compared to **6** for all and three Ena/VASP EVH1 domains, and shed light on how sensibly EVH1 domains react to ligand modifications (Tab. 16.5). Comparing

the free energy change upon binding, table 16.5 shows that the binding strength of the hybrid ligand **2b** was fully restored.

Fig. 12.9.: **Met14 adopted no longer double conformations with ProM-13**
Comparison of Met14 side chain conformations within the asymmetric unit depending on the ligand composition. Each panel (**a**) to (**d**) shows Met14 of two ENAH EVH1 chains (M14A, B, or C) with the according ligand (dark gray) bound nearby. Met14 adopted several rotamer conformations with the bound (**a**) inhibitor **0a** containing ^{4}PP, (**b**) inhibitor **1a** containing ProM-1 or (**c**) inhibitor **6a** containing ProM-9. Modeling only one Met14 conformation (right panel) produced notable difference electron density. (**d**) ProM-13 of inhibitor **7** restricted the conformational freedom of Met14. Even though a second conformation was slightly visible (difference density in left panel), Met14 adopted mainly one conformation. Feature-enhanced (1 σ, blue) and mF0-DFc difference (3 σ, red/green) maps are calculated at (**a**) 1.49 Å, (**b**) 1.09 Å, (**c**) 1.02 Å, (**d**) 1.63 Å.




a
M14B
M14A
b
M14A
M14B
c
M14B
M14A
d
M14C
M14B



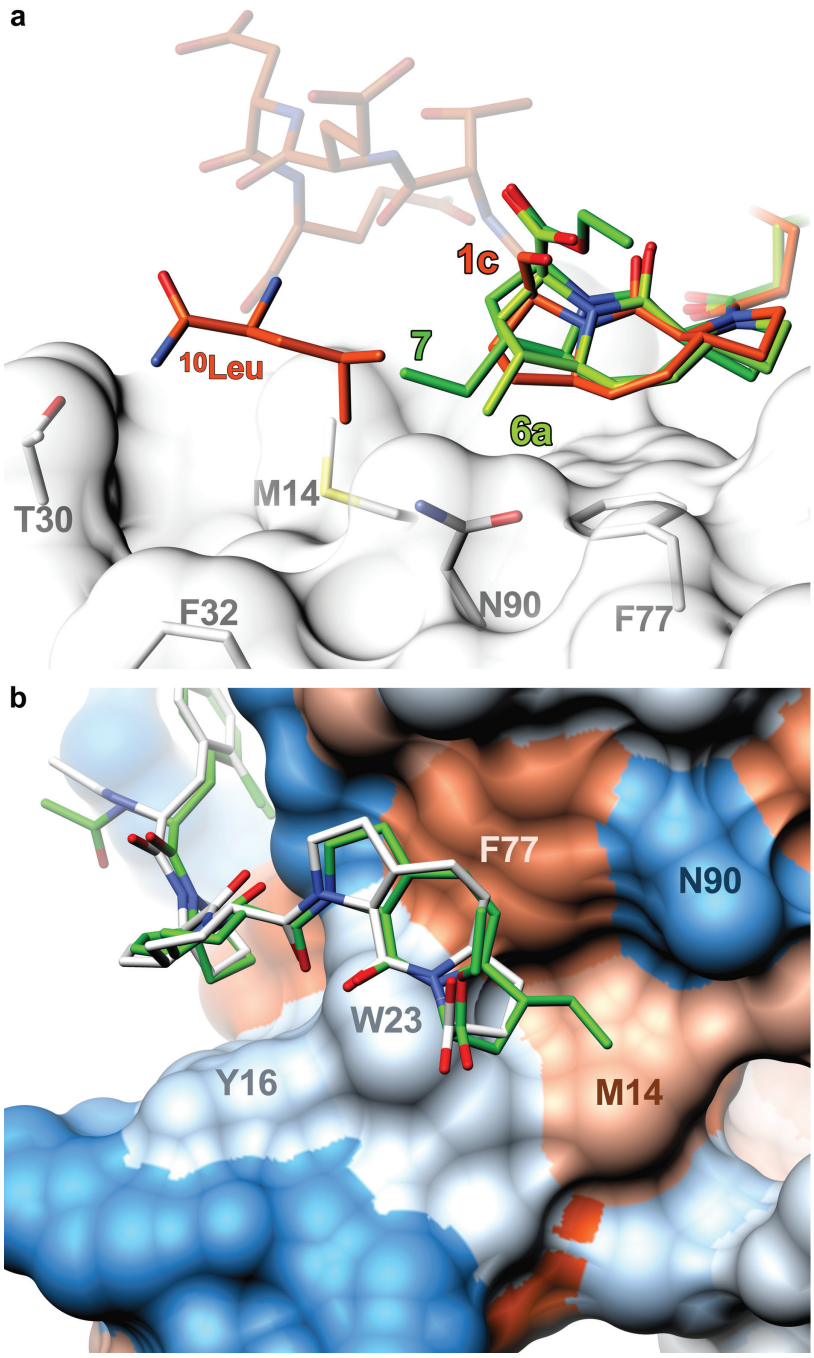


Fig. 12.10.: ***In silico* designed scaffolds ProM-9 and ProM-13 to mimic the hydrophobic interaction of ^{10}Leu**
(a) Superpositions of crystal structures of ENAH EVH1 bound to inhibitors **6a** (ProM-9, pale green), **7** (ProM-13, green) and the peptidic chimera **1c** (orange) revealed little conformational bias among the scaffolds. The new scaffolds ProM-9 and ProM-13 enlarged the coverage capacity towards the interaction site of the terminal ^{10}Leu. **(b)** Superposition of parent inhibitor **2a** (white) and its optimization **7** (green) confirmed successful coverage of the apolar patch M14 and unrestrained orientation of **5**. Solvent accessible surface of ENAH EVH1 is color coded by hydrophobicity (blue-white-red for hydrophilic-hydrophobic).

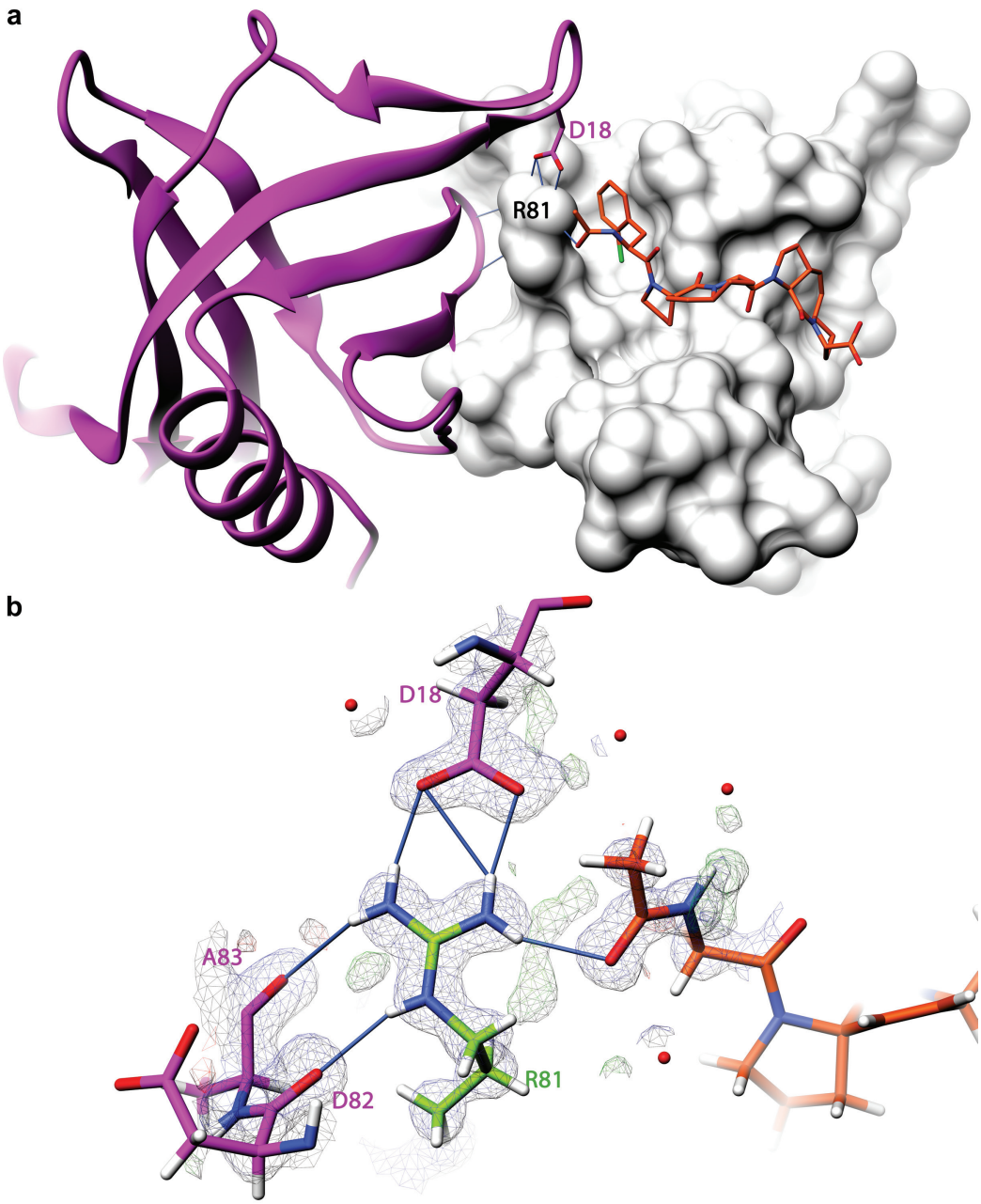
13. Packing artifacts of ENAH EVH1

This chapter describes two unreported packing artifacts that were found in many crystal structures. Especially interesting was the finding that at least ENAH EVH1 showed a putative second binding pocket that bound proline di-peptides over a conserved arginine residue.

13.1. The binding pocket for ^{1}Phe caused by Arg81

The conformation of Arg81 could be biased by a crystal packing artifact that presumes a deep pocket for the bound ligand. Arg81 established a conserved salt bride to the carboxylate group of Asp18 of the neighboring chain in many ENAH EVH1 crystal structures (Fig. 13.1a). The relative orientation of the Arg81 guanidinium group depended on the distances to Asp18 and Asp82 of the neighboring chain. In the majority of analyzed crystal contacts (19 out of 27), Arg81 adopted primary a "side on"[191] conformation to the backbone carbonyl of Asp82 (Fig. 13.1b). Only in the absence of the neighboring Asp82, $N^{\varepsilon}H$ of Arg81 bound the acetylated ligand (as seen in Fig. 14.1). This hydrogen bond pattern between Arg81 and the acetylated ligand was most prominently visible in crystal contacts where Asp18 was too far away to contact Arg81 "end on" with three H-bonds as seen in figure 13.1b. In many structures however, Arg81 adopted a mixture of both mentioned conformations. This caused difference electron densities on the guanidinium group and the acetyl group of the ligand (Fig. 13.1b). The high mobility of Arg81 is in agreement with the crystal structure of EVL EVH1[144] and the NMR solution structure of VASP EVH1,[106] both visible in Fig. 10.2, which supported the hypothesis that Arg81 is presumably more flexible if not involved in the ENAH EVH1-typical salt bridge.

Fig. 13.1.: **The packing artifact of ENAH EVH1 Arg81**
(**a**) Depiction of a single ENAH EVH1 chain (solvent accessible surface in white) could create the illusion of an upright conformation of Arg81 and deep binding pocket for the N-terminal side of the ligand (orange). However, in the crystallographic packing, Arg81 was sandwiched between a neighboring loop (violet) and was involved in an intermolecular "end on"[191] salt bridge to a neighboring Asp18. (**b**) Beside the salt bridge to Asp18, Arg81 was stabilized "side on" either by the carbonyl groups of Asp82 or the acetyl group of the ligand. Clearly visible by the difference electron density map, Arg81 adopted both conformations which offset the acetyl group of the ligand as well. 2mF0-DFc composite (1 σ, blue) and mF0-DFc difference (3 σ, red/green) maps are calculated at 1.0 Å. Hydrogen bonds drawn in blue.


a
D18
R81
b
D18
A83
D82
R81

13.2. Ena/VASP EVH1 has a putative second binding site

Several crystal structures revealed difference electron density over the β-sheet on the back face of the EVH1 domain (Figs. 13.2 and 13.3a). The second binding site was visible initially in low resolution structures, where the difference electron density extended only over 2-3 residues. Neither with feature-enhanced maps[165] calculated at moderate resolution, the ligand orientation could be assigned (SI Fig. 18.2). Only the 1.1 Å resolution complex structure of Ac–[2-Cl-Phe]PP[ProM-1]–OH revealed clear difference signal to place the ligand in correct orientation over Arg47 (Fig. 13.2).

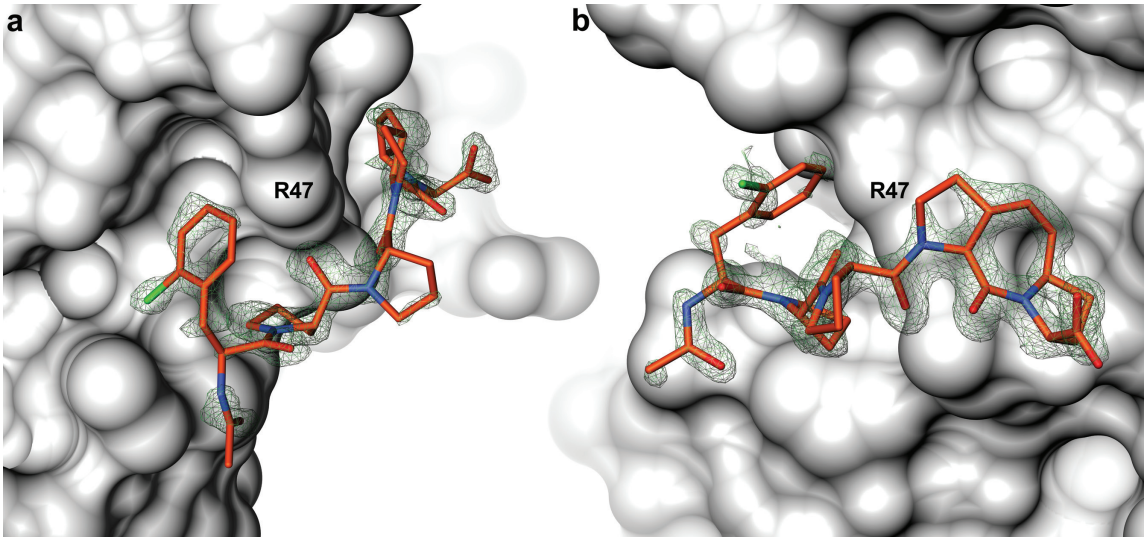


Fig. 13.2.: **OMIT map of ligands bound to the second binding site**
OMIT mF0-DFc difference map calculated at 1.1 Å for the complex structure of ENAH EVH1 bound to Ac–[2-Cl-Phe]PP[ProM-1]–OH, contoured at 3 σ (red/green). The asymmetric unit is composed of two chains, both ligands are shown in the panels. **(a)** The OMIT difference map revealed features of 2-Cl-Phe and the acetyl group on one side of Arg47. **(b)** On the other side of Arg47, unambiguous features of the scaffold like the vinylidene electrons and the carboxylate group were visible.

Interestingly, the relative orientation of both bound ligands agreed with a "wrapping" mechanism reported already for the interactions of N-WASP EVH1/WIP[99] and RanBP EVH1/Ran.[193] Of both mentioned EVH1 domains, the binding mode

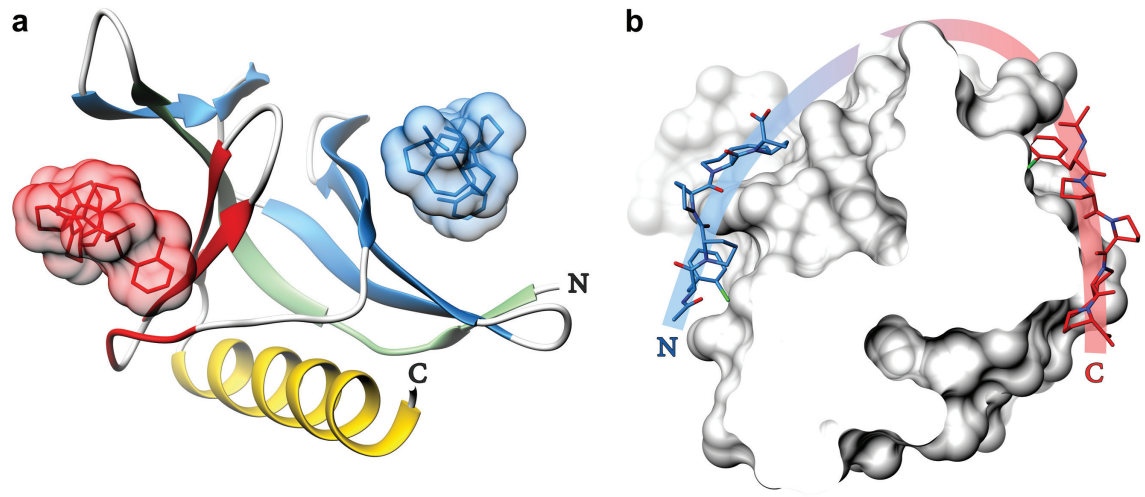


Fig. 13.3.: **ENAH EVH1 binds ligands with both β-sheets**
(a) ENAH EVH1 bound ligands on the canonical β-sheet (red) as well as on the opposite side (blue). Depicted is ENAH EVH1 as ribbon with two bound inhibitors Ac–[2-Cl-Phe]PP[ProM-1]–OH with solvent accessible surfaces. **(b)** The sliced solvent accessible surface of ENAH EVH1 (white) shows both inhibitors (non-canonical blue, canonical red). The matched backbone alignment allows the binding of two successive PRS repeats as found in ActA,[37] vinculin,[192] or zyxin.[76] Extended binding modes over both β-sheets are known already for the N-WASP EVH1 and RanBP EVH1 domains.[99,193]

of N-WASP strongly resembles the interaction found in ENAH EVH1. WIP binds to N-WASP as a linear 25 residue polypeptide in vicinity to the second binding site. Mutations of N-WASP resulting in a severe Wiskott-Aldrich syndrome phenotype suggest a cluster of acidic residues that contacts WIP on the back face of N-WASP (red patches in Fig. 13.4).[99] Within this cluster, Arg76 is the most frequently mutated residue found in the patient population. Every possible single base change missense mutation (Arg to Cys, His, Leu or Pro) has been found.[99] There is strong evidence that mutations of Arg76 impair the direct interaction with WIP.[194] The binding to that crucial Arg was seen the first time in our crystal structures, where ENAH EVH1 Arg47 bound the second inhibitor (Fig. 13.4b) exactly over the acidic cluster. The presence of Arg at this position seems a functional feature, as it is conserved in all three Ena/VASP EVH1 paralogs as well as in both WASP and N-WASP.[194]

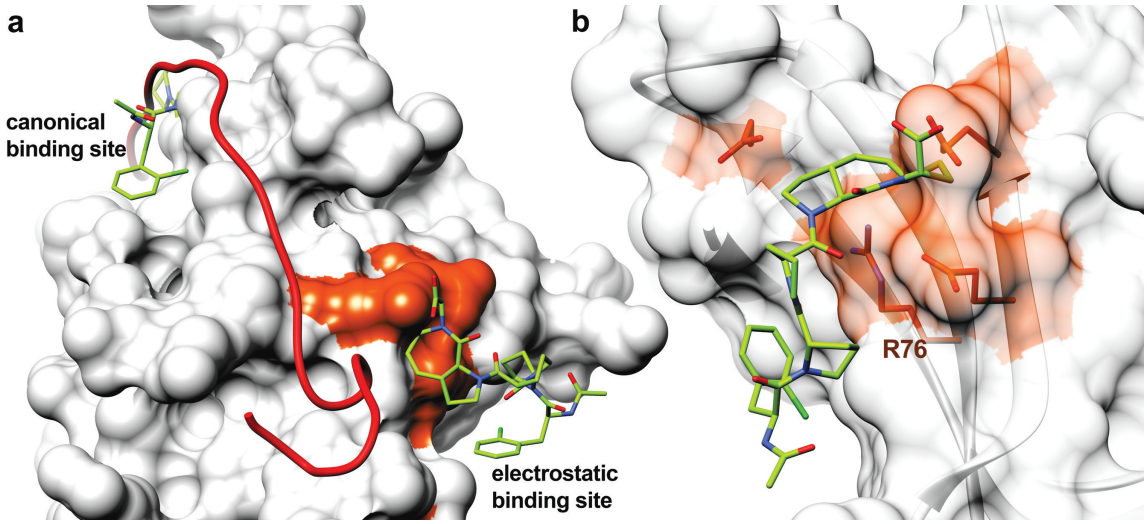


Fig. 13.4.: **N-WASP EVH1 binds WIP at the second binding site**
(a) Solvent-accessible surface of N-WASP EVH1 bound to the 25-mer peptide WIP (red ribbon, PDB code 1MKE). WIP binds to the canonical proline-rich binding site and to a cluster of negatively charged residues on the back face (red patches). Superposition of ENAH EVH1 (chain not shown) reveals that Ac–[2-Cl-Phe]PP[ProM-1]–OH (green) bound to the electrostatic binding site. **(b)** Detailed view of the electrostatic binding site. In central position, Arg76 is postulated to coordinate the acidic residues that bind the conserved basic motif of WIP.[99] Superposition of the EVH1 folds highlights that Ac–[2-Cl-Phe]PP[ProM-1]–OH was bound at the same position by Arg47.

The epitope of ENAH EVH1 around Arg47 showed no similarity to the main binding groove. Most prominent feature and sole interaction site was a polar cluster composed of conserved Glu3, Arg47 and Asn61 which contacted the carbonyl groups of two subsequent prolines (Fig. 13.5a). ^{1}H-^{15}N-HMQC measurements showed perturbation of Glu3, Arg47, and Asn61 backbones only with heavy molar excess of the ActA **wt** 13-mer (Fig. 13.5b). The perturbations in this area presumed that the binding was not an artifact of the ProM scaffolds and estimated an affinity between 1-2 mM (Glu3 and Arg47 backbones) and 4-7 mM (Asn61 side chain and Asn43 backbone). The complete list of calculated affinities can be found in the SI table 20.2.

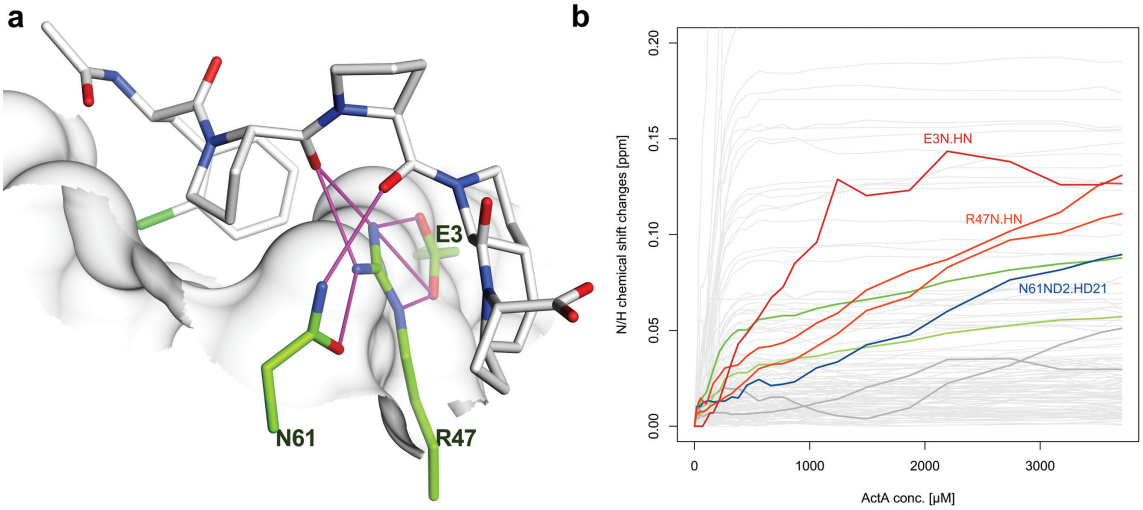


Fig. 13.5.: **^{1}H-^{15}N-HSQC perturbation of the second binding site**
(**a**) The triad Glu3-Arg47-Asn61 binding a di-peptide in PPII conformation to the second binding site of ENAH EVH1. Hydrogen bonds drawn in magenta. (**b**) ^{1}H-^{15}N-HMQC titration experiment with 200 µM ENAH EVH1 and the ActA-derived 13-mer **wt** (Tab. 20.1) revealed perturbations of the triad Glu3-Arg47-Asn61. The perturbations of this epitope are highlighted and suggest an affinity in the lowest micromolar range. Shifts of all remaining assigned peaks are grayed out. Peak shifts that saturated already with 500 µM **wt** belonged to the canonic binding sites.

The weak affinity of the second binding site was supported by isothermal calorimetry measurements of Ac–[2-Cl-Phe][ProM-2][ProM-9]–OH, which was also found in the second binding site of ENAH EVH1. The binding isotherm was fitted with a one-to-one model and the standardized residuals showed normal distribution (Fig. 9.1),

suggesting that the titrated ligand bound to only one site. We concluded that for *in vitro* experiments with monomeric EVH1 domains and near-equimolar ligand concentrations, the second binding site was presumably not significantly occupied and was consequently addressed as packing artifact. The situation might differ in a cellular context where tetrameric Ena/VASP recognizes target proteins that harbor several binding motifs.

14. Additional hot spots not addressed yet

Beside the water bound to W23NH, the complex structures revealed another structurally well ordered water in vicinity of the ligands between the backbones of Lys69 and Arg81 at the position where the strands β5 and 6 of the underlying β-sheet open up (Figs. 14.1 and 3.1). Difference electron density of this water was visible in all structures right after molecular replacement. As discussed in section 13.1, the upright conformation of Arg81 might be a packing artifact that deeply buried the water between sidechains Lys69 and Arg81, giving the impression of a water-filled tunnel (Fig. 14.2). In solution however, the water molecule is presumably much easier accessible from the ligand side than when docked with an ENAH EVH1 crystal structure.

For such docking studies, I used the water network of the 1.01 Å resolution structure of inhibitor **6a** to keep the ligand in position. Docking of chemically feasible substitutions[1] revealed that an ethanol moiety reached the water best and provided the hydrogen bond acceptor and donor properties needed between the β-sheet (Fig. 14.2a). The flexibility of Arg81 would allow to clamp its guanidinium group with a carboxylate group (Fig. 14.2b). Note that at physiological pH, the carboxylate group is de-protonated and would therefore not satisfy both backbone partners of the formerly bound water that we aim to compete off.

[1] personal communication with M. Klein and Dr. A. Soike

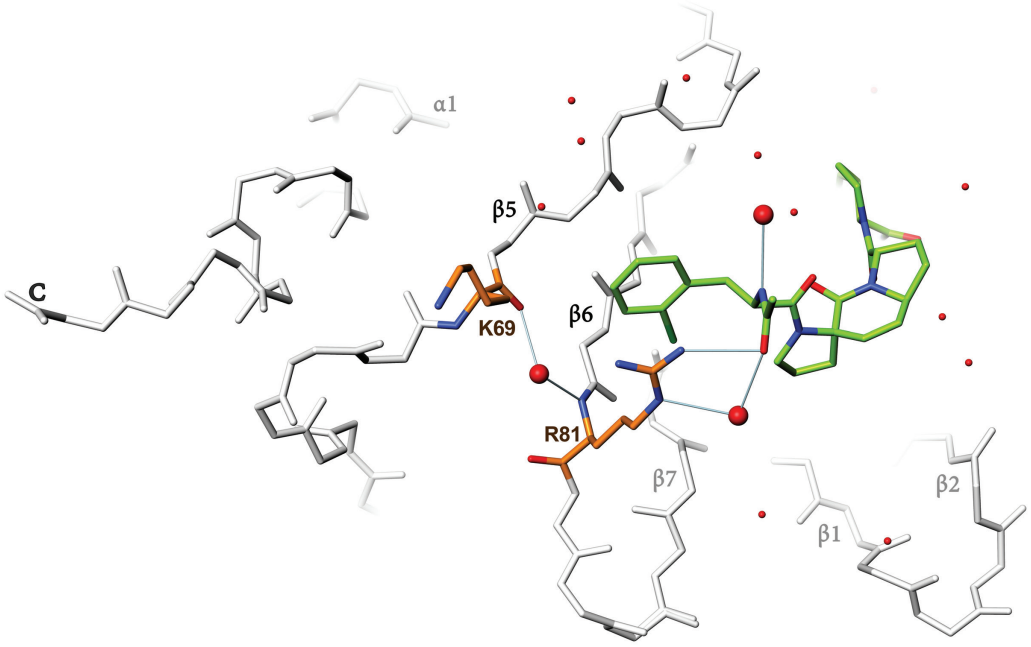


Fig. 14.1.: **A conserved bound water molecule between the backbones of Lys69 and Arg81**
At the N-terminal side of the ligand, three water molecules were visible in most of the structures (bigger red spheres). The most well ordered water was sandwiched between the backbone of Lys69 and Arg81 (orange). At this position, the β-sheet of the main binding groove opens up and an additional water is needed to saturate the backbone H-bond partners of strands β5 and 6. Depicted is Ac–[2-Cl-Phe][ProM-2][ProM-9]–OH (**6a**, green) in complex with ENAH EVH1 (backbone chain trace in white, hydrogen bonds drawn in blue).

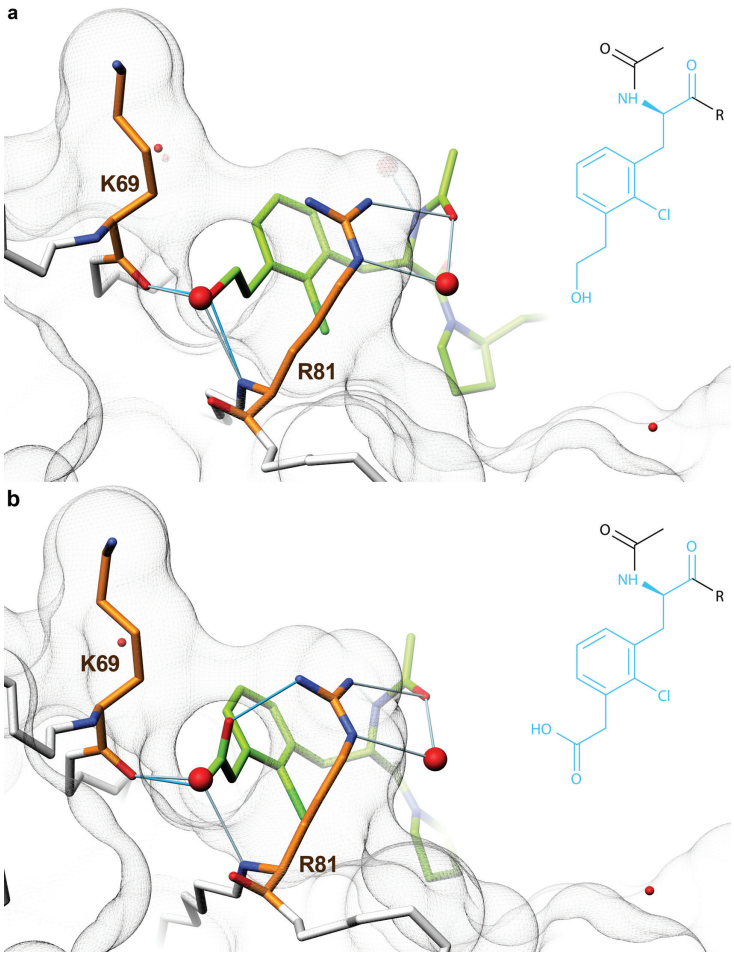


Fig. 14.2.: **Docking study of double-substituted ^{1}Phe**
Docking of a virtual composition based on the 1.01 Å resolution crystal structure of ENAH EVH1 in complex with Ac–[2-Cl-Phe][ProM-2][ProM-9]–OH. (**a**) Of different moieties tested to reach the water molecule, an ethanol finger seemed the most simplistic and chemical feasible solution. (**b**) The Arg81 side chain is presumably much more flexible in solution than suggested by the ENAH EVH crystal structures. This would allow to clamp Arg81 from both sides with a carboxylate group instead of a hydroxyl moiety. Solvent accessible surface of ENAH EVH1 is depicted as transparent mesh, hydrogen bonds drawn in blue.

The only double-substituted derivative of 2-Cl-Phe commercially available was 2-Cl,3-OH-Phe. Docking studies expected no steric clashes but suggested that this moiety was too short to even contact the bound water. Affinity measurements with a modified **wt** 13-mer confirmed that the hydroxyl moiety indeed fit into the void, but no significant gain in affinity resulted from the substitution (Tab. 14.1). Despite extensive trials, this ligand could not be crystallized.

Ac–SFE-(?)-PPPPTEDEL–NH_2		$K_{d,FT}$ [μM]	ΔG [kJ/mol]	$\Delta\Delta G$ [kJ/mol]
wt	Phe	20 (3)	–26.9 (0.3)	(ref)
0b	[2-Cl-Phe]	1.5 (0.3)	–33.3 (0.5)	–6.4 (0.6)
0f	[2-Cl,3-OH-Phe]	1.41 (0.07)	–33.4 (0.1)	–6.5 (0.3)

Tab. 14.1.: **Attempt to contact the bound water in vicinity of 2-Cl-Phe**
Substituting ^{1}Phe of the **wt** 13-mer in *ortho* position with chloride (**0b**) increased ligand affinity drastically. All crystal structures revealed a tightly bound water molecule in vicinity to 2-Cl-Phe. Contacting this water with a hydroxyl group (**0f**) lead to an insignificant increase in binding strength. Two-tailed t-test with α=5%, df=82: t_{crit}=1.99, t_{value}=0.14. Table lists affinities as K_d and ΔG of ligands measured by fluorescence titration using VASP EVH1. Values in parentheses are standard errors.

Part V.

Conclusion

15. Structure-optimized ProMs pave the way for *in vivo* studies with Ena/VASP as antimetastatic target

Interfering with protein-protein interactions remains a highly challenging way to target diseases. In the context of the deadly attribute of cancer, current approaches of metastasis-related drugs are not improving satisfactorily. We propose Ena/VASP as a novel antimetastatic target and designed *in silico* scaffolds that mimic the pentameric peptide sequence responsible for the recruitment of Ena/VASP. The initially found combination of proline-mimicking scaffolds, coined ProM, yielded in a 706 Da, cell-membrane-permeable inhibitor. However, the moderate affinity of the initial inhibitor will restrict the validation in future *in vivo* studies. The modular synthesis strategy of ProM scaffolds allowed optimization of the scaffold that significantly boosted the affinity, while conserving its structural simplicity, low-molecular weight and pharmacological properties. By addressing additional epitopes found in crystal structures of elongated peptides, we optimized the scaffold and boosted the affinity of the initial inhibitor nearly 20-fold. The optimized drug will show the same effects *in vivo* as the initial inhibitor at lower concentrations than 100 μM, thereby paving the way to evaluate Ena/VASP as a novel antimetastatic target.

Part VI.

Acknowledgment

Firstly, I would like to express my sincere gratitude to my advisor Dr. Ronald Kühne for the continuous support of my PhD study, for his patience and trust in my scientific work. His knowledge, encouragement, and intuition was of great help in many situations during the research and writing of this thesis.

Besides my advisor, I would like to thank Prof. Dr. Udo Heinemann for his guidance, support and insightful comments. His second opinion as crystallographer and outside observer was of utmost importance during the project and the writing of the manuscript.

I am immensely grateful to Dr. Yvette Roske who influenced my scientific development the most. Her physical presence, practical guidance and precious support as well as her critical questions during countless fruitful discussions are the reason for my ongoing curiosity in all aspects of macromolecular crystallography.

My sincere thanks go to all actual and former members of our collaborating team. Without their respectful way to interact and support each other, the interdisciplinary project would have been unthinkable. I have great respect for all involved scientists that conducted a piece of research to this drug development project. Every part of the iterative optimization cycle is associated with an enormous effort and asks for experts in their field. The complexity of the project was to communicate new insights in a fruitful way so that knowledge is passed on into the other fields to create there new ideas.

In this context, I am hugely indebted by Prof. Dr. Hans-Günther Schmalz. Without his expertise and the persistence of the chemists in his group, new scaffolds would have remained virtual compounds and never found the way into living organisms.

Part VII.

Supplementary Informations

Abstract of the SI

The SI of the thesis is divided into three main sections. The first section 16 contains tables to access the data from binding studies. The structure-affinity relationships found in the results part are calculated from the following tables and are referred to in the discussion. While the diffraction and refinement statistics of deposited structures are found in the results part, sections 17 and 18 of the SI report in detail expression, purification, crystallization, data processing, and refinement of each model. Solution NMR studies were used to gain insights into the binding mode of VASP EVH1, which could not be crystallized. ^{1}H-^{15}N-HSQC titration experiments were further performed to corroborate additional interaction sites and exclude crystal packing artifacts. The shift perturbation tables are printed in section 20 and include Euclidean distances of roughly 0.1 ppm and further.

16. Affinity measurements of ligands binding to EVH1 domains

In vitro ligand optimization studies are guided by the decreasing molecular weight and increasing affinity of the compound. Hence, the direct approach is to normalize the affinity by the molecular weight, or by the number of heavy atoms building up the molecule.[183] Optimization is then directed towards that composition that shows a significant increased affinity relative to a reference ligand.

16.1. Choosing a suitable measure for affinity

The classical measure of affinity is the dissociation constant K_d. With R being the universal gas constant, T the absolute temperature and K_d being normalized to one mol/l, the change of free molar reaction enthalpy is calculated to:

$$\Delta G = R\,T\,\ln\left(\frac{K_\mathrm{d}}{1\,\mathrm{M}}\right) \tag{16.1}$$

As the free energy ΔG of a reaction is additive, meaning that all m sub-processes sum up to the measured enthalpy,

$$\Delta G = \sum_{i=1}^{m} \Delta G_i \tag{16.2}$$

the over-all measured dissociation constant is calculated in a multiplicative way from its sub-processes:

$$K_\mathrm{d} = 1\,\mathrm{M}\prod_{i=1}^{m}\exp\left[\frac{\Delta G_i}{R\,T}\right] = 1\,\mathrm{M}\prod_{i=1}^{m}\left(\frac{K_{\mathrm{D},i}}{1\,\mathrm{M}}\right) \tag{16.3}$$

Because of this multiplicity of logarithms, two dissociation constants are compared as ratios and their standard deviations are expected to be log-normal distributed. Hence, the K_d carries two very impractical properties that are not suitable to guide ligand optimization: A comparison of more than two ligands based on their dissociation constants is impractical and most importantly does not allow to determine whether the ratios differ significantly.

Instead of the K_d, another measure is used. As already seen in equation (eq.) (16.2), ΔG is additive, and comparing the free molar reaction enthalpies to a reference ligand is highly desirable. From basic considerations, ΔG is assumed to be much closer normal distributed than the K_d. The difference of two changes in free energies,

$$\Delta\Delta G \equiv \Delta G_\mathrm{Lig} - \Delta G_\mathrm{Ref} \tag{16.4}$$

is also normal distributed. Whether such a difference is significant under a certain degree of error (5%) can be computed with a t-test.

16.1.1. Methods to measure ligand affinity

The dissociation constant can be extracted from titration experiments, in which the protein is loaded with an increasing amount of ligand. The increase in ligand-bound protein as a function of totally available ligand is proportional the underlying binding affinity. In the next part we will deduce the formulas used to extract the dissociation constant from titration experiments, extensively discussed in the thesis of Dr. Robert Opitz.[105]

We assume the simplified case in which the protein possesses one binding site for the ligand (one-to-one model).

$$\mathrm{P} + \mathrm{L} \underset{k_\mathrm{off}}{\overset{k_\mathrm{on}}{\rightleftharpoons}} \mathrm{PL} \tag{16.5}$$

Once reached equilibrium, and assuming that neither of the components is processed during the reaction, the law of mass action holds. If the components are ideally diluted, their corresponding activity a can be substituted by the concentration and the dissociation constant K_d computed to:

$$K_\mathrm{d} \equiv \frac{[a_\mathrm{P}]\,[a_\mathrm{L}]}{[a_\mathrm{PL}]} \approx \frac{[\mathrm{P}]\,[\mathrm{L}]}{[\mathrm{PL}]} \tag{16.6}$$

In the equation above, [P] and [L] can be substituted. As the total protein and ligand concentrations are distributed onto a free and bound portion,

$$\begin{aligned} \mathrm{P} &= \mathrm{P_{tot}} - \mathrm{PL} \\ \mathrm{L} &= \mathrm{L_{tot}} - \mathrm{PL} \end{aligned} \tag{16.7}$$

eq. 16.6 is re-written to the law of mass action in a quadratic form:

$$0 = [\mathrm{PL}]^2 - [\mathrm{PL}]\,(\mathrm{P_{tot}} + \mathrm{L_{tot}} + K_\mathrm{d}) + \mathrm{P_{tot}}\,\mathrm{L_{tot}} \tag{16.8}$$

The solution is computed with the quadratic formula

$$[\mathrm{PL}] = -\frac{b}{2} \pm \sqrt{\left(\frac{b}{2}\right)^2 - a}$$

with $b = -(\mathrm{P_{tot}} + \mathrm{L_{tot}} + K_\mathrm{d})$ and $a = \mathrm{P_{tot}}\,\mathrm{L_{tot}}$. From logical considerations, $K_\mathrm{d} > 0$ and $[\mathrm{PL}] < \min(\mathrm{P_{tot}}, \mathrm{L_{tot}})$ only the following root is physically possible:

$$[\mathrm{PL}] = \frac{1}{2}\left(\mathrm{P_{tot}} + \mathrm{L_{tot}} + K_\mathrm{d} - \sqrt{(\mathrm{P_{tot}} + \mathrm{L_{tot}} + K_\mathrm{d})^2 - 4\,\mathrm{P_{tot}}\mathrm{L_{tot}}}\right) \tag{16.9}$$

16.2. Affinity tables

As discussed in Sec. 16.1, the following tables contain both informations, the release of free energy ΔG upon binding as well as the commonly more known dissociation constant K_d. Relationships between two ligand affinities, most importantly the difference in binding energy ($\Delta\Delta G$), are computed from Tab. 16.2, read directly in Tabs. 16.3 and 16.4 and are discussed in Sec. 16.3.

C7S,C64S-VASP EVH1 + Ligand					$K_{d,\,FT}$ [µM]	ΔG [kJ/mol]
Ac–SFE	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	1.3 (0.2)	–33.6 (0.2)
Ac–SFE	[2-CF_3-Phe]	PP	PP	TEDEL–NH_2	1.8 (0.1)	–32.9 (0.1)
Ac–SFE	[2-CH_3-Phe]	PP	PP	TEDEL–NH_2	4.8 (0.6)*	–30.3 (0.3)*
Ac–SFE	[2,6-di-Cl-Phe]	PP	PP	TEDEL–NH_2	1000 (100)*	–17.0 (0.2)*
VASP EVH1 + Ligand					$K_{d,\,FT}$ [µM]	ΔG[kJ/mol]
Ac–SFE	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	1.5 (0.3)	–33.3 (0.5)
Ac–SFE	F	PP	PP	TEDEL–NH_2	20 (3)*	–26.9 (0.3)*
Ac–SFE	[2-Cl,3-OH-Phe]	PP	PP	TEDEL–NH_2	1.41 (0.07)	–33.4 (0.1)

SI Tab. 16.1.: **Fluorescence titration binding studies of substituted Phe** Ligand affinities marked with an asterisks measured by Dr. R. Opitz.

					ENAH EVH1 + Ligand		VASP EVH1 + Ligand	
					$K_{d,\,FT}$ [µM]	ΔG [kJ/mol]	$K_{d,\,FT}$ [µM]	ΔG [kJ/mol]
Ac–SFE	F	PP	PP	TEDEL–NH_2	13 (1)*	–27.9 (0.1)*	19 (2)*	–26.9 (0.3)*
Ac–SFE	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	0.79 (0.05)	–34.8 (0.2)	1.5 (0.3)*	–33.3 (0.5)*
Ac–SFE	[2-Cl-Phe]	PP	PP	TEDDL–NH_2	1.43 (0.09)	–33.4 (0.1)		
Ac–	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	1.25 (0.04)	–33.7 (0.07)		
Ac–	[2-Cl-Phe]	PP	PP	TEDDL–NH_2	3.8 (0.2)	–30.9 (0.1)		
Ac–	[2-Cl-Phe]	PP	PP	TEDEA–NH_2	3.5 (0.1)	–31.4 (0.1)		
Ac–	[2-Cl-Phe]	PP	[ProM-1]	TEDEL–NH_2	0.60 (0.04)	–35.5 (0.2)		
Ac–SFE	[2-Cl-Phe]	PP	PP	–OH	4.6 (0.2)	–30.5 (0.1)		
Ac–SFE	F	[ProM-2]	PP	TEDEL–NH_2	12 (1)*	–28.2 (0.3)*		
Ac–SFE	F	[ProM-1]	PP	TEDEL–NH_2	8 (1)*	–29.1 (0.3)*	40 (4)*	–25.1 (0.2)*
Ac–SFE	F	PP	[ProM-1]	TEDEL–NH_2			5.7 (0.3)	–29.9 (0.1)
Ac–SFE	F	[ProM-1]	[ProM-1]	TEDEL–NH_2			19 (6)*	–26.9 (0.8)*
Ac–	F	PP	PP	–OH	500 (70)*	–19 (0.4)*	800 (80)*	–17.7 (0.2)*
Ac–	F	PP	PP	–OEt	153 (8)	–21.8 (0.1)		
Ac–	[2-Cl-Phe]	PP	PP	–OH	9.9 (0.8)	–28.6 (0.2)	13 (2)	–27.8 (0.3)
Ac–	[2-Cl-Phe]	PP	[ProM-1]	–OH	2.4 (0.4)	–32.0 (0.4)	3.2 (0.3)	–31.4 (0.3)
Ac–	[2-Cl-Phe]	[ProM-2]	[ProM-1]	–OH	2.3 (0.2)	–32.2 (0.2)	2.3 (0.1)	–32.2 (0.2)
Ac–	[2-Cl-Phe]	[ProM-2]	[ProM-1]	–OEt	4.1 (0.3)*	–30.8 (0.2)*	6.3 (0.6)*	–29.7 (0.3)*
Ac–	[2-Cl-Phe]	[ProM-2]	[ProM-4]	–OH	3.0 (0.3)	–31.5 (0.2)	1.6 (0.2)	–33.1 (0.4)
Ac–	[2-Cl-Phe]	[ProM-2]	[ProM-3]	–OH	9.7 (0.8)	–28.6 (0.2)		
Ac–SFE	[2-Cl-Phe]	PP	[ProM-1]	TEDEL–NH_2	0.21 (0.05)	–38.1 (0.6)	0.31 (0.05)*	–37.2 (0.4)*
Ac–SFE	[2-Cl-Phe]	[ProM-2]	[ProM-1]	TEDEL–NH_2	0.15 (0.06)*	–39 (1)*	0.25 (0.07)*	–37.7 (0.7)*
Ac–SFE	[2-Cl-Phe]	PP	[ProM-3]	TEDEL–NH_2			1.5 (0.2)*	–33.2 (0.3)*
Ac–SFE	[2-Cl-Phe]	[ProM-1]	[ProM-1]	TEDEL–NH_2			0.3 (0.1)	–37.1 (0.8)
Ac–	[2-Cl-Phe]	PP	[ProM-9]	–OH	0.23 (0.05)	–37.9 (0.5)	0.38 (0.09)	–36.7 (0.5)
Ac–	[2-Cl-Phe]	[ProM-2]	[ProM-9]	–OH	0.12 (0.02)	–39.5 (0.3)	0.28 (0.06)	–37.4 (0.5)
Ac–	[2-Cl-Phe]	[ProM-2]	[ProM-9]	–OEt	0.38 (0.05)*	–36.6 (0.3)*	0.78 (0.09)*	–34.9 (0.3)*
Ac–	[2-Cl-Phe]	[ProM-2]	[ProM-13]	–OEt	0.18 (0.03)*	–38.5 (0.4)*	0.4 (0.1)*	–36.6 (0.6)*

SI Tab. 16.2.: **Fluorescence titration binding studies of ENAH- (left) and VASP EVH1 (right)** Ligand affinities marked with an asterisks measured by Dr. R. Opitz.

VASP EVH1 + reference ligand					→	VASP EVH1 + substituted ligand					$\Delta\Delta G_{FT}$ [kJ/mol]
Ac–	F	PP	PP	–OH	→	Ac–	[2-Cl-Phe]	PP	PP	–OH	–10 (0.4)
Ac–	F	PP	PP	TEDEL–NH_2	→	Ac–	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	–6.5 (0.3)
Ac–SFE	F	PP	PP	TEDEL–NH_2	→	Ac–SFE	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	–6.4 (0.6)
Ac–SFE	F	PP	[ProM-1]	TEDEL–NH_2	→	Ac–SFE	[2-Cl-Phe]	PP	[ProM-1]	TEDEL–NH_2	–7.3 (0.4)
Ac–SFE	F	[ProM-2]	[ProM-1]	TEDEL–NH_2	→	Ac–SFE	[2-Cl-Phe]	[ProM-2]	[ProM-1]	TEDEL–NH_2	–6.9 (0.7)
Ac–	[2-Cl-Phe]	PP	PP	–OH	→	Ac–	[2-Cl-Phe]	PP	[ProM-1]	–OH	–3.6 (0.4)
Ac–SFE	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	→	Ac–SFE	[2-Cl-Phe]	PP	[ProM-1]	TEDEL–NH_2	–3.9 (0.6)
Ac–SFE	F	PP	PP	TEDEL–NH_2	→	Ac–SFE	F	PP	[ProM-1]	TEDEL–NH_2	–3.0 (0.3)
Ac–	F	PP	PP	–OH	→	Ac–	[2-Cl-Phe]	PP	[ProM-1]	–OH	–13.6 (0.5)
Ac–SFE	F	PP	PP	TEDEL–NH_2	→	Ac–SFE	[2-Cl-Phe]	PP	[ProM-1]	TEDEL–NH_2	–10.3 (0.5)
Ac–	[2-Cl-Phe]	PP	[ProM-1]	–OH	→	Ac–	[2-Cl-Phe]	[ProM-2]	[ProM-1]	–OH	–0.4 (0.6)
Ac–	[2-Cl-Phe]	PP	[ProM-9]	–OH	→	Ac–	[2-Cl-Phe]	[ProM-2]	[ProM-9]	–OH	–0.7 (0.7)

SI Tab. 16.3.: **Incorporation of a building block and its effect on ΔG when binding to VASP EVH1**
Every line represents a comparison of two ligands with the difference of ΔG ($\Delta\Delta G$) derived from fluorescence titration (FT) or CVM titration experiments and calculated relative to the reference ligand. The correlations hold to a great extend for all EVH1 domains, printed here are examples with VASP EVH1.

ENAH EVH1 + reference ligand					→	ENAH EVH1 + substituted ligand					$\Delta\Delta G_{FT}$ [kJ/mol]
Ac–	F	PP	PP	–OH	→	Ac–	[2-Cl-Phe]	PP	PP	–OH	–10.3 (0.5)
Ac–SFE	F	PP	PP	TEDEL–NH_2	→	Ac–SFE	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	–6.9 (0.2)
Ac–	[2-Cl-Phe]	PP	PP	–OH	→	Ac–	[2-Cl-Phe]	PP	[ProM-1]	–OH	–3.4 (0.4)
Ac–	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	→	Ac–	[2-Cl-Phe]	PP	[ProM-1]	TEDEL–NH_2	–1.8 (0.2)
Ac–SFE	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	→	Ac–SFE	[2-Cl-Phe]	PP	[ProM-1]	TEDEL–NH_2	–3.3 (0.6)
Ac–SFE	F	PP	PP	TEDEL–NH_2	→	Ac–SFE	[2-Cl-Phe]	PP	[ProM-1]	TEDEL–NH_2	–10.2 (0.6)
Ac–SFE	F	PP	PP	TEDEL–NH_2	→	Ac–SFE	F	[ProM-2]	PP	TEDEL–NH_2	–0.3 (0.3)
Ac–	[2-Cl-Phe]	PP	PP	–OH	→	Ac–	[2-Cl-Phe]	PP	PP	TEDEL–NH_2	–5.1 (0.2)
Ac–	[2-Cl-Phe]	PP	[ProM-1]	–OH	→	Ac–	[2-Cl-Phe]	PP	[ProM-1]	TEDEL–NH_2	–3.5 (0.4)

SI Tab. 16.4.: **Incorporation of a building block and its effect on ΔG when binding to ENAH EVH1**
Every line represents a comparison of two ligands with the difference of ΔG ($\Delta\Delta G$) derived from fluorescence titration (FT) or CVM titration experiments and calculated relative to the reference ligand. The correlations hold to a great extend for all EVH1 domains, printed here are examples with ENAH EVH1.

	2b	**2**	**6**	**7**
VASP EVH1				
ΔG [kJ/mol]	–35.7 (0.2)	–28.7 (0.6)	–33.5 (0.3)	–36.0 (0.3)
ΔH [kJ/mol]	–27.5 (0.3)	–20.6 (0.1)	–21.6 (0.3)	–25.8 (0.4)
$-T\,\Delta S$ [kJ/mol]	–8.2 (0.5)	–8.1 (0.5)	–11.9 (0.6)	–10.2 (0.6)
ENAH EVH1				
ΔG [kJ/mol]	–36.9 (0.2)	–29.2 (0.2)	–34.9 (0.2)	–37.2 (0.3)
ΔH [kJ/mol]	–30.9 (0.4)	–34.5 (0.8)	–26.0 (0.3)	–31.1 (0.4)
$-T\,\Delta S$ [kJ/mol]	–6.0 (0.5)	+5 (1)	–9.0 (0.4)	–6.1 (0.6)
EVL EVH1				
ΔG [kJ/mol]	–37.6 (0.05)	–30.0 (0.4)	–35.3 (0.3)	–37.5 (0.3)
ΔH [kJ/mol]	–26.4 (0.4)	–19.8 (0.7)	–20.7 (0.2)	–26.5 (0.4)
$-T\,\Delta S$ [kJ/mol]	–11.2 (0.4)	–10 (1)	–14.6 (0.4)	–10.9 (0.6)

SI Tab. 16.5.: **Detailed listing of the ITC measurements of cell-membrane-permeable inhibitors**
The affinities of ligand compositions with optimized scaffolds are listed column wise, starting with the peptidic chimera **2b** and the parent inhibitor **2**. Listed are enthalpy and entropic energy at 25 °C calculated from affinity measurements for each of the three Ena/VASP EVH1 paralogs.

16.3. Independent epitopes allow to optimize ligand locally

A closer look at the above tables reveals that the ligand consisted of segments that contributed independently to the measured binding affinity. This is not surprising as EVH1 domains display a rather rigid underlying interaction site which recognizes ActA-derived peptides and ProM-containing chimeras with the same binding mode. Furthermore, none of the ligands forms intramolecular interactions, which opened the possibility to observe local affinity-relationships independent of the ligand composition. As visible in table 16.2, the tendencies held to a great extend also among the three EVH1 domains. If two or more scaffolds were combined, their respective contribution to the free energy loss simply added up (Tabs. 16.3 and 16.4). The independent contribution of every moiety to ΔG held for short ligands, flanking epitopes, and peptide-scaffold chimeras:

- Introducing 2-Cl-Phe to Ac–^{1}FPPPP–OH (–10.1 kJ/mol) and subsequent exchange of ^{4}PP by ProM-1 (–3.6 kJ/mol) produced the same total free energy loss as the direct comparison of Ac–FPPPP–OH with Ac–[2-Cl-Phe]PP[ProM-1]–OH (–13.7 kJ/mol, VASP EVH1). These ligands bound canonically to ENAH EVH1 and boosted ΔG separately by –9.6 (2-Cl-Phe) and –3.4 kJ/mol (ProM-1), while the direct comparison yielded –13 kJ/mol as well.
- Incorporating 2-Cl-Phe in Ac–SFE^{1}FPPPPTEDEL–NH_2(–6.4 kJ/mol) and subsequent exchange of ^{4}PP by ProM-1 (–3.9 kJ/mol) summed up to the same overall energy loss as when incorporating ProM-1 first (–3 kJ/mol) followed by exchanging Phe with 2-Cl-Phe (–7.3 kJ/mol). The direct comparison between the wild type ligand and the double-substituted ligand Ac–SFE[2-Cl-Phe]PP-[ProM-1]TEDEL–NH_2 yielded in –10.3 kJ/mol as well (VASP EVH1).
- Subsequent exchange of ^{4}PP by ProM-1 and C-terminal elongation of Ac–[2-Cl-Phe]PPPP–OH by –TEDEL–NH_2 residues (or vice versa) either way added up to totally –6.9 kJ/mol. Direct comparison of Ac–[2-Cl-Phe]PPPP–OH and Ac–[2-Cl-Phe]PP[ProM-1]TEDEL–NH_2 also yielded in $\Delta\Delta G$ of –6.9(0.3) kJ/mol (ENAH EVH1).

Generally speaking, exchange of ^{1}Phe by 2-Cl-Phe and ^{4}PP by ProM-1 added around 6.5 kJ/mol and 3.5 kJ/mol to the free energy loss independently. It has been

noted already in Opitz *et al.*[141] that the introduction of ProM-2 for ^{2}PP does not influence the affinity measured by FT. The neutral effect of ProM-2 on the ΔG balance still held true in inhibitors containing ProM-9 as replacement for ^{4}PP. There is no reason to assume that exchange of ^{2}PP for ProM-2 would affect ΔG differently in ProM-12- and ProM-13-containing ligand composition. These two scaffolds were synthesized only in ProM-2-containing ligand compositions.

17. Detailed expression and purification protocols

The supplementary informations of the purification protocols were written for the interested reader. To reproduce the crystallization experiments, the following sections 17.2 and 17.3 each report the experimental details needed to purify the constructs that crystallized during the thesis.

17.1. Expression protocol of Ena/VASP EVH1 domains

One Shot BL21(DE3) electrocompetent cells were thawed on ice and mixed with 0.2 µl of vector, transformed in 1 mm-slit cuvettes with 1.8 kV and put in prewarmed SOC medium for an hour, followed by plating out a small volume onto a 2×YT agar plate containing 0.1 mg/ml ampicillin and 2% (m/v) glucose. After over night incubation at 37 ℃, all single colonies from the streaked plate were picked with sterilized papers and pooled into 2×YT medium prior to equal distribution into prewarmed 2×YT and 0.3 mg/ml ampicillin. After reaching OD_{600} 0.4-0.5, cell cultures were cooled down on ice to 18 ℃, induced with 0.1 mM IPTG and fresh ampicillin was added to 1.5 mg/ml. Expression took place over night at 18 ℃ in unbaffled flasks filled with not more than one liter. Cells were harvested and pellets slightly re-suspended in concentrated PBS buffer (3xPBS, 5mM EDTA, pH 7.0).

This protocol was only changed in the end of the thesis. For the expression of the VASP $EVH1_{min/9mut}$ constructs, an enriched 2×YT medium with trace metals and vitamins was used.[195,196] Thereby *E. Coli* cultures were grown at 37 ℃, gently spun down and transferred into fresh enriched medium (18 ℃) before induction at OD600

0.1 with 0.1 mM IPTG. Constructs were expressed for up to 60 h at 18 ℃, cells harvested and resuspended in 3xPBS for storage at -20 ℃.

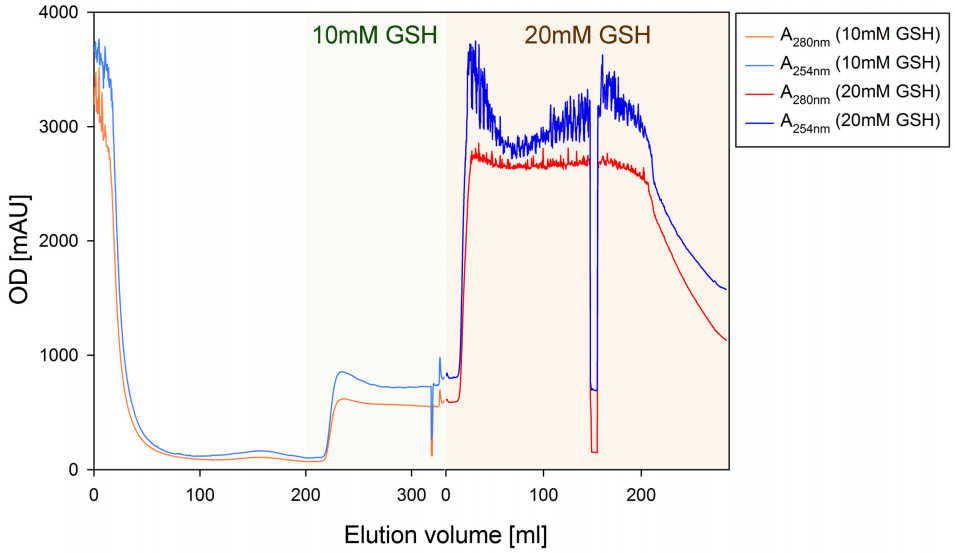
17.2. Purification protocol of ENAH EVH1

Following example shows the purification of ENAH EVH1 which yielded in around 60 mg pure ENAH EVH1. Preceding to the load of any sample into the Äkta purifier, the tubing system, injection valve, F2 outlet and F3 feed tube were washed with both, the elution and running buffers to prevent contamination of the eluting sample. As for the load via the P960 sample pump, its inlet was connected to a syringe mantle or sedimentation funnel. This setup assured an air-free loading of the entire sample onto the affinity columns. Buffers were always made at room temperature (RT), 0.2 µm sterile filtered and finally cooled down to 4 ℃ on ice while blowing in nitrogen gas.

Bacterial pellets from two liters 2×YT medium stored at -20 ℃ were thawed under stirring. 5 µl benzonase, 3 COmplete Tables and $MgCl_2$ was added to final concentration of 1.8 mM. Pellet was resuspended in maximal 40 ml glutathion running buffer. Cells were sonificated for half an hour at 50% intensity and then regularly checked before white foam indicated complete disruption of the cells. Tube was then brought to RT and pH lifted from 6.6 to 7.2 with NaOH, and stirred for half an hour for digestion of DNA by benzonase. Cell lysate was then centrifuged at 13 ℃, 22000 rpm for half an hour (JA25.50 rotor), 90 ml supernatant was passed through 0.2 µm syringe filters and loaded with the P960 sample pump on an equilibrated fast flow glutathion matrix with 1 ml/min at 4 ℃.

Column was washed with 200 ml, 1.5 ml/min. First attempt to elute with 10 mM GSH, 2 mM EDTA, 5 mM β-mercaptoethanol (β-ME), 25 mM NaPi pH 7.0 failed as shown in the chromatogram 17.1. Therefore a new buffer with 20 mM GSH, 2 mM EDTA, 5 mM β-ME, 25 mM NaPi pH 7.0 was made and column eluted with 2 ml/min directly in Schott bottle. The pool of over 250 ml was then prepared for over night cleavage of the GST fusion protein by thrombin. The rule of thumb was to use at least 12 µl thrombin per liter 2×YT grown bacterias or 50 ml eluate. As the expression efficiency was drastically enhanced during the thesis, 100 µl thrombin and additionally 5 mM β-ME were added to the 250 ml eluate. Flask was flushed with nitrogen and

cleavage took place over night in a prewarmed water bath at 27 °C. According to SDS-PAGE, none of the purified batches were cut completely with this method (Fig. 17.2, lane ①).

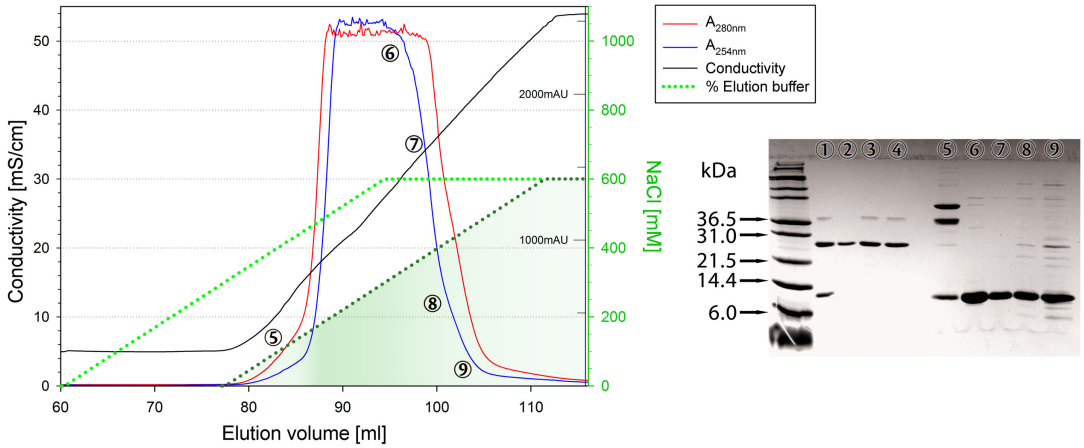


SI Fig. 17.1.: **Wash-out and elution of GST-ENAH EVH1 from glutathion fast flow matrix**
10 mM GSH were not sufficient to detach the GST fusion protein from the column. 20 mM GSH eluted the protein in a large pool from the column. Later purification batches used up to 100 mM GSH to compress the elution volume.

After thrombin digestion, the sample was loaded on a SP sepharose column, equilibrated with 20 mM NaCl, 5 mM β-ME and 25 mM NaPi at pH 7.5 This column was extremely efficient to separate EVH1 from uncut samples, thrombin and GSH as all remaining components exhibited pIs far below the EVH1 domain (pI GSH: 2.8, pI GST 6.1) and did not bind to the matrix. Successful binding of ENAH EVH1 to the matrix was monitored by the conductivity during equilibration and load[1]. The filtered sample was loaded with the P960 sample pump with 2 ml/min. After the column was washed with 3.5 column volumes (CV) running buffer, the bound EVH1 domain was detached from the matrix with a short NaCl gradient over 2 CV with 1 ml/min to 60% elution buffer (1 M NaCl, 5 mM β-ME, 25 mM NaPi pH 7.5). ENAH EVH1 detached with 150 mM NaCl and 16 mS/cm (Fig. 17.2). The pool of 8 ml containing 90 mg of

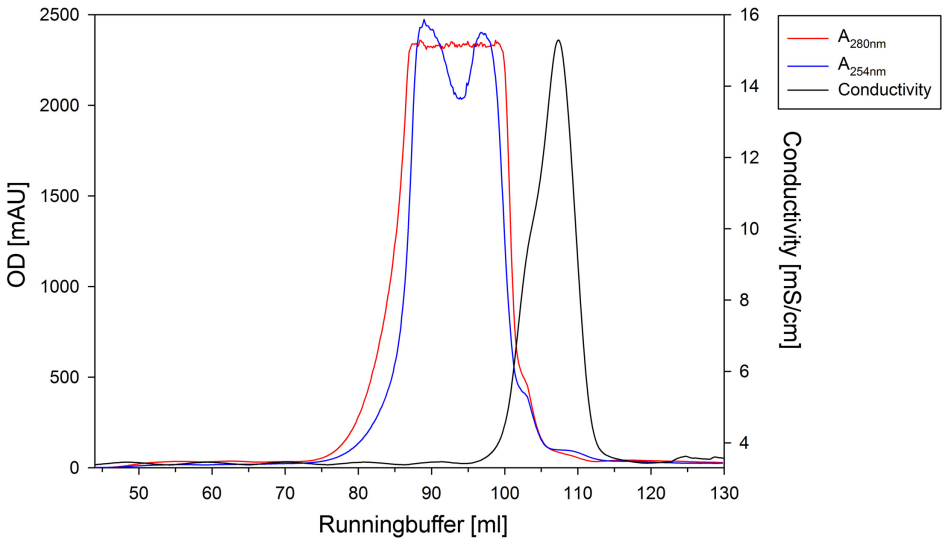
[1]Running buffer displayed typically 5.1 mS/cm, and reached up to 7.5 mS/cm during load

EVH1 domain was reduced to 4 ml with prewashed 20 ml Vivaspin (5 kDa cut-off) at 8000 rcf, 20 °C.



SI Fig. 17.2.: **Elution of ENAH EVH1 from SP sepharose cation exchange matrix** ENAH EVH1 detached with a NaCl gradient from 20 to 620 mM NaCl over 2CV (0-60% elution buffer, green dotted line). ENAH EVH1 detached with around 150 mM NaCl and a conductivity of 16 mS/cm. Right panel shows the according SDS-PAGE: ① loaded sample, ②-④ flow through samples (after 20, 150, and 250 ml), ⑤ shoulder peak start, ⑥-⑦ out of scale peak (middle and end), ⑧ shoulder peak end, ⑨ peak end. The flow through samples contain GST (26 kDa) and the fusion protein GST-ENAH EVH1 (39 kDa). The fusion protein bound more weakly to the matrix and detached in the beginning of the elution.

As final step the sample was loaded with 1 ml/min via superloop on a Superdex 75 size exclusion column, equilibrated with crystallization buffer (20 mM NaCl, 5 mM TCEP, 10 mM HEPES pH 7.0).



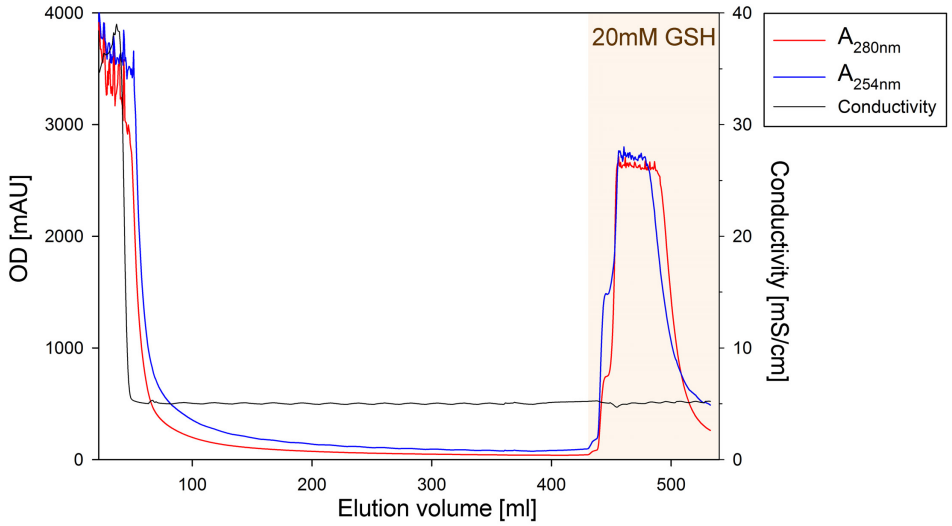
SI Fig. 17.3.: **Chromatogram of ENAH EVH1 passing a Superdex 75**
ENAH EVH1 passed the Superdex 75 size exclusion column shortly before the rise in conductivity indicated the elution of the solutes. The two different fractions suggested by the absorbance at 254 nm could not be confirmed in SDS-PAGE.

17.3. Purification protocol of EVL EVH1 dAT/tail

Initial attempts to purify EVL- and VASP-EVH1 with the above mentioned protocol failed as neither domain bound to the SP sepharose matrix.

The protocol was adapted to facilitate the binding process to the cation exchange matrix and stabilize EVL EVH1 during cleavage with thrombin. The glutathion buffers were accordingly changed to 20 mM NaCl, 5 mM EDTA, 5 mM β-ME, 20 mM NaPi pH 7.0 and 20 mM GSH, 20 mM NaCl, 5 mM β-ME, 20 mM NaPi pH 7.0, which increased the conductivity from 3.4 to 5.1 mS/cm but enabled a rather clean cleavage over night. The GST-fused construct was loaded on the glutathion column, washed with 400 ml, and detached with 20 mM GSH at 1 ml/min. 50 ml were pooled and cleaved with 50 µl thrombin. To ensure reductive environment during the cleavage over night at 27 °C, 10 mM of fresh β-ME was added.

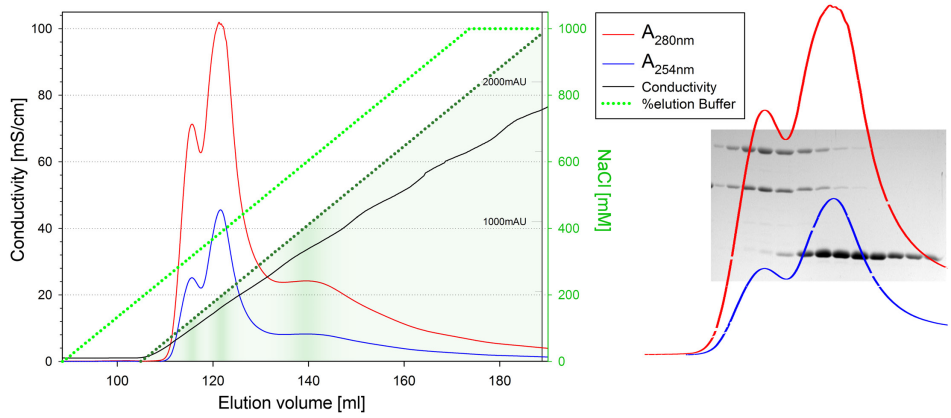
For the SP sepharose, the buffer was adapted to 5 mM β-ME and 10 mM NaPi at pH6.5. The column ran with one unit lower pH and minimal phosphate concentration to maximize the negative charge on the surface of EVL EVH1 while staying above the calculated pI of GST. As the glutathion elution buffer contained totally 60 mM



SI Fig. 17.4.: **Wash-out and elution of GST-EVL EVH1 from glutathion fast flow matrix**
GST-fusion protein was detached with 20 mM GSH and 75 ml of the eluate was pooled.

of charged ions, the sample was diluted under stirring 1:10 (v/v) with ice-cold SP sepharose running buffer. The dilution yielded in an ion concentration less than the equilibrated column and ensured binding to the matrix.

The running buffer itself displayed a conductivity of 1.0 mS/cm and never exceeded 1.5 mS/cm during the load of EVL EVH1. 500 ml were loaded with 5ml/min and washed out with 4.5 CV at 3 ml/min to base line conductivity of 1.0 mS/cm. EVL EVH1 was detached with a gradient over 5 CV, 1 ml/min to 100% elution buffer (1 M NaCl, 5 mM β-ME, 10 mM NaPi pH 6.5). The gradient was not as steep as with ENAH EVH1 (300 mM NaCl/CV) to separate bound GST from EVL EVH1. GST and the fusion protein detached with 120 mM NaCl and 9.8 mS/cm, while the EVH1 domain eluted at around 200 mM NaCl and 16 mS/cm (Fig. 17.5). Note that ENAH EVH1 detached with the same conductivity of 16 mS/cm and similar ionic strength (170 mM NaCl, 25 mM NaPi) but bound to the matrix even in the presence of these ions. The final re-buffering via superdex 75 into the crystallization buffer remained the same for EVL EVH1.



SI Fig. 17.5.: **Elution of EVL EVH1 from SP sepharose cation exchange matrix** EVL EVH1 was detached by NaCl gradient from 0 to 100 mM NaCl over 5 CV (0-100% elution buffer, green dotted line). The same dotted line is re-printed in dark green and shifted 1 CV to see the exact amount of NaCl under the elution profile. EVL EVH1 detached with 200 mM NaCl and a conductivity of 16 mS/cm. Right panel: Fractions of the elution are checked by SDS-PAGE gel. The superposition highlights that GST and GST-fusion protein detached together.

18. ENAH EVH1 complex structures needed for a quantitative structure-affinity relationship study

The supplementary informations for the crystal structures were written for the interested reader. To reproduce the crystallization experiments, the sections 18.1 to 18.15 each report the experimental details needed to reproduce the crystallization experiments of every complex structure. Within these sections, important indicators[156] during space group determination, data reduction, molecular replacement, and refinement for every recorded data set can be found.

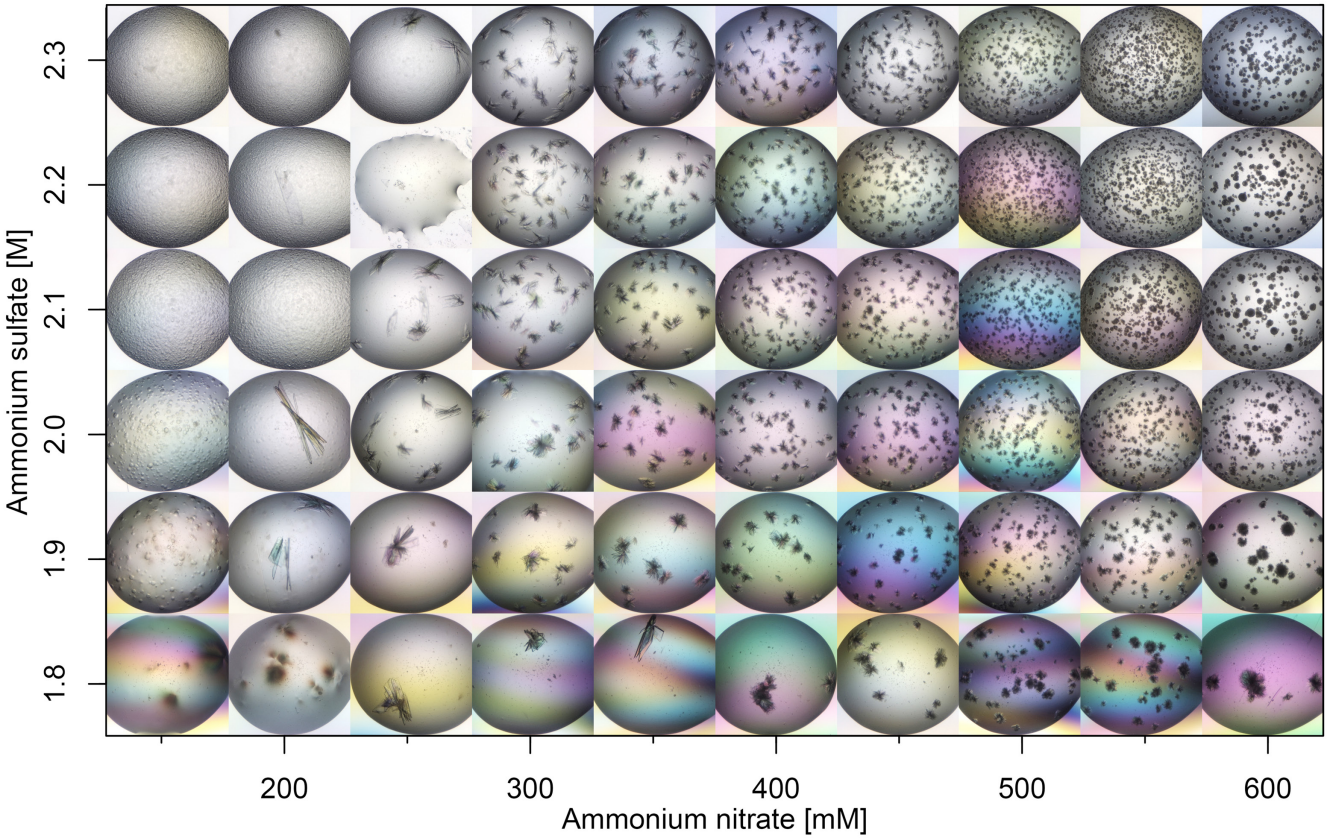
18.1. Ac–[2-Cl-Phe]PPPP–OH in complex with ENAH EVH1

Initial experiments were conduced with 6× molar ligand excess and final concentration of 15 mg/ml. Initial hits, needle tufts were found in three conditions of the AmSO4 suite at 2.2 M $(NH_4)_2SO_4$. Dozens of additives among which NH_4NO_3, KBr, ammonium acetate and ammonium citrate, lithium acetate, but also glycerol created needles as long as the whole diameter of the droplet. Even though impressive, none of the needles showed a clean diffraction pattern and no data set was recorded.

Fine-screening 1.4-2.8 M $(NH_4)_2SO_4$ against 100-300 mM $LiNO_3$, NH_4NO_3, NaSCN and KBr at 20 °C yielded in a cluster of thin plates (2.2 M $(NH_4)_2SO_4$, 200 mM

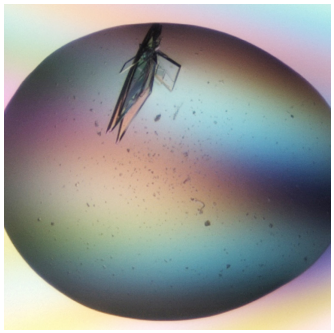
NH_4NO_3) that diffracted to 1.6 Å. Molecular replacement in space group C2, as proposed by MOLREP during data acquisition, was not successful. Re-indexing into correct space group P1 lowered the over-all completeness to only 86%.

The nearly dry droplets were diluted with 1 µl ice-cold water, the crystals crushed and transferred into the protein-ligand sample. Fine-screening 1.8-2.5 M $(NH_4)_2SO_4$ against a wide range of 50-600 mM NH_4NO_3 yielded in tangly crystals and massive crystal shower in nearly 50 conditions between 2.0 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3 up to 2.5 M $(NH_4)_2SO_4$, 500 mM NH_4NO_3. The best data set originated from a single tangly grown crystal (1.8 M $(NH_4)_2SO_4$, 350 mM NH_4NO_3) and was refined with a high-resolution of 1.5 Å.



SI Fig. 18.1.: **Fine-screen of Ac–[2-Cl-Phe]PPPP–OH bound to ENAH EVH1**
Shown are the conditions between 1.8-2.3 M $(NH_4)_2SO_4$ against 150-600 mM NH_4NO_3 on day 35. The screen shows the typical behavior of ENAH EVH1. The solubility line is not sharp but drifts from low-salt conditions to an amorphous regime where spherolites are formed. Such structures dissolved reversibly if nucleation events caused consumption of protein (2.0 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3). In this amorphous regime above 2.1 M $(NH_4)_2SO_4$, NH_4NO_3 tends to cause over-nucleation (right half of upper rows). The ionic strength of NH_4NO_3 first increased the solubility and blocked nuclei to grow (top middle). With increasing concentration, NH_4NO_3 reduced solubility (lowest row) and caused precipitation at the spherolites. The transition into precipitate was different for every ligand but happened generally above 2.8 M $(NH_4)_2SO_4$. The nucleation zone is found between 1.8-2.0 M $(NH_4)_2SO_4$, where ENAH EVH1 sensibly reacted to NH_4NO_3. Fine-screens with other ligands confirmed that the metastable zone rapidly shifts from 180 to 400 mM NH_4NO_3 as the $(NH_4)_2SO_4$ concentration is decreased by 0.2 M to 1.8 M conditions. Single crystals from seed-in experiments were obtained even with 1.6 M $(NH_4)_2SO_4$ if 600-800 mM NH_4NO_3 is used to salt-out the protein. Crystals grown so far from the nucleation zone were usually of good quality.

10 mg/ml ENAH EVH1
6× Ac–[2-Cl-Phe]PPPP–OH
2.2 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3
300/300 nl drop ratio
35 days



11 mg/ml ENAH EVH1
6× Ac–[2-Cl-Phe]PPPP-OH
1.8 M $(NH_4)_2SO_4$, 350 mM NH_4NO_3
200/200 nl drop ratio
35 days

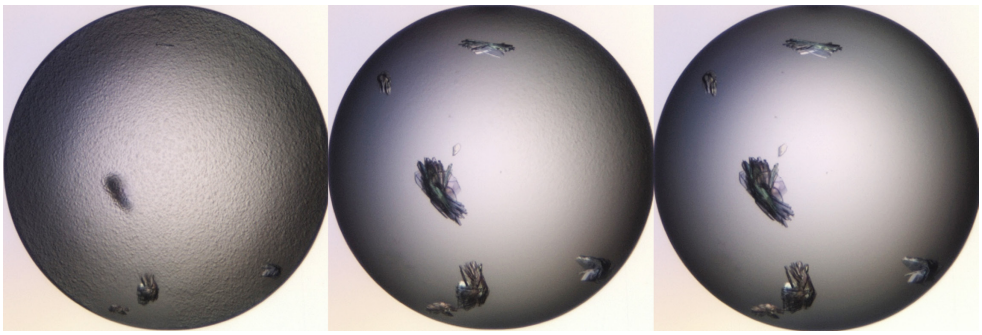
Data set was indexed with space group P1, optimized for acceptable geometry parameters (rmsd spots 1.38 px, spindle 0.23°) and $\langle I/\sigma \rangle_{\text{asymtotic}}$ (ISa) value of 26 and cut at a high-resolution limit of 1.49 Å.

Phaser found a single solution with two chains per asymmetric unit and ended with top log-likely hood gain (LLG) of 6200 and translation function Z-score (TFZ) of 20. Xtriage suggested a pseudo-meroheral twinning with twofold twin operator (-h,-l,-k). Based on the extended Hamilton R-value ratio test [197,198] ran by the PDB_REDO server, [171] the model was refined with TLS groups and individual isotropic atomic displacement parameters (ADPs). The twin law was incorporated with a twin fraction of 0.18 and paired refinement was performed with local NCS restraints. The model was deposited with the accession code 5N91 and refined to an R_{free} value of 20%.

The whole fine-screen plate (partly visible in fig. 18.1) was harvested with ice-cold mother liquor (1.8 M $(NH_4)_2SO_4$, 350 mM NH_4NO_3) and crushed in a Seed Bead tube. Aliquots of 5 and 10 µl were flushed with nitrogen gas, flash-frozen in liquid nitrogen and stored at -80 °C.

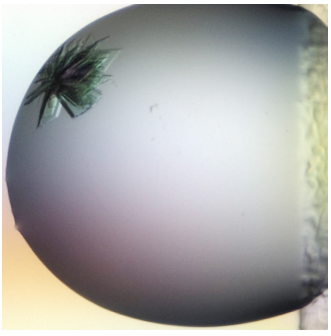
18.2. Ac–[2-Cl-Phe]PP[ProM-1]–OH in complex with ENAH EVH1

Initial experiments were conduced with 6× molar ligand excess and final concentration of 15 mg/ml. Three conditions at 2.2 M $(NH_4)_2SO_4$ with 200 mM salt additives yielded in tangly grown crystals and massive nucleation. Even though the speed varied, the nucleation events caused the underlying gel-like precipitate to dissolve. The crystals from one condition (2.2 M $(NH_4)_2SO_4$, 200 mM KNO_3) were picked after three weeks and diffracted to 1.5 Å.



2.2 M $(NH_4)_2SO_4$, 200 mM KNO_3. Days 3, 7, 21

For fine-screen experiments, the concentration was reduced to 10 mg/ml. Screening 1.4-2.8 M $(NH_4)_2SO_4$ against 100-300 mM KNO_3 as well as 0.1-2.6 M $(NH_4)_2SO_4$ against different pH (MES pH 5.5-6.5, TRIS pH 7.5-8.5) at 20 °C yielded in a tangly grown crystal (1.8 M $(NH_4)_2SO_4$, 200 mM KNO_3) that diffracted to 1.1 Å.



10 mg/ml ENAH EVH1
6× Ac–[2-Cl-Phe]PP[ProM-1]–OH
1.8 M $(NH_4)_2SO_4$, 200 mM KNO_3
300/300 nl drop ratio
7 days

XDS suggested beside space group P1 also C2 (quality of fit 8.7). POINTLESS found neither twinning nor a 2-fold axis (R_{meas} 78% along k). Data set was indexed in

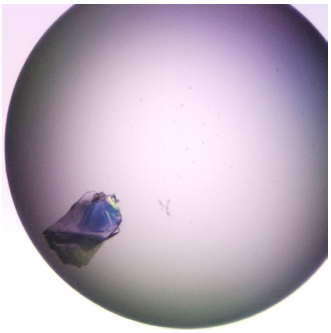
space group P1 and optimized to good geometrical parameters (rmsd spots 0.42 px, spindle 0.03°) and high ISa value of 36. Phaser found a single solution and ended with top LLG 18700, TFZ 20. Xtriage then revealed the pseudo-meroheral twin law (-h,-l,-k). For a triclinc cell with $\beta \cong \gamma \approx 90°$, the orientation of the twin operator was rather untypical and pointed towards a packing in which twinning might interfered with an NCS operator.[199] The twin law reflected tightly packed layers of ENAH EVH1, which broke at the interfaces between them.

Both, the incorporation of the twin operator (-h,-l,-k) with a twin fraction of 0.04 or NCS-based refinement relaxed the model substantially and lowered R_{free} slightly. However, to distinguish between twinning and NCS, R_{free} was a poor model quality indicator as the both models were indistinguishable and the estimated twin fraction was very low.[200] Paired refinement showed that the usage of NCS without the twin operator produced the least overfit model and most detailed difference electron density, improved geometric parameters and decreased R factor gap.[155] The model was deposited with the accession code 5N9C.

18.3. Ac–[2-Cl-Phe]PP[ProM-1]–NH_2 in complex with ENAH EVH1

Initial experiments were conducted with 6× molar ligand excess and final concentration of 15 mg/ml. The complex behaved uniquely in the crystallization screens and produced crystals in only two low-precipitant conditions: The crystal found in 1.6 M $(NH_4)_2SO_4$, 500 mM LiCl at 20 °C diffracted to 1.8 Å and was harvested after 35 days. The second condition (1.6 M $(NH_4)_2SO_4$, 500 mM NaCl) yielded in crystals after 77 days with a resolution limit of 3.3 Å.

Indexing the data set was difficult as the crystal grew twinned and XDS often picked the wrong lattice. Choosing a spot range near the end of the experiment allowed proper integration of the whole data set. Depending on the chosen cell, XDS proposed monoclinic C2 over trigonal crystal systems due to its lower R_{meas} (9.8 compared to 30%), or orthorhombic C222 over monoclinic C2 due to β close to 90°. These were all known space groups in which ENAH EVH1 crystallized during the thesis. However, none of these crystal systems managed to successfully integrate the data set.



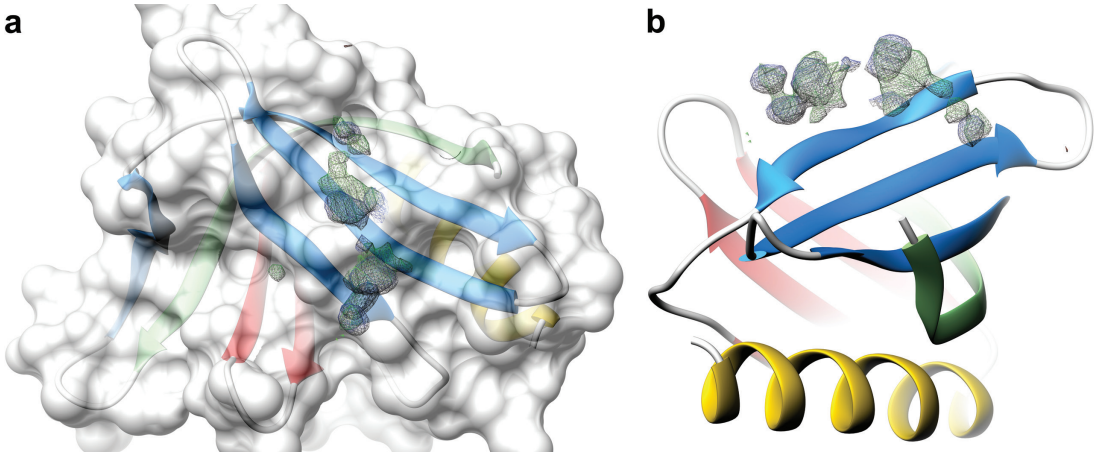
15 mg/ml ENAH EVH1
6× Ac–[2-Cl-Phe]PP[ProM-1]–NH_2
1.6 M $(NH_4)_2SO_4$, 500 mM LiCl
300/300 nl drop ratio
35 days

Integration in monoclinic C2 cell yielded in acceptable standard deviations of spots and spindle positions (0.48 px, 0.04°) and high ISa of 33. Not forcing β to 90° allowed an unbiased look into the systematic absences along *c* and revealed a 3-fold screw axis. As suggested by POINTLESS, re-indexing to $P3_221$ yielded in much better statistics (rmsd spots 0.39 px, spindle 0.04°) and improved ISa value to 44.

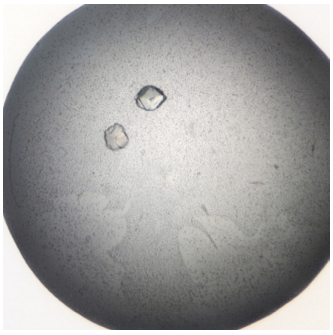
Molecular replacement with Phaser yielded in top LLG 2200, TFZ 19 and clean experimental electron maps in which the ligand was visible right away. Based on the extended Hamilton R-value ratio test[197,198] ran by the PDB_REDO server,[171] the model was refined with isotropic atomic displacement parameters (ADPs) and only one TLS group as proposed by PDB_REDO. Paired refinement with the TLS groups calculated by PHENIX[201] was discontinued after two rounds, as these structures showed no apparent improvement in the electron density map except for a slightly lowered R_{free}. The best model refined without simulated annealing nor NCS at R_{free} 22% and was deposited with the accession code 5N9P.

18.4. Ac–[2-Cl-Phe][ProM-1][ProM-1]–OH in complex with ENAH EVH1

Initial experiments were conduced with 1.8× molar ligand excess and final concentration of 11 mg/ml. The ligand caused problems when diluting in crystallization buffer and formed an insoluble pellet that was dissolved in DMSO. Of this organic phase, 5% (v/v) was added to the crystallization sample and incubated over night at 4 °C. The AmSO4 suite produced among other hits two crystals (2.2 M $(NH_4)_2SO_4$, 200 mM NH_4Cl) that diffracted to 1.5 Å.



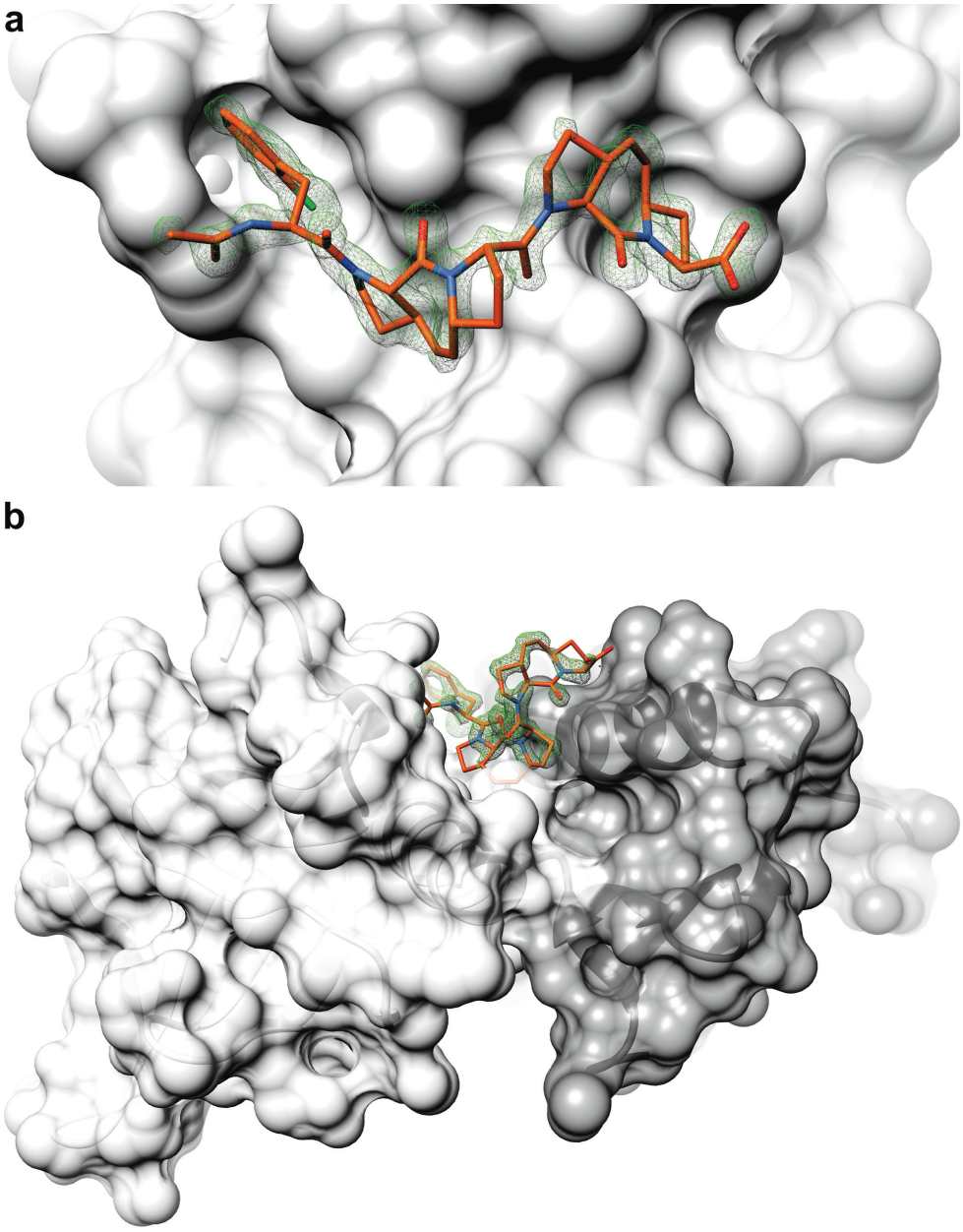
SI Fig. 18.2.: **Ac–[2-Cl-Phe]PP[ProM-1]–NH_2 bound to the second binding site** mF0-DFc difference density map contoured at 3 σ (red/green) and feature-enhanced map contoured at 1 σ (blue). Both calculated at 1.8 Å. Color coding of ENAH EVH according Fig. 3.1. **(a)** The orientation of Ac–[2-Cl-Phe]PP[ProM-1]–NH_2 is not assignable. **(b)** The side-view reveals single resolved pyrrolidine rings.



11 mg/ml ENAH EVH1
1.8× Ac–[2-Cl-Phe][ProM-1][ProM-1]–OH
2.2 M $(NH_4)_2SO_4$, 200 mM NH_4Cl
300/300 nl drop ratio
21 days

Data set was indexed with Laue group ***P m m m*** and optimized to satisfying geometrical parameters (rmsd spots 0.72 px, spindle 0.10°) and ISa value of 22. Unmerged intensities of axial reflections were scored by POINTLESS. According to the probabilities for screw axes (*a*: 0.005, *b*: 0.918, *c*: 0.985), the lattice group frame $P22_12_1$ of the unmerged data set was re-indexed with XDS to $P2_12_12$. Reflection indices were transformed using the re-indexing operator (k,l,h) proposed by POINTLESS. Solvent content analysis suggested 4 chains with a Matthews coefficient[202] (V_M) of 2.32 $Å^3$/Da, which was typical for the packing of ENAH EVH1. Phaser found one solution and ended with top LLG 12400, TFZ 58. Refinement of the Phaser solution produced stunningly clean experimental electron maps in which the ligands, SO_4^{2-}

and water molecules were visible right away. Xtriage reported a pseudo-merohedral twin operator (k,h,-l). Incorporation of the twin law with a twin fraction of 0.11 increased R_{free} without relaxing geometric parameters and was therefore not incorporated into the refinement protocol. Based on the extended Hamilton R-value ratio test[197,198] ran by the PDB_REDO server,[171] the model was refined with isotropic atomic displacement parameters (ADPs) and NCS. R_{free} decreased further after replacing the TLS groups calculated by PHENIX[201] with the optimized range proposed by PDB_REDO. Simulated annealing was used during model rebuilding but was stopped for the final refinement cycles which halted at R_{free} of 20%. The model was deposited with the accession code 5NAJ.

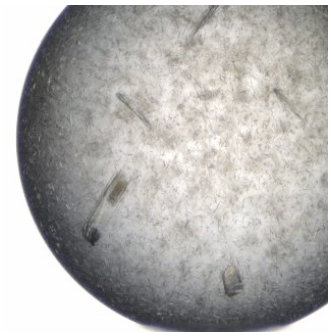


SI Fig. 18.3.: mF0-DFc difference density map contured at 3 σ showing missing density of the inhibitors. **(a)** Ac–[2-Cl-Phe][ProM-1][ProM-1]–OH bound in the canonical binding groove. **(b)** Difference map of an inhibitor containing a diastereoisomer of ProM-1, bound between two neighboring ENAH EVH1 chains (white and dim gray with transparent solvent-accessible surfaces). At this stage, the chirality of the enantiomeric ProM-1 could already be determined. Map calculated after refining the Phaser solution against the merged data set.

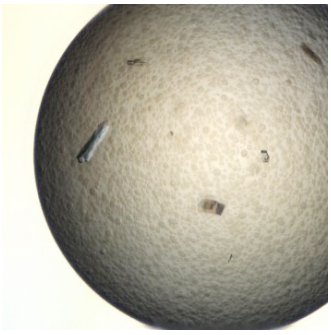
18.5. Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH in complex with ENAH EVH1

The first crystallization experiments with ENAH EVH1 were performed with saturated $(NH_4)_2HPO_4$, as reported by Prehoda *et al.*.[174] Screening 15-80% saturated $(NH_4)_2HPO_4$ against several additives and pH was unsuccessful. In a different attempt, the remaining protein sample was titrated with the Quiagen AmSO4 suite. 10 mg/ml ENAH EVH1 crystallized right away in not less than 7 conditions at 20 °C between 2.2 and 2.4 M $(NH_4)_2SO_4$. The conditions produced highly polarizing plates growing between a shower of micro crystals and tiny needles. The best packed crystal (2.2 M $(NH_4)_2SO_4$, 200 mM Na formate) showed a resolution limit of roughly 1.95 Å. The initial data set was indexed in space group $P2_1$ but refinement stalled at a rather large R value gap of 7% with $rmsd_{bond}$ 0.02 Å and $rmsd_{angle}$ 1.8°. Even though fully refined (R_{free} 32%), this structure was not deposited.

The AmSO4 screen was repeated with 15 mg/ml ENAH EVH1 and ligand excess reduced to 3×. The increased protein concentration tilted the system far into the precipitation regime and only two conditions produced small, irregularly grown crystals. One crystal (2.2 M $(NH_4)_2SO_4$, 200 mM NH_4Br) diffracted to 1.7 Å.



10 mg/ml ENAH EVH1
6× Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH
2.2 M $(NH_4)_2SO_4$, 200 mM Na formate
300/300 nl drop ratio
21 days



15 mg/ml ENAH EVH1
3× Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH
2.2 M $(NH_4)_2SO_4$, 200 mM NH_4Br
300/300 nl drop ratio
21 days

The crystal structure of ENAH EVH1 in complex with the initial inhibitor Ac–[2-Cl-Phe][ProM-1][ProM-2]–OH represents the only model that was solved entirely by the programs of the CCP4 suite. The model was deposited with the accession code 4MY6.

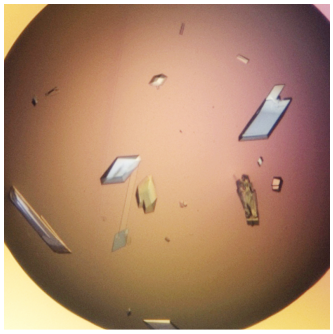
18.6. Ac–[2-Cl-Phe][ProM-2][ProM-3]–OH in complex with ENAH EVH1

Initial experiments were conduced with 2.3× molar ligand excess and final concentration of 11 mg/ml. The AmSO4 suite produced in one single condition (2.2 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3) star-shaped rods. After three weeks, two crystals were harvested and diffracted to 1.4 Å.

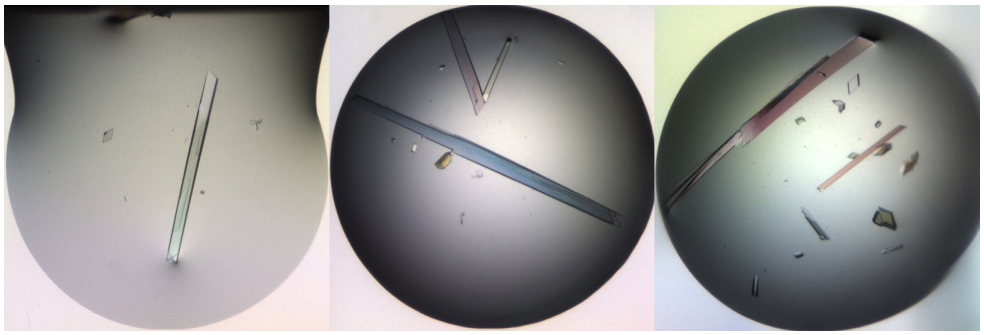
11 mg/ml ENAH EVH1
2.3× Ac–[2-Cl-Phe][ProM-2][ProM-3]–OH
2.2 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3
300/300 nl drop ratio
21 days

Indexing was unambiguous in space group P2 with one 2-fold rotational symmetry and β=90.6°. Molecular replacement of one ENAH EVH1 chain ended with a top LLG of 3700 and structure refined to an R_{free} of 19.4%.

As mentioned in the results, ProM-3 revealed a never-seen ligand-mediated crystal contact. To investigate whether the worse K_d of ProM-3-containing ligands also involved an altered arrangement of water molecules around the ligand, micro seeding experiments were performed with seeds grown in space group P1 (Ac–[2-Cl-Phe]PPPP–OH cocrystals). This aimed to force the ensemble to grow in a packing in which ProM-3 would be solvent exposed. The fine-screen was adapted to 1.6-2.3 M $(NH_4)_2SO_4$ against 180-360 mM NH_4NO_3 yielded in huge crystals grown at 20 ℃ within 3-7 days spreading over 30 conditions (a selection shown in the lower panel of the following picture).



11 mg/ml ENAH EVH1
2.3× Ac–[2-Cl-Phe][ProM-2][ProM-3]–OH
2.1 M $(NH_4)_2SO_4$, 180 mM NH_4NO_3
200/300/100 nl drop ratio
7 days

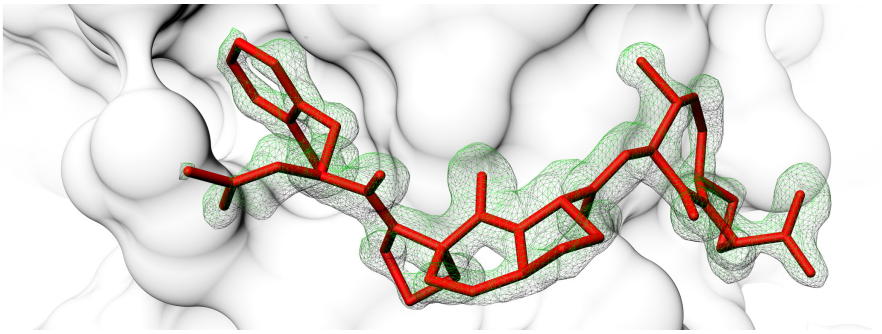


2.1 M $(NH_4)_2SO_4$ with 200, 220 and 240 mM NH_4NO_3

Indexing of several crystals showed that the seeds did not manage to alter the crystal growth in space group P1 as intended. All measured crystals were packed unambiguously in space group P2. The best crystal (2.1 M $(NH_4)_2SO_4$, 180 mM NH_4NO_3) diffracted to 1.15 Å. Molecular replacement by Phaser with one chain per asymmetric unit yielded in top LLG 12000. The two crystals were perfectly isomorphous with a maximal deviation of 0.6% along axis *b*. Both data sets were merged for paired refinement. Phaser was run with the same search model and found a single solution with top LLG 12700.

Unsurprisingly, refinement of the merged data set including data out to 1.15 Å yielded in worse statistics. Paired refinement against the merged data was run at more sensible high-resolution of 1.23 Å with $CC_{1/2}$ value of 58.4% in the outermost shell. Based on the extended Hamilton R-value ratio test[197,198] ran by the PDB_REDO server,[171] anisotropic B-factor models were refined against both, the single 1.15 Å data set as well as against the merged data set cut at 1.23 Å. No twinning was detected. Paired refinement against both data sets at the respective 1.15 and 1.23 Å resolutions was continued over totally 7 cycles of rebuilding and refinements. The model of the unmerged data set was more accurate, as it reached 1.3% lower

R_{free} at the high resolution of 1.23Å. The model was deposited with the accession code 5NBF.



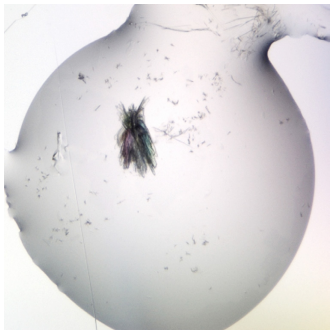
SI Fig. 18.4.: mF0-DFc difference density map contured at 3 σ showing missing density of Ac–[2-Cl-Phe][ProM-2][ProM-3]–OH. Map calculated after refining the Phaser solution against the merged data set.

18.7. Ac–[2-Cl-Phe][ProM-2][ProM-4]–OH in complex with ENAH EVH1

Initial experiments were conduced with 1.8× molar ligand excess and final concentration of 11 mg/ml. The ligand caused problems when diluting in crystallization buffer and formed an insoluble pellet that was dissolved in DMSO. Of this organic phase, 5% (v/v) was added to the crystallization sample and incubated over night at 4 °C. The AmSO4 suite produced in few conditions micro crystals among which 2.2 M $(NH_4)_2SO_4$ and 200 mM NH_4NO_3. None of the crystals were suitable for data acquisition.

Micro crystals grown in the two conditions of the initial screen were pooled into the remaining protein-ligand mixture and a new fine-screen was designed. Screening 1.8-2.5 M $(NH_4)_2SO_4$ against 50-600 mM NH_4NO_3 yielded in micro crystals in the low salt conditions and spherolites above 250 mM NH_4NO_3. Conditions below 2.0 M $(NH_4)_2SO_4$ stayed clear until the twenty first day and started then to produce dense needle tufts and tangly grown crystals. One condition (1.9 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3) produced a tangly crystal at 20 °C that started growing after three weeks and was harvested before the next image was taken. From this crystal only the debris is visible in the following picture.

11 mg/ml ENAH EVH1
1.8× Ac–[2-Cl-Phe][ProM-2][ProM-4]–OH
2.2 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3
200/200 nl drop ratio
14 days



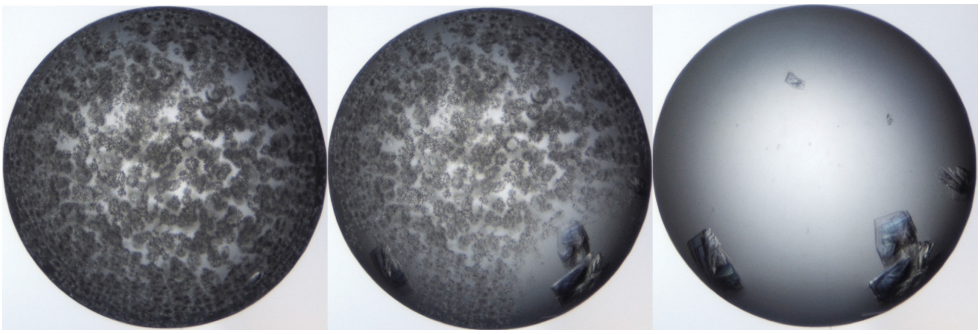
11 mg/ml ENAH EVH1
1.8× Ac–[2-Cl-Phe][ProM-2][ProM-4]–OH
1.9 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3
200/200 nl drop ratio
35 days

The fine-screen yielded in a data set that was processed in space group P1 at a resolution limit of 1.4 Å and final ISa of 36. Phaser found a single solution and ended with top LLG 5000, TFZ 17. POINTLESS found no twin fraction (α 0.06) but Xtriage suggested a possible pseudo-merohedral twinning. Twin operator (-h,-l,-k) with a twin fraction of 0.15 was included into the refinement and lowered R_{free} 3% compared to untwinned refinement. Based on the Hextended Hamilton R-value ratio test [197,198] ran by the PDB_REDO server, [171] the model was refined with isotropic ADPs and the TLS groups calculated by PHENIX. [201]. The model converged much slower than expected and halted around R_{free} of 25% with comparatively high ADPs. It was deposited with the accession code 5NCF.

18.8. Ac–WPPPPTEDEL–NH_2 in complex with ENAH EVH1

Ligands with elongated C-terminus crystallized the first time in the AmSO4 suite, as well. Using 15 mg/ml ENAH EVH1, such 10-meric compositions displayed similar phase boundaries as the short pentamers and formed spherolites and amorphous precipitate at 2.2 M $(NH_4)_2SO_4$, which started to clear out around 21 days with

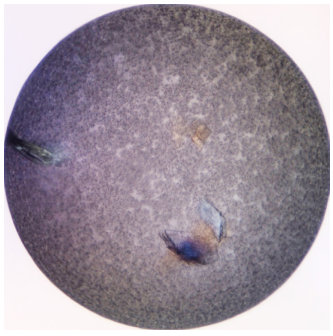
the first nucleation events. By day 77 most precipitate dissolved and small tangly needles grew in different additives.



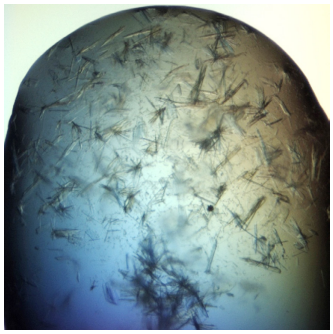
2.2 M $(NH_4)_2SO_4$, 200 mM KBr. Days 77, 100, 150

From this screen at 20 °C, the first harvested crystals diffracted not further than 4 Å. Two data sets were recorded with crystals that grew slower and appeared after 35 days: 2.2 M $(NH_4)_2SO_4$, 200 mM $NaNO_3$ with a resolution limit of 2.90 Å, and 2.2 M $(NH_4)_2SO_4$, 200 mM NaCl diffracting to 2.70 Å. A third condition (2.2 M $(NH_4)_2SO_4$, 200 mM LiCl) produced tangly crystals that grew much later and were harvested after 100 days. The third crystal diffracted to 2.84 Å. The concentration of ENAH EVH1 was reduced to 11 mg/ml to work against over-nucleation. Fine-screening 1.4-2.8 M $(NH_4)_2SO_4$ against 100-300 mM $LiNO_3$, NH_4NO_3, NaSCN, and KBr yielded only in tangled needles and plate tufts around 2.0 M. Neither repetition of the AmSO4 screen with reduced concentration yielded in crystals with a diffraction power further than 2.7 Å, nor did repetitions of $(NH_4)_2SO_4$ fine-screens with 11 mg/ml ENAH EVH1 at 27 °C.

All three data sets were indexed in space group C2. The cell dimensions suggested 6 ENAH EVH1 chains per asymmetric unit (Vm 2.35 $Å^3$/Da). The two weaker data sets (2.90 and 2.84 Å) were perfectly isomorphous and refined with good geometrical parameters and ISa values (rmsd spots 0.81-1.26 px, spindle 0.1-0.2°, ISa 36-43). Merging allowed to push the resolution limit to 2.7 Å and $CC_{1/2}$ value of 59.0%. Phaser found a single solution for the merged data set and ended with top LLG 3800, TFZ 24. The unmerged 2.84 Å data set ended with top LLG 410, TFZ 24. Three solutions with Z-scores up to 21 were rejected for failing packing test. Even though the initial electron density maps for the merged data set looked significantly better, paired refinement was continued. Xtriage suggested for both data



15 mg/ml ENAH EVH1
6× Ac–WPPPPTEDEL–NH_2
2.2 M $(NH_4)_2SO_4$, 200 mM LiCl
300/300 nl drop ratio
100 days



15 mg/ml ENAH EVH1
6× Ac–WPPPPTEDEL–NH_2
2.2 M $(NH_4)_2SO_4$, 200 mM NaCl
300/300 nl drop ratio
100 days

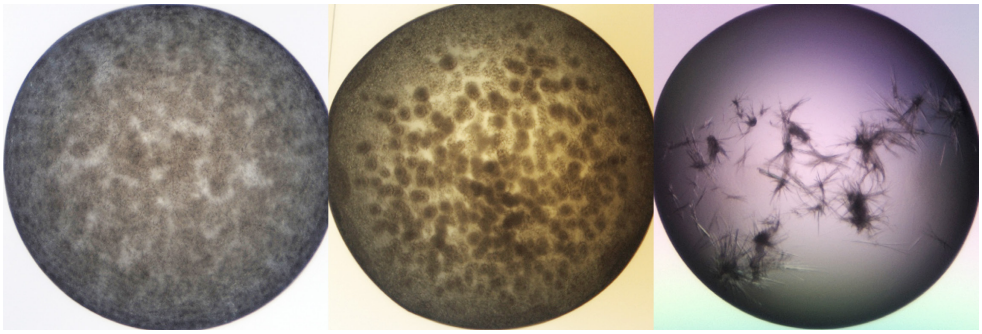
sets pseudo-merohedral twinning, possibly along an NCS axis. Twin operator (-h,-k,l) was included into paired refinements and improved R_{free} by 3.3% during the model building.

The best diffracting crystal (2.70 Å) deviated along axis *c* nearly 11%. XDS refined the data set to satisfying geometrical parameters (rmsd spots 0.72 px, spindle 0.11°) and ISa value of 28. Phaser found a single solution and ended with top LLG 3200, TFZ 22. The merged data set, its stronger single data, as well as the non-isomorphous data were used for a paired refinement over 7 consecutive cycles at the sensible resolution limits. A comparison of the models calculated at equal high-resolution cut-offs revealed that the merged data set converged to nearly 3% lower R_{free} and was deposited with the accession code 5NC7.

18.9. Ac–[2-Cl-Phe]PPPPTEDEL–NH_2 in complex with ENAH EVH1

Quiagen AmSO4 initial screen at 15 mg/ml yielded in light amorphous precipitation similar to Ac–WPPPPTEDEL–NH_2 initial experiments. Here, the conditions around

2.2 M $(NH_4)_2SO_4$ produced only few nucleation events and the spherolite-like precipitate remained unchanged until the last observation (150 days). Nucleation events happening in old droplets and low $(NH_4)_2SO_4$ concentrations (1.6 M, after three weeks) produced irregularly growing structures composed of needles.



2.2 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3, 100 days
2.2 M $(NH_4)_2SO_4$, 200 mM K_3-citrate, 100 days
1.6 M $(NH_4)_2SO_4$with 500 mM LiCl, 100 days

The concentration of ENAH EVH1 was reduced to 12 mg/ml and three crystallization experiments were set up at 27 °C, screening 1.4-2.8 M $(NH_4)_2SO_4$ against 100-300 mM $LiNO_3$, NH_4NO_3, NaSCN, and KBr, plus a plate with 1.8-2.5 M $(NH_4)_2SO_4$ against 50-600 mM NH_4NO_3; as well as a diluted AmSO4 initial screen. ENAH EVH1 crystallized after two weeks in more than 13 high-salt conditions from 1.8 up to 2.4 M $(NH_4)_2SO_4$ and five crystals were harvested after a month (no pictures taken). The best condition (2.2 M $(NH_4)_2SO_4$ against 500 mM NH_4NO_3) produced crystals diffracting beyond 1.8 Å.

Data set was initially indexed with space group $I2_12_12_1$. Optimization in XDS yielded in a low ISa value of 13 and unacceptable 1.76 px rmsd deviation for spot positions. Phaser found 2 solutions, the best with top LLG 590, TFZ 7.2. L-test suggested twinning and POINTLESS proposed I222. However, re-indexing in I222 did not improve ISa either. Data set was re-indexed with the correct space group C2 and optimized to significantly better rmsd deviations (spots 1.27 px, spindle 0.14°) and high ISa value of 29. Data set was cut at a high-resolution limit of 1.58 Å and Phaser found a single solution, ending with top LLG 5300, TFZ 19.8.

Xtriage proposed a pseudo-merohedral twinning with the twofold twin law (-h-2×l,-k,l). Incorporation of the twin operator into the refinement cleaned up the difference density map noticeably and lowered R_{free} marginally by 0.6% with a twin faction of

0.11. Paired refinement towards the end suggested however to remove the twin law and use NCS-based refinement instead. The model was deposited with the accession code 5NC2.

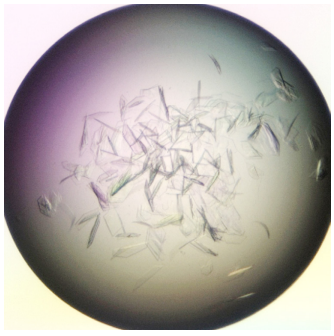
18.10. Ac–[2-Cl-Phe]PP[ProM-1]TEDEL–NH_2 in complex with ENAH EVH1

Two additional fine-screens were prepared which covered 1.6-2.3 M $(NH_4)_2SO_4$ against 180-360 mM and 360-800 mM NH_4NO_3. The micro seeds consisted of a combination of space groups C2 (Ac–[2-Cl-Phe]PPPPTEDEL–NH_2, 1.6 Å resolution) and P1 (Ac–[2-Cl-Phe]PPPP–OH, 1.5 Å resolution), 1:1 (v/v) diluted with ice-cold high-salt mother liquor (2.2 M $(NH_4)_2SO_4$, 500 mM NH_4NO_3). Both fine-screens were titrated as replicas and stored at 20 and 27 °C. The $(NH_4)_2SO_4$ concentration of the fine-screen was lowered because the mother liquor of the seeds brought additional salt into the droplet. This setup put the crystallization conditions above 2.0 M $(NH_4)_2SO_4$ close to the solubility line of the protein or even started right in the metastable zone. Most of these droplets produced massive crystal shower from the first hour and equilibrated within two days.

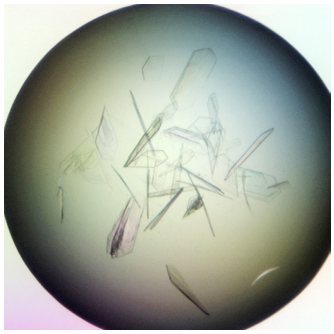
Eleven crystals were harvested on the third day. Increasing the temperature significantly worked against nucleation and irregular growth, two qualities from which earlier attempts to crystallize C-terminally elongated ligands suffered. The best crystal (1.7 M $(NH_4)_2SO_4$, 800 mM NH_4NO_3) grew at 27 °C and diffracted to 1.45 Å.

Also Ac–[2-Cl-Phe]PP[ProM-1]TEDEL–NH_2 lead to a crystal packed in space group C2 with cell dimensions similar to the other two TEDEL-containing ligand complexes. The comparable low ISa value of 13 but good standard deviations (0.93 px, 0.10°) pointed towards twinned data rather than poor refined geometry parameters or wrong space group assignment. However, POINTLESS found only a twin fraction of 0.03.

Phaser found a single solution yielding in a top LLG of 5800. Structure was refined with TLS and NCS restraints to an R_{free} value of 20%. Xtriage suggested a possible pseudo-merohedral twinning with twin law (-h-2×l,-k,l). Twin operator was included into the refinement, however the twin fraction of 0.06 had no positive effect, neither for the difference density map nor on R_{free}. Instead, NCS-based



15 mg/ml ENAH EVH1
4× Ac–[2-Cl-Phe]PP[ProM-1]TEDEL–NH_2
1.7 M $(NH_4)_2SO_4$, 800 mM NH_4NO_3
200/300/100 nl drop ratio
3 days
20 °C



15 mg/ml ENAH EVH1
4× Ac–[2-Cl-Phe]PP[ProM-1]TEDEL–NH_2
1.7 M $(NH_4)_2SO_4$, 800 mM NH_4NO_3
200/300/100 nl drop ratio
3 days
27 °C

refinement yielded in significant better statistics and twin operator was not considered for the final refinement cycles. The model was deposited with the accession code 5ND0.

18.11. Ac–[2-Cl-Phe][ProM-2][ProM-12]–OMe in complex with ENAH EVH1

Initial experiments at 15 mg/ml were prepared with 5× ligand excess at 20 and 27 °C, using micro seeds from Ac–[2-Cl-Phe]PPPP–OH in a 1:6 (v/v) dilution. Several fine-screens covering totally 1.3-2.0 M $(NH_4)_2SO_4$ against 50-110 mM NH_4NO_3 yielded in two crystals grown in 2.0 M $(NH_4)_2SO_4$, 220 mM NH_4NO_3 at 20 °C. Indexing revealed ice-related problems, even though the data sets diffracted to 1.6-1.5 Å. The two crystals were perfectly isomorphous with a maximal deviation of 0.1% along axis *c*. Merging the data allowed to push the resolution limit to 1.42Å. Further fine-screening yielded in a third data set of a crystal grown in 1.9 M $(NH_4)_2SO_4$, 140 mM NH_4NO_3. The third crystal showed a smaller cell with a maximal deviation of 2.5% along axis *a*.

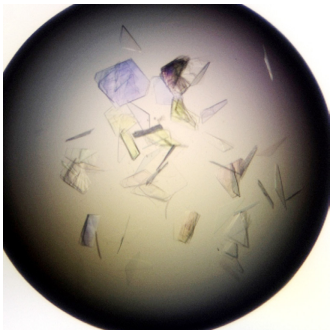
All data set were indexed in C2. Paired refinement was run with four different structure factors: the merged data of the first two crystals, the single data of the third crystal, and two merged data sets using either the first or the second cell dimensions. Phaser found only single solutions and ended with top LLG between 4000-6000 and TFZ 17-21. As the TLS groups calculated by PHENIX[201] caused a reduced Hamilton R ratio,[197,198] PDB_REDO[171] accordingly proposed refinement of the model with one TLS group per chain to prevent over-fitting. The models were refined with isotropic atomic displacement parameters and simulated annealing was stopped towards the end of rebuilding. Paired refinement and rebuilding of four models was continued for 7 cycles. Refinement at the same high-resolution cut-off showed that refinement against the merged data set of the isomorphous crystals produced the most accurate model which was deposited with the accession code 5NDU.

The fine-screen experiments containing Ac–[2-Cl-Phe][ProM-2][ProM-12]–OMe in complex with ENAH EVH1 were used to store ProM-12-containing seeds grown in space group C2. Thick or tangly grown crystals were washed in ice-cold 1.9 M $(NH_4)_2SO_4$, 450 mM NH_4NO_3 by subsequently transferring them in three droplets of 100 µl. This transfer aimed to wash away all unbound ligand and ensured that the seed material contained only little ligand. Crystals pooled in the last droplet were transferred into Seed Bead tube on ice. Droplets with crystal shower were washed with 20 µl of the same solution before the pipet tip was used to detach the crystals from the crystallization plate. A minimal amount of volume was sucked in and brought into the Seed Bead tube. Aliquots of 5 and 10 µl were flushed with nitrogen gas, flash-frozen in liquid nitrogen and stored at -80 °C.

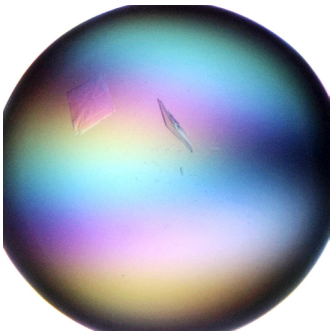
18.12. Ac–[2-Cl-Phe][ProM-2][ProM-12]–OH in complex with ENAH EVH1

Initial experiments at 11 mg/ml were prepared with micro seeds from Ac–[2-Cl-Phe]-[ProM-2][ProM-12]–OMe in a 1:3 (v/v) dilution for the fine-screens and 1:6 (v/v) for the AmSO4 initial screen. Screening in $(NH_4)_2SO_4$ concentrations against a broad range of 140-800 mM NH_4NO_3 yielded in dozens of tangly crystals at 20 °C. Five crystals grown in 1.7 M $(NH_4)_2SO_4$, 180 mM NH_4NO_3 and 1.6 M $(NH_4)_2SO_4$, 260 mM

NH_4NO_3 were harvested after 2 days and diffracted to around 1.7 Å, most of them showing highly twinned latices.



11 mg/ml ENAH EVH1
2.5× Ac–[2-Cl-Phe][ProM-2][ProM-12]–OH
1.7 M $(NH_4)_2SO_4$, 180 mM NH_4NO_3
200/300/100 nl drop ratio
2 days
20 °C



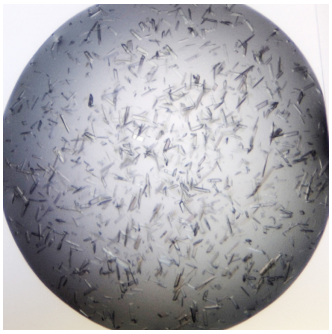
11 mg/ml ENAH EVH1
2.5× Ac–[2-Cl-Phe][ProM-2][ProM-12]–OH
1.6 M $(NH_4)_2SO_4$, 260 mM NH_4NO_3
200/300/100 nl drop ratio
2 days
20 °C

From the first condition, the acquired data set was processed unambiguously in C222 and optimized to acceptable geometry parameters (rmsd spots 0.91 px, spindle 0.11°). The high-resolution cut-off was pushed beyond $CC_{1/2}$ 0.60 as the data set was 100% complete at the outer shell between 1.7-1.8 Å and contained substantial information ($\langle I/\sigma \rangle_{mrgd}$ 1.46). Guided by the highest ISa value of paired CORRECT runs at 1.60-1.69 Å, the diffraction limit was set to 1.65 Å ($CC_{1/2}$ 0.43 in outermost shell) and yielded in an ISa value of 19.

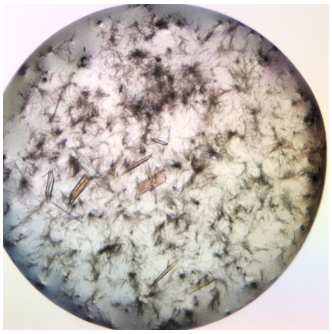
Phaser found a single solution with one chain per asymmetric unit and ended with top LLG 2400 and TFZ 4.6. The low TFZ as well as the results of the L-test indicated that too the high-resolution cut-off was too optimistic. Based on the Hamilton R-value ratio test [198] ran by the PDB_REDO server, [171] the model was refined with isotropic ADPs and the TLS groups calculated by PHENIX [201]. Simulated annealing was stopped for the final refinement cycles and R_{free} converged to a good value of 21%. The model was deposited with the accession code 5NCP.

18.13. Ac–[2-Cl-Phe]PP[ProM-9]–OH in complex with ENAH EVH1

Quiagen AmSO4 initial screen was conduced with 3.7× molar ligand excess and final concentration of 15 mg/ml. Few conditions produced tangly crystals at 2.2 M $(NH_4)_2SO_4$ in combination with 200 mM NH_4NO_3, 200 mM KNO_3, 200 mM $LiNO_3$ or 200 mM NaBr. All measured crystals were twinned. Fine-screening 1.4-2.8 M $(NH_4)_2SO_4$ against 100-300 $LiNO_3$, NaBr, and KBr, respectively, as well as 1.8-2.5 M $(NH_4)_2SO_4$ against 100-300 NH_4NO_3 yielded in massive but tangly grown plates. One condition (2.0 M $(NH_4)_2SO_4$, 300 mM KBr) produced single, polarizing plates. Two crystals grown at 20 °C were harvested and diffracted to 1.65 Å. KBr screening was discontinued after other complexes also grew in an amorphous regime that produced irregularly growing needle tufts and spherolites around 1.9-2.1 M $(NH_4)_2SO_4$ with KBr (see image).



15 mg/ml ENAH EVH1
3.7× Ac–[2-Cl-Phe]PP[ProM-9]–OH
2.2 M $(NH_4)_2SO_4$, 200 mM NaBr
300/300 nl drop ratio
21 days
20 °C



15 mg/ml ENAH EVH1
2.5× Ac–[2-Cl-Phe]PP[ProM-9]–OH
2.0 M $(NH_4)_2SO_4$, 300 mM KBr
300/300 nl drop ratio
35 days
20 °C

Data set was indexed with space group C222 and optimized for satisfying geometry parameters (rmsd spots 0.98 px, spindle 0.14°). The data set was cut at a high-resolution of 1.65 Å. Both XDS and POINTLESS evaluated possible twofold screw axis along *c* due to 5 missing reflections. However, both programs proposed C222 and

reported no twinning. Phaser found a single solution with two chains per asymmetric unit and ended with top LLG 5300 and TFZ 20. Paired refinement showed that optimal combination consisted of TLS-based refinement without NCS nor annealing steps. The model refined to an R_{free} value of 19% and was deposited with the accession code 5NBX.

18.14. Ac–[2-Cl-Phe][ProM-2][ProM-9]–OH in complex with ENAH EVH1

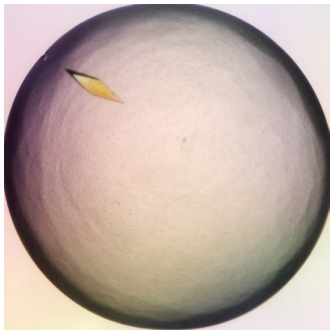
Quiagen AmSO4 initial screen was conduced with 2.3× molar ligand excess and final concentration of 11 mg/ml. The ligand caused problems when diluting in crystallization buffer and formed an insoluble pellet that was dissolved in DMSO. Of this organic phase, 5% (v/v) was added to the crystallization sample and incubated over night at 4 ℃.

The crystals grown at 20 ℃ in 2.2 M $(NH_4)_2SO_4$ and 200 mM of NH_4NO_3 or KCl, respectively, belong to the most beautiful and best diffracting crystals ever recorded during the thesis. Scattering power up to 1 Å (2.2 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3) made subsequent fine-screening obsolete.

Due to the crystal alignment, axial reflections along $h00$ were absent. Even though POINTLESS indicated only one screw axis ($00l$) to be significant, space groups $P2_122_1$ as well as $P22_12_1$ showed much lower probability than $P2_12_12_1$. Data set was indexed with space group $P2_12_12_1$ and optimized for satisfying geometry parameters (rmsd spots 0.80 px, spindle 0.16°) and ISa value of 26. The data set was cut at a high-resolution limit of 1.01 Å. Neither POINTLESS nor Xtriage suggested twinning.

Molecular replacement was performed in all possible spacegroups within the point group P222. Phaser found 4 solutions with two chains per asymmetric unit and ended with top LLG 170, TFZ 11.6. A large non-origin Patterson peak originated from translational NCS. Based on the extended Hamilton R-value ratio test [197,198] ran by the PDB_REDO server, [171] a model with individual anisotropic ADPs was refined. For the last cycles of refinement, simulated annealing and NCS restraints were stopped, which yielded in optimal convergence of R_{free} to a satisfying value of

11 mg/ml ENAH EVH1
2.3× Ac–[2-Cl-Phe][ProM-2][ProM-9]–OH
2.2 M $(NH_4)_2SO_4$, 200 mM NH_4NO_3
300/300 nl drop ratio
21 days
20 °C



11 mg/ml ENAH EVH1
2.3× Ac–[2-Cl-Phe][ProM-2][ProM-9]–OH
2.2 M $(NH_4)_2SO_4$, 200 mM KCl
300/300 nl drop ratio
150 days
20 °C

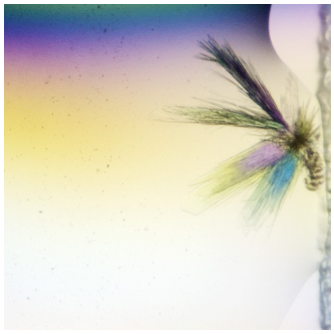
15%. The model was deposited with the accession code 5NCG with an R factor gap of 2.5%.

18.15. Ac–[2-Cl-Phe][ProM-2][ProM-13]–OEt in complex with ENAH EVH1

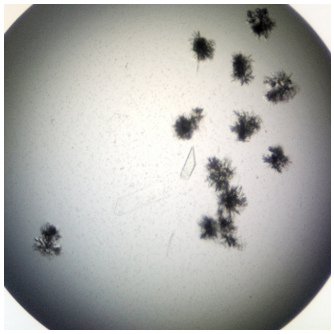
Initial experiments were conduced with 1.6× molar ligand excess and final concentration of 15 mg/ml using the Quiagen AmSO4 initial screen. Due to heavy precipitation, the experiment was repeated with a diluted AmSO4 initial screen. Of this second screen, few crushed crystals were used in a subsequent fine-screen of 1.8-2.5 M $(NH_4)_2SO_4$ against 50-600 mM of NH_4NO_3. This experiment yielded in spherolites and tangly grown needle tufts between 1.8-2.1 M $(NH_4)_2SO_4$ and 400-600 mM NH_4NO_3 at 20 °C.

Three crystals (1.8 M $(NH_4)_2SO_4$, 450 mM NH_4NO_3) were measured and diffracted to roughly 1.7 Å. One recorded data set was indexed and processed unambiguously in space group P1. Phaser found a single solution with 4 chains per asymmetric unit and ended with top LLG 5400 and TFZ 14.

Needle tufts and solid structures were harvested into Seed Bead tube with ice-cold buffer (1.8 M $(NH_4)_2SO_4$, 450 mM NH_4NO_3). Aliquots of 5 and 10 µl were flushed with nitrogen gas, flash-frozen in liquid nitrogen and stored at -80 °C.



15 mg/ml ENAH EVH1
1.6× Ac–[2-Cl-Phe][ProM-2][ProM-13]–OEt
1.8 M $(NH_4)_2SO_4$, 450 mM NH_4NO_3
300/300 nl drop ratio
21 days
20 °C



15 mg/ml ENAH EVH1
1.6× Ac–[2-Cl-Phe][ProM-2][ProM-13]–OEt
1.8 M $(NH_4)_2SO_4$, 360 mM NH_4NO_3
200/300/100 nl drop ratio
3 days
20 °C

For further fine-screening, the $(NH_4)_2SO_4$ was reduced and NH_4NO_3 increased and combined with micro seeding (1:4 (v/v) dilution). Seeding and reducing $(NH_4)_2SO_4$ did not work against the amorphous growth. One condition (1.8 M $(NH_4)_2SO_4$, 360 mM NH_4NO_3) produced a single crystal that was harvested after 3 days. The cell dimensions of both data sets were isomorphous. Processing and optimization of the second data set revealed significant problems in resolution shells above 2.0 Å. Therein $CC_{1/2}$ declined constantly to 65% towards 1.7 Å but completeness and spot intensities dropped to 36% and 0.8 respectively, suggesting that the detector distance was too close. The diffraction images confirmed that edge-to-edge distance allowed only 2.0 Å resolution. The second data set was cut at resolution limit of 2.0 Å and merged into the 1.63 Å data set. Phaser found a single solution and yielded in top LLG 5400 and TFZ 14. Even though the experimental maps of the merged data set looked slightly worse, refinement was conducted over 7 cycles. Paired refinements converged very slow and halted around 27%. Refinement and rebuilding was discontinued after 8 cycles.

Additional microseeding experiments at 20 °C in low-salt conditions with 20 mg/ml ENAH EVH1 did not fully overcome the tangly growth. However, the slow growth speed lead to significantly less twinned crystals that were harvested after 200 days. Surprisingly, these conditions produced crystals unambiguously packed in space group P2. Of the two data sets that were recorded, the crystal fragment grown in 1.8 M $(NH_4)_2SO_4$, 260 mM NH_4NO_3 optimized to good geometry parameters (rmsd spots 0.42 px, spindle 0.02°) and ISa value of 21 and diffracted to a high-resolution limit of 1.22 Å. Phaser found a single solution and yielded in top LLG 6109 and TFZ 65. The second crystal grew in even lower salt concentrations of 1.6 M $(NH_4)_2SO_4$, 320 mM NH_4NO_3. Data set was optimized to good geometry parameters (rmsd spots 0.61 px, spindle 0.09°) and ISa value of 19 and cut at a high-resolution limit of 1.51 Å. Phaser found a single solution and yielded in top LLG 4254 and TFZ 56. The first data set showed ice-related problems while the latter data were complete but diffracted only to 1.5 Å. The weaker data was merged into the 1.22 Å data set, which allowed to push the completeness in the low resolution shell between 6.3-5.6 Å from 67.5 to 92.8%. The single solution of phaser with this data set reached top LLG 6575 and TFZ 69, suggesting that significant data was added. Indeed, the following paired refinement revealed that the model based on the merged data converged to the lowest R_{free}. The model was deposited with the accession code 5NEG.

19. Attempts to crystallize EVL EVH1 and VASP EVH1

19.1. Ac–[2-Cl-Phe]PP[ProM-9]–OH in complex with EVL EVH1

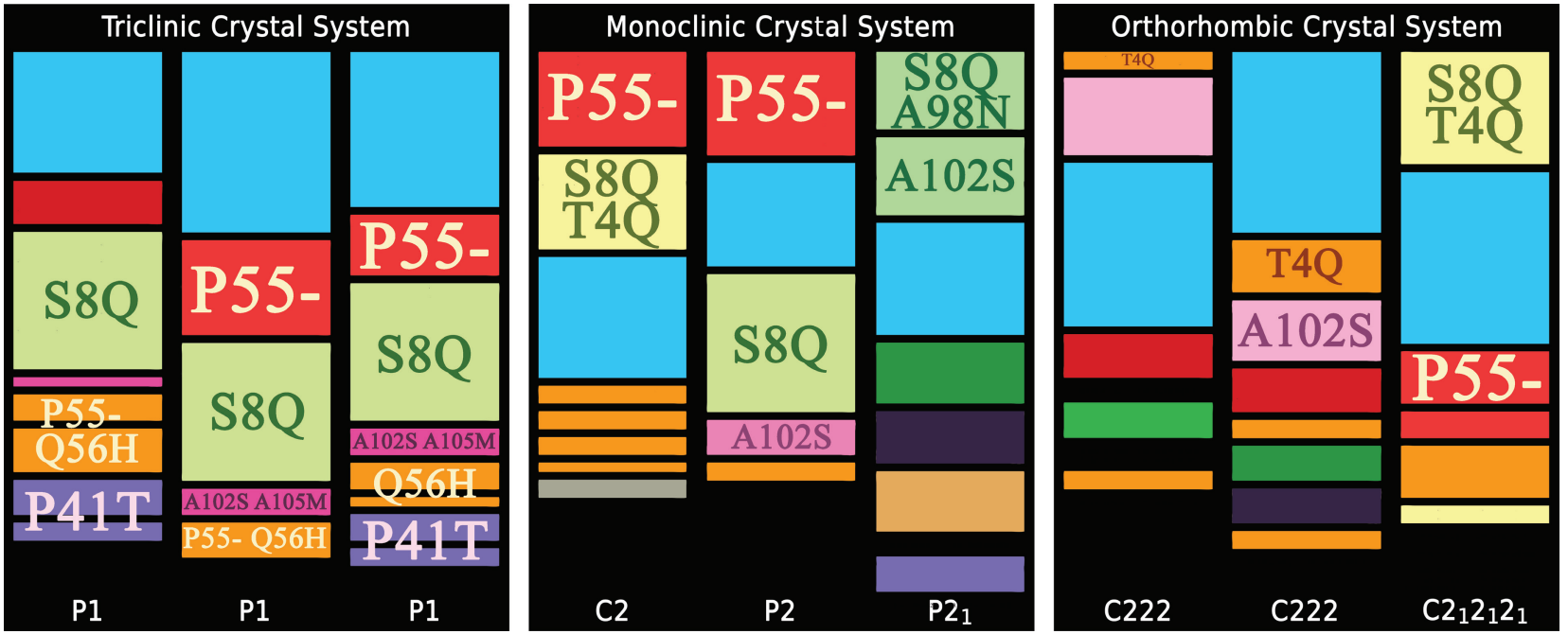
Most of the above mentioned inhibitor compositions were used for cocrystallization experiments with the EVL EVH1 constructs as well. However, crystallization of EVL EVH1 turned out to be difficult and worked only with the inhibitor Ac–[2-Cl-Phe]PP[ProM-9]–OH in complex with EVL EVH1 $_{\Delta AT,tail}$ (Sec. 8.1.1).

The fine-screening yielded in two data sets diffracting to 1.84 and 1.55Å. However, the latter data set showed severe problems with ice-ring sensitive areas. As the two crystals were satisfyingly isomorphous with a maximal deviation of 3.8% along axis c, data sets were merged and scaled with either the cells of the data sets and additionally with a mixed cell dimension. Either of these three cell dimensions allowed a high-resolution cut-off at roughly 1.6Å ($CC_{1/2}$ 0.50). However, paired refinement showed that the models of the merged data sets converged significantly slower than the single 1.84Å data set. After placing the inhibitors, the merged data sets were refined at 1.84Å for comparison and showed R_{free} of 34-36%, approximately 10% above the unmerged models. All models were refined with isotropic atomic displacement parameters (ADPs). Comparison of the models was difficult as the TLS groups calculated by PHENIX[201] caused for some models a reduced Hamilton R ratio.[197] PDB_REDO[171] accordingly refined with one TLS group per chain to prevent over-fitting. For the 1.55Å single data as well as the merged data with this cell, 9 TLS groups were used. The best model reached R_{free} of 25% but was not deposited.

19.2. Design of new VASP EVH1 constructs

As many binding studies were performed with VASP EVH1, it was the main crystallization target in the beginning and again towards the end of the thesis. Motivated by articles reviewing protein engineering for crystallization,[150,151] the so-far crystallized ENAH- and EVL EVH1 complexes were launched on the PDBePISA sever to calculate the macromolecular interfaces.[203] Crystal contacts were superpositioned, examined, classified, and scored upon similarity. The graphical representation (Fig. 19.1) revealed that several crystal contacts seemed to be of crucial importance and were established in all solved structures, independent of the lattice system. On the other hand, a certain lattice system was not defined by a unique set of contacts. As it tuned out, the relative orientation of ENAH EVH1 domains created similar contacts even within the same space group. The aim of the scoring in figure 19.1 was not to find correlations between the crystal contacts and the space group, but to find point mutations of VASP EVH1 which might have abolished crystal growth.

Knowing the indispensable crystal contacts enabled a direct comparison of the sequences of VASP EVH1 and ENAH EVH1 at these crucial positions in the protein fold. To introduce the least mutations while opening as many contact-combinations as possible, three VASP EVH1 mutations were designed. A serine to glutamine mutation at position 8 (S8Q, VASP EVH1 $_{\mathrm{S8Q}}$) and a minimally mutated (VASP EVH1 $_{\mathrm{min}}$) construct enabling crystal growth with the lowest symmetry. The third construct exhibited 9 mutations (VASP EVH1 $_{\mathrm{9mut}}$) to allow most contacts used in the monoclinic and orthorhombic crystal systems. The sequence alignment is shown in table 19.1. Except for the point mutation (whose forward primer is found in table 8.2), the inserts were ordered with a pEX-A2 vector backbone (eurofins), equipped with BaMHI (5') and XhoI (3') restriction sites, cloned into an empty pGEX-4T1 plasmid (clontech), purified and stored in water at -20 °C. While the expression of VASP EVH1 $_{\mathrm{S8Q}}$ was comparable to the initial construct, expression levels of VASP EVH1 $_{\mathrm{min}}$ and VASP EVH1 $_{\mathrm{9mut}}$ were notably lower. Extensive crystallization with VASP EVH1 $_{\mathrm{S8Q}}$ did not yield in any altered behavior in the screens. Neither rising the temperature nor using cross-seeding with ENAH EVH1 crystals was successful. VASP EVH1 $_{\mathrm{min}}$ was used in three fine-screens but did not crystallize.



SI Fig. 19.1.: **Scoring of the crystal contacts of 9 ENAH EVH1 structures**
Interfaces of each structure (in rows) found by PDBePISA sever and grouped by the crystal systems. The crystal contacts were color-coded according to their similarity. The involved amino acids within the crystal contact were compared with the sequence of VASP EVH1. This revealed point mutations of VASP EVH1 residues enabling the crystal contact found in ENAH EVH1 crystals. Crystal contacts without a mentioned point mutation are established by amino acids conserved between ENAH and VASP.

Construct	Sequence	
VASP EVH1 $_{ini}$	MSETVICSSRATVMLYDDGNKRWLPAGTGPQAFSRVQIYH	40
VASP EVH1 $_{S8Q}$	MSETVICQSRATVMLYDDGNKRWLPAGTGPQAFSRVQIYH	40
VASP EVH1 $_{min}$	MSETVICQSRATVMLYDDGNKRWLPAG-GPQAFSRVQIYH	39
VASP EVH1 $_{9mut}$	MSEQVICQSRATVMLYDDGNKRWLPAG-GPQAFSRVQIYH	38
ENAH EVH1 $_{ini}$	MSEQSICQARAAVMVYDDANKKWVPAG-GSTGFSRVHIYH	48
	*** **.:**:**:***.**:*:*** *. .****:***	
VASP EVH1 $_{ini}$	NPTANSFRVVGRKMQPDQQVVINCAIVRGVKYNQATPNFH	80
VASP EVH1 $_{S8Q}$	NPTANSFRVVGRKMQPDQQVVINCAIVRGVKYNQATPNFH	80
VASP EVH1 $_{min}$	NTTANSFRVVGRKMQPDQQVVINCAIVRGVKYNQATPNFH	79
VASP EVH1 $_{9mut}$	NTTANSFRVVGRKMQ-DHQVVINCAIVRGVKYNQATPNFH	78
ENAH EVH1 $_{ini}$	HTGNNTFRVVGRKIQ-DHQVVINCAIPKGLKYNQATQTFH	78
	:. *:********:* *:********* :*:****** .**	
VASP EVH1 $_{ini}$	QWRDARQVWGLNFGSKEDAAQFAAGMASALEALEG	115
VASP EVH1 $_{S8Q}$	QWRDARQVWGLNFGSKEDAAQFAAGMASALEALEG	115
VASP EVH1 $_{min}$	QWRDARQVWGLNFGSKEDAAQFAAGMASALEAL--	112
VASP EVH1 $_{9mut}$	QWRDARQVWGLNFGSKEDANQFASGMMSALEAL--	111
ENAH EVH1 $_{ini}$	QWRDARQVYGLNFGSKEDANVFASAMMHALEVL--	111
	********:********** **:.* ***.*	

SI Tab. 19.1.: **Comparison of VASP and ENAH EVH1 protein sequences**
The mutation variants of VASP EVH1 are listed in between both wild type sequences and include 1 to 9 of totally 35 mutations.

20. Resonance peak tables of ^{1}H-^{15}N-HSQC experiments

The peak tables of the ^{1}H-^{15}N-HSQC measurements are listed here in the end of the supplementary part. The data are not mentioned further in the discussion, but appear color-coded on the surface of ENAH EVH1 in the results and the discussion.

As mentioned in methods part 8.5, the ligand excess was calculated via nominal concentrations. Hence, the following table include M-value, K_{d}-assumption and the maximal protein saturation ($[\mathrm{PL}]/\mathrm{P}_{\mathrm{tot}_{\mathrm{max}}}$) on which the titration data was based. For a better overview, only the maximal changes in chemical shift ($\Delta\delta_{\mathrm{TOTAL}}$) are given in ppm.

Ligand	Ac–SFEFPPPPTEDEL–NH_2		
Target	ENAH EVH1	ENAH EVH1	EVL EVH1
$[PL]/P_{tot\,max}$	98%	99.5% (18.5× excess)	98%
K_d [µM]	10	10	7.3
M	0.8	0.8	0.8

Residue	$\Delta\delta_{TOTAL}$	Residue	$\Delta\delta_{TOTAL}$	Residue	$\Delta\delta_{TOTAL}$
W23NE-HE	0.43	R81N-HN	0.54	?-?	0.51
A75N-HN	0.18	W23NE-HE	0.44	W23NE-HE	0.42
Q81N-HN	0.15	W23N-HN	0.33	W23HN-N	0.41
H80N-HN	0.14	A73N-HN	0.23	?-?	0.38
N92ND2-HD22	0.12	Q79N-HN	0.19	A74HN-N	0.37
V26N-HN	0.12	H78N-HN	0.17	Q73HN-N	0.25
W80NE-HE	0.11	N90ND2-HD22	0.17	?-?	0.24
R47N-HN	0.10	V24N-HN	0.16	?-?	0.24
V86N-HN	0.10	N90ND2-HD21	0.16	H79HN-N	0.20
D17N-HN	0.09	V86N-HN	0.15	D18HN-N	0.18
A26N-HN	0.09	N44N-HN	0.15	D17HN-N	0.17
G88N-HN	0.09	W80N-HN	0.14	R11HN-N	0.16
N71N-HN	0.08	N71N-HN	0.14	V24HN-N	0.15
W23N-HN	0.08	E3N-HN	0.13	G89HN-N	0.14
A83N-HN	0.08	D17N-HN	0.13	W81HN-N	0.13
		G88N-HN	0.13	D83HN-N	0.12
		R47N-HN	0.13	V87HN-N	0.12
		W80NE-HE	0.12	?-?	0.11
		N61ND2-HD22	0.11	T77HN-N	0.11
		A26N-HN	0.11	Q81HN-N	0.10
		F91N-HN	0.10	V66HN-N	0.09
		N43N-HN	0.09	A98HN-N	0.09
		K69N-HN	0.09	?-?	0.09
		N61ND2-HD21	0.09		
		G31N-HN	0.09		
		A63N-HN	0.09		

SI Tab. 20.1.: **^{1}H-^{15}N-HSQC titration experiments of the ActA-derived 13-mer Ac–SFEFPPPPTEDEL–NH_2 binding to ENAH- and VASP EVH1**
Ligand was measured twice on ENAH EVH1 to detect the chemical shift changes of the second binding site (Sec. 13.2). Listed are the changes in chemical shift of the backbone (N-HN) and side chain resonance for the given residue in ppm. For a better overview, only changes in chemical shifts with an Euclidean distance of roughly 0.1 ppm and further are listed.

Peak	$K_{d,0}$ [µM]	$K_{d,M}$ [µM]	M	Pr(>F)	$\Delta\delta_{TOTAL}$ [ppm]
R81N.HN		7.0	0.76	4e-5	0.54
W23NE.HE	140	no fit			0.44
W23N.HN	54	K_d<0			0.33
A73N.HN		3.3	0.68	2e-9	0.23
Q79N.HN		4.3	0.66	2e-10	0.19
H78N.HN		2.3	0.69	5e-11	0.18
N90ND2.HD22		8.7	0.74	9e-6	0.17
V24N.HN		3.1	0.69	3e-11	0.16
N90ND2.HD21		1.3	0.76	4e-6	0.16
V86N.HN	77	K_d<0			0.15
W80N.HN	49	K_d<0			0.15
N44N.HN	98		M>1		0.15
E3N.HN	990			2e-7	0.14
N71N.HN	53		M>1	0.5	0.14
D17N.HN		6.7	0.74	8e-7	0.13
G88N.HN		5.9	0.70	6e-7	0.13
R47N.HN	2300	no fit			0.13
W80NE.HE		1.0	0.56	1e-6	0.13
N61ND2.HD22	4700	no fit			0.11
A26N.HN		1.4	0.66	4e-10	0.11
F91N.HN		20	0.79	7e-3	0.010
X..G31NNH		2.7	0.74	4e-6	0.094
K69N.HN		3.4	0.65	4e-9	0.090
N43N.HN	4700	no fit			0.090
N61ND2.HD21	7300	no fit			0.089
A63N.HN	220	no fit			0.088
S29N.HN	390	no fit			0.078
G50N.HN	210		M>1		0.075
Y70N.HN		12	0.76	4e-4	0.075
N90N.HN		6.2	0.75	1e-4	0.072
K22N.HN		4.7	0.68	1e-5	0.067
T76N.HN		2.0	0.62	2e-8	0.066
N71ND2.HD21	Kd<0		M>1		0.066
F77N.HN		11	0.76	1e-4	0.065
R51N.HN	82	no fit			0.064
D82N.HN	73		0.95	0.8	0.062
Q85NE2.HE22		1.0	0.71	1e-09	0.061
V58N.HN	410	no fit			0.057
A11N.HN	27		M>1		0.057
Y16N.HN	31	K_d<0			0.056
V15N.HN		0.74	0.66	2e-08	0.056
T74N.HN	40	K_d<0			0.052

SI Tab. 20.2.: **Affinity of Ac–SFEFPPPPTEDEL–NH_2 binding to ENAH EVH1 based on ^{1}H-^{15}N-HSQC titration**
Perturbations were fitted with equation 16.9, with and without M-value ($K_{d,M}$, $K_{d,0}$). Unreasonable fits (K_d< 0 and M > 1) were excluded and the remaining data compared by ANOVA. Depending on the significance level 0.05, the better estimate of K_d was chosen.

Ligand	Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH		
Target	ENAH EVH1	VASP EVH1	EVL EVH1
$[PL]/P_{tot\,max}$	98%	96%	98%
$K_{d,\,FT}$ [µM]	2.3 (0.2)	2.7 (0.7)	1.4 (0.2)

Residue	$\Delta\delta_{TOTAL}$	Residue	$\Delta\delta_{TOTAL}$	Residue	$\Delta\delta_{TOTAL}$
W23NE-HE	0.85	W23NE-HE	0.78	W23NE-HE	0.86
Q79N-HN	0.27	A75N-NH	0.38	W23N-NH	0.34
A73N-HN	0.25	W23N-NH	0.28	Q4HN-N	0.2
W80N-HN	0.19	Q81N-NH	0.21	Q80HN-N	0.24
V86N-HN	0.16	N78N-NH	0.2	?-?	0.21
N90ND2-HD21	0.15	V70N-NH	0.18	V87HN-N	0.2
H78N-HN	0.14	Y16N-NH	0.17	R82HN-N	0.2
N90ND2-HD22	0.14	V67N-NH	0.16	H79HN-N	0.2
T74N-HN	0.11	W82NE-HE	0.16	N72HN-N	0.18
K69N-HN	0.11	Q87N-NH	0.16	G89HN-N	0.16
G88N-HN	0.1	W82N-NH	0.16	D18HN-N	0.14
Y16N-HN	0.1	F93N-NH	0.15	R85HN-N	0.14
F91N-HN	0.09	N63N-NH	0.14	D17HN-N	0.14
A26N-HN	0.09	N73N-NH	0.13	?-?	0.12
		A26N-NH	0.12	T77HN-N	0.11
		N92ND1-HD1	0.12	?-?	0.11
		L15N-NH	0.11	V15HN-N	0.11
		G69N-NH	0.11	A84HN-N	0.11
		V36N-NH	0.1	A98HN-N	0.1
		T76N-NH	0.09	Q73HN-N	0.1
				F92HN-N	0.09
				I26HN-N	0.09
				D83HN-N	0.09

SI Tab. 20.3.: **^{1}H-^{15}N-HSQC titration experiments of the ligand Ac–[2-Cl-Phe][ProM-2][ProM-1]–OH binding to all three EVH1 paralogs**
Listed are the changes in chemical shift of the backbone (N-HN) and side chain resonance for the given residue in ppm. For a better overview, only changes in chemical shifts with an Euclidean distance of roughly 0.1 ppm and further are listed.

Ligand	Ac–WPPPPTEDEL–NH_2		Ac–SFE[2-Cl-Phe][ProM-2][ProM-1]TEDEL–NH_2	
Target	ENAH EVH1		VASP EVH1	
$[PL]/P_{tot\,max}$	98% (6× excess)		98.6%	
$K_{d,\,FT}$ [µM]	20.5		0.6	
M	0.7		0.8	
	Residue	$\Delta\delta_{TOTAL}$	Residue	$\Delta\delta_{TOTAL}$
	W23NE-HE	0.61	W23NE-HE	0.56
	?-?	0.58	V67N-NH	0.18
	Q79N-HN	0.32	Y16N-NH	0.15
	A73N-HN	0.32	D17N-NH	0.14
	A83N-HN	0.28	W82NE-HE	0.14
	W23N-HN	0.26	G69N-NH	0.12
	R81N-HN	0.26	V36N-NH	0.11
	H78N-HN	0.22	F93N-NH	0.11
	N71N-HN	0.20	N73N-NH	0.11
	V24N-HN	0.19	G94N-NH	0.10
	G88N-HN	0.19	R22N-NH	0.09
	N44N-HN	0.18	Q31NE1-HE1	0.09
	Y70N-HN	0.18	T76N-NH	0.09
	W80NE-HE	0.16		
	W80N-HN	0.16		
	N90ND2-HD22	0.16		
	?-?	0.15		
	N90ND2-HD21	0.15		
	V86N-HN	0.15		
	A26N-HN	0.14		
	D17N-HN	0.14		
	G31N-NH	0.14		
	V15N-HN	0.12		
	F91N-HN	0.10		
	K22N-HN	0.09		
	Q72NE2-HE22	0.09		
	Y16N-HN	0.09		
	R51N-HN	0.09		
	Q85NE2-HE22	0.09		
	T76N-HN	0.09		

SI Tab. 20.4.: **^{1}H-^{15}N-HSQC titration experiments of the ActA derived ligands**
Trp-containing 10-mer titrated to 6-times excess to detect changes in chemical shifts at the second binding site. The ProM-containing hybrid (**2b**) was measured to confirm the shifts reported by Ball *et al.*[106] For a better overview, only the Euclidean distances further than 0.1 ppm are listed.

21. Index of abbreviations

Abl	Abelson tyrosine-protein kinase
ActA	Actin assembly-inducing protein
ADP	atomic displacement parameter
Arp	Actin-related protein
ANOVA	analysis of variance
β-ME	2-mercaptoethanol
Cdc42	Cell division control protein 42 homolog
ΔG	(Gibbs) free energy of reaction
$\Delta\Delta G$	difference of two Gibbs free energies
FAK	Focal adhesion kinase
EDTA	ethylendiaminetetraacetic acid disodium salt dihydrate
ENAH	Protein enabled homolog
EVH	Ena/VASP homology
EVL	Ena/VASP-like protein
F-actin	fibrous actin
Fmoc	Fluorenylmethoxycarbonyl
FT	fluorescence titration
GSH	L-glutathion reduced
GST	Glutathione-(S)-Transferase
GYF	glycine-tyrosine-phenylalanine
^{1}H-^{15}N-HSQC	heteronuclear single quantum correlation experiment
IPTG	isopropyl-β-D-thiogalactoside
ISa	$\langle I/\sigma \rangle_{asymtotic}$
ITC	isothermal calorimetry
K_d	dissociation constant
LLG	log-likely hood gain
Lpd	Lamellipodin, Ras-associated and pleckstrin homology domains-containing protein 1 (RAPH1)

continued on next page

mGluR	Metabotropic glutamate receptor
NaPi	sodium phosphate
NCS	non-crystallographic symmetry
NPF	Nucleation promoting factor
OD_{600}	optical density at 600 nm per cm
PBS	Phosphate-buffered saline
PDB	protein data bank
pI	isoelectric point
PI3K	Phosphatidylinositol-4:5-bisphosphate 3-kinase
PPII	poly-*L*-proline type II helix
PRD	PRS-recognizing protein domain
ProM	Proline mimetic scaffold
PRS	proline-rich sequences
Ran	GTP-binding nuclear protein Ran
RanBP	Ran-binding protein
RAS	GTPase rat sarkoma
rmsd	root-mean-square deviation
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SH3	Src-homology 3
SOC	Super Optimal Broth, 20 mM glucose
Src	Proto-oncogene c-sarcoma
TCEP	Tris(2-carboxyethyl)phosphin
TLS	translation/libration/screw parameters
TFZ	translation function Z-score
VASP	Vasodilator-stimulated phosphoprotein
VCA	verprolin homology, connecting and acidic regions
WASP	Wiskott-Aldrich-syndrome protein
WAVE	WASP family verprolin-homologous protein
WIP	WAS/WASL-interacting protein

Bibliography

[1] P. S. Steeg, Nat Rev Cancer **16**, 201 (2016).

[2] S. McGuire, *World Cancer Report 2014*, volume 7, World Health Organization, International Agency for Research on Cancer Geneva, Switzerland, 2016.

[3] I. J. Fidler and S. Paget, Nat. Rev. Cancer **3**, 453 (2003).

[4] P. Friedl and S. Alexander, Cell **147**, 992 (2011).

[5] H. Paz, N. Pathak, and J. Yang, Oncogene **33**, 4193 (2014).

[6] C. Albiges-Rizo, O. Destaing, B. Fourcade, E. Planus, and M. R. Block, J. Cell. Sci. **122**, 3037 (2009).

[7] T. Shibue, M. W. Brooks, and R. A. Weinberg, Cancer Cell **24**, 481 (2013).

[8] R. L. Siegel, K. D. Miller, and A. Jemal, CA Cancer J Clin **65**, 5 (2015).

[9] A. Jemal et al., CA Cancer J Clin **55**, 10 (2005).

[10] A. Huttenlocher and A. R. Horwitz, Cold Spring Harb Perspect Biol **3**, a005074 (2011).

[11] M. Vicente-Manzanares, C. K. Choi, and A. R. Horwitz, J. Cell. Sci. **122**, 199 (2009).

[12] J. L. Sottnik et al., Clin. Exp. Metastasis **30**, 569 (2013).

[13] B. Zhou et al., Mol. Cancer Res. **12**, 143 (2014).

[14] W. Guo et al., Cell **126**, 489 (2006).

[15] M. A. Dechantsreiter et al., J. Med. Chem. **42**, 3033 (1999).

[16] J. S. Desgrosellier and D. A. Cheresh, Nat. Rev. Cancer **10**, 9 (2010).

[17] M. P. Playford and M. D. Schaller, Oncogene **23**, 7928 (2004).

[18] N. P. Shah et al., Am. J. Hematol. **90**, 1060 (2015).

[19] W. P. Mason, Neuro-oncology **17**, 634 (2015).

[20] M. Campone et al., Ann. Oncol. **23**, 610 (2012).

[21] L. M. Coussens, B. Fingleton, and L. M. Matrisian, Science **295**, 2387 (2002).

[22] M. Egeblad and Z. Werb, Nat. Rev. Cancer **2**, 161 (2002).

[23] A. Schultze and W. Fiedler, Expert Opin Investig Drugs **19**, 777 (2010).

[24] J. R. Infante et al., J. Clin. Oncol. **30**, 1527 (2012).

[25] B. Y. Lee, P. Timpson, L. G. Horvath, and R. J. Daly, Pharmacol. Ther. **146**, 132 (2015).

[26] G. Carmona et al., Oncogene (2016).

[27] A. L. Law et al., J. Cell Biol. **203**, 673 (2013).

[28] L. Blanchoin, R. Boujemaa-Paterski, C. Sykes, and J. Plastino, Physiol Rev **94**, 235 (2014).

[29] M. Krause, E. W. Dent, J. E. Bear, J. J. Loureiro, and F. B. Gertler, Annu Rev Cell Dev Biol **19**, 541 (2003).

[30] L. A. Cameron, P. A. Giardini, F. S. Soom, and J. A. Theriot, Nat Rev Mol Cell Biol **1**, 110 (2000).

[31] R. Fischer, N. Zekert, and N. Takeshita, Molecular Microbiology **68**, 813 (2008).

[32] T. D. Pollard and G. G. Borisy, Cell **112**, 453–465 (2003).

[33] B. J. Wallar and A. S. Alberts, Trends Cell Bio **13**, 435 (2003).

[34] T. D. Pollard, Annu Rev Biophys Biomol Struct **36**, 451 (2007).

[35] M. A. Chesarone, A. G. DuPage, and B. L. Goode, Nat Rev Mol Cell Biol **11**, 62 (2010).

[36] A. Mogilner and G. Oster, Biophys J **71**, 3030 (1996).

[37] K. Niebuhr et al., EMBO J **16**, 5433 (1997).

[38] R. C. May et al., Curr Biol **9**, 759 (1999).

[39] J. E. Bear and F. B. Gertler, J Cell Sci **122**, 1947 (2009).

[40] V. Noireaux et al., Biophys. J. **78**, 1643 (2000).

[41] A. Kawska et al., Proc Natl Acad Sci U S A **109**, 14440 (2012).

[42] B. J. Nolen et al., Nature **460**, 1031 (2009).

[43] S. Kurisu and T. Takenawa, Genome Biol **10**, 226.1 (2009).

[44] Z. Chen et al., Nature **468**, 533 (2010).

[45] Y. Kano, J. D. Cook, J. E. Lee, and M. Ohh, Semin. Cell Dev. Biol. (2016).

[46] R. Rohatgi et al., Cell **97**, 221 (1999).

[47] S. Eden, R. Rohatgi, A. V. Podtelejnikov, M. Mann, and M. W. Kischner, Nature **418**, 790 (2002).

[48] S. K. Mitra and D. D. Schlaepfer, Curr. Opin. Cell Biol. **18**, 516 (2006).

[49] M. Jacob et al., Cell. Signal. **21**, 1308 (2009).

[50] G. Feldmann et al., Cancer Res. **70**, 4460 (2010).

[51] H. Merk et al., Oncotarget **7**, 6088 (2016).

[52] D. Breitsprecher et al., EMBO J. **30**, 456 (2011).

[53] S. Romero et al., Cell **119**, 419 (2004).

[54] L. Haviv et al., Proc Natl Acad Sci U S A **103**, 4906 (2006).

[55] A. Schirenbeck et al., Proc Natl Acad Sci U S A **103**, 7694 (2006).

[56] C. Bachmann, L. Fischer, U. Walter, and M. Reinhard, J. Biol. Chem. **274**, 23549 (1999).

[57] X. J. Chen et al., Dev Cell **30**, 569 (2014).

[58] H. Okada et al., PLoS One **7**, e37035 (2012).

[59] S. Havrylenko et al., Mol. Biol. Cell **26**, 55 (2015).

[60] M. Barzik, L. M. McClain, S. L. Gupton, and F. B. Gertler, Mol Biol Cell **25**, 2604 (2014).

[61] Y. H. Han et al., J. Biol. Chem. **277**, 49877 (2002).

[62] C. Sarmiento et al., J Cell Biol **180**, 1245 (2008).

[63] F. B. Gertler, K. Niebuhr, M. Reinhard, J. Wehland, and P. Soriano, Cell **87**, 227 (1996).

[64] J. E. Bear et al., Cell **101**, 717 (2000).

[65] E. Lafuente et al., Dev Cell **7**, 585 (2004).

[66] M. Krause et al., Dev Cell **7**, 571 (2004).

[67] A. J. Ridley, Cell **145**, 1012 (2011).

[68] P. Beli, D. Mascheroni, D. Xu, and M. Innocenti, Nat Cell Biol **10**, 849 (2008).

[69] C. Bilancia et al., Dev Cell **28**, 394 (2014).

[70] H. N. Higgs and T. D. Pollard, J. Biol. Chem. **274**, 32531 (1999).

[71] L. Blanchoin et al., Nature **404**, 1007 (2000).

[72] J. B. Marchand, D. A. Kaiser, T. D. Pollard, and H. N. Higgs, Nat Cell Biol **3**, 76 (2001).

[73] M. Innocenti et al., Nat. Cell Biol. **6**, 319 (2004).

[74] Z. Ding et al., Oncogene **33**, 2065 (2014).

[75] M. Reinhard, M. Rudiger, B. M. Jockusch, and U. Walter, FEBS **399**, 103 (1996).

[76] B. Drees et al., J. Biol. Chem. **275**, 22503 (2000).

[77] A. S. Menzies et al., J. Neurosci. **24**, 8029 (2004).

[78] S. L. Gupton and F. B. Gertler, Dev. Cell **18**, 725 (2010).

[79] A. V. Kwiatkowski et al., Neuron **56**, 441 (2007).

[80] E. T. Roussos et al., Breast Cancer Res. **12**, R101 (2010).

[81] D. Hanahan and R. A. Weinberg, Cell **144**, 646 (2011).

[82] D. S. Zuzga et al., Int. J. Cancer **130**, 2539 (2012).

[83] C. V. Rajadurai et al., J. Cell Biol. **214**, 719 (2016).

[84] M. J. Oudin et al., Cancer Discov **6**, 516 (2016).

[85] R. L. Klemke, Curr. Opin. Cell Biol. **24**, 662 (2012).

[86] F. Di Modugno et al., Clin Cancer Res **12**, 1470 (2006).

[87] L. Dertsiz et al., Thorax **60**, 576 (2005).

[88] L. D. Hu, H. F. Zou, S. X. Zhan, and K. M. Cao, Oncol. Rep. **19**, 1015 (2008).

[89] F. Di Modugno et al., PLoS One **5**, e15852 (2010).

[90] P. Tucker, I. Evans, and W. Wood, Dis Model Mech **4**, 126 (2011).

[91] J. E. Bear et al., Cell **109**, 509 (2002).

[92] M. Reinhard, T. Jarchau, and U. Walter, Trends Biochem Sci **26**, 243 (2001).

[93] M. Reinhard, K. Jouvenal, D. Tripier, and U. Walter, Proc Natl Acad Sci U S A **92**, 7956 (1995).

[94] T. Chakraborty et al., EMBO J **14**, 1314 (1995).

[95] S. Goswami et al., Clin. Exp. Metastasis **26**, 153 (2009).

[96] F. Gertler and J. Condeelis, Trends Cell Biol. **21**, 81 (2011).

[97] F. Di Modugno et al., Proc. Natl. Acad. Sci. U.S.A. **109**, 19280 (2012).

[98] M. J. Oudin et al., Clin. Exp. Metastasis **33**, 249 (2016).

[99] B. F. Volkman, K. E. Prehoda, J. A. Scott, F. C. Peterson, and W. A. Lim, Cell **111**, 565 (2002).

[100] F. C. Peterson and B. F. Volkman, Front Biosci (Landmark Ed) **14**, 833 (2009).

[101] J. T. Nguyen et al., Chem. Biol **7**, 463 (2000).

[102] H. Li and D. Lawrence, Chem Biol **12**, 905 (2005).

[103] L. J. Ball, T. Jarchau, H. Oschkinat, and U. Walter, FEBS **513**, 45 (2002).

[104] M. Zettl and M. Way, Curr. Biol. **12**, 1617 (2002).

[105] R. Opitz, *Inhibition von Protein-Protein-Wechselwirkungen: Peptidomimetika für VASP-EVH1*, PhD thesis, Freie Universität Berlin, Fachbereich Biologie, Chemie, Pharmazie, 2012.

[106] L. J. Ball et al., EMBO J **19**, 4903 (2000).

[107] M. Barzik et al., J Mol Biol **309**, 155 (2001).

[108] B. K. Kay, M. P. Williamson, and P. Sudol, Faseb J **14**, 231 (2000).

[109] M. Kuemin, S. Schweizer, C. Ochsenfeld, and H. Wennemers, J Am Chem Soc **131**, 15474 (2009).

[110] P. M. Cowan, S. McGavin, and A. C. North, Nature **176**, 1062 (1955).

[111] F. Rabanal, M. D. Ludevid, M. Pons, and E. Giralt, Biopolymers **33**, 1019 (1993).

[112] M. P. Williamson, Biochem. J. **297 (Pt 2)**, 249 (1994).

[113] A. A. Adzhubei and M. J. E. Sternberg, J Mol Biol **229**, 472 (1993).

[114] B. J. Stapley and T. P. Creamer, Protein Science **8**, 587 (1999).

[115] G. Rubin, M. Yandell, and S. Lewis, Sci **287**, 2204 (2000).

[116] D. E. Scott, A. R. Bayly, C. Abell, and J. Skidmore, Nat Rev Drug Discov **15**, 533 (2016).

[117] B. Aillard et al., Org Biomol Chem **13**, 4562 (2015).

[118] D. Witter, S. Famiglietti, J. Cambier, and A. Castelhano, Bioorg Med Chem Lett **17**, 3137 (1998).

[119] P. Seneci et al., Bioorg. Med. Chem. **17**, 5834 (2009).

[120] H. T. Yu et al., Cell **76**, 933 (1994).

[121] A. B. Sparks et al., Proc Natl Acad Sci U S A **93**, 1540 (1996).

[122] M. Reinhard et al., EMBO J **14**, 1583 (1995).

[123] N. M. Mahoney, P. A. Janmey, and S. C. Almo, Nat Struct Biol **4**, 953 (1997).

[124] C. E. Schutt, J. C. MYSLIK, M. D. ROZYCKI, N. C. W. GOONESEKERE, and U. LINDBERG, Nature **365**, 810 (1993).

[125] H. I. Chen and M. Sudol, Proc Natl Acad Sci U S A **92**, 7819 (1995).

[126] M. J. Macias et al., Nature **382**, 646 (1996).

[127] K. Nishizawa, C. Freund, J. Li, G. Wagner, and E. L. Reinherz, Proc Natl Acad Sci U S A **95**, 14897 (1998).

[128] C. Freund, V. Dotsch, K. Nishizawa, E. L. Reinherz, and G. Wagner, Nat Struct Biol **6**, 656 (1999).

[129] O. Pornillos, S. L. Alam, D. R. Davis, and W. I. Sundquist, Nat Struct Biol **9**, 812 (2002).

[130] O. Pornillos et al., EMBO J **21**, 2397 (2002).

[131] S. Feng, J. K. Chen, H. Yu, J. A. Simon, and S. L. Schreiber, Science **266**, 1241 (1994).

[132] D. Golemi-Kotra et al., J Am Chem Soc **126**, 4 (2004).

[133] A. Hodges and A. Schepartz, J Am Chem Soc **129**, 11024 (2007).

[134] J. Holtzman, K. Woronowicz, D. Golemi-Kotra, and S. A, Biochemistry **46**, 13541 (2007).

[135] J. H. Holtzman, *Remodeling miniature proteins that target EVH1 domains with high affinity and specificity*, PhD thesis, Yale University, 2008.

[136] J. Zimmermann et al., J Biol Chem **278**, 36810–36818 (2003).

[137] C. Hunke, T. Hirsch, and J. Eichler, Chembiochem **7**, 1258 (2006).

[138] J. Nguyen, C. Turck, F. Cohen, R. Zuckermann, and W. Lim, Sci **282**, 2088 (1998).

[139] J. Morken, T. Kapoor, S. Feng, F. Shirai, and S. L. Schreiber, J Am Chem Soc **120**, 30 (1998).

[140] J. Zaminer et al., Angew Chem Int Ed Engl **49**, 7111 (2010).

[141] R. Opitz et al., Proc Natl Acad Sci U S A **112**, 5011 (2015b).

[142] C. Reuter et al., Chemistry **21**, 8464 (2015).

[143] V. Hack et al., Angew Chem Int Ed Engl **52**, 9539 (2013).

[144] A. A. Fedorov, E. Fedorov, F. Gertler, and S. C. Almo, Nat Struct Biol **6**, 661 (1999).

[145] E. Sayers et al., Nucleic Acids Res **37**, D5 (2009).

[146] D. Thierry-Mieg and J. Thierry-Mieg, Genome Biol **7**, S12.1 (2006).

[147] C. Reuter, J. Neudörfl, and S. HG, Acta Crystallogr Sect E Struct Rep **E66** (2010).

[148] P. Huy, J. Neudörfl, and S. HG, Org. Lett. **13**, 216 (2011).

[149] M. Beyermann, K. Fechner, J. Furkert, E. Krause, and M. Bienert, J Med Chem **39**, 3324 (1996).

[150] Z. Derewenda, Acta Crystallographica **66**, 604 (2010).

[151] D. Fusco, J. Headd, D. Simone, J. Wang, and P. Charbonneau, Soft Matter **10**, 290 (2014).

[152] W. Kabsch, J Appl Cryst **26**, 795 (1993).

[153] K. Diederichs, Acta Crystallogr D Biol Crystallogr **66**, 733 (2010).

[154] K. Diederichs, 2vb1, `http://strucbio.biologie.uni-konstanz.de/xdswiki/index.php/2VB1`, 2011.

[155] P. A. Karplus and K. Diederichs, Sci **336**, 1030 (2012).

[156] P. A. Karplus and K. Diederichs, Curr Opin Struct Biol **22**, 60 (2015).

[157] K. Diederichs, *Crystallographic Data and Model Quality*, Springer New York, 2016.

[158] A. J. McCoy et al., J Appl Crystallogr. **40**, 658 (2007).

[159] E. Potterton, P. Briggs, M. Turkenburg, and E. Dodson, Acta Crystallogr D Biol Crystallogr **59**, 1131 (2003).

[160] P. Johannesson, Moloc: using prolog for conceptual modelling, 1991, http://www.moloc.ch.

[161] A. A. Lebedev et al., Acta Crystallogr D Biol Crystallogr **68**, 431 (2012).

[162] A. W. Schüttelkopf and D. M. F. van Aalten, Acta Crystallogr D Biol Crystallogr **60**, 1355 (2004).

[163] P. R. Gerber, J Comput Aided Mol Des **12**, 37 (1998).

[164] L. F. Ten Eyck, Meth. Enzymol. **115**, 324 (1985).

[165] P. V. Afonine et al., Acta Crystallogr. D Biol. Crystallogr. **71**, 646 (2015).

[166] G. N. Murshudov, A. A. Vagin, and E. J. Dodson, Acta Crystallogr D Biol Crystallogr **53**, 240 (1997).

[167] P. D. Adams et al., Acta Crystallogr D Biol Crystallogr **66**, 213 (2010).

[168] P. Emsley and K. Cowtan, Acta Crystallogr D Biol Crystallogr **60**, 2126 (2004).

[169] V. Chen et al., Acta Crystallogr D Biol Crystallogr **66**, 12 (2010).

[170] I. Davis et al., Nucleic Acids Res **35**, W375 (2007).

[171] R. P. Joosten, F. Long, G. N. Murshudov, and A. Perrakis, IUCrJ **1**, 213 (2014).

[172] D. G. Goddard, T. D. andKneller, Sparky 3, University of California, San Francisco.

[173] H. N. Heine, *Peptidmimetika an Zellulosemembranen – SPOT-Synthese und Screening kombinatorischer Peptoid-Bibliotheken*, PhD thesis, Mathematisch-Naturwissenschaftlichen Fakultät I der Humboldt-Universität zu Berlin, 2000.

[174] K. E. Prehoda, D. J. Lee, and W. A. Lim, Cell **97**, 471 (1999).

[175] W. C. Hamilton, Acta Crystallogr **18**, 502 (1965).

[176] A. T. Brünger, Nature **355**, 472 (1992).

[177] A. M. Petros et al., J. Med. Chem. **49**, 656 (2006).

[178] M. Müller, *Charakterisierung und Validierung von Ena/Vasodilatator stimulierendes Phosphoprotein (VASP) Homologie 1 (EVH1)-Domänen-Inhibitoren*, PhD thesis, Freie Universität Berlin, Fachbereich Biologie, Chemie, Pharmazie, 2016.

[179] F. Sayyed and C. Suresh, Phys Chem Chem Phys **115**, 5660 (2011).

[180] B. Bochicchio and A. M. Tamburro, Chirality **14**, 782 (2002).

[181] A. A. Adzhubei, M. J. Sternberg, and A. A. Makarov, J. Mol. Biol. **425**, 2100 (2013).

[182] T. Creamer and M. Campbell, Adv Protein Chem **62**, 263 (2002).

[183] C. Abad-Zapatero, Expert Opin Drug Discov **2**, 469 (2007).

[184] B. Rost, J Struct Biol **134**, 204 (2001).

[185] D. W. A. Buchan, F. Minneci, T. C. O. Nugent, K. Bryson, and D. T. Jones, Nucleic Acids Res **41**, 340 (2013).

[186] S. Pistor, T. Chakraborty, U. Walter, and J. Wehland, Curr. Biol. **5**, 517 (1995).

[187] G. A. Smith, J. A. Theriot, and D. A. Portnoy, J Cell Biol **135**, 647 (1996).

[188] G. D. Aurora, R andRose, Protein Sci **7**, 21 (1998).

[189] L. J. Ball, R. Kuhne, J. Schneider-Mergener, and H. Oschkinat, Angew Chem Int Ed Engl **44**, 2852 (2005).

[190] J. Beneken et al., Neuron **26**, 143 (2000).

[191] J. B. Mitchell, J. M. Thornton, J. Singh, and S. L. Price, J. Mol. Biol. **226**, 251 (1992).

[192] N. P. J. Brindle, M. R. Holt, J. E. Davies, C. J. Price, and D. R. Critchley, Biochem J **318**, 753 (1996).

[193] I. R. Vetter, C. Nowak, T. Nishimoto, J. Kuhlmann, and A. Wittinghofer, Nature **398**, 39 (1999).

[194] D. M. Stewart, L. Tian, and D. L. Nelson, J. Immunol. **162**, 5019 (1999).

[195] F. Studier, Protein Expr Purif **41**, 207 (2005).

[196] A. Sivashanmugam et al., Protein Sci **18**, 936 (2009).

[197] A. Bacchi, V. S. Lamzin, and K. S. Wilson, Acta Crystallogr. D Biol. Crystallogr. **52**, 641 (1996).

[198] E. A. Merritt, Acta Crystallogr. D Biol. Crystallogr. **68**, 468 (2012).

[199] A. A. Lebedev, A. A. Vagin, and G. N. Murshudov, Acta Crystallogr. D Biol. Crystallogr. **62**, 83 (2006).

[200] G. N. Murshudov, Appl. Comput. Math **10**, 250 (2011).

[201] P. V. Afonine, A. Urzhumtsev, R. W. Grosse-Kunstleve, and P. D. Adams, Computational Crystallography Newsletter **1**, 24 (2010).

[202] B. W. Matthews, J. Mol. Biol. **33**, 491 (1968).

[203] E. Krissinel and K. Henrick, J Mol Biol **372**, 774 (2007).