**This is a postprint of an article published in
Strijowski, U., Hirsch, T., Quintilla, A., Wenzel, W., Eichler, J.
Structural and functional mimicry of the binding site of hYAP-WW domain
for proline-rich ligands
(2007) International Journal of Peptide Research and Therapeutics, 13 (1-
2), pp. 245-250.**

Published in: *International Journal of Peptide Research and Therapeutics* **13** (1-2), 245-250, 2007.

Structural and Functional Mimicry of the Binding Site of hYAP-WW Domain for Proline-Rich Ligands

Ulf Strijowski¹, Tatjana Hirsch¹, Aina Quintillá², Wolfgang Wenzel² and Jutta Eichler^{1*}

¹Helmholtz Centre for Infection Research, Braunschweig, Germany.

² Research Centre Karlsruhe, Institute for Nanotechnology, Karlsruhe, Germany.

*To whom correspondence should be addressed: Jutta Eichler, Helmholtz Centre for Infection Research, 38124 Braunschweig, Germany. Tel: +49-531-6181-3110, Fax: +49-531-6181-3499; email: jutta.eichler@helmholtz-hzi.de

Running Title: WW domain binding site mimetic

ABSTRACT

Based on the 3D structure of the WW domain of human Yes-associated protein (hYAP-WW) in complex with a proline-rich peptide ligand, we have designed and synthesized a cyclic peptide covering a fragment of hYAP-WW that contains its primary contact residues for the interaction with the ligand. This peptide was found to specifically recognize a proline-rich ligand for hYAP-WW. Its conformation was calculated using molecular dynamics simulation, based on long-range NOEs identified by NMR spectroscopy, and indicates an arrangement of primary contact residues similar to hYAP-WW.

KEY WORDS: WW domain, structure-based design, proteinmimetics, biomimetic synthesis, proline-rich ligands

INTRODUCTION

Specific interactions between proteins and their ligands are the molecular basis of essentially all biological processes. The exploration of these interactions at the molecular and atomic level is an important step towards the modulation of protein function through controlled interference with the underlying binding events. The design and generation of molecules capable of mimicking conformationally defined binding and/or functional sites of natural proteins, represents a promising strategy for the exploration and understanding of protein structure and function. Such proteinmimetics are also useful tools for a range of biomedical applications, including the inhibition of protein-ligand interactions (Eichler 2005).

WW domains, which owe their name to the presence of two highly conserved tryptophan residues in their sequences, are relatively short (35-45 amino acid residues) protein interaction domains in various signaling and regulatory proteins, which are involved in signaling pathways underlying several human diseases (Chen and Sudol 1995). Similar to SH3 (Yu et al. 1994), EVH1 (Niebuhr et al. 1997) and GYF (Nishizawa et al. 1998) domains, WW domains recognize proline-rich ligand motifs in other proteins (Ball et al. 2005, Kay et al. 2000). NMR analysis of complexes of the WW domain of human Yes-associated protein (hYAP-WW) with proline-rich peptide ligands revealed that the ligand binding site of hYAP-WW is located within the $\beta 2$ and $\beta 3$ strands of a three-stranded antiparallel β -sheet in the C-terminal part of the protein (Fig. 1) (Macias et al. 1996, Pires et al. 2001). The side chains of hydrophobic residues within this sequence stretch (Y28, L30, and W39) form a hydrophobic patch on the protein surface, and constitute, along with H32, the primary contact residues of hYAP-WW for its interaction with proline-rich ligands. The importance of these residues for the interaction with the ligand was also confirmed using arrays of synthetic hYAP-WW variants, which addressed each position of the protein sequence individually (Toepert et al. 2001, Toepert et al. 2003).

As part of our ongoing efforts aimed at mimicking conformationally defined protein binding sites (Eichler et al. 2004, Hunke et al. 2006, Franke et al. 2007), the goal of this study was to mimic the binding site of hYAP-WW by a synthetic peptide presenting a fragment (27-40) of the protein, which covers two strands of the β -sheet that contain the primary contact residues (Y28, L30, H32 and W39). In order to assess the influence of a conformational constraint on the binding behaviour of this peptide, it was synthesized in cyclic, as well as in linear shape.

MATERIALS AND METHODS

Peptide Synthesis

Peptides (see Table 1 for sequences) were synthesized by automated solid-phase synthesis, as previously described (Franke et al. 2004). Briefly, peptides were generated as C-terminal amides on polyoxyethylene-grafted polystyrene resin, to which the Rink amide linker was attached (TentaGel S RAM resin). Fmoc-amino acids were coupled as HOBt esters (5 eq., 2 x 1 hour), followed by capping with acetic anhydride/pyridine/DMF (1:2:3) for 15 min. In WWpep, WWpep* and WWpep**, an ivDde-side chain protected lysine residue was added to the C-terminus of the sequence, and the N-terminus was succinylated using 10eq. of succinic anhydride in 5% DIEA/DMF. After removal of the ivDde group (5% hydrazine in DMF, 7 x 5 min.), the peptides were cyclized on the resin through amide formation between the succinyl group at the N-terminus and the ϵ -amino group of the C-terminal lysine residue, using 2 eq. PyBOP/HOBt and 4 eq. DIEA (2 x 4 hours). Peptides were cleaved from the resin using a mixture of TFA, DCM, water and triisopropylsilane (70:20:5:5, 2 x 3 hours), precipitated in a cold 1:1 mixture of t-butylmethyl ether and cyclohexane, extracted with water, and lyophilized. Crude peptides were purified by preparative RP-HPLC (Column: Nucleosil RP18, 250 x 10 mm; flow rate: 3 ml/min; gradient: 5-95% MeCN in H₂O, both containing 0.1% TFA, in 60 min; UV detection: 220 nm).

NMR Spectroscopy

WWpep** was dissolved at 5 mg/mL in H₂O/D₂O, 10:1. Two-dimensional ¹H NMR Spectra (COSY, TOCSY, mixing time 60 ms, and NOESY, mixing time 500 ms) were recorded using a Bruker AVANCE DMX 600 instrument (SFO1: 600.13 MHz, water suppression, locked on Deuterium signal of solvent, external Calibration). Spectra were processed with XWINNMR (Bruker), and signals were assigned using the program SPARKY (T.D. Goddard and D.G. Kneller, SPARKY 3, University of California, San Francisco). Peaks of NOE spectra were integrated and volumes calibrated towards the volume of the H β^{proR} -H β^{proS} -crosspeak of selected amino acids.

Molecular Dynamics Simulation

Starting from a structurally unbiased cyclic conformation, all-atom explicit water simulations of WWpep** were performed using the GROMACS molecular dynamics package (Berendsen et al. 1995, Lindahl et al. 2001). The system was solvated and equilibrated for 20 ps at 300K, 1atm, and then run for 1 ns using NMR-derived constraints to guide the simulation. The following constraints were used: Lys41H δ 1-Y28H β 1 (3.4 Å), Lys41H ϵ 1-Y28H β 2 (3.2 Å), Lys41H ϵ 1-Y28HN (3.8 Å), Q40H γ 1-F29H α (3.9 Å) and T37H α -N31HN (3.9 Å), in which Lys41 represents the C-terminally added lysine residue used for cyclization *via* its ϵ -amino group.

Binding Assay

hYAP-WW, hYAP-WW*, WWpep and WWpep*, respectively (100 μ L; 10, 5 and 2.5 μ M in phosphate buffer (0.1M, pH 7.2), were coated in duplicate overnight at 4°C to the wells of a 96-well sulfhydryl binding assay plate (Corning). Unspecific binding was blocked with 1% BSA in phosphate buffer containing 0.01% Tween 20, for one hour, followed by

incubation with ligand solution (Biotin-Ahx-EYPPYPPPPYPSG-NH₂, 100 µL, 15 µM in 0.1% BSA/phosphate buffer containing 0.01% Tween 20) for 90 min. Plate-bound ligand was detected using an anti-biotin-HRP conjugate (1 hour, dilution: 1:1000 in phosphate buffer containing 0.01% Tween 20), followed by developing the plate in the dark with 100 µL OPD solution (1 mg/mL in 0.03% H₂O₂/water). The reaction was stopped after 10 min. by adding 50 µL 4 N sulfuric acid, and absorbances (ODs) were read at 490 nm using a plate reader (Spectramax 250).

RESULTS AND DISCUSSION

Complete hYAP-WW, a cyclic peptide covering its binding site (WWpep), a linear analog of WWpep, analogs of hYAP-WW and WWpep, in which three of the primary contact residues were replaced by alanine (hYAP-WW* and WWpep*), a truncated analog of WWpep lacking the C-terminal Ahx and cysteine residues (WWpep**), as well as a labeled proline-rich ligand for hYAP-WW, (see Table 1 for peptide sequences) were synthesized using automated Fmoc/tBu-based solid phase synthesis, and purified by preparative HPLC (Fig. 2). hYAP-WW and derived peptides were equipped with an additional cysteine residue for site-selective, covalent immobilization on assay plates. The ligand represents residues 742-754 of the p53-binding protein 2 (Naumovski and Cleary 1996), and was biotinylated for detection in the binding assay.

Table 1. Sequences of Synthesized Peptides

Peptide	Sequence
hYAp-WW	H-Cys-Ahx ^a -VPLPAGWEMAKTSSGQRYFLNHIDQTTTWQDPRKAMLS-NH ₂
hYAP-WW*	H-Cys-Ahx ^a -VPLPAGWEMAKTSSGQRAFANHIDQTTTAQDPRKAMLS-NH ₂
WWpep	<i>cyclo</i> [succ ^b -RYFLNHIDQTTTWQ-Lys]-Ahx ^a -Cys-NH ₂
WWpep*	<i>cyclo</i> [succ ^b -RAFANHIDQTTTAQ-Lys]-Ahx ^a -Cys-NH ₂
WWpep**	<i>cyclo</i> [succ ^b -RYFLNHIDQTTTWQ-Lys]-NH ₂
linear WWpep	Ac-RYFLNHIDQTTTWQ-Ahx-Cys-NH ₂
Ligand	Biotin-Ahx-EYPPYPPPPYPSG-NH ₂

^aAhx, -NH-(CH₂)₅-CO-; ^bsucc, -CO-(CH₂)₂-CO-

Binding affinities to the proline-rich ligand of hYAP-WW, WWpep and its linear analog, as well as hYAP-WW* and WWpep* were compared in a direct binding assay. As shown in Figure 3, WWpep retains approximately 20% of the affinity of hYAP-WW, while the linear analog of WWpep has no affinity to the ligand, suggesting that cyclization stabilizes conformations required for binding to the proline-rich ligand, likely by bringing into spatial proximity primary contact residues that are on opposite ends of the sequence of WWpep (Y28 and W39). Furthermore, replacement of three primary contact residues in hYAP-WW (hYAP-WW*), as well as in WWpep (WWpep*), with alanine resulted in complete loss of affinity to the ligand, suggesting a similar binding specificity of the complete domain and the mimetic peptide.

WWpep** was structurally characterized using two-dimensional ^1H NMR spectroscopy. In addition to the NOEs that were expected based on sequential proximity, such as $\text{HN}(i+1)\text{-H}\alpha(i)$, $\text{H}\alpha(i+1)\text{-H}\alpha(i)$, $\text{HN}(i+1)\text{-HN}(i)$, several long-range NOEs between spin systems of sequentially distant amino acids were observed, in particular between the $\text{C}\alpha$ proton of F29 and a $\text{C}\gamma$ proton of Q40, as well as between the amide proton of N31 and the $\text{C}\alpha$ proton of T37 (Fig. 4). Using these long-range NOEs as constraints, a conformation of WWpep** was calculated by molecular dynamics simulation. Conformations that obey all constraints at various time-intervals during the simulation are in agreement within a heavy atom RMSD of 1.5 Å (Fig. 5, left). Although the obtained structure does not involve a beta sheet (Fig. 5, right), the distances between, and arrangement of, the primary contact residues (Y28, L30, W39 and H32) are similar to those in hYAP-WW (Table 2), with the exception of the distance between Y28 and W39, which is almost twice as long in WWpep**. This similarity in presentation of contact residues is likely to be the structural basis of the ability of the mimetic peptide to interact with the proline-rich ligand.

Table 2. Distances between the C α of sequentially distant contact residues in the hYAP-WW – ligand complex and in WWpep**.

Contact residues	Distance in the hYAP-WW – Ligand Complex ^a [Å]	Distance in WWpep**[Å]
Y28-H32	12.6	10.6
Y28-W39	4.2	8.3
L30-W39	9.0	8.5
H32-W39	14.3	14.1

^a pdb entry 1k9q.

CONCLUDING REMARKS

Based on the 3D structure of a hYAP-WW – ligand complex, cyclic peptides (WWpep and WWpep**) were generated, which present the hYAP-WW binding site for proline-rich ligands. WWpep was found to interact with a proline-rich ligand for hYAP-WW, while the respective linear peptide had no affinity to the ligand. Furthermore, replacement of the primary contact residues in hYAP-WW and WWpep with alanine resulted in complete loss of affinity to the ligand, indicating a similar binding specificity of both molecules. Based on long-range NOEs detected in NMR spectra of WWpep**, a conformation was modeled using molecular dynamics simulations, in which the primary contact residues are presented in an arrangement similar to hYAP-WW. These results illustrate the applicability of synthetic peptides as functional and structural mimetics of conformationally defined protein binding sites. Ultimately, such synthetic mimics, as well as their substitution analogs, can be used as probes for the understanding of molecular events underlying human diseases.

ACKNOWLEDGEMENTS

We thank Christel Kakoschke and Victor Wray for assistance with NMR spectroscopy. This work was supported by the BioFuture Program (Grant No. 0311882) of the German Federal Department of Education and Research (BMBF), the German National Science Foundation (DFG We 1863/10-2) and the Kurt Eberhard Bode Foundation. We acknowledge a grant from the Barcelona Supercomputer Center.

REFERENCES

- Ball, L.J., Kuehne, R., Schneider-Mergener, J. and Oschkinat, H. (2005) Recognition of proline-rich motifs by protein-protein-interaction domains. *Angew.Chem.Int.Ed.* **44**, 2852-2869.
- Berendsen, J.C.D., van der Spoel, D. and van Gunsteren, W. (1995) A message-passing parallel molecular dynamics implementation, *Comp. Phys. Comm.* **91**, 43-56.
- Chen, H. and Sudol, M. (1995) The WW Domain of Yes-Associated Protein Binds a Proline-Rich Ligand that Differs from the Consensus Established for Src Homology 3-Binding Modules. *Proc.Natl.Acad. Sci.USA* **92**, 7819-7823.
- Eichler, J. (2005) Synthetic Peptide Arrays and Peptide Combinatorial Libraries for the Exploration of Protein-Protein Interactions and the Design of Protein Inhibitors. *Comb.Chem.High Throughput Screening* **8**, 135-143.
- Eichler, J. (2004) Rational and Random Strategies for the Mimicry of Discontinuous Protein Binding Sites. *Protein Peptide Lett.* **11**, 281-290.
- Franke, R., Doll, C., Wray V. and Eichler, J. (2004) Loops on Loops: Generation of Complex Scaffolded Peptides Presenting Multiple Cyclic Fragments. *Org.Biomol.Chem.* **2**, 2847-2851.
- Franke, R., Hirsch, T., Overwin, H. and Eichler, J. (2007) Synthetic Mimicry of the CD4 Binding Site of HIV-1 gp120 for the Design of Novel Immunogens. *Angew.Chem.* in press.
- Hunke, C., Hirsch, T. and Eichler, J. (2006) Structure-Based Synthetic Mimicry of Discontinuous Protein Binding Sites: Inhibitors of the Interaction of Mena EVH1 Domain with Proline-Rich Ligands. *ChemBioChem* **7**, 1258-1264.
- Kay, B.K., Williamson, M.P. and Sudol, M. (2000) The importance of being proline: the interaction of proline-rich motifs in signaling proteins with their cognate domains. *FASEB J.* **14**, 231-241.
- Lindahl E., Hess, B. and van der Spoel, D. (2001) A package for molecular simulation and trajectory analysis, *J. Mol. Mod.* **7**, 306-317.

Macias, M.J, Hyvönen, M., Baraldi, E., Schultz, J., Sudol, M., Saraste, M. and Oschkinat, H. (1996) Structure of the WW domain of a kinase-associated protein complexed with a proline-rich peptide. *Nature* **382**, 646-649.

Naumovski, L. Cleary, M.L. (1996) The p53-binding protein 53BP2 also interacts with Bcl2 and impedes cell cycle progression at G2/M. *Mol.Cell.Biol.* **16**, 3884-3892.

Niebuhr, K., Ebel, F., Frank, R., Reinhard, M., Domann, E., Carl, U.D., Walter, U., Gertler, F. B., Wehland, J. and Chakraborty, T. (1997) A novel proline-rich motif present in ActA of *Listeria monocytogenes* and cytoskeletal proteins is the ligand for the EVH1 domain, a protein module present in the Ena/VASP family. *EMBO J.* **16**, 5433-5444.

Nishizawa, K., Freund, C., Li, J., Wagner, G. and Reinhardt, E. I. (1998) Identification of a proline-binding motif regulating CD2-triggered T lymphocyte activation. *Proc.Natl.Acad.Sci.USA* **95**, 14897-14902.

Pires, J.R., Taha-Nejad, F., Toepert, F., Ast, T., Hoffmuller, U., Schneider-Mergener, J., Kuehne, R., Macias, M.J. and Oschkinat, H. (2001) Solution structures of the YAP65 WW domain and the variant L30K in complex with the peptides GTPPPPYTVG, N-(n-octyl)-GPPPY and PLPPY and the application of peptide libraries reveal a minimal binding epitope. *J.Mol.Biol.* **314**, 1147-1156.

Toepert, F., Pires, J., Landgraf, C., Oschkinat, H. and Schneider-Mergener, J. (2001) Synthesis of an array of 837 variants of the hYAP WW protein domain. *Angew.Chem.Int.Ed.* **40**, 897-900.

Toepert, F., Knaute, T., Guffler, S., Jose, R., Pires, J., Matzdorf, T., Oschkinat, H. and Schneider-Mergener, J. (2003) Combining SPOT Synthesis and native chemical ligation to create large arrays of WW protein domains. *Angew.Chem.Int.Ed.* **42**, 1136-1140.

Yu, H., Chen, J.K., Feng, S., Dalgarno, D.C., Brauer, A.W. and Schreiber, S.L. (1994) Structural basis for the binding of proline-rich peptides to SH3 domains. *Cell* **76**, 933-945.

Figure 1. NMR-Structure of a complex of hYAP-WW with the peptide ligand GTPPPPYTVG (Macias et al., 1996). The fragment shown in darker shade represents the peptide synthesized to mimic the hYAP-WW binding site (27RYFLNHIDQTTTWQ40), presenting the primary contact residues of hYAP-WW.

Figure 2. HPLC chromatograms (conditions: PLRP-S 100A column 50 x 2.1 mm; flow rate: 400 μ L/min; Gradient: 5 – 65 % acetonitrile/water, both containing 0.1 % TFA, in 15 min; detection: 214 nm) and ESI mass spectra of purified hYAP-WW (top, M = 4578) and WWpep** (bottom, M = 2032).

Figure 3. Binding of the ligand (Biotin-Ahx-EYPPYPPPPYPSG-NH₂) to plate-bound hYAP-WW and derived peptides, respectively. See Materials and Methods for experimental detail.

Figure 4. Long-range NOEs in the 2D ¹H NMR spectra of WWpep**. Left: Sections of the NOESY spectrum. Right: Graphical presentation of the cyclic peptide. Sequence-independent NOEs are marked by squares (left) and double arrows (right), respectively. Residue numbering was kept consistent with hYAP-WW.

Figure 5. Conformation of WWpep**, obtained through molecular dynamics simulation using long-range NOEs as constraints. Left: Overlay of six conformations. Right: Presentation of primary contact residues (numbering was kept consistent with hYAP-WW).