Received: 24 September 2020 | Revised: 18 October 2020 | Accepted: 30 October 2020

DOI: 10.1002/pro.3988

ARTICLE

THE PROTEIN SOCIETY WILEY

Study of the TEAD-binding domain of the YAP protein from animal species

Yannick Mesrouze | Fedir Bokhovchuk | Marco Meyerhofer | Catherine Zimmermann | Patrizia Fontana | Dirk Erdmann | Patrick Chène

Disease Area Oncology, Novartis Institutes for Biomedical Research, Basel, Switzerland

Correspondence
Patrick Chène, Novartis, WSJ 386 4.02.01, CH-4002, Basel, Switzerland.
Email: patrick_chene@yahoo.com

Abstract

The Hippo signaling pathway, which plays a central role in the control of organ size in animals, is well conserved in metazoans. The most downstream elements of this pathway are the TEAD transcription factors that are regulated by their association with the transcriptional coactivator YAP. Therefore, the creation of the binding interface that ensures the formation of the YAP:TEAD complex is a critical molecular recognition event essential for the development/survival of many living organisms. In this report, using the available structural information on the YAP:TEAD complex, we study the TEAD-binding domain of YAP from different animal species. This analysis of more than 400 amino acid sequences reveals that the residues from YAP involved in the formation of the two main contact regions with TEAD are very well conserved. Therefore, the binding interface between YAP and TEAD, as found in humans, probably appeared at an early evolutionary stage in metazoans. We find that, in contrast to most other animal species, several Actinopterygii species possess YAP variants with a different TEAD-binding domain. However, these variants bind to TEAD with a similar affinity. Our studies show that the protein identified as a YAP homolog in *Caenorhabditis elegans* does not contain the TEAD-binding domain found in YAP of other metazoans. Finally, we do not identify in non-metazoan species, amino acid sequences containing both a TEAD-binding domain, as in metazoan YAP, and WW domain(s).

KEYWORDS

Hippo pathway, metazoan, protein evolution, TEAD, YAP

Statement: YAP and TEAD form a key complex in the Hippo pathway. The study of the TEAD-binding domain of YAP of more than 400 animals species allows to identify the key residues involved in the interaction with TEAD. Using this information, this analysis was extended to the characterization of the interaction of TEAD-binding domain of the YAP protein from different animal species and TEAD.

1 | INTRODUCTION

The Hippo signaling pathway plays an important role during development, cell proliferation, regeneration, and tissue homeostasis.[1–3] In humans, this pathway consists

of a core kinase cascade involving MST1/2 and LATS1/2 that regulates the phosphorylation of the Yes-associated protein (YAP) and of its paralog the transcriptional coactivator with PDZ-binding motif (TAZ) protein.[4–6] When the Hippo pathway is on, YAP is phosphorylated and remains in the cytoplasm, but when the pathway is off, YAP becomes unphosphorylated and translocates into the nucleus, where it associates with the TEAD (TEA/ATTS domain) transcription factors. This association leads to TEAD activation, which regulates the expression of different genes (e.g., *CTGF* or *Cyr61*). Since the Hippo pathway has a beneficial role in stimulating tissue repair and regeneration following injury, designing drugs that modulate its activity could be of value in regenerative medicine.[7,8] The Hippo pathway is also deregulated in several cancers[5,9,10] and the development of inhibitors of the YAP:TEAD interaction is foreseen as a possible avenue to generate new anticancer drugs.[11,12]

Owing to its important biological function, the YAP protein is considered a key effector of the Hippo pathway[13]; and YAP homologs are already present not only in ancient metazoans,[14] but also in some non-metazoan species.[15] The human YAP protein contains several domains/motifs: TEAD-binding domain, one or two WW domains, coiled-coil and PDZ motifs.[16,17] The TEAD-binding domain (TBD) is intrinsically disordered in solution,[18,19] but upon binding to TEAD it adopts a well-defined conformation, and the determination of the structure of YAP in complex with TEAD has revealed at the atomic level how these two proteins interact with each other.[20,21] The TBD is about 50 residues long and binds to TEAD via three different secondary structure elements: a β-strand, an α-helix, and an Ω-loop (Figure 1a). Each of these elements interacts with a distinct interface at the surface of TEAD, and experimental data suggest that the two main contact regions are the α-helix and the Ω-loop.[21,22] However, peptides mimicking these two elements have a rather weak affinity for TEAD ($YAP^{61–74}$ (α-helix) $K_d > 150\ \mu M$[22]; YAP^{84-99} (Ω-loop) $K_d \sim 4\ \mu M$[23]), but when they are connected by a loop (linker) in the TBD of YAP ($YAP^{61–99}$), the overall affinity is significantly increased ($K_d \sim 60\ nM$[22]). Structure–function studies have also mapped the residues from the α-helix and the Ω-loop that are important for the interaction with TEAD.[20,21,24,25]

The structural and functional information that has been gained in recent years on the YAP:TEAD binding interface provides detailed knowledge of how these two proteins interact with each other. Using this information, we analyze in this report the TBD of YAP homologs from over 400 animal species. We show that the key residues identified in the TBD of human YAP are present in *Trichoplax adhaerens*, suggesting that the YAP:TEAD binding interface was established early on in multicellular organisms. We also reveal the unique feature of several Actinopterygii species that possess YAP variants with a different TBD. Finally, we study the interaction between human TEAD and peptides mimicking the TBD of evolutionary distant YAP orthologs.

2 | MATERIAL AND METHODS

2.1 | BLASTp searches

The Basic Local Alignment Search Tool (BLAST) was used (BLASTp; https://blast.ncbi.nlm.nih.gov/Blast.cgi; default algorithm parameters; maximum target sequences: 5000) to identify putative YAP homologs in protein sequences from animal species deposited at the non-redundant

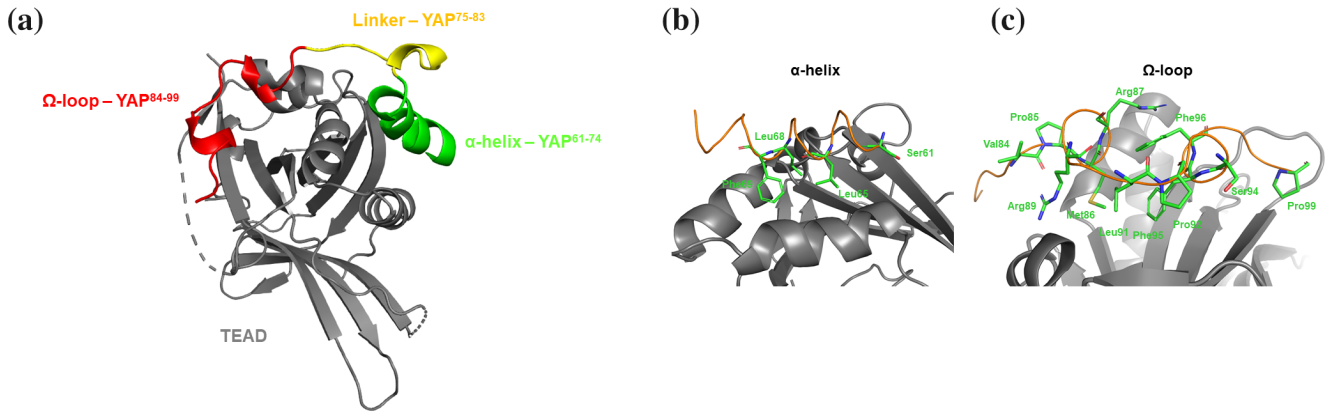

FIGURE 1 Structure of YAP:TEAD complex. (a) Overall structure. The different secondary structure elements of human YAP^{61-99} are colored green (α-helix), yellow (linker) and red (Ω-loop). TEAD is gray. (b and c) α-helix and Ω-loop binding interfaces. The YAP residues from the α-helix (b) and from the Ω-loop (c) region mentioned in the text are represented by green sticks. The main chain of YAP is colored orange and TEAD gray. This figure was drawn from the PDB structure code 3KYS with PyMol (Schrödinger Inc., Cambridge, MA)

protein sequences (nr) database (National Center for Biotechnology Information; Taxid: 33208; February 2020). The query sequence—SETDLEALFNAVMNPKTANVPQTVPM RLRKLPDSFFKPP—was the region 61–99 from human YAP (α-helix: linker:Ωloop from the TBD of YAP; Seq. Id. P46937). The sequences obtained from this search (Supporting Information) were grouped by species, the putative TBD was localized in each of them and only the sequences containing at least one WW domain (defined as W-x_{21}-W, where x = any amino acid) in addition to the TBD were considered putative YAP proteins. TAZ homologs (not included in this study) were identified using the following two criteria: the absence of a proline residue at the position corresponding to Pro85 in human YAP and the presence of a tryptophan residue at the position corresponding to Met86 in human YAP.[22] Within each species, the regions corresponding to the TBD were compared, and each unique TBD sequence was compiled into a single file (Figure S4). TEAD homologs were identified in the sequences from *Anabas testudineus* deposited at the nr-database (taxid: 64144) by a BLAST search against the amino acid sequence corresponding to the YAP-binding domain of human TEAD4, $TEAD4^{217–434}$ (Seq. Id. Q15561).

2.2 | Synthetic peptides

The synthetic peptides (both *N*-acetylated and C-amidated) were purchased from Biosynthan (Germany). The purity (>90%) and the chemical integrity of the peptides was determined by liquid chromatography–mass spectrometry (LC–MS) from 10 mM stock solutions in 90:10 (vol/vol) dimethyl sulfoxide: water. The peptide derived from *Caenorhabditis elegans* YAP was dissolved in 50:50 (vol/vol) acetonitrile: (water + 1 mM TCEP) to minimize the potential oxidation of the cysteine residue present in its sequence. The concentration of the stock solutions was determined by HPLC and the solubility of the peptides was measured with a NEPHELOstar (BMG LABTECH, Germany).

2.3 | Protein cloning, expression, and purification

The YAP-binding domain of human TEAD4, $TEAD4^{217–434}$ (Seq. Id. Q15561), was obtained as previously described.[26] The DNA encoding for TEAD from *Anabas testudineus* (residues 211–428; Seq. Id. XP_026221540.1) was codon optimized for *Escherichia coli* expression and synthesized by GeneArt (ThemoFisher Scientific, Waltham, MA). The coding region was PCR amplified with Q5 High Fidelity 2x Master Mix (New England Biolabs, Ipswich, MA). The sense oligonucleotide encoding the gene from amino acid 211 was designed to contain an AviTag and an additional LguI site for cloning (5′-AAAGGAAAAAAGCTCTTCA CCGGGTTTGAACGACATCTTCGAAGCTCAGAAGATC-GAATGGCACGAGGGTGGCGGTAGTGGTGGTGGCTC-TAGAAGCATTGGCACCACCAAAC-3′). The antisense oligonucleotide encoded up to amino acid position 428 and comprised an additional LguI site for cloning purposes (5′-TTTCCTTTTGCTCTTCGTTATTCTTTAACCAGACGATA-AATATGATGCTG-3′). The PCR product was purified with the ReliaPrep™ DNA Clean-up and Concentration system (Promega, Madison, WI) and cloned into the in-house vector pLAF71 containing a His-Tag and an LguI-cassette for T2S cloning (Type IIS, StarGate, IBA Lifesciences, Germany). The cloning reaction with equal amounts of fragment and vector, SapI (New England Biolabs, Ipswich, MA), and T4 Ligase (Rapid DNA ligation kit, Sigma-Aldrich, St. Louis, MI), was performed at 37°C for 30 min followed by temperature cycle ligation between 10°C and 30°C for 90 min. The ligation product was transformed into *E. coli* DH5α. Protein expression and purification was done as for $TEAD4^{217–434}$. The purity and the molecular weight of the proteins were verified by LC–MS (Figure S1).

2.4 | TR-FRET and Surface Plasmon Resonance

The potency (IC_{50}) and affinity (K_d) for TEAD4 of the different peptides were measured in a TR-FRET assay and by Surface Plasmon Resonance (SPR), respectively, as previously described.[23,25] The SPR methodology used to measure the affinity of the YAP-derived peptides with TEAD from *Anabas testudineus* was the same as for human TEAD4. Representative inhibition curves (TR-FRET) and sensorgrams (SPR) are shown in Figures S2 and S3, respectively.

3 | RESULTS AND DISCUSSION

3.1 | Sequence analysis of the TEAD-binding domain of the YAP protein from animal species

A BLASTp search against the region corresponding to the TBD of human YAP ($YAP^{61–99}$) in the genomes of animal species identified 1873 entries. From this initial hit list, 492 unique sequences belonging to 415 different animal species were obtained. Chordata (224 species) and Arthropoda (164 species) are the most represented in the dataset, while only a limited number of species from other phyla (e.g., Mollusca, Cnidaria ...) were found.

Sequence logos were created for these three groups of species in the regions corresponding to the α-helix ($YAP^{61–74}$) and the Ω-loop ($YAP^{84–99}$) of human YAP (Figure 2). The sequence logos are similar between the three groups of species, but a higher variability is observed in the "other phyla" group. Overall, this indicates that the two binding interfaces—α-helix and Ω-loop—of the TBD of YAP are well conserved among the different species studied.

In the α-helix region, the three residues $Leu65^{Hs}$ (^{Hs} refers to YAP from *Homo sapiens*), $Leu68^{Hs}$ and $Phe69^{Hs}$ are the most conserved. These amino acids form a LxxLF motif that binds to a hydrophobic cleft at the surface of TEAD (Figure 1b)[20,21] and their mutation to alanine destabilizes the YAP:TEAD interaction by more than 1.5 kcal/mol.[25] In just a few sequences, a methionine is at position-68^{Hs}. There is more variability at the other positions in the α-helix, but $Ser61^{Hs}$, located at the N-terminus (Figure 1b), is well conserved and is replaced by a threonine in a subset of Arthropoda species (mainly Diptera). This shows that a small and phosphorylatable side chain is favored at this position. Hao et al. have identified $Ser61^{Hs}$ as a potential site for LATS1 phosphorylation.[28] Conserved negatively charged residues are present at the N-terminus of the α-helix (colored red in Figure 2), while positively charged residues (colored blue in Figure 2) are located at the N-terminus of the Ω-loop. Therefore, the TBD of YAP forms a kind of dipole with its two binding sites harboring opposite charges. Ionic strength has a mild effect on the YAP:TEAD interaction,[29] so this spatial distribution of charges could be relevant for other aspects of YAP function, for example, to adopt specific conformations in solution.

Several residues of the Ω-loop region are well conserved (Figures 1c and 2). Zhang et al. have suggested that $Val84^{Hs}$ at the N-terminus of the Ω-loop has a shielding effect on the folding of this region[24] and the presence of this residue dramatically increases the affinity of peptides mimicking the Ω-loop of YAP.[23] $Val84^{Hs}$ is often replaced by a leucine and in some cases by an isoleucine or a lysine. $Pro85^{Hs}$, $Pro92^{Hs}$ and $Pro99^{Hs}$ are present in virtually all the sequences. $Pro85^{Hs}$ is important for maintaining the local structure at the N-terminus of the Ω-loop.[22,24] $Pro92^{Hs}$ is probably required for the appropriate folding of the Ω-loop and its mutation to alanine destabilizes the YAP:TEAD interaction by more than 3 kcal/mol.[29] The role of $Pro99^{Hs}$ in the formation of the YAP:TEAD complex is

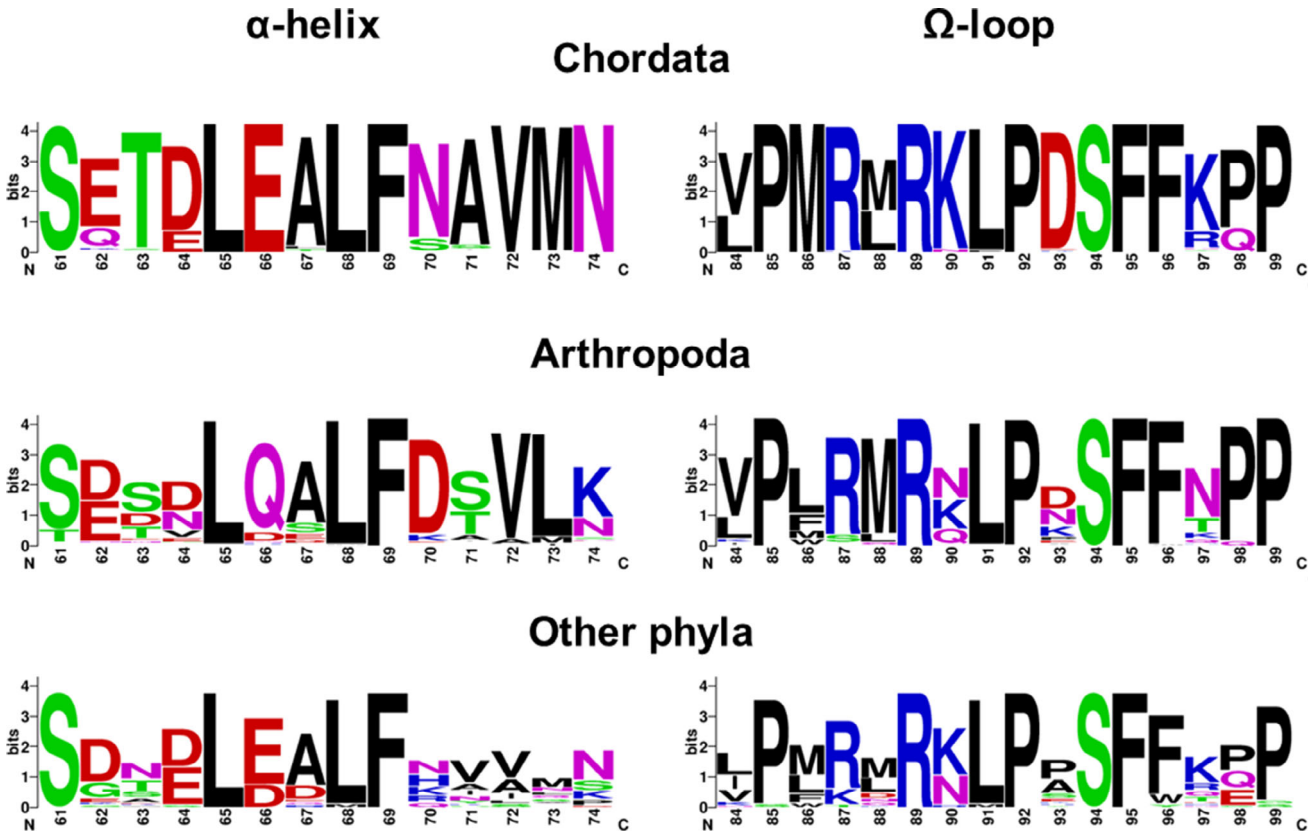

FIGURE 2 Protein logos of the α-helix and Ω-loop regions of the TEAD-binding domain of YAP. The amino acid sequences of the TBD of YAP from Chordata, Arthropoda and "other phyla" (see text) species have been aligned. Protein logos of the regions corresponding to the α-helix ($YAP^{61–74}$) and the Ω-loop ($YAP^{84–99}$) of human YAP have been generated by WebLogo (https://weblogo.berkeley.edu/)[27]

unclear and its mutation to alanine has only a moderate effect on binding.[24] Met86^{Hs}, Leu91^{Hs}, and Phe95^{Hs} form a hydrophobic core that may contribute to the stabilization of the bound Ω-loop, and these residues make hydrophobic interactions with TEAD (Figure 1c). Met86^{Hs} is replaced in some sequences by a leucine or a phenylalanine that should make similar interactions with TEAD. In the sequences from Coleoptera species, a tryptophan is found at position-86^{Hs} and this aromatic residue is present at the same position in TAZ.[22] Leu91^{Hs} is replaced by a phenylalanine in some Actinopterygii (Cyprinodontiformes) species in agreement with a study of synthetic peptides that mimic YAP, showing that bulkier amino acids (e.g., cyclobutylalanine) could be present at position-91^{Hs}.[30] Phe95^{Hs} is conserved in all the sequences, revealing its importance for the interaction with TEAD. Phe96^{Hs} has a particular role in the formation of the YAP:TEAD complex because it does not directly interact with TEAD (Figure 1c). It is located at the top of the hydrophobic core formed by Met86^{Hs}, Leu91^{Hs}, and Phe95^{Hs}, shielding it from solvent. Its mutation to alanine destabilizes binding by more 3 kcal/mol.[25] This residue is quite well conserved, but in a few sequences it is replaced by a tryptophan. It has been shown that the presence of a larger aromatic residue at position-96^{Hs} (e.g., 1-naphtylalanine) enhances the affinity of YAP[30] and that FAM181A, which also binds to TEAD via an Ω-loop, contains a tryptophan at this position.[23] Phe96^{Hs} also makes a π-cation interaction with Arg87^{Hs}, and this interaction is thought to contribute to the stabilization of the bound Ω-loop (Figure 1c).[24,30] Arg87^{Hs} is quite well conserved, but in Coleoptera, where a tryptophan is present at position-86^{Hs}, Arg87^{Hs} is often replaced by a serine. Arg89^{Hs} and Ser94^{Hs} are two critical residues at the YAP:TEAD binding interface (Figure 1c). Arg89^{Hs} forms a salt bridge with an aspartic residue located at the surface of TEAD, while Ser94^{Hs} makes hydrogen bonds with a tyrosine and glutamic acid from TEAD. The mutation of Arg89^{Hs} or Ser94^{Hs} to alanine destabilizes the YAP:TEAD complex by more than 3 kcal/mol.[25,31,32] These two residues are conserved in all the sequences, confirming their important contribution to the formation of the YAP:TEAD complex.

The linker region, residues 75-83^{Hs}, connects the α-helix and the Ω-loop. The amino acids from the linker make few contacts with TEAD[20,21] and this region of the TBD of YAP is probably flexible. There is a lower sequence conservation in the linker than in the two binding interfaces (Figure S4). However, within the same group of species, the amino acid sequence is relatively well conserved. The number of residues in the linker varies from 4 amino acids (e.g., *Trichoplax adhaerens*) to 15 amino acids (e.g., cnidarian; Figure S4). The amino acid content and length of the linker are more variable among the species belonging to the "other phyla" group. Even if the linker does not contribute directly to binding, it has a role in the interaction. For example, TAZ and YAP have a different linker, and their swap between the TBDs reduces the affinity for TEAD.[22] Chen et al. have also shown that a PxxΦP motif (x = any amino acids, Φ = hydrophobic side chain; the prolines correspond to Pro81^{Hs} and Pro85^{Hs}) present in the C-terminus of the linker is important for the interaction with TEAD and for transforming activity.[20] As proline residues can have an effect on the dynamic and the conformation of loops/linkers (e.g., References 33–35), we looked in greater detail at the presence of this residue in the region corresponding to YAP 75-85^{Hs} (to include Pro85^{Hs} from the PxxΦP motif). While there is little variation in the number and position of proline residues in the sequences from closely related species, this is not the case when looking at the whole dataset. For example, five proline residues are found in *Acropora digitifera* or *A. millepora* while there is only one in *Galendromus occidentalis* or *Varroa destructor*. While most of the sequences contain a PxxΦP motif, several do not, for example, *Pentromyzon marinus*, *Poecilia formosa*, and *Aplysia californica*. This suggests that the presence of a PxxΦP motif is not always required in the TBD of YAP proteins (see also below).

3.2 | Some species contain YAP variants that have a different TEAD-binding domain

In agreement with earlier findings showing that several spliced forms of YAP can exist in one species (e.g., in human[36]), our BLAST search identified more than one YAP sequence in the majority of the vertebrate species (Supporting Information). In most of the cases, these variants have an identical TBD, but they also differ in their TBD in several Actinopterygii species (Figure S4). Since similar variants are often found in closely related species and because they usually differ by several residues, sequencing errors cannot explain this observation.

The difference in amino acids between these YAP variants occurs not only in the linker region, but also in the α-helix and Ω-loop. We usually identified two TBD variants in the same species, but up to three were found in Salmoniformes (*Oncorhynchus* and *Salmo*; Figure S4). However, in several species (e.g., *Danio rerio*) all the YAP variants share the same TBD as in most of the vertebrates. Chen et al. have shown that paralogs of genes from the Hippo pathway, including YAP, were first identified during evolution in fishes[37] and the presence of several paralogs of YAP in these organisms is probably linked to the whole-

genome duplication events that occurred in fish ancestors.[38] Therefore, the YAP variants with a distinct TBD may have originated during these duplication events. This unique feature prompted us to ask the following question: have the YAP variants from the same species that bear a different TBD a different affinity for TEAD?

3.3 | Interaction between human TEAD4 and peptides mimicking the TEAD-binding domain of YAP variants from different Actinopterygii species

The amino acid differences in the α-helix and Ω-loop regions between the YAP variants from the same Actinopterygii species occur at positions that are not essential for the interaction with TEAD (Figures S4 and S2) suggesting that they should not trigger significant changes in binding affinity. However, the linker regions can be different both in the content and in the position of proline residues (called hereafter proline motif) and we identified six different proline motifs: P-x_5-P-x_3-P (as in humans), P-x_5-P-P-x_2-P, P-x_5-P-P-x_3-P, P-x_6-P-x_3-P, P-x_6-P-P-x_2-P and P-x_{10}-P (x = any amino acid) among the Actinopterygii species studied. As mentioned above, a difference in the proline content of the linker may affect its dynamic/conformation and as a consequence the interaction with TEAD. To explore this possibility, eight TBDs belonging to four different fish species were selected to include the six proline motifs identified in Actinopterygii (Table 1). Synthetic peptides corresponding to these TBDs were synthetized and their potency compared to the one of the corresponding human YAP peptides. The peptides were tested in a TR-FRET assay for their ability to inhibit the interaction between $YAP^{60–100}$ and $TEAD4^{217–434}$.[23] All the peptides dose-dependently reduce the TR-FRET signal, showing that they compete with $YAP^{60–100}$ for binding to TEAD (Figure S2). The slope values measured from the inhibition curves are close to 1 for all the peptides, and the half-maximal inhibitory concentrations (IC_{50}) measured are similar for each peptide pair belonging to the same species (less than a threefold difference; Table 1). The two peptides derived from *Sander lucioperca* are the most potent, but altogether all the peptides derived from fish YAP have a similar potency (∼ 10 to 60 nM), which is comparable to that measured with the peptide derived from *H. sapiens* (∼ 60 nM). *Poecilia formosa*_2, which does not contain a PXXΦP motif, has a potency similar to that of all the other peptides (Table 1). Hence, the absence of the PXXΦP motif in the TBD of *P. formosa* does not preclude an efficient interaction with TEAD. Altogether, our findings show that neither the presence of different proline motifs nor the variation in amino acids in the sequence of these eight peptides significantly changes their potency.

3.4 | Interaction with a TEAD protein from the same Actinopterygii species

The previous findings might be misleading because the experiments were conducted with heterologous binding partners: YAP peptides derived from Actinopterygii species and human TEAD. The possibility of these peptides showing a different behavior if their potency is measured with a TEAD protein from the same species can therefore

TABLE 1 Potency of peptides mimicking YAP from different Actinopterygii species. The potency (IC_{50}) of the synthetic peptides was measured in a TR-FRET assay. The secondary structure adopted by human YAP once bound to TEAD is indicated (α-helix, linker, and Ω-loop). The proline residues present in the different motifs are underlined. x = any amino acid. The values represent the average IC_{50} and the corresponding standard error of $n \geq 2$ independent experiments. *Homo sapiens* residues 61–99 from Seq. Id. P46937; *Anabas testudineus*_1 residues 21–59 from Seq. Id. XP_026225556.1; *Anabas testudineus*_2 residues 21–60 from Seq. Id. XP_026227259.1, *Esox lucius*_1 residues 21–59 from Seq. Id. XP_012991334.1, *Esox lucius*_2 residues 23–62 from Seq. Id. XP_028977077.1; *Poecilia formosa*_1 residues 21–59 from Seq. Id. XP_007570007.1; *Poecilia formosa*_2 residues 21–60 from Seq. Id. XP_007556469.1; *Sander lucioperca*_1 residues 21–59 from Seq. Id. XP_031153389.1; *Sander lucioperca*_2 residues 21–60 from Seq. Id. XP_031176101.1

Species	Sequence αααααααααααααααα111111111ΩΩΩΩΩΩΩΩΩΩΩΩΩΩΩΩ	Proline motif	IC_{50} (nM)
Homo sapiens	SETDLEALFNAVMNPK-TANVPQTVPMRLRKLPDSFFKPP	P-x_5-P-x_3-P	60 ± 2
*Anabas testudineus*_1	SETDLEALFNAVMNPK-NTIVPPSVPMRLRKLPDSFFKPP	P-x_5-P-P-x_2-P	41 ± 2
*Anabas testudineus*_2	SQTELEALFNAVMNPSKEARHPASLPMRMRKLPDSFFRQP	P-x_6-P-x_3-P	24 ± 2
*Esox lucius*_1	SETDLEALFNAVMNPK-NTIVPPSVPMRMRKLPDSFFTPP	P-x_5-P-P-x_2-P	64 ± 3
*Esox lucius*_2	SQTDLEALFNAVINPKTATLPPSSLPMRMRKLPDSFFRQP	P-x_5-P-P-x_3-P	25 ± 1
*Poecilia formosa*_1	SQTDLELLFNSVMNPK-SSNVPASLPMRMRNLPDSFFKPP	P-x_5-P-x_3-P	65 ± 3
*Poecilia formosa*_2	SNTELEALFNAVMNPGKAARQTQSVPMRMRKFPDSFFKPP	P-x_{10}-P	39 ± 3
*Sander lucioperca*_1	SETDLEALFNAVMNPK-NTIVPPSVPMRMRKLPDSFFKPP	P-x_5-P-P-x_2-P	12 ± 1
*Sander lucioperca*_2	SQTELEALFSAVMNPSKSARQPPSLPMRMRKLPDSFFRQP	P-x_6-P-P-x_2-P	12 ± 1

not be excluded. To check this hypothesis, we studied the interaction between YAP-derived peptides and TEAD from *Anabas testudineus* (species chosen in an arbitrary manner). A BLASTp search identified different TEAD homologs in the genome of this species, and an amino acid sequence alignment shows that their putative YAP-binding domain has a good sequence similarity with the one of human TEAD4 (Figure S5). The YAP-binding domain of the TEAD variant from *A. testudineus*, which has the lowest sequence similarity with human TEAD4 (region 211–428; Seq. Id. XP_026221540.1; Figure S5), was cloned and expressed in *Escherichia coli*. The purified protein is acylated (Figure S3) as previously observed for human TEAD,[26,39,40] and the acylation site present in human TEAD4, Cys367, is conserved in TEAD from *A. testudineus* (Cys361; Figure S5). The N-biotinylated-Avi-tagged TEAD proteins were immobilized on sensor chips, and the binding of the *A. testudineus* and human YAP-derived peptides was measured by Surface Plasmon Resonance. The peptides bind in a dose-dependent manner to human and *A. testudineus* TEAD, and the maximal signal measured at equilibrium (R_{max}^{eq}) is close to the calculated maximal feasible signal (R_{max}^{th}; Figure S3). The two *A. testudineus* YAP-derived peptides bind to *A. testudineus* TEAD with a similar affinity (dissociation constants measured at equilibrium, K_d^{eq}; Table 2) revealing that, even when YAP and TEAD are from the same species, the differences observed between the TBDs have little effect on the YAP:TEAD interaction. In line with the high amino acid sequence similarity between human and *Anabas* TEAD (Figure S5), the three peptides bind equally well to these two proteins (Table 2). This suggests that the YAP:TEAD binding interface is well conserved between these two species.

3.5 | Study of the TEAD-binding domain of YAP from non-metazoan and metazoan species

Using the knowledge gained from our analysis of the TBD of YAP from a variety of animal species, we next studied in greater detail the interaction between human TEAD and peptides mimicking the TBD of selected YAP orthologs that are evolutionarily distant from human YAP.

A YAP ortholog has been found in the very basal metazoan, *Trichoplax adhaerens*,[14] and our BLAST search also identified in this species a sequence that bears a LxxLF motif and the key residues present in the Ω-loop of human YAP (Figures S4 and 3a). The main differences with the human sequence are in the linker, which does not contain a PxxΦP motif, lacks $Pro75^{Hs}$ and is much shorter. The potency, $IC_{50} \sim 900$ nM (Table 3), of a peptide mimicking the TBD of YAP from *T. adhaerens* shows that it competes efficiently with human YAP for binding to human TEAD. Therefore, YAP from *T. adhaerens* is able to recognize the YAP-binding site present at the surface of human TEAD. Nevertheless, the potency of this peptide is lower than the potency measured with the other peptides tested in this study (Table 1). As the key residues present in the α-helix and Ω-loop of human YAP are also found in YAP from *T. adhaerens*, the lower potency of the latter could be due to its linker region being significantly different.

Hilman and Gat did not identify a YAP homolog in *Caenorhabditis* species and suggested the loss of YAP in the nematode lineage,[14] but more recently Isawa et al. described the presence of a YAP homolog in *Caenorhabditis elegans*.[41] We did not find a sequence corresponding to YAP in the genome of *C. elegans* from BLASTp searches in the nr-database (Figure S4) or in the database used by Iwasa et al. (query: residues 61–99 from human YAP; http://www.wormbase.org/). Therefore, we utilized the sequence of the TBD of YAP from *C. elegans* provided by Iwasa et al. to make an alignment with the corresponding region from human YAP (Figure 3a). Several residues involved in the interaction between human YAP and TEAD are not conserved in the region corresponding to the α-helix. $Leu65^{Hs}$ and $Phe69^{Hs}$ could be replaced by an isoleucine and a proline that may affect the formation/stability of the α-helix found at the location of $Ser61^{Hs}$. In the Ω-loop region, while $Leu91^{Hs}$, $Pro92^{Hs}$ and $Ser94^{Hs}$ seem to be present, $Val84^{Hs}$, $Pro85^{Hs}$, $Met86^{Hs}$, and $Arg89^{Hs}$ are missing. $Phe95^{Hs}$ and $Phe96^{Hs}$ are replaced by tyrosine residues and $Pro99^{Hs}$ is not conserved or located more toward the C-terminus. Altogether,

TABLE 2 Affinity of YAP mimetics for *Anabas testudineus* and human TEAD. The N-biotinylated-Avitagged TEAD proteins were immobilized on sensor chips, and the affinity of the peptides (see Table 1 for the amino acid sequence) were measured at equilibrium (K_d^{eq}). The values represent the average K_d^{eq} and the corresponding standard error of $n \geq 2$ independent experiments

TEAD protein	*Homo sapiens* K_d^{eq} (nM)	*Anabas testudineus*_1 K_d^{eq} (nM)	*Anabas testudineus*_2 K_d^{eq} (nM)
Anabas testudineus	103 ± 6	99 ± 6	60 ± 6
Homo sapiens (TEAD4)	81 ± 5	71 ± 7	42 ± 2

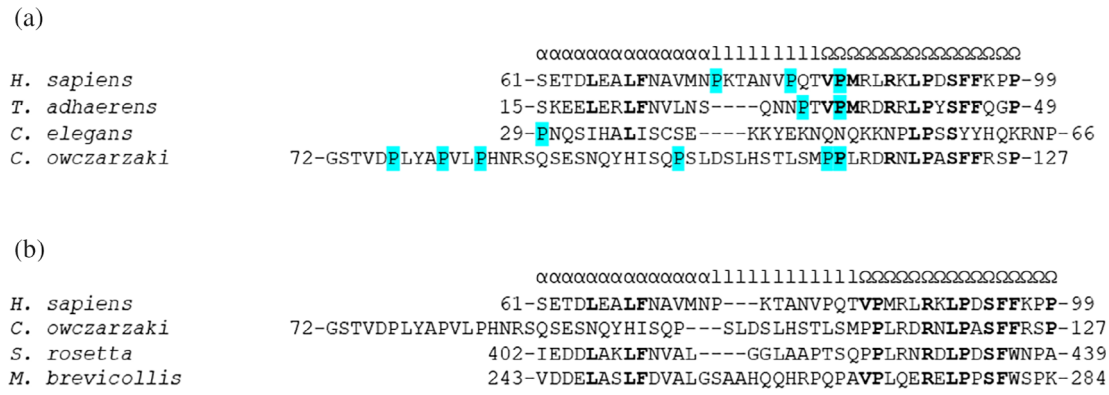

FIGURE 3 Amino acid sequence of the TEAD-binding domain of YAP from metazoan and non-metazoan species. (a) The sequences of the TBD of YAP from *Trichoplax adhaerens* (Seq. Id. XP_002108065.1), *Caenorhabditis elegans* (Seq. Id. F13E6.4 [NP_001369894.1]) and *Capsaspora owczarzaki* (Seq. Id. JN202490.1) have been manually aligned to the corresponding region of human YAP (Seq. Id. P46937). The sequences from *C. elegans* and *C. owczarzaki* have been extended at their N-terminus to take into account the alignments proposed by Iwasa et al.[41] and Ikmi et al.,[17]respectively. The proline residues (including Pro85^{Hs}) outside of the Ω-loop region are highlighted in cyan. (b) The sequence of the TBD of human YAP has been manually aligned with the corresponding region of the sequences from *C. owczarzaki*, *Salpingoeca rosetta* (Seq. Id. XP_004994687.1) and *Monosiga brevicollis* (Seq. Id. A9UXI0_MONBE). The residues involved in the YAP:TEAD interaction (see text) are indicated in bold in the human sequence and when conserved in the other sequences. Dashes represent missing residues. The secondary structure adopted by human YAP bound to TEAD is indicated (α-helix, linker, and Ω-loop)

TABLE 3 Potency of peptides derived from *Trichoplax adhaerens*, *Caenorhabditis elegans*, and *Capsaspora owczarzaki* YAP. The potency (IC_{50}) of synthetic peptides mimicking YAP from *T. adhaerens* (residues 15–49 from Seq. Id. XP_002108065.1), *C. elegans* (residues 30–66 from Seq. Id. F13E6.4 [NP_001369894.1]) and *C. owczarzaki* (residues 85–127 from Seq. Id. JN202490.1) was measured in a TR-FRET assay. The values represent the average IC_{50} and the corresponding standard error of $n \geq 2$ independent experiments. n.a.m. = no activity measured

Species	Sequence	IC_{50} (nM)
Trichoplax adhaerens	15-SKEELERLFNVLNSQNNPTVPMRDRRLPYSFFQGP-49	880 ± 70
Caenorhabditis elegans	30-NQSIHALISCSEKKYEKNQNQKKNPLPSSYYHQKRNP-66	n.a.m.
Capsaspora owczarzaki	85-HNRSQSESNQYHISQPSLDSLHSTLSMPPLRDRNLPASFFRSP-127	107000 ± 9000

this shows that the TBD of YAP from *C. elegans* identified by Isawa et al. is significantly different from the one present in other metazoans. The potency of a peptide mimicking the TBD of YAP from *C. elegans* was so low that it could not be measured in the TR-FRET assay (Table 3) suggesting that it has a very weak affinity for the YAP-binding site present at the surface of human TEAD. Since Iwasa et al. have reported an interaction between YAP from *C. elegans* and EGL-44 (*C. elegans* homolog of TEAD),[41] it is possible that the YAP-binding site from EGL-44 is very different from the one present in human TEAD and/or that YAP from *C. elegans* and EGL-44 interact via additional binding interfaces.

Sebé-Pedrós et al. found a YAP-like sequence in the genome of the amoeboid holozean *Capsaspora owczarzaki*, and they proposed a premetazoan origin of the Hippo signaling pathway.[15] Since we did not identify a YAP homolog from a BLASTp search in the genome of *C. owczarzaki* (query: residues 61–99 from human YAP; nr-database taxids: 595528 and 192,875), we used the sequence provided by Sebé-Pedrós et al. There is a good sequence similarity with human YAP in the Ω-loop region, and the main difference is the presence of a proline at position-84^{Hs} (Figure 3a). However, the sequence similarity is low in the region corresponding to the α-helix. There is no LxxLF motif, but Phe69^{Hs} could be replaced by an isoleucine and Leu68^{Hs} by a histidine. In the vestigial-like proteins, which bind to TEAD in the same region, Leu68^{Hs} is replaced by a histidine that forms two hydrogen bonds with TEAD.[42,43] Leu65^{Hs} from the LxxLF motif is not conserved in *C. owczarzaki*. Since a LxxLF motif is separated from the Ω-loop by 54 amino acids in FAM181B,[23] we looked for the presence of this motif in regions located more toward the N-terminus of YAP from *C. owczarzaki*, but we did not find any. Ikmi et al. have proposed that the residues 72-GSTVDPLYAPVL-83 form an α-helix in the TBD of YAP from *C. owczarzaki*.[17] As such, this α-helix, which is 12 residues long, would contain two prolines, Pro77^{Co} and Pro81^{Co} (^{Co} refers to YAP from *C. owczarzaki*). Since prolines are usually considered to be "α-helix breaker"

residues (e.g., References 44, 45, ...), we hypothesized that, if an α-helix would be present in this region, it could be located between the residues 85^{Co} and 99^{Co} because prolines are found at positions 77^{Co}, 81^{Co}, 84^{Co}, and 100^{Co}. Therefore, we measured the potency of a peptide mimicking the region 85–127 of the sequence identified by Sebé-Pedrós et al. The potency of this peptide is low, $IC_{50} \sim 110$ μM (Table 3), showing that it has a weak affinity for human TEAD, but it is similar to that measured with the isolated Ω-loop of human YAP (lacking $Val84^{Hs}$).[22,23] Therefore, the peptide derived from *C. owczarzaki* may interact with TEAD only via an Ω-loop, as suggested by our sequence analysis, which did not identify a LxxLF motif in this protein.

YAP orthologs have also been identified in two additional non-metazoan species: *Salpingoeca rosetta* and *Monosiga brevicollis*.[15,17] We did not find a region similar to the TBD from metazoan YAP in the sequences provided by Sebé-Pedros et al. for *S. rosetta* and *M. brevicollis*, although they contain one or two WW domains (Figure S6). The alignment reported by Ikmi et al. also shows a low sequence similarity between the TBD of YAP from *S. rosetta*/*M. brevicollis* and the TBD of YAP from metazoan species.[17] To identify potential YAP homologs in these non-metazoan species, we conducted a BLASTp search against the residues 85–127 of YAP from *C. owczarzaki* (sequence Figure 3a) in the nr-database (*S. rosetta* taxids: 86017 and 946362; *M. brevicollis* taxids: 81525, 81824, 431895, and 487148). A sequence from *S. rosetta* was found (Seq. Id. XP_004994687.1, Figure S6), but no hits were obtained for *M. brevicollis*. Further BLASTp searches (queries: residues 85–127 from *C. owczarzaki* and residues 402–439 from *S. rosetta*, sequences Figure 3) using the browser for protist genomes, Ensembl Protists (taxid: *Monosiga brevicollis* MX1, https://protists.ensembl.org/index.html),[46] enabled us to identify a sequence from *M. brevicollis* (Seq. Id. A9UXI0_MONBE, Figure S6). These sequences from *S. rosetta* and *M. brevicollis* have a good sequence similarity with the TBD present in metazoan YAP (Figure 3b), but none of them contains a WW domain (Figure S6). To summarize, in the three non-metazoan species—*C. owczarzaki*, *S. rosetta*, and *M. brevicollis*–we were unable to identify a protein sequence containing both a TBD formed of a LxxLF motif and an Ω-loop, as is found in metazoan YAP and one or two WW domains (defined as W-x_{21}-W, x = any amino acid).

4 | CONCLUSION

The formation of a stable interface between YAP and TEAD is essential for the function of the Hippo pathway. In this report, we show that the key residues present in the two main contact regions—the α-helix and the Ω-loop—of the TBD of human YAP are well conserved among metazoans. The presence of these residues in species such as *T. adhaerens* suggests that these binding interfaces appeared at an early stage in the evolution of metazoans. Our study also shows that the highest amino acid variability is found in the linker region connecting the α-helix and the Ω-loop. Therefore, the linker, which is required for a tight interaction between YAP and TEAD, is more permissive to the effect of mutations as illustrated by the similar affinity for TEAD of the YAP variants from Actinopterygii species that possess TBDs with different linkers. The absence of the PxxΦP motif[20] in the YAP protein from different species and our results with *Poecilia formosa* suggest that this motif is not required for binding to TEAD. However, the last proline of this motif, $Pro85^{Hs}$, is highly conserved and is required for an efficient interaction with TEAD[22,24] and YAP function.[47] The number and position of prolines in the linker is more variable in metazoan YAP than previously observed,[17] and the presence of specific proline motifs may not help in tracking the evolution of the structure of the TBD of YAP. Hilman and Gatt did not identify YAP in nematodes,[14] Iwasa et al. described a YAP protein in *C. elegans* that interacts with EGL-44, the homolog of TEAD in this species,[41] and we show that the putative TBD present in the protein found by Iwasa et al. is significantly different from that of other metazoans. A hypothesis to reconcile these observations is that the protein described by Iwasa et al. is not a YAP ortholog, but that it may have the same biological function in *C. elegans*. A more detailed characterization of this protein should help to check this hypothesis. We did not identify in three non-metazoan species any protein sequences containing both a TBD and one or two WW domains, as are found in metazoan YAP. However, we found sequences that have a good sequence similarity with the Ω-loop region from YAP but that lack the LxxLF motif and/or the WW domain. Nevertheless, these results obtained from a limited number of non-metazoans do not preclude that some unicellular species possess a YAP protein similar to that found in metazoans.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Yannick Mesrouze: Investigation; methodology. **Fedir Bokhovchuk:** Investigation; methodology. **Marco Meyerhofer:** Investigation; methodology. **Catherine Zimmermann:** Investigation; methodology. **Patrizia Fontana:** Investigation; methodology. **Dirk Erdmann:** Formal analysis; supervision; writing-original draft.

Patrick Chene: Conceptualization; data curation; investigation; supervision; writing-original draft.

ORCID

Patrick Chène https://orcid.org/0000-0002-6010-9169

REFERENCES

1. Fu V, Plouffe SW, Guan KL. The Hippo pathway in organ development, homeostasis, and regeneration. Curr Opin Cell Biol. 2017;49:99–107.
2. Wang Y, Yu A, Yu FX. The Hippo pathway in tissue homeostasis and regeneration. Protein Cell. 2017;8:349–359.
3. Wang J, Liu S, Heallen T, Martin JF. The Hippo pathway in the heart: Pivotal roles in development, disease, and regeneration. Nat Rev Cardiol. 2018;15:672–684.
4. Liu AM, Wong KF, Jiang X, Qiao Y, Luk JM. Regulators of mammalian Hippo pathway in cancer. Biochim Biophys Acta. 2012;1826:357–364.
5. Zhang K, Qi HX, Hu ZM, et al. YAP and TAZ take center stage in cancer. Biochemistry. 2015;54:6555–6566.
6. Meng Z, Moroishi T, Guan KL. Mechanisms of Hippo pathway regulation. Genes Dev. 2016;30:1–17.
7. Santucci M, Vignudelli T, Ferrari S, et al. The Hippo pathway and YAP/TAZ-TEAD protein-protein interaction as targets for regenerative medicine and cancer treatment. J Med Chem. 2015;58:4857–4873.
8. Moya IM, Castaldo SA, Van den Mooter L, et al. Peritumoral activation of the Hippo pathway effectors YAP and TAZ suppresses liver cancer in mice. Science. 2019;366:1029–1034.
9. Pobbati AV, Hong W. Emerging roles of TEAD transcription factors and its coactivators in cancers. Cancer Biol Ther. 2013;14:390–398.
10. Zanconato F, Cordenonsi M, Piccolo S. YAP/TAZ at the roots of cancer. Cancer Cell. 2016;29:783–803.
11. Gibault F, Coevoet M, Sturbaut M, et al. Toward the discovery of a novel class of YAP-TEAD interaction inhibitors by virtual screening approach targeting the YAP-TEAD protein-protein interface. Cancer. 2018;10:E140.
12. Holden JK, Cunningham CN. Targeting the Hippo pathway and cancer through the TEAD family transcription factors. Cancer. 2018;10:E81.
13. Hansen CG, Moroishi T, Guan KL. YAP and TAZ: A nexus for Hippo signaling and beyond. Trends Cell Biol. 2015;25:499–513.
14. Hilman D, Gat U. The evolutionary history of YAP and the Hippo/YAP pathway. Mol Biol Evol. 2011;28:2403–2417.
15. Sebé-Pedros A, Zheng Y, Ruiz-Trillo I, Pan D. Premetazoan origin of the Hippo signaling pathway. Cell Rep. 2012;1:13–20.
16. Wang K, Degerny C, Xu M, Yang XJ. YAP, TAZ, and Yorkie: A conserved family of signal-responsive transcriptional coregulators in animal development and human disease. Biochem Cell Biol. 2009;87:77–91.
17. Ikmi A, Gaertner B, Seidel C, Srivastava M, Zeitlinger J, Gibson MC. Molecular evolution of the Yap/Yorkie proto-oncogene and elucidation of its core transcriptional program. Mol Biol Evol. 2014;31:1375–1390.
18. Tian W, Yu J, Tomchick DR, Pan D, Luo X. Structural and functional analysis of the YAP-binding domain of human TEAD2. Proc Natl Acad Sci USA. 2010;107:7293–7298.
19. Feichtinger M, Sára T, Platzer G, et al. ^{1}H, ^{13}C, ^{15}N resonance assignment of human YAP 50–171 fragment. Biomol NMR Assign. 2018;12:179–182.
20. Chen L, Chan SW, Zhang XQ, et al. Structural basis of YAP recognition by TEAD4 in the Hippo pathway. Genes Dev. 2010;24:290–300.
21. Li Z, Zhao B, Wang P, et al. Structural insights into the YAP and TEAD complex. Genes Dev. 2010;24:235–240.
22. Hau JC, Erdmann D, Mesrouze Y, et al. The TEAD4-YAP/TAZ protein-protein interaction: Expected similarities and unexpected differences. Chembiochem. 2013;14:1218–1225.
23. Bokhovchuk F, Mesrouze Y, Delaunay C, et al. Identification of FAM181A and FAM181B as new interactors with the TEAD transcription factors. Protein Sci. 2020;29:509–520.
24. Zhang Z, Lin Z, Zhou Z, et al. Structure-based design and synthesis of potent cyclic peptides inhibiting the YAP-TEAD protein-protein interaction. ACS Med Chem Lett. 2014;5:993–998.
25. Mesrouze Y, Bokhovchuk F, Meyerhofer M, et al. Dissection of the interaction between the intrinsically disordered YAP protein and the transcription factor TEAD. Elife. 2017;6:e25068.
26. Mesrouze Y, Meyerhofer M, Bokhovchuk F, et al. Effect of the acylation of TEAD4 on its interaction with co-activators YAP and TAZ. Prot Sci. 2017;26:2399–2409.
27. Crooks GE, Hon G, Chandonia JM, Brenner SE. WebLogo: A sequence logo generator. Genome Res. 2004;14:1188–1190.
28. Hao Y, Chun A, Cheung K, Rashidi B, Yang X. Tumor suppressor LATS1 is a negative regulator of oncogene YAP. J Biol Chem. 2008;283:5496–5509.
29. Bokhovchuk F, Mesrouze Y, Meyerhofer M, et al. An early association between the alpha-helix of the TEAD binding domain of YAP and TEAD drives the formation of the YAP:TEAD complex. Biochemistry. 2020;59:1804–1812.
30. Furet P, Salem B, Mesrouze Y, et al. Structure-based design of potent linear peptide inhibitors of the YAP-TEAD protein-protein interaction derived from the YAP omega-loop sequence. Bioorg Med Chem Lett. 2019;29:2316–2319.
31. Mesrouze Y, Bokhovchuk F, Izaac A, et al. Adaptation of the bound intrinsically disordered protein YAP to mutations at the YAP:TEAD interface. Protein Sci. 2018;27:1810–1820.
32. Bokhovchuk F, Mesrouze Y, Izaac A, et al. Molecular and structural characterization of a TEAD mutation at the origin of Sveinsson's chorioretinal atrophy. FEBS J. 2019;286:2381–2398.
33. George RA, Heringa J. An analysis of protein domain linkers: Their classification and role in protein folding. Protein Eng. 2003;15:871–879.
34. Krieger F, Moglich A, Kiefhaber T. Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains. J Am Chem Soc. 2005;127:3346–3352.
35. Papaleo E, Saladino G, Lambrughi M, Lindorff-Larsen K, Gervasio FL, Nussinov R. The role of protein loops and linkers in conformational dynamics and allostery. Chem Rev. 2016;116:6391–6423.
36. Gaffney CJ, Oka T, Mazack V, et al. Identification, basic characterization and evolutionary analysis of differentially spliced mRNA isoforms of human YAP1 gene. Gene. 2012;509:215–222.

37. Chen Y, Han H, Seo G, et al. Systematic analysis of the Hippo pathway organization and oncogenic alteration in evolution. Sci Rep. 2020;10:3173.
38. Glasauer SM, Neuhauss SC. Whole-genome duplication in teleost fishes and its evolutionary consequences. Mol Genet Genomics. 2014;289:1045–1060.
39. Chan PY, Han X, Zheng B, et al. Autopalmitoylation of TEAD proteins regulates transcriptional output of the Hippo pathway. Nat Chem Biol. 2016;12:282–289.
40. Noland CL, Gierke S, Schnier PD, et al. Palmitoylation of TEAD transcription factors is required for their stability and function in Hippo pathway signalling. Structure. 2016;24:179–186.
41. Iwasa H, Maimaiti S, Kuroyanagi H, et al. Yes-associated protein homolog, YAP-1, is involved in the thermotolerance and aging in the nematode *Caenorhabditis elegans*. Exp Cell Res. 2013;319:931–945.
42. Pobbati AV, Chan SW, Lee I, Song H, Hong W. Structural and functional similarity between Vgll1-TEAD and YAP-TEAD complexes. Structure. 2012;20:1135–1140.
43. Mesrouze Y, Hau JC, Erdmann D, et al. The surprising features of the TEAD4-Vgll1 protein-protein interaction. Chembiochem. 2014;15:537–542.
44. Blaber M, Zhang XJ, Matthews BW. Structural basis of amino acid α helix propensity. Science. 1993;260:1637–1640.
45. Fujiwara K, Toda H, Ikeguchi M. Dependence of a-helical and b-sheet amino acid propensities on the overall protein fold type. BMC Struct Biol. 2012;12:18.
46. Howe KL, Contreras-Moreira B, De Silva N, et al. Ensembl genomes 2020-enabling non-vertebrate genomic research. Nucleic Acids Res. 2020;48:D689–D695.
47. Wu S, Liu Y, Zheng Y, Dong J, Pan D. The TEAD/TEF family protein scalloped mediates transcriptional output of the Hippo growth-regulatory pathway. Dev Cell. 2008;14:388–398.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Mesrouze Y, Bokhovchuk F, Meyerhofer M, et al. Study of the TEAD-binding domain of the YAP protein from animal species. *Protein Science*. 2021;30:339–349. https://doi.org/10.1002/pro.3988