


Biochemical properties of VGLL4 from *Homo sapiens* and Tgi from *Drosophila melanogaster* and possible biological implications

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

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

Correspondence

Patrick Chène, Disease Area Oncology,
Novartis Institutes for Biomedical
Research, WSJ 386 4.02.01, CH-4002
Basel, Switzerland.
Email: patrick_chene@yahoo.com

Abstract

The TEAD (Sd in drosophila) transcription factors are essential for the Hippo pathway. Human VGLL4 and drosophila Tgi bind to TEAD/Sd via two distinct binding sites. These two regions are separated by few amino acids in VGLL4 but they are very distant from each other in Tgi. This difference prompted us to study whether it influences the interaction with TEAD4/Sd. We show that the full-length VGLL4/Tgi proteins behave as intrinsically disordered proteins. They have a similar affinity for TEAD4/Sd revealing that the length of the region between the two binding sites has little effect on the interaction. One of their two binding sites (high-affinity site) binds to TEAD4/Sd 100 times more tightly than to the other site, and size exclusion chromatography experiments reveal that VGLL4/Tgi only form trimeric complexes with TEAD4/Sd at high protein concentrations. In solution, therefore, VGLL4/Tgi may predominantly interact with TEAD4/Sd via their high-affinity site to create dimeric complexes. In contrast, when TEAD4/Sd molecules are immobilized on sensor chips used in Surface Plasmon Resonance experiments, one VGLL4/Tgi molecule can bind simultaneously with an enhanced affinity to two immobilized molecules. This effect, due to a local increase in protein concentration triggered by the proximity of the immobilized TEAD4/Sd molecules, suggests that *in vivo* VGLL4/Tgi could bind with an enhanced affinity to two nearby TEAD/Sd molecules bound to DNA. The presence of two binding sites in VGLL4/Tgi might only be required for the function of these proteins when they interact with TEAD/Sd bound to DNA.

KEYWORDS

hippo pathway, scalloped, TEAD, Tgi, transcription factor, VGLL4

1 | INTRODUCTION

The TEAD (TEA/ATTS domain) transcription factors are the most distal elements of the Hippo pathway, which is essential in organ morphogenesis.^{1–4} These proteins can

bind to DNA but they are unable to stimulate transcription on their own. Their transcriptional activity is regulated via their association with different proteins such as YAP (Yes-associated protein), TAZ (Transcriptional co-activator with PDZ motif) or VGLL1-3 (Vestigial-like,

VGLL1-3 are paralogs; Vestigial [Vg] is the *Drosophila* ortholog).^{5–8} All these proteins bind to an overlapping area at the surface of TEAD (Scalloped [Sd] in *Drosophila*). The current structural data show that the TEAD-binding domain (TBD) of YAP and of its paralog TAZ is formed of a β -strand, an α -helix and an Ω -loop, while the TBD of VGLL1 contains only a β -strand and an α -helix.^{9–12} The “hot spot” for the interaction with TEAD is the Ω -loop in YAP/TAZ and the α -helix in VGLL1.^{10,11,13,14} Recently, it has been found that the Sd/TEAD-binding domain of Vg and VGLL2 contains a β -strand: α -helix region, but also an Ω -loop.¹⁵ The α -helix from the TEAD/Sd-binding domain of VGLL1/Vg contains an amino acid motif, Φ -D/E-D/E-H-F (Φ hydrophobic residue), that is characteristic for this protein family.⁵ This motif is located in a conserved region called Tondu domain (TDU).⁵ A similar motif, L-x-x-L/M-F (x: variable amino acid), is present in the α -helix of the TBD of the YAP proteins,¹⁶ but the α -helices from VGLL1/Vg and YAP do not interact exactly in the same manner with this transcription factor.^{15,17} Since the YAP/TAZ and VGLL1-3/Vg proteins contain only one TBD, it is not surprising to observe in the published X-ray structures that they form a 1:1 complex with TEAD^{9–11,15} (but see Ref. 12). Nevertheless, the current data have been generated with fragments of these proteins; more complex oligomers might be formed with the full-length proteins and in the presence of DNA.

In contrast to the VGLL1-3 proteins, the members of the VGLL4 (Tondu-domain-containing growth inhibitor [Tgi] is the *Drosophila* ortholog) family possess two TDU, suggesting that they may form with TEAD complexes with a molecular ratio higher than 1:1. Indeed, Jiao et al.¹⁸ have shown by gel filtration and dynamic light scattering that VGLL4 and TEAD form a complex with a 1:2 molecular ratio. The structure of the VGLL4:TEAD complex reveals how these two proteins interact at atomic resolution.¹⁸ The TBD of VGLL4 is formed of an α -helix (helix-1), a short linker, a β -strand and two additional α -helices (helix-2 and helix-3). A Φ -D/E-D/E-H-F motif is present in helix-1 and in helix-2. Jiao et al. reported that helix-1 binds to a different TEAD molecule than the β -strand:helix-2:helix-3 region with the conserved Φ , H and F residues from helix-1 facing TEAD (these interactions are not visible in the structure currently deposited at the Protein Data Bank (PDB code 4ln0; Figure S1).¹⁸ The β -strand forms an intramolecular β -sheet with two TEAD molecules (Figure S1). Helix-2 binds to TEAD in a manner similar to that of the α -helix of the TEAD/Sd-binding domain from VGLL1/Vg with the conserved Φ , H and F residues located at the binding interface (Figure 1). Finally, helix-3 folds back on helix-2 and makes additional interactions with TEAD (Figure 1).

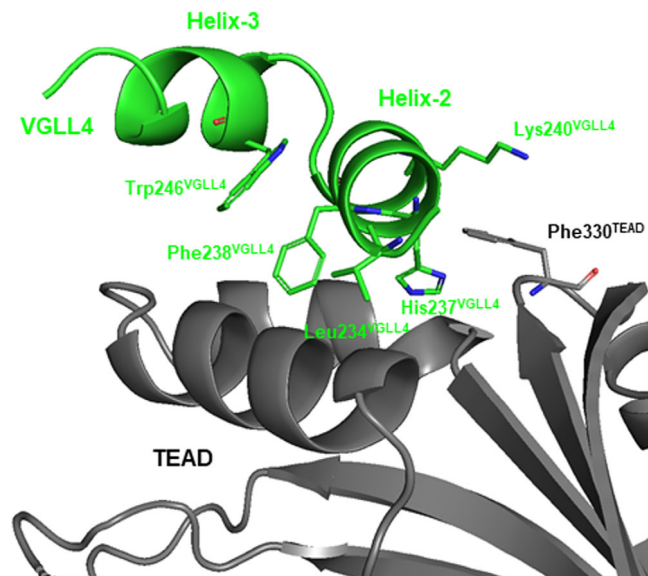


FIGURE 1 Structure of the VGLL4:TEAD4 complex. The region located at the interface between helix-2:helix-3 (in green) and TEAD (in grey) is represented. The three residues Leu234^{VGLL4}, His237^{VGLL4}, and Phe238^{VGLL4} from the Φ -D/E-D/E-H-F motif are indicated. Trp246^{VGLL4} present in helix-3 and Lys240^{VGLL4} that makes a π -cation interaction with Phe330^{TEAD} are also shown. The figure was drawn from the structure of the VGLL4:TEAD complex (PDB 4ln0) with PyMOL (Schrödinger Inc., Cambridge, MA)

In VGLL4 and Tgi, the two Φ -D/E-D/E-H-F motifs are distant from each other by 23 and 224 residues, respectively. This substantial difference in the structure of these two proteins prompted us to study in various biochemical assays whether VGLL4 and Tgi interact in a similar fashion with TEAD4 and Sd.

2 | MATERIALS AND METHODS

2.1 | Synthetic peptides

The synthetic peptides (both N-acetylated and C-amidated) were purchased from Biosynthan (Germany). The two peptides VGLL4^{206–220} (Ac-DPVVEEHFRRSLGKN-NH₂) and VGLL4^{230–256} (Ac-SVSITGSVDDHFAKALGDTWLQIKAAK-NH₂) are derived from the sequence of human VGLL4 (UniProtKB Q14135, Isoform 1). The two Tgi peptides Tgi^{129–143} (Ac-MCDIDEHFRRSLGEN-NH₂) and Tgi^{354–378} (Ac-FTKTEASVDDHFAKALGETWKKLQG-NH₂) are derived from the sequence of Tgi from *Drosophila melanogaster* (UniProtKB Q8IQJ9). The purity 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.

2.2 | Protein cloning, expression, and purification

Human TEAD4 (region 217–434) and Sd from *D. melanogaster* (region 223–440) were purified as previously described.^{15,19} The amino acid sequences of human VGLL4 (UniProtKB Q14135, amino acids 2–290) and Tgi (UniProtKB Q8IQJ9, amino acid 2–382) from *D. melanogaster* were back-translated into an *Escherichia coli* codon-optimized DNA sequence using the GeneArt online ordering tool. The coding sequence was extended by *LguI* recognition site adaptors and the DNA synthesized by GeneArt (Thermo Fisher, Switzerland). The DNA fragments encoding VGLL4 and Tgi were subcloned by T2S restriction enzyme cloning²⁰ into pET derived vectors providing an N-terminal His₆-tag, a glycine-serine spacer, a lipoyl-solubilizing tag (in-house design), and a second glycine-serine spacer, followed by a HRV3C protease cleavage site. The VGLL4 expression construct comprised an additional Streptag II following the His-tag. In brief, 200 ng of vector and 200 ng of synthetic DNA insert were incubated in buffer B (Fermentas, Waltham, MA), supplemented with 1 mM ATP (Fermentas, Waltham, MA), 1 mM DTT (Fermentas, Waltham, MA), 5 U *LguI* (Thermo Fisher, Switzerland), and 1 μ L T4 DNA Ligase (Roche, Switzerland) for 1 cycle of 30 min at 37°C followed by 30 cycles (30 min at 37°C, 1 min at 10°C, 1 min at 30°C, [total time: 90 min]). The final expression constructs were confirmed by Sanger sequencing.

The expression plasmids were transformed into *E. coli* NiCo21 (DE3) cells (New England Biolabs, Ipswich, MA) and recombinant proteins expressed in 1 L LB Medium with Kanamycin (25 μ g/mL, Invitrogen, Carlsbad, CA) in Erlenmeyer flasks upon induction with 0.2 mM isopropyl β -D-1-thiogalactopyranoside (IPTG, PanReac Applichem, Germany) at 18°C, 220 rpm, for 16 h.

The bacterial cells from two liters expression volume were resuspended in 80 mL lysis buffer (50 mM Tris pH 8.0, 1 M NaCl, 2 mM MgCl₂, 1 mM TCEP, 20 mM imidazole, 0.1% Tween 20, 5% glycerol), supplemented with cComplete EDTA free protease inhibitor (1 tablet/50 mL buffer; Roche, Switzerland) and TurboNuclease (20 μ L/50 mL buffer; Sigma-Aldrich, St Louis, MI). The cells were lysed by three passages through a high-pressure homogenizer (Avestin Emulsiflex C3) at 800–1000 bar and the lysate centrifuged for 40 min at 40,000 g (Sorvall Lynx, F20-12x50) to remove insoluble material.

The cell lysate was loaded (flow 2 mL/min) onto two HisTrap HP 1 mL columns (Immobilized metal affinity chromatography [IMAC], Cytiva, Marlborough, MA) mounted on an ÄKTA Pure 25 FPLC system. Unbound material was washed away with 10 column volumes

(CV) Buffer 1 (50 mM Tris pH 8.0, 1 M NaCl, 2 mM MgCl₂, 1 mM TCEP, 20 mM imidazole, 0.1% Tween 20, 5% glycerol), followed by 10 CV of Buffer 2 (50 mM Tris pH 8.0, 300 mM NaCl, 2 mM MgCl₂, 1 mM TCEP, 5% glycerol) supplemented with 20 mM imidazole. Bound protein was eluted with a 0%–100% linear gradient Buffer 2 with 250 mM imidazole over 10 CV in 1 mL fractions.

VGLL4 and Tgi protein pools were treated with His₆-MBP-3C protease (produced in-house) during dialysis overnight at 5°C in a Slide-A-Lyzer cassette (MWCO 3.5 kDa, Thermo Fisher, Switzerland) against 2 L of Buffer 3 (50 mM Tris pH 8.0, 300 mM NaCl, 2 mM MgCl₂, 1 mM TCEP, 5% glycerol), or Buffer 4 (50 mM Tris pH 8.0, 225 mM NaCl, 2 mM MgCl₂, 1 mM TCEP, 5% glycerol, 0.05% Tween 20).

The proteins were subjected to a reverse IMAC purification step on two HisTrap HP 1 mL columns (flow 2.0 mL/min; Cytiva, Marlborough, MA) with 10 CV of Buffer 3 and 4, respectively, followed by 10 CV of Buffer 2 supplemented with 20 mM imidazole, where the cleaved, untagged proteins eluted and were collected in 1 mL fractions. The VGLL4 and Tgi protein solutions were concentrated with Amicon Ultra 15 concentrators (10 kDa MWCO, Millipore, Burlington, MA) and polished by size exclusion chromatography with Buffer 5 (50 mM HEPES pH 7.4, 100 mM KCl, 0.25 mM TCEP, 1 mM EDTA, 0.05% Tween 20) on Superdex 75 10/300 (Cytiva, Marlborough, MA) or ProteoSEC Dynamic 16/60 3–70 HR (ProteinArk, UK) columns, respectively. The 0.5 mL fractions containing the recombinant protein were pooled and snap frozen on dry ice.

The purity, concentration, and identity of the proteins was determined by RP-HPLC and LC–MS (Figure S3).

2.3 | Fluorescence thermal shift assay

The proteins (2 μ M) were diluted in 50 mM HEPES, 100 mM KCl, 0.25 mM TCEP, 1 mM EDTA, 2% (vol/vol) DMSO, pH 7.4, containing 2 \times SYPRO Orange dye (ThermoFisher Scientific, Waltham, MA). The protein solutions were then added to 384-well, thin-walled Hard-Shell PCR microplates (BioRad, Hercules, CA) that were covered by optically clear adhesive seals to prevent evaporation. Measurements were carried out with a CFX384 Real-Time PCR Detection System (BioRad, Hercules, CA). The temperature was increased from 20 to 95°C at 1°C/30 s, and the fluorescence intensity was measured with the excitation and emission filters set to 465 and 590 nm, respectively. The data were analyzed with the CFX Manager software (BioRad, Hercules, CA).

2.4 | Circular dichroism

The proteins were dialyzed in 20 mM phosphate buffer pH 7.4, 100 mM KF, 0.25 mM TCEP and diluted in this buffer to approximately 0.2 mg/mL. CD spectra were recorded on a J-815 spectropolarimeter (Jasco, France) using 1 mm path length quartz cells (instrument setup: “standard” sensitivity, 0.5 nm bandwidth, 10 nm/min scanning, 2 s digital integration time, 1 nm step resolution). The measuring chamber was maintained in nitrogen (17 L/min). Each spectrum was recorded as an average of four scans to reduce noise. After baseline correction, the mean residue ellipticity was calculated $[\theta]_{MRW,\lambda} = MRW \cdot \theta_{\lambda} / 10 \cdot d \cdot c$, where MRW is the mean residue weight ($MRW = M / [N - 1]$), M is the molecular mass of the protein, N is the number of amino acids in the protein, θ_{λ} is the observed ellipticity (degrees) at wavelength λ , d is the path length (cm) of the cell, and c is the protein concentration (g/mL).

2.5 | Size exclusion chromatography-multiangle light-scattering

Proteins were mixed in a 1:2 stoichiometric ratio (Tgi:Sd or VGLL4:TEAD4), concentrated with an Amicon Ultra 15 concentrator (10 kDa MWCO, Millipore, Burlington, MA) and the final protein concentration was checked by RP-HPLC analysis. The protein samples were next analyzed at two different concentrations by analytical Size Exclusion Chromatography coupled to a Multi-Angle Light-Scattering detector (SEC-MALS).²¹ Briefly, an Agilent 1100 series HPLC system (Agilent, Santa Clara, CA) with an Optilab T-rEX refractive index detector and a DAWN Heleos-II light-scattering detector (Wyatt, Santa Barbara, CA) was equipped with a Superdex 200 Increase 5/150 GL column (Cytiva, Marlborough, MA). The column was equilibrated with 50 mM Tris, 100 mM NaCl, 2 mM MgCl₂, 1 mM TCEP, pH 8.0 and the protein samples (60 μ L) were injected onto the column. Data collection and processing was performed with the ASTRA software (v. 7.1; Wyatt, Santa Barbara, CA). The protein concentration was calculated using the refractive index signal (assuming $dn/dc = 0.185$). Molecular mass moments at peak maximum were calculated using the Zimm fit method.²²

2.6 | Surface plasmon resonance

All the experiments were carried out with a Biacore T200 optical biosensor and Series S sensor Chip SA (Cytiva, Marlborough, MA). The chips were washed three times

with 1 M NaCl/50 mM NaOH. The N-Avitagged proteins (TEAD4²¹⁷⁻⁴³⁴ and Sd²²³⁻⁴⁴⁰) were injected at a flow rate of 5 mL/min in SPR immobilization buffer (50 mM HEPES, 100 mM KCl, 0.25 mM TCEP, 1 mM EDTA, 0.05% (vol/vol) Tween 20, 0.05% (wt/vol) BSA, pH 7.4). The experiments were performed at 25°C with a flow rate of 50 mL/min in SPR running buffer (SPR immobilization buffer containing 2% (vol/vol) DMSO). The tested analytes were diluted in SPR running buffer. After baseline equilibration with a series of buffer blanks, a DMSO correction series was performed from 1% to 3%. Each cycle consisted of an injection phase of the analytes and a dissociation phase. All data were referenced for a blank streptavidin reference surface and blank injections of running buffer to minimize the influence of baseline drift upon binding.

The sensorgrams obtained with the synthetic peptides mimicking VGLL4 and Tgi were globally fitted with a 1:1 interaction model using the Biacore T200 evaluation software (Cytiva, Marlborough, MA). For the experiments carried out with the full-length proteins, the signal measured for 5 s before the dissociation phase was extracted from the Biacore device and fitted with GraphPad Prism (v. 9.0.1; GraphPad Software, San Diego, CA) using a one-site ($\text{signal} = B_{\text{max}} \cdot [L] / (K_d + [L]) + \text{Bkcgd}$) or a two-sites ($\text{signal} = B_{\text{max}}^{\text{Hi}} \cdot [L] / (K_d^{\text{Hi}} + [L]) + B_{\text{max}}^{\text{Lo}} \cdot [L] / (K_d^{\text{Lo}} + [L]) + \text{Bkcgd}$) binding model (Signal: signal measured, see above; [L]: [VGLL4] or [Tgi]; K_d : equilibrium dissociation constant; K_d^{Hi} and K_d^{Lo} : equilibrium dissociation constants of the high and low-affinity sites, respectively; B_{max} : signal at saturation; $B_{\text{max}}^{\text{Hi}}$ and $B_{\text{max}}^{\text{Lo}}$: signal at saturation for the high and low-affinity sites, respectively; Bkcgd: background signal (offset) at [L] = 0). The results obtained with the two models were compared using the Akaike's Information Criterion.²³

The immobilization levels were between 9 and 14 RU for the experiments carried out at low immobilization levels with Tgi; 48 and 70 RU for the experiments carried out at high immobilization levels with Tgi; 8 and 12 RU for the experiments carried out at low immobilization levels with VGLL4; 47 and 63 RU for the experiments carried out at high immobilization levels with VGLL4.

3 | RESULTS AND DISCUSSION

3.1 | Amino acid sequence analysis

Since the two Φ -D/E-D/E-H-F motifs are significantly more distant from each other in the amino acid sequence of Tgi from *D. melanogaster* than in human VGLL4 (Figure S2), we considered whether the presence of a long linker between the two binding sites is specific to

Tgi from this insect species or if homologs from other insect species have the same feature. 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 Tgi homologs in the protein sequences from insect species deposited at the non-redundant protein sequences database (National Centre for Biotechnology Information; Taxids 6,960 and 50,557; November 2020). The query sequence was the region 347–382 from Tgi (SSTNISIFTKTEASVDDHFAKALGETWKKLQGGHKE, UniProtKB Q8IQJ9), which contains the putative β -strand:helix-2:helix-3 region of this protein. This search led to 288 entries (Supplementary Material S1) from which we identified 159 unique sequences belonging to 156 different insect species (Figure S2). The Φ -D/E-D/E-H-F motifs were localized in each sequence and to increase the likelihood that they are present in regions similar to helix-1 and helix-2:helix-3, we used two structural features. In VGLL1, an arginine residue located two amino acids after the phenylalanine of the Φ -D/E-D/E-H-F motif makes an important π -cation interaction with a phenylalanine from TEAD.¹⁷ Since a similar interaction is observed in the Vg:Sd complex,¹⁵ but also between Lys240^{VGLL4} and Phe330^{TEAD} in the VGLL4:TEAD complex (Figure 1),¹⁸ we looked for the presence of an arginine/lysine at this position in each sequence. These charged residues are present in all sequences identified in the BLASTp search, and interestingly an arginine is always found after the Φ -D/E-D/E-H-F motif in the region corresponding to helix-1, while this is a lysine in the region corresponding to helix-2 (Figure S2). Trp246^{VGLL4} is located in the middle of helix-3 and it makes key hydrophobic interactions in the VGLL4:TEAD complex (Figure 1).¹⁸ All sequences possess a tryptophan residue at a position similar to that of Trp246^{VGLL4} in VGLL4 (Figure S2). Altogether, this suggests that the sequences identified in the BLASTp search contain regions similar to helix-1 and helix-2:helix-3 from VGLL4.

We next counted the number of amino acids between the two Φ -D/E-D/E-H-F motifs in each of these protein sequences (Figure S2). This number, which varies from 24 (*Acromyrmex echinator*) to 314 (*Ceratitis capitata*), is more constant within each insect order: Coleoptera 44 to 51, Diptera 128 to 314 (except *Contarinia nasturtii*), Hemiptera 27 to 61, Hymenoptera 24 to 35 or Lepidoptera 52–65. This shows that the number of amino acids between both motifs is quite variable and that Tgi from *D. melanogaster* is not the only protein with a large number of residues between its two motifs. Nevertheless, the diversity in the length of the region connecting the two motifs is intriguing. This region, particularly when it is

long, could have other functions in addition to connecting both motifs. This is indeed the case in Tgi from *D. melanogaster* where it contains PPxY (x: any amino acid) sequences, the recognition motif for WW domains, which are important for the function of this protein.^{24,25} PPxY sequences are also present in Tgi proteins from other *Drosophila* species, but they are absent in the sequences of other Diptera species such as for example *Anopheles sinensis*, *A. stephensi* and *A. darlingi* (Seq. Id. KFB47446.1, XP_035898942.1 and ETN57905.1, respectively) (Supplementary Material S1). This indicates that the region between the Φ -D/E-D/E-H-F motifs may enable Tgi to acquire functions required in a species-specific cellular context, while still binding to Sd. *In vivo* studies of more VGLL4/Tgi homologs may help in gaining a better understanding of the role of this region in the function of these proteins.

The BLASTp search also reveals that all the sequences from Lepidoptera species contain three Φ -D/E-D/E-H-F motifs (Figure S2). The additional motif looks similar to a region corresponding to helix-2:helix-3. The role (if any) of this additional region is unknown and further studies may help finding out whether three Φ -D/E-D/E-H-F motifs are required for the function of Tgi from Lepidoptera.

3.2 | Characterization of Tgi and VGLL4

The previous findings are puzzling and raise several questions on the biological role and properties of the proteins corresponding to the sequences identified in our BLASTp search. Nevertheless, the function of Tgi from *D. melanogaster* and its interaction with Sd have been clearly established.^{24,25} This protein is very different from human VGLL4 because it contains a much larger number of amino acids between its two Φ -D/E-D/E-H-F motifs. Therefore, we decided to examine whether this difference in the structure of these proteins affects their interaction with TEAD4 or Sd. To facilitate this study, the full-length Tgi and VGLL4 proteins were cloned, expressed in *E. coli* and purified to homogeneity (Figure S3).

Tgi and VGLL4 were subjected to a fluorescence thermal shift assay (FTSA) to determine their thermal stability. In such an assay, the fluorescence of a dye, SYPRO orange, is measured at different temperatures in the presence of the protein. In solution, the fluorescence of SYPRO orange is quenched by water molecules, but when it interacts with hydrophobic regions, it dissociates from the water molecules and its fluorescence increases. Therefore, when the temperature is raised during FTSA, the globular proteins that unfold expose the residues of their hydrophobic core to the dye, triggering an increase

in fluorescence. Under our experimental conditions, the level of fluorescence measured with Tgi and VGLL4 does not significantly change when the temperature increases (Figure 2(a)). This behavior is similar to that observed with the intrinsically disordered YAP fragment, YAP⁵⁰⁻¹⁷¹, but it is different to that of the well-folded TEAD4 and Sd proteins (Figure 2(a)).

We next studied the structure of Tgi and VGLL4 in solution by circular dichroism (CD). The spectra obtained with these two proteins show that there is no significant change in molar ellipticity in the 210–230 nm region (Figure 2(b)). This is similar to the spectrum obtained with YAP⁵⁰⁻¹⁷¹, but a marked change in ellipticity is observed in the spectra of TEAD4 and Sd (Figure 2(b)).

The absence of bands in the 210–230 nm region in the spectra of Tgi/VGLL4 suggest that—under our experimental conditions—no stable β -sheets or α -helices are present in these proteins.

Altogether, the study of the thermal stability and of the structure in solution of VGLL4/Tgi suggest that they behave more as intrinsically disordered proteins than globular proteins.

3.3 | Binding to Sd and TEAD4

To study the VGLL4:TEAD4 and Tgi:Sd interactions, TEAD4/Sd were immobilized on sensor chips and the

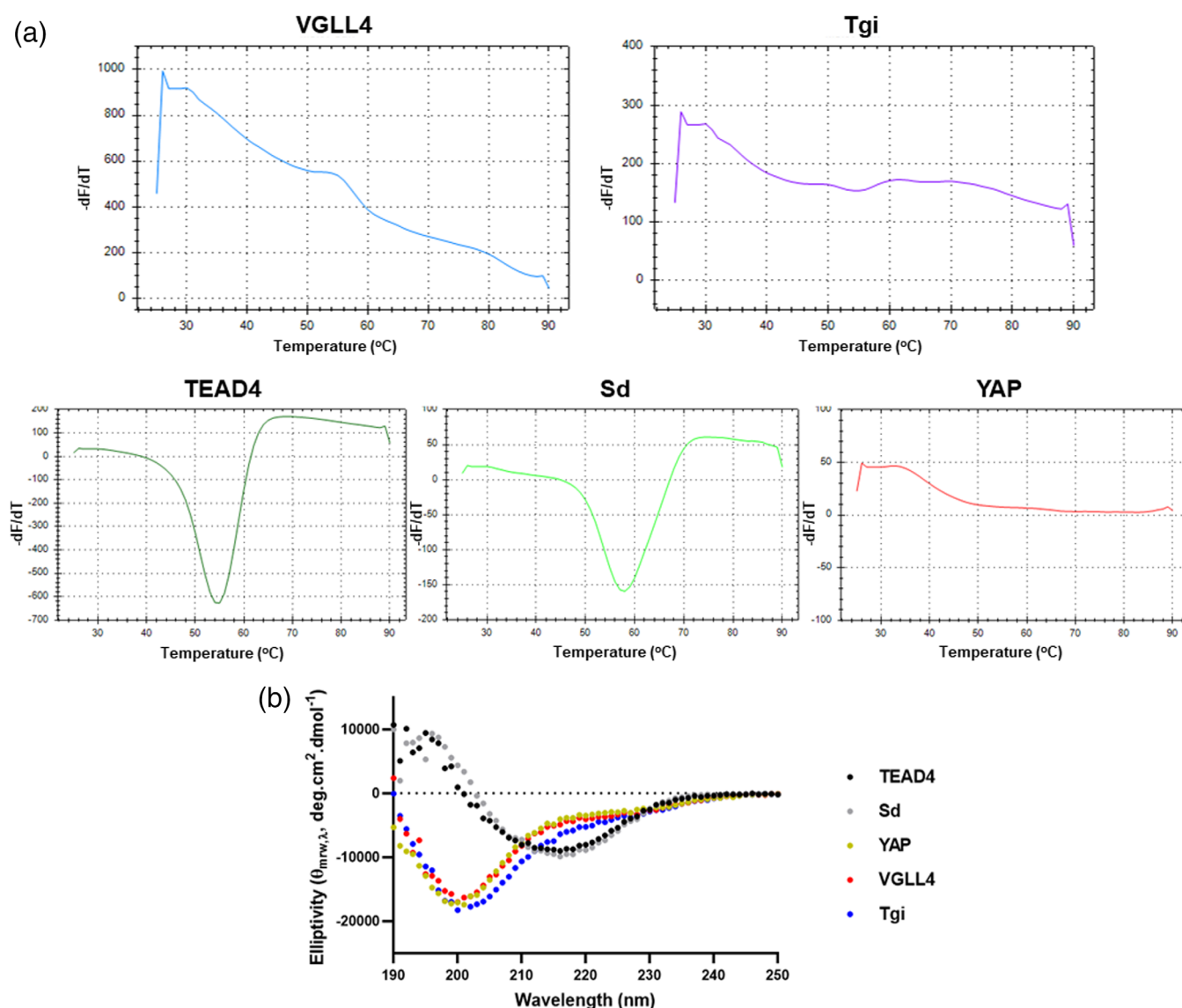


FIGURE 2 Fluorescence thermal shift assay and circular dichroism. (a) Fluorescence thermal shift assay. The proteins (2 μ M) were heat-denatured in the presence of SYPRO orange, and the fluorescence of the dye was recorded between 25 and 95°C. The figure represents the plots of the first derivative of the fluorescence emission (in relative fluorescence unit [RFU]) as a function of temperature. (b) Circular dichroism. The different proteins (0.2 mg/mL) were analyzed by CD as described in the Material and Methods section (Section 2)

affinity (K_d) of VGLL4/Tgi was determined by Surface Plasmon Resonance (SPR) (Figure S4). In a first step, synthetic peptides corresponding to the helix-1 and to the β -strand:helix-2:helix-3 regions of VGLL4/Tgi were designed using the sequences and the structural data provided by Jiao et al.¹⁸ (Figure 3(a)). The affinity of the peptides corresponding to the same binding site of VGLL4 and Tgi is similar (Figure 3(a)). As expected from the structure of the VGLL4:TEAD complex, the peptides mimicking the β -strand:helix-2:helix-3 region have a higher affinity for TEAD4/Sd than the peptides mimicking helix-1 (Figure 3(a)). The K_d of the VGLL4/Tgi peptides was compared with the data published on peptides derived from proteins of the VGLL1-3/Vg family. The affinity of the peptides mimicking helix-1 is 4–6 times lower than that of a peptide corresponding to the α -helix of the TBD from VGLL1¹⁴ (Figure 3(a)). The absence in VGLL4/Tgi of a residue corresponding to Ile37^{VGLL1}, which contributes to the VGLL1:TEAD interaction,¹⁷ and the presence of Pro207^{VGLL4}, which may destabilize the formation/stabilization of a α -helix in VGLL4, could explain the lower affinity of the VGLL4/Tgi peptides. The K_d of the peptides mimicking the β -strand:helix-2:helix-3

region of VGLL4/Tgi is very similar to that of the peptides corresponding to the β -strand: α -helix region from the TEAD/Sd-binding domain of VGLL1/VGLL2/Vg^{14,15} (Figure 3(a)). This is particularly surprising because VGLL1/VGLL2/Vg do not possess helix-3, which contributes to the VGLL4:TEAD interaction.¹⁸ Two observations may explain this finding. Jiao et al. have shown that the β -strand has a weak contribution to the formation of the VGLL4:TEAD complex¹⁸ while this secondary structure element is important for the interaction between VGLL1 and TEAD.¹⁴ The region located between the β -strand and the Φ -D/E-D/E-H-F motif is different between VGLL4 and VGLL1 (Figures 3(a) and S5). Therefore, the interactions created by helix-3 with TEAD/Sd and the bound helix-2 may compensate the weaker contacts formed by the region located at the N-terminus of the Φ -D/E-D/E-H-F motif (including the β -strand). This could explain why the peptides mimicking the β -strand:helix-2:helix-3 region of VGLL4/Tgi do not have a higher affinity than the peptides corresponding to β -strand: α -helix region of VGLL1/VGLL2/Vg.

The formation of ternary complexes between VGLL4/Tgi and TEAD4/Sd requires two bimolecular reactions

| (a) | K_d (nM) | α_1 OR α | | |
|----------------------------|-----------------------|------------------------|---|------------|
| | | ←-----→ | | |
| VGLL4 | 54000±4000 | 206- | DPVVEEHFRRSLGKN -220 | |
| Tgi | 33000±3000 | 129- | MCDIDEHFRRSLGEN -143 | |
| VGLL1 ^a | 9000±1000 | 36- | DINSMVDEHFSRALRN -51 | |
| | | β | α_2 OR α | α_3 |
| | | ←-----→ | ←-----→ | ←-----→ |
| VGLL4 | 360±20 | 230- | SVSIT ----- GSVDDHFAKALGDTWLQIKAAK -256 | |
| Tgi | 340±20 | 354- | FTKTE ----- ASVDDHFAKALGETWKKLQG -378 | |
| VGLL1 ^a | 160±20 | 27- | SVIFTYFEGDINSMVDEHFSRALRN | - 51 |
| VGLL2 ^b | 340±30 | 85- | CVLFTYFQGDISSVDEHFSRALS | -108 |
| Vg ^b | 291±4 | 288- | CVVFTNYSGDTASQVDEHFSRALN | -311 |
| (b) | | | | |
| Low immobilization levels | | | | |
| | K_d^{app} (nM) | | | |
| VGLL4 | 300±20 | | | |
| Tgi | 210±20 | | | |
| High immobilization levels | | | | |
| | $K_d^{app, low}$ (nM) | $K_d^{app, high}$ (nM) | | |
| VGLL4 | 370±30 | 8±1 | | |
| Tgi | 470±20 | 2.3±0.1 | | |

FIGURE 3 Affinities for TEAD4 and Sd. (a) The amino acid sequence of the peptides and the secondary structure they adopt in their bound conformation are indicated (VGLL1 from PDB 5z2q;¹¹ Vg from PDB 6y20;¹⁵ VGLL4 from PDB 4ln0¹⁸ - α 1-3: correspond to helix-1-3). The peptide sequences are from human VGLL1 UniProtKB Q99990; human VGLL2 UniProtKB Q8N8G2; human VGLL4 UniProtKB Q14135; *Drosophila melanogaster* Vg UniProtKB Q26366 and *D. melanogaster* Tgi UniProtKB Q8IQJ9. The residues in bold are mentioned in the text. The affinities (K_d) were obtained by surface plasmon resonance. The values are the averages and the standard errors of $n \geq 2$ experiments. a. K_d taken from¹⁴ and b K_d taken from.¹⁵ (b) The affinities of VGLL4/Tgi for TEAD4/Sd immobilized at low and high levels are indicated. The values represent the averages and the standard errors of $n \geq 2$ experiments. $K_d^{app, low}$ and $K_d^{app, high}$ correspond to the apparent binding affinities of the low and high-affinity sites, respectively

(Figure 4(a)). The large difference in affinity measured with the peptides mimicking the two binding sites of VGLL4/Tgi, suggests that, in solution and at low protein concentration, VGLL4/Tgi should predominantly bind to TEAD/Sd via the high-affinity site (β -strand:helix-2:helix-3) to form binary complexes. However, if higher protein concentrations are present, a ternary complex could be created via the interaction of the low-affinity site (helix-1). To check this hypothesis, we conducted SEC-MALS (Size Exclusion Chromatography-Multiangle Light-Scattering) experiments. In these studies, VGLL4/Tgi were incubated with TEAD4/Tgi at two different protein concentrations. At the lowest protein concentration, dimeric complexes are predominantly formed, while at much higher protein concentrations ternary complexes are also present (Figure 5). This shows that—in solution—the formation of a ternary complex between VGLL4/Tgi and TEAD4/Sd requires the presence of high protein concentrations.

The protein surface density (immobilization level) on sensor chips can be controlled in SPR experiments. At very low immobilization levels, the TEAD/Sd molecules are distant from each other and the probability for a VGLL4/Tgi molecule to interact with two of them should be reduced. Since the region corresponding to helix-1 has a very low affinity, the K_d measured in such conditions should be similar to the affinity determined with the peptides mimicking the β -strand:helix-2:helix-3 region (high-affinity site). At higher immobilization levels, there is a

greater probability that a VGLL4/Tgi molecule, already bound via the high-affinity site to an immobilized TEAD4/Sd molecule, will interact via its low-affinity site with a second nearby immobilized molecule before it dissociates from the sensor chips (Figure 4(b)). This proximity effect (avidity^{26,27}) should enhance the apparent overall affinity. SPR experiments were conducted at two different immobilization levels of TEAD4/Sd (Figure S4). However, one should keep in mind that these measurements are governed by the properties of the sensor chips' surface,²⁸ the buffer flow in the measurement cell²⁹ and may result from concomitant binding mechanisms. Therefore, we shall report them as apparent affinities, K_d^{app} . The SPR data were fitted using a one-site and a two-sites binding model (see Material and Methods for details). The one-site binding model best fitted the data obtained at low immobilization levels, while the two-sites binding model gave superior results at higher immobilization levels (Figure S4). The K_d^{app} values measured for VGLL4/Tgi at low immobilization levels (Figure 3(b)) are similar to the K_d measured with the peptides mimicking the β -strand:helix-2:helix-3 region (Figure 3(a)). This suggests that under these experimental conditions, VGLL4/Tgi bind to immobilized TEAD/Sd essentially via the high-affinity site. The affinities measured at higher immobilization levels are strikingly different. A $K_d^{app,high}$ (high-affinity site) in the low nanomolar range and a $K_d^{app,low}$ (low-affinity site) in the triple-digit nanomolar range were determined (Figure 3(b)). This shows that the

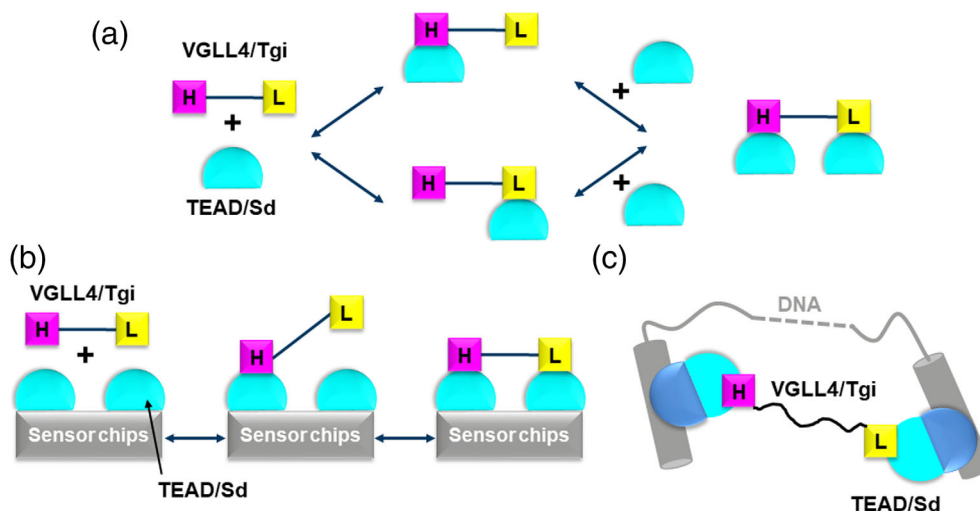


FIGURE 4 Schematic representation of the binding mechanism of VGLL4/Tgi to TEAD/Sd. (a). Proposed binding mechanism in solution. (b). Proposed binding mechanism in SPR experiments at high immobilization levels. TEAD/Sd are represented in light blue. The magenta (H) and yellow (L) squares correspond to the high and low-affinity sites of VGLL4/Tgi, respectively. The grey rectangles represent part of the sensor chips used in SPR where TEAD/Sd are immobilized. (c). Hypothetical interaction between VGLL4/Tgi with TEAD/Sd bound to DNA. The DNA-binding domain of TEAD/Sd is represented in dark blue. The grey cylinders correspond the TEAD/Sd DNA-binding elements. The dotted lines in the schematic representation of DNA indicate that the distance between the TEAD/Sd DNA-binding elements may vary

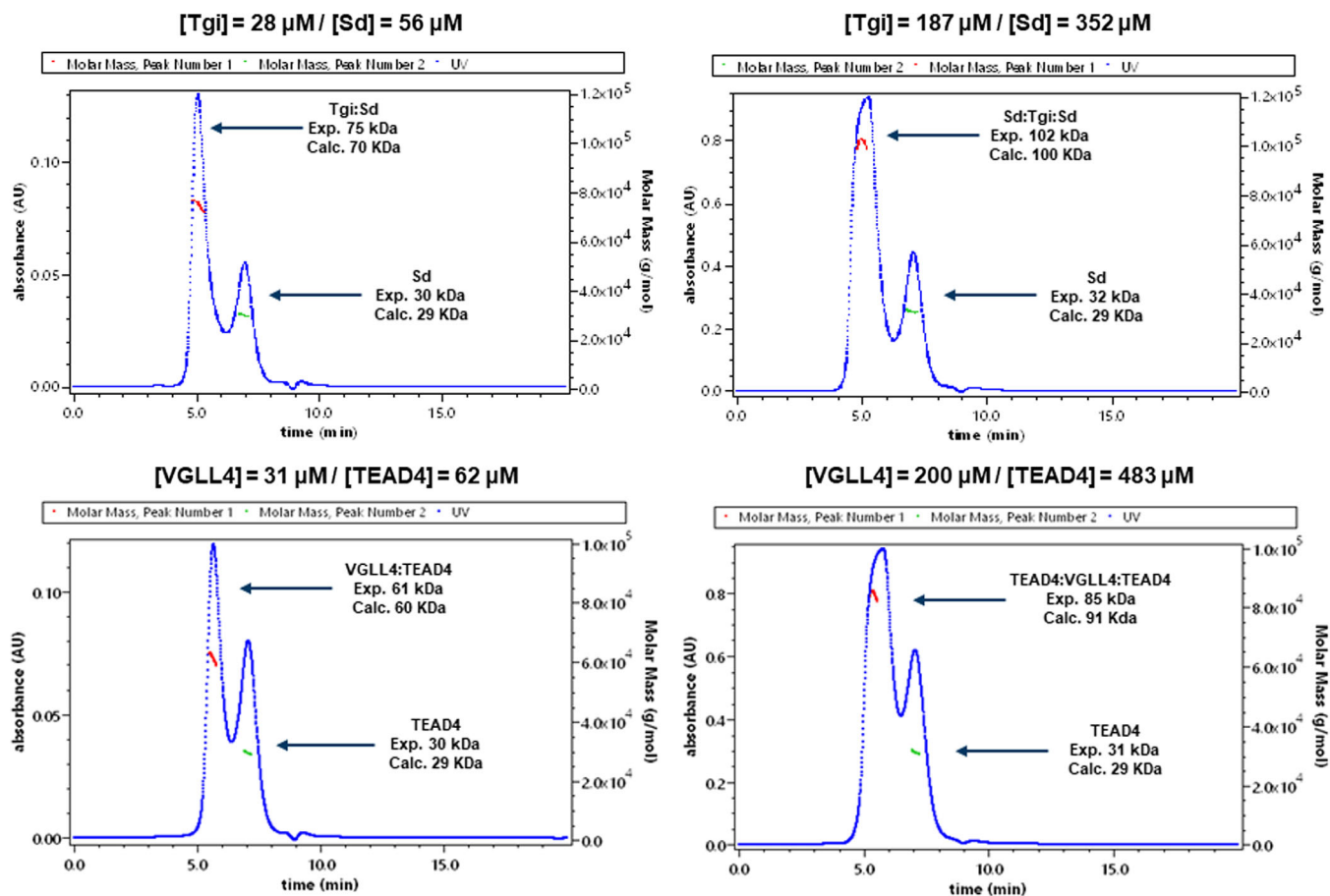


FIGURE 5 Size exclusion chromatography/multiangle light-scattering (SEC-MALS) experiments. VGLL4/Tgi and TEAD4/Sd were preincubated at the indicated concentrations and subjected to size exclusion chromatography (see Material and Methods for details [Section 2]). The figure represents the chromatograms from different SEC-MALS experiments. The mass of the proteins or protein complexes determined experimentally (Exp.) or calculated (Calc.) are indicated

affinity of VGLL4/Tgi is enhanced when the density in TEAD4/Sd molecules is increased at the surface of the sensor chips and that, in such conditions, one VGLL4/Tgi molecule interacts with two TEAD4/Sd molecules. Jiao et al., who have shown that VGLL4 binds to two TEAD molecules, have also measured the affinity of VGLL4 (probably residues 203–256) for immobilized TEAD4 by Biolayer Interferometry, but they report a single K_d value of 6.8 nM.¹⁸

4 | CONCLUSIONS

Overall, our results show that—under our experimental conditions—VGLL4 and Tgi behave in a very similar fashion. Therefore, using our tools and assays, we do not observe a significant effect of the length of the region between the two Φ -D/E-D/E-H-F motifs on the interaction with TEAD4/Sd. The difference in affinity between the two binding sites present in VGLL4/Tgi is so large, more

than 100-fold, that in solution and at low concentrations, these proteins should predominantly form dimeric complexes with TEAD/Sd. Therefore, in such conditions, VGLL4/Tgi would interact via their high-affinity TDU with these transcription factors in a similar fashion to the proteins from the VGLL1-3/Vg family, which contain only one TDU. However, our results suggest an additional possibility in the way VGLL4/Tgi and TEAD/Sd might interact in cells. Several molecules of the DNA-binding domain of TEAD can bind to the same DNA fragment containing different TEAD binding elements^{30,31} suggesting that, in cells, full-length TEAD molecules could be in proximity once they are bound to DNA. In such a case, VGLL4 may interact with them in a “similar” fashion to that observed in our SPR experiments at high immobilization levels with the DNA taking the role of the sensor chips (Figure 4(c)). The binding of VGLL4 via its high-affinity site to a first TEAD:DNA complex would increase the local concentration, enabling the low-affinity site to interact with another TEAD molecule also bound to DNA. Therefore, the

presence of two TDU domains might be important for the biological function of the proteins from the VGLL4/Tgi family in a DNA context. The variation in the length of the region connecting the Φ -D/E-D/E-H-F motifs amongst the VGLL4/Tgi proteins might be linked to different positions of the TEAD/Sd molecules bound to genomic DNA in different species. The validation/invalidation of these very speculative hypotheses far exceeds the scope of our present study focused on the biochemical characterization of recombinant VGLL4 and Tgi, but we hope that our work will trigger research activities that could help better understand the function of an important tumor suppressor protein such as VGLL4.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Yannick Mesrouze: Investigation; methodology. **Marco Meyerhofer:** Investigation; methodology. **Catherine Zimmermann:** Investigation; methodology. **Patrizia Fontana:** Investigation; methodology. **Dirk Erdmann:** Supervision. **Patrick Chène:** Conceptualization; investigation; supervision.

ORCID

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

REFERENCES

- Davis JR, Tapon N. Hippo signalling during development. *Development*. 2019;146:dev167106.
- Ma S, Meng Z, Chen R, Guan KL. The hippo pathway: Biology and pathophysiology. *Annu Rev Biochem*. 2019;88:577–604.
- Zheng Y, Pan D. The hippo signaling pathway in development and disease. *Dev Cell*. 2019;50:264–282.
- Wu Z, Guan KL. Hippo signaling in embryogenesis and development. *Trends Biochem Sci*. 2021;46:51–63.
- Simon E, Fauchoux C, Zider A, Thézé N, Thiébaud P. From vestigial to vestigial-like: The *Drosophila* gene that has taken wing. *Dev Genes Evol*. 2016;226:297–315.
- Callus BA, Finch-Edmondson ML, Fletcher S, Wilton SD. YAPing about and not forgetting TAZ. *FEBS Lett*. 2019;593:253–276.
- Yamaguchi N. Multiple roles of vestigial-like family members in tumor development. *Front Oncol*. 2020;10:1266.
- Reggiani F, Gobbi G, Ciarrocchi A, Sancisi V. YAP and TAZ are not identical twins. *Trends Biochem Sci*. 2021;46:154–168.
- 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.
- Li Z, Zhao B, Wang P, et al. Structural insights into the YAP and TEAD complex. *Genes Dev*. 2010;24:235–240.
- 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.
- Kaan HYK, Chan SW, Tan SKJ, et al. Crystal structure of TAZ-TEAD complex reveals a distinct interaction mode from that of YAP-TEAD complex. *Sci Rep*. 2017;7:2035.
- 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.
- Mesrouze Y, Hau JC, Erdmann D, et al. The surprising features of the TEAD4-Vgll1 protein-protein interaction. *Chembiochem*. 2014;15:537–542.
- Mesrouze Y, Aguilar G, Bokhovchuk F, et al. A new perspective on the interaction between the vg/VGLL1-3 proteins and the TEAD transcription factors. *Sci Rep*. 2020;10:17442.
- 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.
- Mesrouze Y, Erdmann D, Zimmermann C, Fontana P, Schmelzle T, Chène P. Different recognition of TEAD transcription factor by the conserved β -strand:Loop: α -helix motif of the TEAD-binding site of YAP and VGLL1. *ChemistrySelect*. 2016;1:2993–2997.
- Jiao S, Wang H, Shi Z, et al. A peptide mimicking VGLL4 function acts as a YAP antagonist therapy against gastric cancer. *Cancer Cell*. 2014;25:166–180.
- Mesrouze Y, Meyerhofer M, Bokhovchuk F, et al. Effect of the acylation of TEAD4 on its interaction with co-activators YAP and TAZ. *Protein Sci*. 2017;26:2399–2409.
- Engler C, Kandzia R, Marillonnet S. A one pot, one step, precision cloning method with high throughput capability. *PLoS One*. 2008;3:e3647.
- Tarazona MP, Saiz E. Combination of SEC/MALS experimental procedures and theoretical analysis for studying the solution properties of macromolecules. *J Biochem Biophys Methods*. 2003;56:95–116.
- Wyatt PJ. Light scattering and the absolute characterization of macromolecules. *Anal Chim Acta*. 1993;272:1–40.
- Akaike H. A new look at the statistical model identification. *IEEE Trans Automat Contr*. 1974;19:716–723.
- Guo T, Lu Y, Li P, et al. A novel partner of scalloped regulates hippo signaling via antagonizing scalloped-Yorkie activity. *Cell Res*. 2013;23:1201–1214.
- Koontz LM, Liu-Chittenden Y, Yin F, et al. The hippo effector Yorkie controls normal tissue growth by antagonizing scalloped-mediated default repression. *Dev Cell*. 2013;25:388–401.
- Muller KM, Arndt KM, Pluckthun A. Model and simulation of multivalent binding to fixed ligands. *Anal Biochem*. 1998;261:149–158.
- Vauquelin G, Charlton SJ. Exploring avidity: Understanding the potential gains in functional affinity and target residence time of bivalent and heterobivalent ligands. *Br J Pharmacol*. 2013;168:1771–1785.
- Zhao H, Boyd LF, Schuck P. Measuring protein interactions by optical biosensors. *Curr Protoc Protein Sci*. 2017;88:21–25.
- Suh JL, Watts B, Stuckey JJ, et al. Quantitative characterization of bivalent probes for a dual bromodomain protein, transcription initiation factor TFIID subunit 1. *Biochemistry*. 2018;57:2140–2149.
- Anbanandam A, Albarado DC, Nguyen CT, Halder G, Gao X, Veeraraghavan S. Insights into transcription enhancer factor 1 (TEF-1) activity from the solution structure of the TEA domain. *Proc Natl Acad Sci U S A*. 2006;103:17225–17230.

31. Lee DS, Vonnheim C, Albarado D, Raman CS, Veeraraghavan S. A potential structural switch for regulating DNA-binding by TEAD transcription factors. *J Mol Biol.* 2016;428:2557–2568.

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, Meyerhofer M, Zimmermann C, Fontana P, Erdmann D, Chène P. Biochemical properties of VGLL4 from *Homo sapiens* and Tgi from *Drosophila melanogaster* and possible biological implications. *Protein Science.* 2021;30:1871–1881. <https://doi.org/10.1002/pro.4138>