Divergent evolution of a protein-protein interaction revealed through ancestral sequence reconstruction and resurrection

Louise Laursen[1], Jelena Čalyševa[2,3], Toby J. Gibson[2], and Per Jemth[1]*

[1]Department of Medical Biochemistry and Microbiology, Uppsala University, BMC Box 582, SE-75123 Uppsala, Sweden.

[2]Structural and Computational Biology Unit, European Molecular Biology Laboratory, Heidelberg 69117, Germany.

[3]Collaboration for joint PhD degree between EMBL and Heidelberg University, Faculty of Biosciences

Abstract

The postsynaptic density extends across the postsynaptic dendritic spine with Discs large (DLG) as the most abundant scaffolding protein. DLG dynamically alters the structure of the postsynaptic density, thus controlling the function and distribution of specific receptors at the synapse. PDZ domains make up one of the most abundant protein interaction domain families in animals. One important interaction governing postsynaptic architecture is that between the PDZ3 domain from DLG and cysteine-rich interactor of PDZ3 (CRIPT). However, little is know regarding functional evolution of the PDZ3:CRIPT interaction. Here, we subjected PDZ3 and CRIPT to ancestral sequence reconstruction, resurrection and biophysical experiments. We show that the PDZ3:CRIPT interaction is an ancient interaction, which was present in the last common ancestor of Eukaryotes, and that high affinity is maintained in most extant animal phyla. However, affinity is low in nematodes and insects, raising questions about the physiological function of the interaction in species from these animal groups. Our findings demonstrate how an apparently established protein-protein interaction involved in cellular scaffolding in bilaterians can suddenly be subject to dynamic evolution including possible loss of function.

Introduction

The postsynaptic density extends across the postsynaptic dendritic spine and is composed of receptors, signaling enzymes, cytoskeletal structural elements and cytoplasmic scaffolding proteins. A major role of the postsynaptic density is to stabilize and anchor glutamate receptors such as AMPA and NMDA. Proteins from the Discs large (DLG) family such as Postsynaptic density protein-95 (PSD-95), also called DLG4, is involved in different protein-protein interactions in the postsynaptic density. Here, they control molecular organization and regulate synaptic strength by altering the function and distribution of AMPA receptors at the synapse (Chen et al. 2011). DLG4 contains five folded domains: Postsynaptic density-95/Discs large/Zonula occludens (PDZ)1, PDZ2, PDZ3, Src homology 3 (SH3) and guanylate kinase like (GK). DLG4 is one of four paralogs in vertebrates together with SAP97 (DLG1), PSD-93 (DLG2) and SAP102 (DLG3), respectively. These four proteins arose as a result of two consecutive whole genome duplications in the vertebrate lineage approximately 440 million years ago (McLysaght, Hokamp, and Wolfe 2002; Putnam et al. 2008). The amino acid sequences of the three PDZ domains found in each of these proteins are well conserved. In fact, the identity and similarity are high also for PDZ domains in the corresponding orthologous proteins in evolutionarily distantly related animals such as *Drosophila melanogaster* and the fresh-water polyp *Hydra vulgaris* (Fig. 1a).

The PDZ domain is ancient and based on sequence similarity it can be traced in animals as well as in fungi and plants (Emes and Grant 2011), which contain a permuted variant (Ivarsson et al. 2008). Within the animal kingdom there are examples of genomes containing hundreds of distinct PDZ domains showing that it has been recycled as a protein-protein interaction module many times during evolution following gene or genome duplications. This abundance of PDZ domains in the animal proteome is likely facilitated by its structural architecture allowing easy integration into other proteins (Harris and Lim 2001). PDZ domains consist of 80-100 residues and have a compact globular fold usually containing 5-6 antiparallel β strands and two α helices. However, DLG PDZ3 contains a third α helix at the C-terminus (α_3) (Fig. 1c).

Protein ligands of PDZ domains usually bind via their C-termini to a binding groove in the PDZ domain. This groove is shaped by a conserved motif in the $\beta_1\beta_2$ loop (GLGF or variants of it), the β_2 strand and the α_2 helix. In most cases, the C-terminus of the protein ligand arranges as an anti-parallel β strand in the groove together with the β_2 strand (Fig. 1c). The part of the protein ligand that binds to the PDZ domain is called the PDZ binding motif (PBM). Affinity is primarily dictated by the protein ligand residues denoted P_{-2} and P_0, where P_0 is the C-terminal residue and P_{-2} the third residue from the C-terminus (Fig. 1d) . Early studies led to classification of PBMs as type I (T/S-X-ϕ-COOH), type II (ϕ-X-ϕ-COOH) or type III (D/E-X-ϕ-COOH), based on the nature of the residues at P_{-2} and P_0. However, the classification of PDZ domains only depending on two residues of the binding partner is an over simplification. Not only is the specificity of PDZ-ligand interactions usually dependent on more than P_{-2} and P_0 residues of the ligand (Ernst et al. 2014), but also several properties of the PDZ domains themselves modulate binding, for example, residues outside of the binding groove (Ye et al. 2018) and extensions to the canonical PDZ domain such as α_3 in DLG PDZ3 (Zeng et al. 2016).

One protein ligand for DLG PDZ3 is cysteine-rich interactor of PDZ3 (CRIPT) (Niethammer et al. 1998). This is an interaction that is well studied both from a structural and biophysical point of view, since it was the first structure solved of a PDZ domain with peptide ligand (Doyle 1996). Full affinity of CRIPT (K_d in the low μM range for the *Homo sapiens* CRIPT:DLG4 PDZ3 interaction) is obtained with the six last amino acid residues(Saro et al. 2007; Gianni et al. 2011; Toto et al. 2016). CRIPT has a neurological function and is required for DLG to promote dendrite growth by AMPA activation (Zhang et al. 2008; Beique et al. 2006; Zhang et al. 2017). However, certain animals including the model organisms *D. melanogaster* (where DLG was first identified) and *Caenorhabditis elegans* have CRIPT proteins that lack the classical PDZ-binding motif (Fig. 1b and Supplementary Fig. 1). In *D. melanogaster*, DLG is involved in a range of processes, including synaptic clustering of Shaker potassium channels(Tejedor et al. 1997), junction structure, cell polarity and localization of membrane proteins. Recent work shows that CRIPT clusters next to DLG in the synapse, thereby promoting dendrite growth (Zhang et al. 2017). Furthermore it was reported from knock down experiments that CRIPT is essential for DLG dependent dendrite growth (Zhang et al. 2017). Thus, there is some ambiguity whether CRIPT from, *e.g.*, *D. melanogaster*, binds PDZ3 domains when it lacks the classical PDZ-binding motif and whether the biological action of DLG and CRIPT is partially independent of the PDZ3:CRIPT interaction (Zhang et al. 2017).

Evolutionary biochemistry, *i.e.*, phylogenetic reconstruction of ancestral sequences followed by expression and experimental characterization of the ancient proteins (Fig. 2a-b) is a powerful tool for understanding protein function (Hochberg and Thornton

2017), including protein-protein interactions (Hultqvist et al. 2017; Jemth et al. 2018; Wheeler et al. 2018). To better understand the evolution and hence function of the PDZ3:CRIPT interaction we reconstructed sequences of ancestral time-matched variants of PDZ3 and CRIPT and compared their binding affinity and stability to PDZ3 and CRIPT from seven extant species. We show that the PDZ3:CRIPT interaction has evolved such that the affinity has been maintained or slightly increased in most extant animal phyla (chordates, hemichordates, molluscs and cnidarians) whereas affinity seems to be lost among most nematodes and insects. Our findings raise two questions: Is CRIPT the natural ligand for DLG PDZ3, and is the control of AMPA and NMDA receptors independent of PDZ3:CRIPT interaction, at least in nematodes and insects?

Results

Reconstruction of ancestral sequences.

Protein sequences for DLG family me mbers (including DLG4 and its vertebrate paralogs DLG1, DLG2 and DLG3) and CRIPT proteins were recovered from the NCBI database (NCBI Resource Coordinators). The final set of sequences for ancestral reconstruction consisted of 309 unique sequences of CRIPT from 498 different species and 249 unique sequences of DLG family PDZ3 domains from 324 different species (Supplementary Fig. 1 and 2). The number of different species exceeded the number of unique sequences due to 100% sequence identity between several closely related species. While all of the PDZ3 sequences were from Metazoan (animal) species, CRIPT sequences also included other kingdoms since homologous proteins with high sequence similarity were found in both fungi and plants. Within the animal kingdom, the similarity between homologous sequences of both DLG PDZ3 and CRIPT was very high (Supplementary Fig. 1 and 2). Availability and quality of sequences were variable so the widest most representative dataset was constructed using search methods, sequence alignment, filtering and sorting using Python programming and manual curation. Due to high sequence similarity, posterior probabilities for the maximum likelihood estimate (ML) of reconstructed ancestral sequences were relatively high for most positions in the alignment (Supplementary Data Excel file 1 and 2). We were able to reconstruct PDZ3 from five historical time points, namely from the common ancestor of all extant bilaterians, deuterostomes, protostomes, and hexapods (insects), respectively, as well as from the common ancestor of all extant vertebrates containing four DLG paralogs (Fig. 1a). The PDZ3 from the latter ancestor is denoted 1R PDZ3, since the four DLG paralogs arose as a result of two consecutive whole genome duplications called 1R and 2R, respectively (McLysaght, Hokamp, and Wolfe 2002; Putnam et al. 2008). The C-terminus, *i.e.*, the PBM of CRIPT was reconstructed at eight historical time points, namely from the common ancestors of all extant eukaryotes, opisthokonts (animals and fungi), eumetazoans (animals except sponges), bilaterians, deuterostomes, protostomes, vertebrates (1R) and hexapods, respectively (Fig. 1b and Supplementary Fig. 3). In addition to the ancient ones, we expressed and characterized DLG PDZ3 and CRIPT variants from the following extant species representing different distantly related animal groups: the non-bilaterian animal *Hydra vulgaris* (a cnidarian), *Octopus bimaculoides* (California two-spot octopus, a mollusc and protostome) , *Loa loa* ("eye worm", a nematode or roundworm causing the disease loiasis, protostome), *Dendroctonus ponderosae* (mountain pine beetle, an insect, protostome), *Drosophila melanogaster* (fruit fly, insect, protostome), *Saccoglossus kowalevskii* (acorn worm, a hemichordate, deuterostome) and *Homo sapiens* (representing chordates, deuterostome) (Fig. 2 and Supplementary Fig. 3). We use "native interaction" to denote interactions between time- and species matched PDZ3:CRIPT interactions. Based on sequence alignment and structure prediction, it appears that all ancestral and extant DLG PDZ3 domains have a similar structure with the usual PDZ fold consisting of six β strands and two α helices, but also including a third C-terminal α helix (α_3), present in the crystal structures of human DLG4 PDZ3 (Doyle et al. 1996) (Supplementary Fig. 4).

Evolution of PDZ3:CRIPT interaction

From the sequence alignment and phylogeny of the C-terminal residues in CRIPT (Supplementary Fig. 1) we can follow the evolution of the PBM (Supplementary Fig. 3). Among viridiplantae, plants and mosses show a striking and complete conservation of a type 1 motif (YKQSNV), while green algae have a polar residue at position P_0. The majority of animal phyla including vertebrates, annelids and molluscs (except Euthyneura, i.e., snails and slugs) have CRIPT with a classical type 1 PBM. However, fungi and arthropods (including insects, spiders and crustaceans) appear to lack a PBM, although there are a few exceptions. Interestingly, CRIPT from the cnidarian *H. vulgaris* ends with a Cys (Fig. 1b).

In accordance with the alignment and phylogeny, CRIPT from the last common ancestor of all extant Eukaryotes contained a classical type 1 PBM (YKQSSV) (Supplementary Fig. 3). A type 1 PBM was likely also present in the ancestor of all animals but has since mutated in distinct animal phyla and orders. Clearly, the ancestor of bilaterian animals contained a type 1 PBM and so did the ancestors of the two major bilaterian groups, protostomes and deuterostomes. However, the last common ancestor of today's arthropods likely carried a mutated PBM in CRIPT, ending with Ala or Thr, instead of the canonical Val residue.

One intriguing notion is that the C-terminal residues P_{-1} to P_{-5} are highly conserved even among CRIPTs with Ala or Thr at position P_0. This raises the question whether CRIPTs without a canonical type 1 PBM still retain binding to their native DLG PDZ3? To address this question, PDZ3 and CRIPT from seven extant and five ancestral species were subjected to binding experiments using ITC and stopped-flow spectroscopy (Fig. 3 and 4).

The binding of *H. sapiens* CRIPT to *H. sapiens* DLG4 PDZ3 is well characterized (Niethammer et al. 1998; Saro et al. 2007; Gianni et al. 2011; Toto et al. 2016; Doyle et al. 1996), and we report a similar binding affinity (0.4 μM) as previous studies (Fig. 3 and 5,Table 1 and 2). Furthermore, PDZ3 from *H. vulgaris* (with a C-terminal Cys residue), *O. bimaculoides, S. kowalevskii* and the common ancestors of bilaterial animals, protostomes, deuterostomes, and vertebrates (1R), respectively, were all found to bind to their time-matched native CRIPT with low μM affinity (Fig. 3, 4 and 5, Table 1 and 2). On the other hand, PDZ3 domains from *D. melanogaster*, *D. ponderosae* and the ancestor of hexapods bind poorly to their respective native CRIPT, which lacks a type 1 PBM. While the interactions were too weak for stopped-flow spectroscopy, ITC experiments provided a rough estimate of the affinities ($K_d \geq$ 10-30 μM) (Table 2, Fig. 4 and 5). More surprisingly, PDZ3 from the nematode *L. loa* was also found to bind poorly to its native CRIPT even though it contains a type 1 PBM with a C-terminal Val (Fig. 3).

To directly compare each PDZ3 domain with regard to binding of CRIPT with a type 1 PBM, the affinity of CRIPT from the ancestor of all eukaryotes (YKQSSV) was measured with all PDZ3 domains included in the study. The ancestral Eukaryota CRIPT was found to bind to all resurrected and present-day PDZ3 domains, including insect PDZ3 domains, in the low μM range (Fig 3, 4 and 5, Table 1). However, the binding of Eukaryota CRIPT to the nematode *L. loa* PDZ3 was somewhat weaker than for other PDZ3 domains (17 μM), suggesting that *L. loa* PDZ3 has a binding groove with a slightly different character than PDZ3 domains from the other investigated extant and ancestral species. Thus, while nematodes, the sister phylum of arthropods, has retained the type 1 PBM in CRIPT the high affinity of the PDZ3:CRIPT interaction is lost, suggesting that it may not be functional in either group of animals. Another option is that the PDZ3:CRIPT interaction is functional but interactions outside the canonical groove is essential for full affinity among Ecdysozoa as observed for some other PDZ:ligand complexes (Ye, Huang et al. 2018) (Zeng et al. 2016; Li et al. 2014; Pan et al. 2011) or that the PDZ3:CRIPT interaction is posttranslationally regulated (Sundell et al. 2018). Obviously, lower affinity could have evolved as a response to a certain course of events leading to for example high local cellular concentrations of DLG and CRIPT. Finally, a trivial explanation is that the PDZ3 domains used in the experiments were not properly folded. While it is challenging to test intracellular expression levels we attempted to partially address the other issues.

Mutational analysis to resolve the low CRIPT affinity of PDZ3 from *L. loa*

The *L. loa* PDZ3:CRIPT interaction is particularly interesting since it appears as if the PDZ3 domain has lost affinity for CRIPT with type 1 PBM. Sequence alignment (Fig. 1a) and structural prediction of the PDZ3 variants show that *L. loa* PDZ3 differs in several amino acid positions as compared to other PDZ3 domains, which inspired us to check whether interactions outside the canonical groove could be essential for the PDZ3:CRIPT interaction. *L. loa* PDZ3 has several unique residues at positions that previously have been reported to be important for binding affinity (Supplementary Fig. 5), *e.g.* , Gly→Pro at position 335 in the $\beta_2\beta_3$ loop next to the canonical binding groove and Arg→ Gln at position 399 in α_3. The presence of Pro335 and/or Gln399 is rare; they are only found in 12 sequences among the 253 sequences used for the ancestral sequence reconstruction. Furthermore, a previous study showed the significance of the $\beta_2\beta_3$ loop for increasing affinity by stabilization of a salt-bridge formed by Arg 399 in α_3 and Glu 334 in the $\beta_2\beta_3$ loop in *H. sapiens* DLG4 PDZ3 (Mostarda, Gfeller, and Rao 2012). Thus, two single point mutations, P335G and Q399R, were introduced into *L. loa* PDZ3, to test if the amino acid substitutions can explain the weak binding to CRIPT. Indeed, the affinity increased two and three-fold, respectively, towards ancestral Eukaryota CRIPT (Table 1), supporting the notion that the derived substitutions in *L. loa* PDZ3 indeed weakened the interaction with CRIPT.

Global stability of ancestral and extant PDZ3 domains

Reconstructed ancestral proteins often show a higher stability than the present-day proteins they were predicted from. One biological explanation (for ectotherm organisms) is the higher environmental temperatures in the past (Gaucher, Govindarajan, and Ganesh 2008; Risso, Gavira, and Sanchez-Ruiz 2014), but others argue that the higher stability of ancient proteins is an artifact of the sequence reconstruction because of a consensus sequence bias in the approach(Trudeau, Kaltenbach, and Tawfik 2016; Sternke, Tripp, and Barrick 2019; Williams et al. 2006).

We determined the global stability of all selected extant and resurrected ancient PDZ3 variants using urea denaturation experiments monitored by far UV circular dichroism, which detects secondary structure and is a robust probe of global folding (Fig. 6a-d, Table 3). It is clear that all PDZ3 domains in the study are well folded under our experimental conditions and that they unfold by an apparent two state mechanism in equilibrium urea denaturation experiments. Thus, our binding data are not influenced by low stability of ancient or extant proteins. However, because of uncertainties in curve fitting, as described below, we were careful in the interpretation. All resurrected PDZ3 domains were found to be more stable in comparison to extant PDZ3 domains except *H. sapiens* PDZ3. Generally, we observe a slight increase in the urea midpoint for the deuterostome lineage from bilaterian to *H. sapiens* DLG4 PDZ3. However, there is some ambiguity in the actual stability of the PDZ3 variants due to uncertainty in determination of the $m_{\mathrm{D\text{-}N}}$ value, which reflects the change in solvent accessible hydrophobic surface area upon denaturation and which is directly related to the stability: $\Delta G_{D\text{-}N} = [\mathrm{Urea}]_{50\%} \times m_{\mathrm{D\text{-}N}}$. It is fair to assume that all PDZ3 variants of similar sequence length also have a similar $m_{\mathrm{D\text{-}N}}$ value. Therefore, we fitted the data with a shared $m_{\mathrm{D\text{-}N}}$ value, which gives a more robust estimate of the differences in $\Delta G_{D\text{-}N}$ between PDZ3 variants. But we also fitted the data with a free $m_{\mathrm{D\text{-}N}}$ value to get an idea of the error in this parameter. In addition, some PDZ3 variants did not show complete unfolding at the highest urea concentration, resulting in a lower accuracy of the fitted parameters. The described uncertainties are not unique to our study but pertain to all studies where thermodynamic stabilities are compared, including previous studies addressing ancient versus modern proteins.

The α_3 extension increases the stability of PDZ3

Most studies of DLG4 PDZ3 have been performed on a construct based on the original crystal structure (Doyle 1996). However, in DLG4, PDZ3 is part of the supramodule PDZ3-SH3-GK in which an extended α helix (α_3) connects PDZ3 and SH3 (Fig. 7a), but only 7 of the 19 amino acid residues of the linker are present in the constructs investigated in the present work and in previous publications. Recent studies have reported that the entire α_3 and a peptide ligand longer than 6 residues are required for specific PDZ:ligand interactions (Zeng et al. 2016) (Zeng et al. 2018) (Ye et al. 2018). The α_3 extension does not significantly affect the affinity between *H. sapiens* PDZ3-SH3-GK and a 6-mer or 15-mer CRIPT peptide (Laursen et al. 2019). However, as compared to the ancestral bilaterian PDZ3, the α_3 region is highly conserved in *H. sapiens* PDZ3 with only two substitutions, whereas *D. ponderosae*, *D. melanogaster* and *L. loa* PDZ3 have 4, 8 and 9 substitutions, respectively, in the 19 residue long primary structure of α_3 (Fig 7a). We calculated the α helix propensity of isolated α_3 (residues 395-413) for all PDZ3 variants in the present study using the AGADIR software (Munoz and Serrano 1994). The helical propensity varied from 2-17 % (Fig. 7a), suggesting that the stability of α_3 may have changed during evolution. (The helical propensity refers to an isolated peptide but it could correlate with stability in the context of the folded protein.) Interestingly, *L. loa*, *D. melanogaster* and *D. ponderosae* PDZ3 have the highest helix propensity in α_3. Therefore, we decided to express, purify and analyze PDZ3 from these three species with a full length 19 residue α_3 to assess the effect on binding and stability. We observed a slightly higher thermodynamic stability for the longer variants of PDZ3 in comparison to PDZ3 domains without an extended α_3 (Fig 7b-d and Table 3). Similarly to the PDZ3 variants without α_3 extension, the affinities between the extended α_3 PDZ3 variants and their respective native CRIPT ligands (15-mer) were too low to be measured by stopped-flow. But with ITC it was again possible to obtain a rough estimate of the K_d values. The affinity between *D. melanogaster* PDZ3 and CRIPT was not affected by either the α_3 extension or a longer CRIPT peptide (Table 2). Similarly, while the affinity seems to increase 2-fold for the respective native *D. ponderosae* and *L. loa* PDZ3:CRIPT interactions, this effect is within error of the experiment (Table 2). However, the α_3 extension increased the affinity of *L. loa* PDZ3 for the Eukaryota CRIPT significantly, by 10-fold (Table 1), suggesting that the α_3

extension of PDZ3, perhaps even in the context of a supramodule, needs to be considered if a weak interaction is analyzed.

Negative regulation by phosphorylation

Reversible phosphorylation results in positive or negative regulation of protein-protein interactions and therefore plays an important role for cellular signaling. Protein phosphorylation mainly occurs at Ser, Thr and Tyr residues. Thus, with regard to the PDZ3:CRIPT interaction, phosphorylation at P_{-2} (Thr) in CRIPT disables binding to PDZ3 (Sundell et al. 2018). However *D. melanogaster* CRIPT lacks the type 1 motif but has three possible phosphorylation sites, ending with SST. Therefore we tested the hypothesis that the weak interaction between *D. melanogaster* PDZ3 and its native CRIPT is positively regulated by phosphorylation of CRIPT. Phosphorylation of *D. melanogaster* CRIPT at Ser_{-1} did not affect the binding (Table 2), whereas phosphorylation of Ser_{-2} and Thr_0 seem to weaken the affinity further since no binding could be detected by either stopped-flow or ITC. Thus, if anything, the PDZ3:CRIPT interaction is negatively regulated by phosphorylation in *D. melanogaster*.

The conclusions based on resurrected maximum likelihood variants of PDZ3 are robust to sequence uncertainty.

Finally, our historical description of the PDZ3:CRIPT evolution is contingent on accurate native affinities and stabilities for the resurrected proteins. Thus, even if the reconstructed sequences are not correct, what about the experimantally determined properties of the proteins? Different approaches can be applied to check the robustness of results based on the ML variant, for example generating a library of variants of the ML ancestral sequence by point mutations (Yokoyama and Radlwimmer 2001) or using Bayesian sampling to reconstruct a set of sequences by sampling amino acids over the posterior probability distribution (Yang and Rannala 1997). However an easier approach is to reconstruct a "worst-case scenario" variant. Here, one sequence is reconstructed in which every amino acid residue with a posterior probability below a certain threshold, for example 0.8, is replaced by the residue with the second highest probability at that position. Such a variant, denoted AltAll (alternative residue at all positions), represents a very unlikely worst case scenario that can be used to assess how robust conclusions are to errors in the reconstructed ML sequences (Eick et al. 2017; Anderson et al. 2016) simply by comparing the ML and AltAll proteins. If they display the same affinity for a ligand, that affinity is very likely to be the true ancestral affinity (Fig. 1a and Supplementary Fig. 3 and 6).

Thus, we used the AltAll approach as a relatively strong test of robustness with a single protein variant as control. Worst-case scenario AltAll sequences from CRIPT and PDZ3 were reconstructed, expressed and subjected to the same experiments as the ML variants to validate the conclusions based on the resurrected ML proteins. The C-termini of ancestral CRIPT could be reconstructed with high confidence due to the strong conservation observed in extant species. Therefore, AltAll CRIPT variants were only tested in two cases, for the ancestor of hexapods and ancestor of deuterostomes, respectively, both of which had one ambiguous position (Supplementary Fig. 3 and 6). AltAll variants of PDZ3 contained 3-5 substitutions for 1R, Deuterostomia, Protostomia and Bilateria PDZ3, at positions outside the canonical binding groove (Fig. 1a). The primary structure of AltAll for the ancestral

Hexapoda PDZ3 contained 16 substitutions in comparison to the ML sequence. This is because PDZ3 at the Hexapoda node is resurrected from relatively few sequences and from species with a large number of residue substitutions. However, all substitutions were situated outside of the canonical binding groove.

Overall, the AltAll and ML PDZ3 domains showed a similar binding affinity for Eukaryota and native CRIPT (Supplementary Fig. 6 and Supplementary Table 1). Furthermore, secondary structure monitored by circular dichroism were similar for ML and AltAll variants, showing that the folding is robust (Supplementary Fig. 7a-e). We determined the global stabilities by urea denaturation for all AltAll variants (Supplementary Fig. 7f-j, and Supplementary Table 2) and found that they were similar to the corresponding ML variants for 1R, bilaterians and protostomes. The AltAll PDZ3 variant for the ancestor of deuterostomes displayed slightly lower stability compared to the ML variant (Supplementary Fig. 7j and Supplementary Table 2). The curve fitting of AltAll and ML PDZ3 from the ancestor of hexapods suffered from sloping baselines and a short denatured baseline resulting in poorly defined $m_{\text{D-N}}$ values (Supplementary Fig. 7h). However, qualitatively, it is clear that the AltAll variant is less stable than ML, but well folded. Thus, all AltAll variants were properly folded as judged by the CD spectra (Supplementary Fig. 7a-e) and shape of the urea denaturation curves (as reflected in the $m_{\text{D-N}}$ values) (Supplementary Fig. 7f-j and Supplementary Table 2).

In conclusion, the AltAll and ML proteins displayed similar properties with regard to stability and binding, and therefore other proteins in the ensemble of likely variants will most probably do that too, including the true ancestral PDZ3 variants (Eick et al. 2017).

Neuroligin with type 1 PBM is common among deuterostome but not protostome animals

The low affinity of the PDZ3:CRIPT interaction in nematodes and insects inspired us to look at another proposed biological ligand of DLG4 PDZ3, namely neuroligin. Similarly to CRIPT, neuroligin interacts with PDZ3 via its C-terminus at synaptic junctions (Jeong et al. 2019; Irie et al. 1997). To compare the PBMs of neuroligin to those of CRIPT, the NCBI protein sequence database was searched for homologs of *H. sapiens* (UniProt Q8N2Q7) and *C. elegans* (UniProt Q9XTG1) neuroligin-1, as described for CRIPT in the Methods section. *H. sapiens* and *C. elegans* neuroligin sequences were used for the search, because they were both of high quality and represented neuroligin sequences from deuterostomes and protostomes, respectively. The search resulted in 1725 sequences in 437 species – 419 metazoan species and 18 fungal species. Among the neuroligin sequences, type 1 PBM, like the one of CRIPT, was found to be predominant in deuterostomes, while type 2 PBM (ϕ-X-ϕ-COOH) was predominant in protostomes, especially hexapods. Thus, we observe an even more dynamic evolution of the PBM in neuroligin as compared to CRIPT. It appears likely that bilaterian neuroligins recognize different PDZ domains in different species, while most of the neuroligin sequences from fungi and cnidarians – the only species outside of bilaterians that have neuroligin homologs – lack a PBM. However, the lower number of high quality neuroligin sequences makes these results less conclusive than for CRIPT.

Discussion

We have used ancestral sequence reconstruction and resurrection to better understand the connection between structure, function and evolution of the PDZ3:CRIPT interaction and its biological function. Already at the time of the last common ancestor of all extant metazoa, CRIPT contained a type I PBM with high affinity to most present day DLG PDZ3 domains, and likely to many other PDZ domains with type I PBM specificity. This affinity has remained constant for eons among diverging lineages and until today. But, at the beginning of the evolutionary lineage leading to present day arthropods (approximately 550-600 Myr) a point mutation in the C-terminal residue of CRIPT resulted in substantially decreased affinity for DLG PDZ3, and again, likely all PDZ domains with type I specificity. What is the role of CRIPT and how would this mutation influence function? Large regions of CRIPT are highly conserved among extant species (Supplementary Figure 9) and mutation in CRIPT is rare among humans, and associated with rare Mendelian disorders (Leduc et al. 2016). A recent study including *in vivo* experiments indicated an essential role for CRIPT in dendritic growth (Zhang et al. 2017). Experiments were performed in the nematode *C. elegans*, which, like *D. melanogaster*, has a Thr at the C-terminus of CRIPT. A severely mutated CRIPT (deletion of around 130 residues) resulted in abnormalities in dendritic growth, which could be rescued by expression of *H. sapiens* CRIPT. It was not possible to knock down CRIPT in mouse suggesting a non-redundant and critical cell biological function in vertebrate species (Zhang et al. 2017). On the other hand, it has been possible to knock down one or several of the DLG family members because these proteins can compensate for each other (Bonnet et al. 2013; Howard et al. 2010). In light of their experiments, and the lack of a type I

PBM in the C-terminus of CRIPT from *C. elegans* and *D. melanogaster*, Zhang et al suggested that the critical biological action of CRIPT could be partially independent of the PDZ3:CRIPT interaction in these animals. In the present paper we corroborate this idea by showing that PDZ3 domains from extant species of nematode and insect phyla bind their native CRIPT with significantly lower affinities than ancestral PDZ3:CRIPT pairs and those from extant vertebrates, even when CRIPT contains a large hydrophobic residue at the C-terminus, such as Leu in *L. loa* CRIPT. (There are also no obvious non-canonical internal PBMs in CRIPT.) The fact that PDZ3 domains from nematodes and insects have retained the affinity for type I PBM reinforces this notion.

By necessity, many protein interaction studies are performed with domains cut out from highly modular, often large proteins, which are hard to work with in pure form. However, effects from regions outside the binding surface or domain are increasingly considered in protein-protein interaction studies. Indeed, extensions of PDZ domains are associated with several potential functions: 1) protein dynamics-based modulation of target binding affinity, 2) provision of binding sites for macro-molecular assembly, 3) structural integration of multi-domain modules, and 4) expansion of the ligand-binding pocket (Wang et al. 2010). We analyzed the α3 extension in PDZ3 that possibly extends the ligand binding surface beyond the canonical binding groove. Interestingly, while this α3 extension had only minor effects on the low affinity native PDZ3:CRIPT interactions (*L. loa*, *D. melanogaster* and *D. ponderosae*), it indeed had a significant effect on binding to the high affinity ancestral eukaryote CRIPT peptide to *L. loa* PDZ3, demonstrating high sequence dependence in the interaction, upon addition of α3. The sequence dependence was also reflected in the stabilities of PDZ3 with and without α_3, where only *L. loa* PDZ3 displayed a clear stabilization in presence of the helix. With this in mind we note that PDZ3 stabilities vary across present day animals and that ancestral versions of PDZ3 were likely a little more stable. However, all PDZ3 variants were well folded and thus fit for its purpose, binding to another protein. Slow destabilization of domains over evolutionary time is consistent with random mutations, neutral for function and not detrimental as long as folding is correct and stability sufficient.

Materials and Methods

To collect protein sequences, the NCBI database (NCBI Resource Coordinators) was searched for CRIPT and DLG sequences using search terms and the BLAST tool (Altschul et al. 1990) with the *H. sapiens* DLG4 PDZ3 sequence. Sequences were then aligned using Clustal Omega software (Sievers et al. 2011) and filtered. Non-redundant sequences with information on species of origin, homologous to *H. sapiens* DLG4 PDZ3 or *H. sapiens* CRIPT, respectively, were kept. The final sets of sequences were aligned using Clustal Omega and manually curated. The resulting alignments were used to infer the ancestral sequences of DLG family PDZ3 domains and of CRIPT. For this sequence reconstruction, species trees were downloaded using CommonTree tool from Taxonomy Browser in the NCBI database. For the sequences, which are identical for several species, the lowest common taxonomic rank was used in the tree. The taxonomy trees were modified before the ancestral reconstruction: For every clade branching out into more than 2 daughter clades on one branch, clades were split into smaller clades that only had two daughter clades at the time due to software requirements. The PDZ3 tree was adjusted according to DLG family homology; due to evidence of genome duplication in early chordates, vertebrate sequences were split into DLG1, DLG2, DLG3 and DLG4 clades (Putnam et al. 2008) (McLysaght, Hokamp, and Wolfe 2002). In the resulting CRIPT and PDZ3 trees, species names were replaced with sequence names to match the identifiers in the alignment. The RAxML software (Stamatakis 2014) was used to infer ancestral sequences for PDZ3 and CRIPT. Because of low posterior probabilities for some positions in the reconstructed ML sequences, we reconstructed AltAll sequences for each ML sequence. In the AltAll sequences the amino acid residues with the second highest probabilities were included for each position where the probability of the ML residue was below 0.8.

Protein expression and Purification. cDNA encoding DLG PDZ3 from the following extant species were cloned into a modified pRSET vector (Invitrogen) and transformed into *Escherichia coli* BL21(DE3) pLys cells (Invitrogen) for expression: *H. sapiens* (DLG4), *S. kowalevskii*, *D. melanogaster*, *D. ponderosae*, *L. loa*, *O. bimaculoides* and *H. vulgaris*. Similarly, reconstructed cDNA encoding the following ancestral DLG PDZ3 (both ML and AltAll variants) were expressed: 1R, Deuterostomia, Bilateria, Hexapoda and Protostomia. The expression construct contained an N-terminal 6xHis-tag. Cells were grown in LB medium at 37°C and overexpression of protein was induced with 1 mM Isopropyl-β-D-thiogalactopyranoside overnight at 18°C. Cells were harvested by centrifugation at 4°C and the pellet re-suspended in 50 mM Tris pH 7.8, 10 % glycerol and stored at -20°C. Pellets were thawed and sonicated 2 x 4 min followed by centrifugation. The supernatant was filtered and added to a pre-equilibrated (50 mM Tris, pH 7.8, 10% glycerol) Nickel Sepharose Fast Flow column (GE Healthcare). Proteins were eluted with 250 mM imidazole and dialysed into 50 mM Tris pH 7.8, 10% glycerol buffer overnight. Proteins were then loaded onto a Q sepharose column for further purification and eluted with a 500 mM NaCl gradient in 50 mM Tris, pH 7.8, 10% glycerol. Protein purity and identity were analyzed by SDS-PAGE and MALDI-TOF mass spectrometry.

Circular dichroism and urea denaturation experiments. Experiments were performed in 50 mM sodium phosphate, pH 7.45, 21 mM (I = 150) at 10°C on a JASCO J-1500 spectrapolarimeter. Far-UV spectra for all PDZ3 domains were recorded from 260 nm to 200 nm with 10 μM protein (the average of 16 scans). Thermodynamic stability of PDZ3 domains was determined by urea denaturation (0 - 8.1 M with 0.3 M steps) and monitoried by molar ellipticity at 222 nm with 22 μM PDZ3. Data were fitted to a two-state unfolding process using both free fitting of each parameter or a shared $m_{\text{D-N}}$ to obtain the midpoint (denoted $[\text{Urea}]_{50\%}$,where 50% of the protein is folded).

$$F = \frac{(\alpha_N + \beta_N \cdot [Urea]) + (\alpha_D + \beta_D \cdot [Urea]) \cdot e^{\frac{m_{D-N} \cdot ([Urea] - [Urea]_{50\%})}{R \cdot T}}}{1 + e^{\frac{m_{D-N} \cdot ([Urea] - [Urea]_{50\%})}{R \cdot T}}}$$

F is the observed spectroscopic signal; α_N is the signal of the native state in 0 M urea; β_N is the slope of the native baseline; α_D and β_D are the corresponding parameters for the denaturated state (Clarke et al. (1993)).

Kinetic experiments

All kinetic experiments were performed in 50 mM sodium phosphate, pH 7.45, 21 mM KCl (I = 150) at 10°C. Binding and dissociation experiments were performed at SX-17 MV stopped-flow spectrometer (Applied Photophysics) and carried out as described previously (Chi et al. (2010)). Briefly, the association rate constant (k_{on}) was obtained from binding experiments. PDZ3 (1 µM) was mixed rapidly with dansyl labelled (D) CRIPT peptide (concentration range: 2 to 20 µM). Excitation was done at 345 nm and emission was recorded above 420 nm using a long-pass filter). The change in fluorescence signal obtained from D-CRIPT was monitored over time and fitted to a single exponential function to deduce the observed rate constant (k_{obs}). Observed rate constants were plotted against the corresponding D-CRIPT concentration. Data were fitted to an equation for a reversible bimolecular association reaction (Malatesta 2005) to extract the association rate constant without considering pseudo first order conditions.

$$k_{obs} = (k_{on}^2([PDZ]_0 - [CRIPT]_0)^2 + k_{off}^2 + 2k_{on}k_{off}([PDZ]_0 + [CRIPT]_0 + [PDZ]_0))^{0.5}$$

The dissociation rate constant (k_{off}) was obtained from a separate displacement experiment. An equilibrium complex was pre-formed between PDZ3 and D-CRIPT (2:10 µM) and mixed rapidly with high excess of unlabeled CRIPT (100, 150 and 200 µM, respectively). The change in fluorescence signal obtained from dansyl-labelled CRIPT was monitored over time and fitted to a single exponential function to deduce the observed rate constant (k_{obs}). The average of the three k_{obs} values is reported as the dissociation rate constant k_{off}.

ITC Binding experiments

All isothermal titration calorimetry (ITC) binding experiments were performed in 50 mM sodium phosphate pH 7.45, 21 mM KCl (I = 150) at 25°C on an iTC200 microcalorimeter (Malvern Instruments). PDZ3 and CRIPT peptide were dialysed against the same buffer to minimize artifacts from buffer mismatch in the titrations. CRIPT peptide was titrated 16 times into the cell containing PDZ3. For every titration the heat release decreased from the binding of CRIPT to PDZ3, due to complex formation and PDZ3 saturation. Change in heat over time was integrated to kcal/mol and fitted to a one site binding model to obtain binding stoichiometry (n), equilibrium binding constant (K_d) and the enthalphy change upon binding (ΔH).

Acknowledgements

We thank Eva Andersson for technical assistance. This project was funded by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 675341 (to T.G. and P.J.) and by the Swedish Research Council grant no. 2016-04965 (to P.J.).

Author contributions

L.L. and P.J. conceived and designed the project. L.L performed all experiments and analyzed the data. J.Č. performed all the bioinformatic work including the ancestral sequence reconstruction. L.L., J.Č., and P.J. wrote the initial draft. All authors contributed to preparation of the paper and interpretation of data.

Competing interests

The authors declare no competing interests.

Data availability

The authors declare that data supporting the findings of this study are available within the paper and its supplementary information file, including probability tables for reconstruction of CRIPT and PDZ3 sequences. Additional raw data are available from corresponding authors upon request.

References

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. 1990. 'Basic local alignment search tool', *J Mol Biol*, 215: 403-10.

Anderson, D. P., D. S. Whitney, V. Hanson-Smith, A. Woznica, W. Campodonico-Burnett, B. F. Volkman, N. King, J. W. Thornton, and K. E. Prehoda. 2016. 'Evolution of an ancient protein function involved in organized multicellularity in animals', *Elife*, 5: e10147.

Beique, J. C., D. T. Lin, M. G. Kang, H. Aizawa, K. Takamiya, and R. L. Huganir. 2006. 'Synapse-specific regulation of AMPA receptor function by PSD-95', *Proc Natl Acad Sci U S A*, 103: 19535-40.

Bonnet, S. A., D. S. Akad, T. Samaddar, Y. Liu, X. Huang, Y. Dong, and O. M. Schluter. 2013. 'Synaptic state-dependent functional interplay between postsynaptic density-95 and synapse-associated protein 102', *J Neurosci*, 33: 13398-409.

Chen, X., C. D. Nelson, X. Li, C. A. Winters, R. Azzam, A. A. Sousa, R. D. Leapman, H. Gainer, M. Sheng, and T. S. Reese. 2011. 'PSD-95 is required to sustain the molecular organization of the postsynaptic density', *J Neurosci*, 31: 6329-38.

Chi, C. N., A. Bach, M. Gottschalk, A. S. Kristensen, K. Strømgaard, and P. Jemth. 2010. 'Deciphering the kinetic binding mechanism of dimeric ligands using a potent plasma-stable dimeric inhibitor of postsynaptic density protein-95 as an example', *J Biol Chem*, 285: 28252-60.

Clarke, J., A. M. Hounslow, M. Bycroft, and A. R. Fersht. 1993. 'Local breathing and global unfolding in hydrogen exchange of barnase and its relationship to protein folding pathways', *Proc Natl Acad Sci U S A*, 90: 9837-41.

Doyle, D. A., A. Lee, J. Lewis, E. Kim, M. Sheng, and R. MacKinnon. 1996. 'Crystal structures of a complexed and peptide-free membrane protein-binding domain: molecular basis of peptide recognition by PDZ', *Cell*, 85: 1067-76.

Eick, G. N., J. T. Bridgham, D. P. Anderson, M. J. Harms, and J. W. Thornton. 2017. 'Robustness of Reconstructed Ancestral Protein Functions to Statistical Uncertainty', *Mol Biol Evol*, 34: 247-61.

Emes, R. D., and S. G. Grant. 2011. 'The human postsynaptic density shares conserved elements with proteomes of unicellular eukaryotes and prokaryotes', *Front Neurosci*, 5: 44.

Ernst, A., B. A. Appleton, Y. Ivarsson, Y. Zhang, D. Gfeller, C. Wiesmann, and S. S. Sidhu. 2014. 'A structural portrait of the PDZ domain family', *J Mol Biol*, 426: 3509-19.

Gaucher, E. A., S. Govindarajan, and O. K. Ganesh. 2008. 'Palaeotemperature trend for Precambrian life inferred from resurrected proteins', *Nature*, 451: 704-7.

Gianni, S., S. R. Haq, L. C. Montemiglio, M. C. Jurgens, A. Engstrom, C. N. Chi, M. Brunori, and P. Jemth. 2011. 'Sequence-specific long range networks in PSD-95/discs large/ZO-1 (PDZ) domains tune their binding selectivity', *J Biol Chem*, 286: 27167-75.

Harris, B. Z., and W. A. Lim. 2001. 'Mechanism and role of PDZ domains in signaling complex assembly', *J Cell Sci*, 114: 3219-31.

Hochberg, G. K. A., and J. W. Thornton. 2017. 'Reconstructing Ancient Proteins to Understand the Causes of Structure and Function', *Annu Rev Biophys*, 46: 247-69.

Howard, M. A., G. M. Elias, L. A. Elias, W. Swat, and R. A. Nicoll. 2010. 'The role of SAP97 in synaptic glutamate receptor dynamics', *Proc Natl Acad Sci U S A*, 107: 3805-10.

Hultqvist, G., E. Aberg, C. Camilloni, G. N. Sundell, E. Andersson, J. Dogan, C. N. Chi, M. Vendruscolo, and P. Jemth. 2017. 'Emergence and evolution of an interaction between intrinsically disordered proteins', *Elife*, 6.

Irie, M., Y. Hata, M. Takeuchi, K. Ichtchenko, A. Toyoda, K. Hirao, Y. Takai, T. W. Rosahl, and T. C. Sudhof. 1997. 'Binding of neuroligins to PSD-95', *Science*, 277: 1511-5.

Ivarsson, Y., C. Travaglini-Allocatelli, M. Brunori, and S. Gianni. 2008. 'Folding and misfolding in a naturally occurring circularly permuted PDZ domain', *J Biol Chem*, 283: 8954-60.

Jemth, P., E. Karlsson, B. Vogeli, B. Guzovsky, E. Andersson, G. Hultqvist, J. Dogan, P. Guntert, R. Riek, and C. N. Chi. 2018. 'Structure and dynamics conspire in the evolution of affinity between intrinsically disordered proteins', *Sci Adv*, 4: eaau4130.

Jeong, J., S. Pandey, Y. Li, J. D. Badger, 2nd, W. Lu, and K. W. Roche. 2019. 'PSD-95 binding dynamically regulates NLGN1 trafficking and function', *Proc Natl Acad Sci U S A*, 116: 12035-44.

Laursen, L., E. Karlsson, S. Gianni, and P. Jemth. 2019. 'Functional interplay between protein domains in a supramodular structure involving the postsynaptic density protein PSD-95', *J Biol Chem*.

Leduc, M. S., Z. Niu, W. Bi, W. Zhu, I. Miloslavskaya, T. Chiang, H. Streff, J. R. Seavitt, S. A. Murray, C. Eng, A. Chan, Y. Yang, and S. R. Lalani. 2016. 'CRIPT exonic deletion and a novel missense mutation in a female with short stature, dysmorphic features, microcephaly, and pigmentary abnormalities', *Am J Med Genet A*, 170: 2206-11.

Li, Y., Z. Wei, Y. Yan, Q. Wan, Q. Du, and M. Zhang. 2014. 'Structure of Crumbs tail in complex with the PALS1 PDZ-SH3-GK tandem reveals a highly specific assembly mechanism for the apical Crumbs complex', *Proc Natl Acad Sci U S A*, 111: 17444-9.

Malatesta, F. 2005. 'The study of bimolecular reactions under non-pseudo-first order conditions', *Biophys Chem*, 116: 251-6.

McLysaght, A., K. Hokamp, and K. H. Wolfe. 2002. 'Extensive genomic duplication during early chordate evolution', *Nat Genet*, 31: 200-4.

Mostarda, S., D. Gfeller, and F. Rao. 2012. 'Beyond the binding site: the role of the beta(2)-beta(3) loop and extra-domain structures in PDZ domains', *PLoS Comput Biol*, 8: e1002429.

Munoz, V., and L. Serrano. 1994. 'Elucidating the folding problem of helical peptides using empirical parameters', *Nat Struct Biol*, 1: 399-409.

Niethammer, M., J. G. Valtschanoff, T. M. Kapoor, D. W. Allison, R. J. Weinberg, A. M. Craig, and M. Sheng. 1998. 'CRIPT, a novel postsynaptic protein that binds to the third PDZ domain of PSD-95/SAP90', *Neuron*, 20: 693-707.

Pan, L., J. Chen, J. Yu, H. Yu, and M. Zhang. 2011. 'The structure of the PDZ3-SH3-GuK tandem of ZO-1 protein suggests a supramodular organization of the membrane-associated guanylate kinase (MAGUK) family scaffold protein core', *J Biol Chem*, 286: 40069-74.

Putnam, N. H., T. Butts, D. E. Ferrier, R. F. Furlong, U. Hellsten, T. Kawashima, M. Robinson-Rechavi, E. Shoguchi, A. Terry, J. K. Yu, E. L. Benito-Gutierrez, I. Dubchak, J. Garcia-Fernandez, J. J. Gibson-Brown, I. V. Grigoriev, A. C. Horton, P. J. de Jong, J. Jurka, V. V. Kapitonov, Y. Kohara, Y. Kuroki, E. Lindquist, S. Lucas, K. Osoegawa, L. A. Pennacchio, A. A. Salamov, Y. Satou, T. Sauka-Spengler, J. Schmutz, I. T. Shin, A. Toyoda, M. Bronner-Fraser, A. Fujiyama, L. Z. Holland, P. W. Holland, N. Satoh, and D. S. Rokhsar. 2008. 'The amphioxus genome and the evolution of the chordate karyotype', *Nature*, 453: 1064-71.

Risso, V. A., J. A. Gavira, and J. M. Sanchez-Ruiz. 2014. 'Thermostable and promiscuous Precambrian proteins', *Environ Microbiol*, 16: 1485-9.

Saro, D., T. Li, C. Rupasinghe, A. Paredes, N. Caspers, and M. R. Spaller. 2007. 'A thermodynamic ligand binding study of the third PDZ domain (PDZ3) from the mammalian neuronal protein PSD-95', *Biochemistry*, 46: 6340-52.

Sievers, F., A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Soding, J. D. Thompson, and D. G. Higgins. 2011. 'Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega', *Mol Syst Biol*, 7: 539.

Stamatakis, A. 2014. 'RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies', *Bioinformatics*, 30: 1312-3.

Sternke, M., K. W. Tripp, and D. Barrick. 2019. 'Consensus sequence design as a general strategy to create hyperstable, biologically active proteins', *Proc Natl Acad Sci U S A*, 116: 11275-84.

Sundell, G. N., R. Arnold, M. Ali, P. Naksukpaiboon, J. Orts, P. Guntert, C. N. Chi, and Y. Ivarsson. 2018. 'Proteome-wide analysis of phospho-regulated PDZ domain interactions', *Mol Syst Biol*, 14: e8129.

Tejedor, F. J., A. Bokhari, O. Rogero, M. Gorczyca, J. Zhang, E. Kim, M. Sheng, and V. Budnik. 1997. 'Essential role for dlg in synaptic clustering of Shaker K+ channels in vivo', *J Neurosci*, 17: 152-9.

Toto, A., S. W. Pedersen, O. A. Karlsson, G. E. Moran, E. Andersson, C. N. Chi, K. Stromgaard, S. Gianni, and P. Jemth. 2016. 'Ligand binding to the PDZ domains of postsynaptic density protein 95', *Protein Eng Des Sel*, 29: 169-75.

Trudeau, D. L., M. Kaltenbach, and D. S. Tawfik. 2016. 'On the Potential Origins of the High Stability of Reconstructed Ancestral Proteins', *Mol Biol Evol*, 33: 2633-41.

Wang, C. K., L. Pan, J. Chen, and M. Zhang. 2010. 'Extensions of PDZ domains as important structural and functional elements', *Protein Cell*, 1: 737-51.

Wheeler, L. C., J. A. Anderson, A. J. Morrison, C. E. Wong, and M. J. Harms. 2018. 'Conservation of Specificity in Two Low-Specificity Proteins', *Biochemistry*, 57: 684-95.

Williams, P. D., D. D. Pollock, B. P. Blackburne, and R. A. Goldstein. 2006. 'Assessing the accuracy of ancestral protein reconstruction methods', *PLoS Comput Biol*, 2: e69.

Yang, Z., and B. Rannala. 1997. 'Bayesian phylogenetic inference using DNA sequences: a Markov Chain Monte Carlo Method', *Mol Biol Evol*, 14: 717-24.

Ye, F., Y. Huang, J. Li, Y. Ma, C. Xie, Z. Liu, X. Deng, J. Wan, T. Xue, W. Liu, and M. Zhang. 2018. 'An unexpected INAD PDZ tandem-mediated plcbeta binding in Drosophila photo receptors', *Elife*, 7.

Yokoyama, S., and F. B. Radlwimmer. 2001. 'The molecular genetics and evolution of red and green color vision in vertebrates', *Genetics*, 158: 1697-710.

Zeng, M., Y. Shang, Y. Araki, T. Guo, R. L. Huganir, and M. Zhang. 2016. 'Phase Transition in Postsynaptic Densities Underlies Formation of Synaptic Complexes and Synaptic Plasticity', *Cell*, 166: 1163-75 e12.

Zeng, M., F. Ye, J. Xu, and M. Zhang. 2018. 'PDZ Ligand Binding-Induced Conformational Coupling of the PDZ-SH3-GK Tandems in PSD-95 Family MAGUKs', *J Mol Biol*, 430: 69-86.

Zhang, L., A. M. Jablonski, J. Mojsilovic-Petrovic, H. Ding, S. Seeholzer, I. P. Newton, I. Nathke, R. Neve, J. Zhai, Y. Shang, M. Zhang, and R. G. Kalb. 2017. 'SAP97 Binding Partner CRIPT Promotes Dendrite Growth In Vitro and In Vivo', *eNeuro*, 4.

Zhang, L., J. Schessl, M. Werner, C. Bonnemann, G. Xiong, J. Mojsilovic-Petrovic, W. Zhou, A. Cohen, P. Seeburg, H. Misawa, A. Jayaram, K. Personius, M. Hollmann, R. Sprengel, and R. Kalb. 2008. 'Role of GluR1 in activity-dependent motor system development', *J Neurosci*, 28: 9953-68.

Figure Legends:

Fig. 1

Extant and reconstructed ancient DLG PDZ3 and CRIPT sequences. a Alignment of DLG PDZ3 sequences. The sequence numbering of extant PDZ3 domains refers to the full-length DLG protein from the respective species. For simplicity, we use numbering of PDZ3 according to the human DLG4 sequence throughout the paper. In the alignment of ancestral sequences (green background), residues with posterior probabilities lower than 0.5 (red) and higher than 0.5 but lower than 0.8 (yellow) are highlighted. The secondary structure is from the crystal structure of DLG4 PDZ3 depicted in panel c with the corresponding color code. **b** Alignment of C-terminal CRIPT sequences. **c** Crystal structure of human DLG4 PDZ3 in complex with a CRIPT peptide (YKQTSV) (PDB ID: 5HEB). Secondary structure elements are marked in accordance with the alignment in panel a. The structure was visualised using UCSF Chimera software. **d** Illustration of three common C-terminal PDZ binding motif types. X, any amino acid residue; Φ, hydrophobic residue. Previous studies have shown that human DLG4 PDZ3 prefers a type I motif with a Val at position P_0.

Fig. 2

Overview of setup to investigate the evolution of the PDZ3:CRIPT interaction using evolutionary biochemistry. a Ancestral sequence reconstructions of CRIPT and PDZ3, respectively, were performed based on multiple sequence alignments and a species tree. In the reconstruction, 309 unique sequences of CRIPT from 498 different species and 249 unique sequences of PDZ3 from the DLG protein family from 324 different species were included. **b** Expression and purification of selected extant and reconstructed ancestral protein variants. **c** The PDZ3:CRIPT interaction was analyzed by isothermal titration calorimetry and stopped-flow spectroscopy. **d** Global stability of PDZ3 domains were determined by circular dichroism-monitored urea denaturation experiments.

Fig. 3

Characterization of the PDZ3:CRIPT interaction from extant species with a type 1 PBM in CRIPT. a-d Examples of ITC experiments for native PDZ3:CRIPT interactions for four extant species: *H. sapiens*, *S. kowalewski*, *L. loa* and *O. bimaculoides*. **e-l** Binding kinetics from stopped-flow experiments for the same native interactions as in panel **a-d**, and also including experiments with a reconstructed CRIPT from the ancestor of all eukaryotes ($\text{CRIPT}_{\text{Eukaryota}}$). **e-h** Observed rate constants were obtained from kinetic binding traces (examples shown in panels **i-l**) and plotted as function of CRIPT concentration at a constant concentration of PDZ3 (1 μM) to estimate the association rate constant k_{on}. **i-l** Examples of kinetic binding traces from stopped-flow experiments using 4 μM CRIPT and 1 μM PDZ3. The fitted line (red) is a single exponential from which the observed rate constant k_{obs} was obtained. k_{off} was determined in a separate experiment as explained in the Methods section, and K_{d} calculated as $k_{\text{off}}/k_{\text{on}}$. Experiments were performed at 25°C for ITC and at 10°C for stopped-flow.

Fig. 4

Characterization of the PDZ3:CRIPT interaction from extant species without a type 1 PBM in CRIPT. a-c Examples of ITC experiments for native PDZ3:CRIPT interactions for three extant species (*H. vulgaris*, *D. melanogaster* and *D. ponderosae*, respectively). **d-i** Binding kinetics from stopped-flow experiments for the native *H. vulgaris* interaction and also experiments with the respective PDZ3 and a reconstructed CRIPT from the ancestor of all eukaryotes ($\mathrm{CRIPT}_{\mathrm{Eukaryota}}$). **d-f** Observed rate constants were obtained from kinetic binding traces (examples shown in panels **g-i**) and plotted as function of CRIPT concentration at a constant concentration of PDZ3 (1 μM) to estimate the association rate constant k_{on}. **g-i** Examples of kinetic binding traces from stopped-flow experiments using 4 μM CRIPT and 1 μM PDZ3. The fitted line (red) is a single exponential from which the observed rate constant k_{obs} was obtained. k_{off} was determined in a separate experiment as explained in the Methods section, and K_{d} calculated as $k_{\mathrm{off}}/k_{\mathrm{on}}$. Experiments were performed at 25°C for ITC and at 10°C for stopped-flow.

Fig. 5

Evolution of affinity for the PDZ3:CRIPT interaction

Simplified species tree with nodes for selected ancestral and extant species used in the study. The native time-matched affinity for each interaction is shown in colour code text at each node and the affinity between each PDZ3 and Eukaryota CRIPT is shown in black text. For the native interaction in the deuterostome node, two CRIPT peptides were tested due to low posterior probability for Ser at position P_{-2}. K_d values were determined by stopped-flow experiments at 10°C.

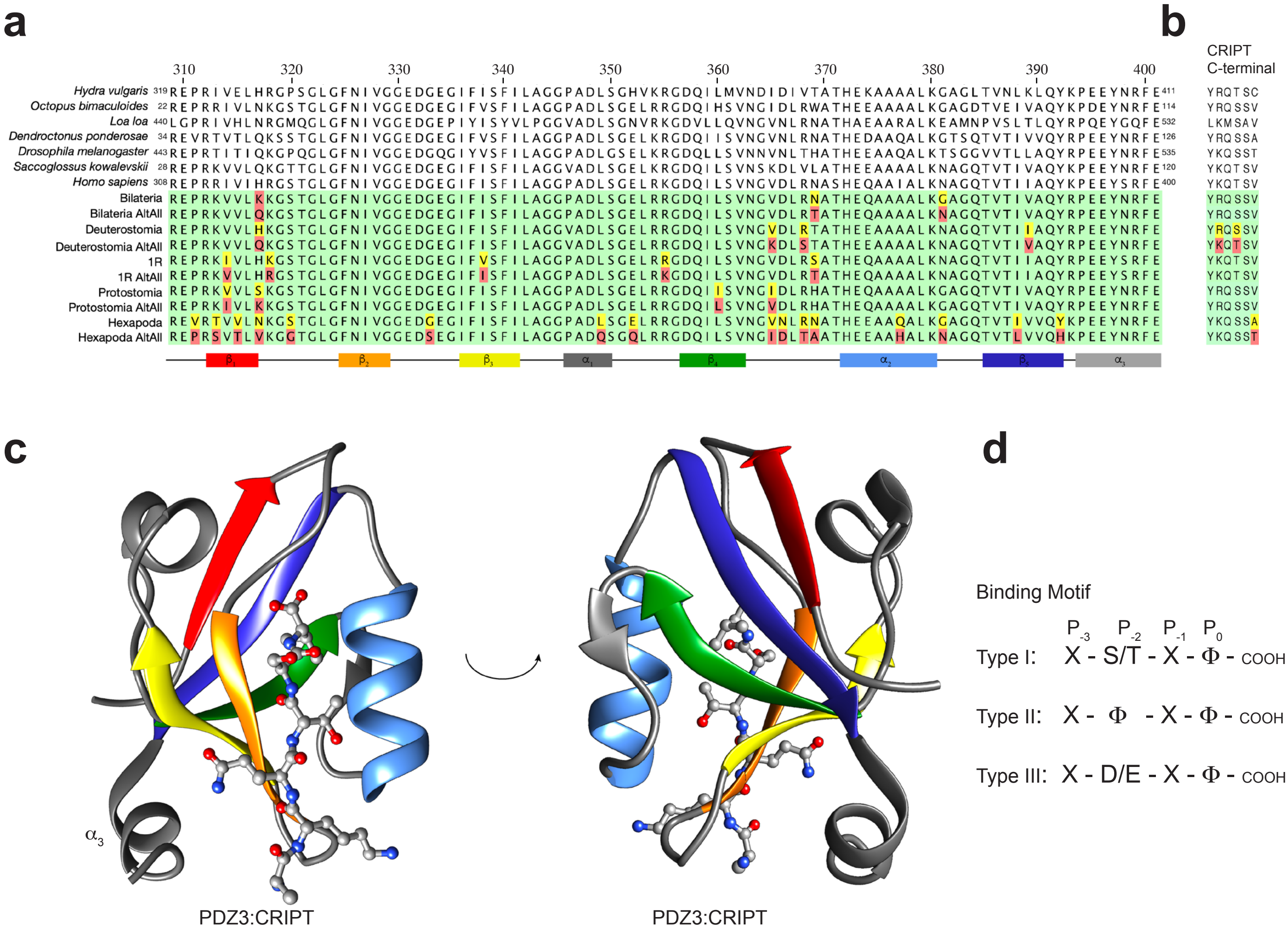
Fig. 6

Secondary structure and global stability of extant and resurrected ancestral PDZ3 domains. **a-d** Urea denaturation (0-8.1 M) of extant and ancestral PDZ3 variants, as monitored by circular dichroism at 222 nm. Data were fitted to a two-state model for protein (un)folding. See Table 3 for fitted parameters. To facilitate comparison, the urea denaturation curve of *H. sapiens* PDZ3 (grey) is present in all graphs. **e-h**) Secondary structure content analysed by circular dichroism between 200-260 nm. Each spectrum is an average of 5 individual scans measured at 10°C in 50 mM sodium phosphate, pH 7.45, 21 mM KCl (I = 150 mM).

Fig. 7

Properties of PDZ3 domains with an extended α_3. a Schematic illustration of DLG4 with the PDZ1, PDZ2, PDZ3, SH3 and GK domains together with a sequence alignment of the extended α_3 helix (19 aa) for the PDZ3 variants from extant and ancestral species included in the study. Alpha helix propensity was calculated by the AGADIR software at 283K, I = 150 mM, pH 7.45. **b-d** Urea denaturation experiments of PDZ3 from *D. melanogaster*, *L. loa* and *D. ponderosae*, respectively, with and without the α_3 extension, as monitored by circular dichroism at 222 nm. See Table 3 for fitted parameters.

FIGURE 1



				CRIPT C-terminal
Hydra vulgaris	319	REPRIVELHRGPSGLGFNIVGGEDGEGIFISFILAGGPADLSGHVKRGDQILMVNDIDIVTATHEKAAAALKGAGLTVNLKLQYKPEEYNRFE	411	YRQTSC
Octopus bimaculoides	22	REPRRIVLNKGSTGLGFNIVGGEDGEGIFVSFILAGGPADLSGELRRGDQIHSVNGIDLRWATHEEAAAALKGAGDTVEIVAQYKPDEYNRFE	114	YRQSSV
Loa loa	440	LGPRIVHLNRGMQGLGFNIVGGEDGEPIYISYVLPGGVADLSGNVRKGDVLLQVNGVNLRNATHAEAARALKEAMNPVSLTLQYRPQEYGQFE	532	LKMSAV
Dendroctonus ponderosae	34	REVRTVTLQKSSTGLGFNIVGGEDGEGIFVSFILAGGPADLSGELKRGDQIISVNGVNLRNATHEDAAQALKGTSQTVTIVVQYRPEEYNRFE	126	YRQSSA
Drosophila melanogaster	443	REPRTITIQKGPQGLGFNIVGGEDGQGIYVSFILAGGPADLGSELKRGDQLLSVNNVNLTHATHEEAAQALKTSGGVVTLLAQYRPEEYNRFE	535	YKQSST
Saccoglossus kowalevskii	28	REPRKVVLQKGTTGLGFNIVGGEDGEGIFISFILAGGVADLSGELKRGDQILSVNSKDLVLATHEEAAAALKNAGQTVTIVAQYKPEEYNRFE	120	YKQTSV
Homo sapiens	308	REPRRIVIHRGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRKGDQILSVNGVDLRNASHEQAAIALKNAGQTVTIIAQYKPEEYSRFE	400	YKQTSV
Bilateria		REPRKVVLKKGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRRGDQILSVNGVDLRNATHEQAAAALKGAGQTVTIVAQYRPEEYNRFE		YRQSSV
Bilateria AltAll		REPRKVVLQKGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRRGDQILSVNGVDLRTATHEQAAAALKNAGQTVTIVAQYRPEEYNRFE		YRQSSV
Deuterostomia		REPRKVVLHKGSTGLGFNIVGGEDGEGIFISFILAGGVADLSGELRRGDQILSVNGVDLRTATHEEAAAALKNAGQTVTIIAQYRPEEYNRFE		YRQSSV
Deuterostomia AltAll		REPRKVVLQKGSTGLGFNIVGGEDGEGIFISFILAGGVADLSGELRRGDQILSVNGKDLSTATHEEAAAALKNAGQTVTIVAQYRPEEYNRFE		YKQTSV
1R		REPRKIVLHKGSTGLGFNIVGGEDGEGIFVSFILAGGPADLSGELRRGDQILSVNGVDLRSATHEQAAAALKNAGQTVTIIAQYRPEEYSRFE		YKQTSV
1R AltAll		REPRKVVLHRGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRKGDQILSVNGVDLRTATHEQAAAALKNAGQTVTIIAQYRPEEYSRFE		YKQTSV
Protostomia		REPRKVVLSKGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRRGDQIISVNGIDLRHATHEQAAAALKGAGQTVTIVAQYRPEEYNRFE		YRQSSV
Protostomia AltAll		REPRKIVLKKGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRRGDQILSVNGVDLRHATHEQAAAALKGAGQTVTIVAQYRPEEYNRFE		YRQSSV
Hexapoda		REVRTVVLNKGSTGLGFNIVGGEDGEGIFISFILAGGPADLSGELRRGDQILSVNGVNLRNATHEEAAQALKGAGQTVTIVVQYKPEEYNRFE		YKQSSA
Hexapoda AltAll		REPRSVTLVKGGTGLGFNIVGGEDSEGIFISFILAGGPADQSGQLRRGDQILSVNGIDLTAATHEEAAHALKNAGQTVTLVVQHKPEEYNRFE		YKQSST

FIGURE 2

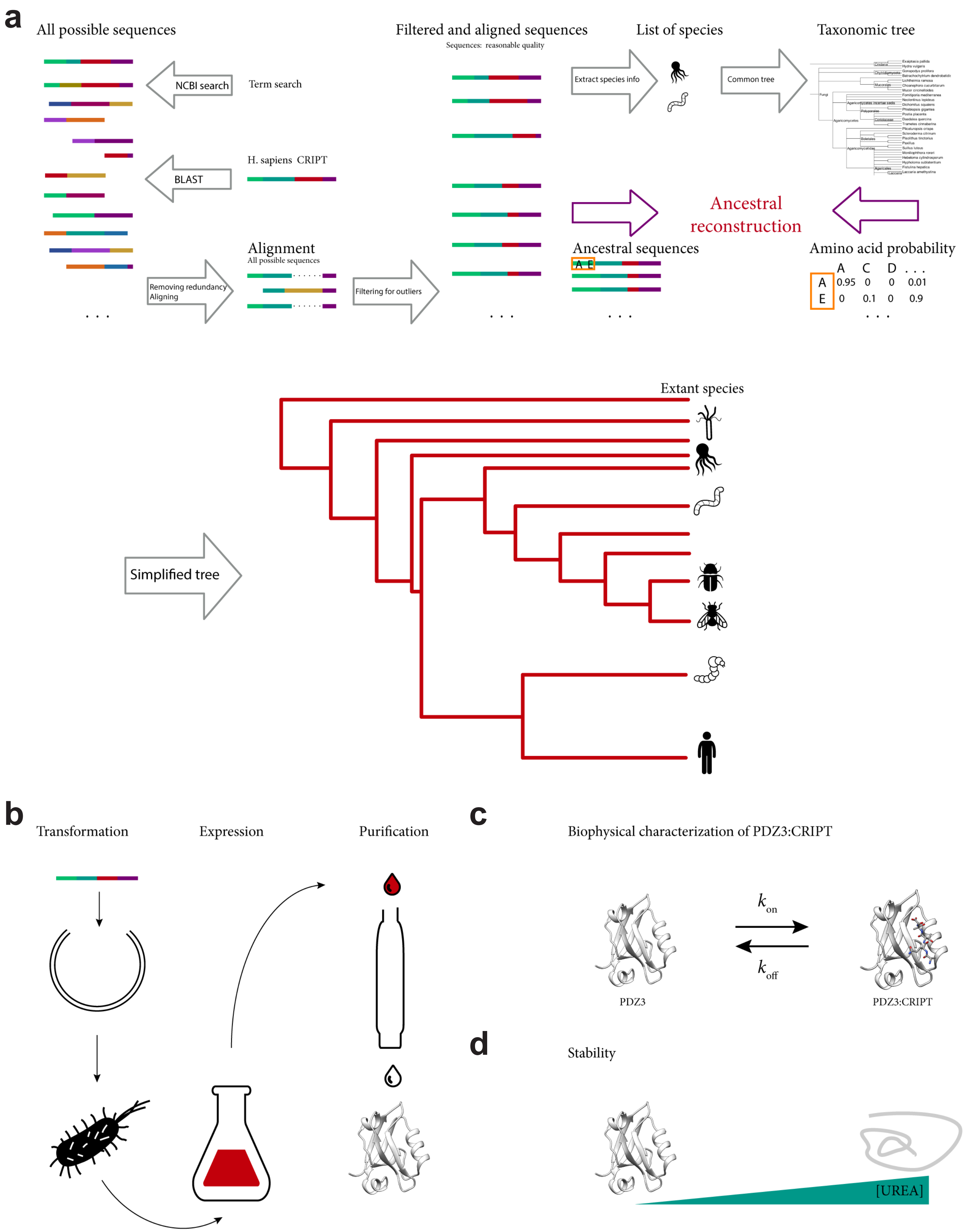

a
All possible sequences
NCBI search
Term search
BLAST
H. sapiens CRIPT
Removing redundancy
Aligning
Alignment
All possible sequences
Filtering for outliers
Filtered and aligned sequences
Sequences: reasonable quality
Extract species info
List of species
Common tree
Taxonomic tree
Cnidaria
Exaiptasia pallida
Hydra vulgaris
Chytridiomycota
Gonapodya prolifera
Batrachochytrium dendrobatidis
Mucorales
Lichtheimia ramosa
Choanephora cucurbitarum
Mucor circinelloides
Fungi
Fomitiporia mediterranea
Neolentinus lepideus
Agaricomycetes incertae sedis
Dichomitus squalens
Polyporales
Phlebiopsis gigantea
Postia placenta
Agaricomycetes
Coriolaceae
Daedalea quercina
Trametes cinnabarina
Plicaturopsis crispa
Scleroderma citrinum
Boletales
Pisolithus tinctorius
Paxillus
Suillus luteus
Agaricomycetidae
Moniliophthora roreri
Hebeloma cylindrosporum
Hypholoma sublateritium
Agaricales
Fistulina hepatica
Laccaria
Laccaria amethystina
Ancestral reconstruction
Ancestral sequences
A E
Amino acid probability
A C D . . .
A 0.95 0 0 0.01
E 0 0.1 0 0.9
. . .
Extant species
Simplified tree
b
Transformation
Expression
Purification
c
Biophysical characterization of PDZ3:CRIPT
k_{on}
k_{off}
PDZ3
PDZ3:CRIPT
d
Stability
[UREA]

FIGURE 3

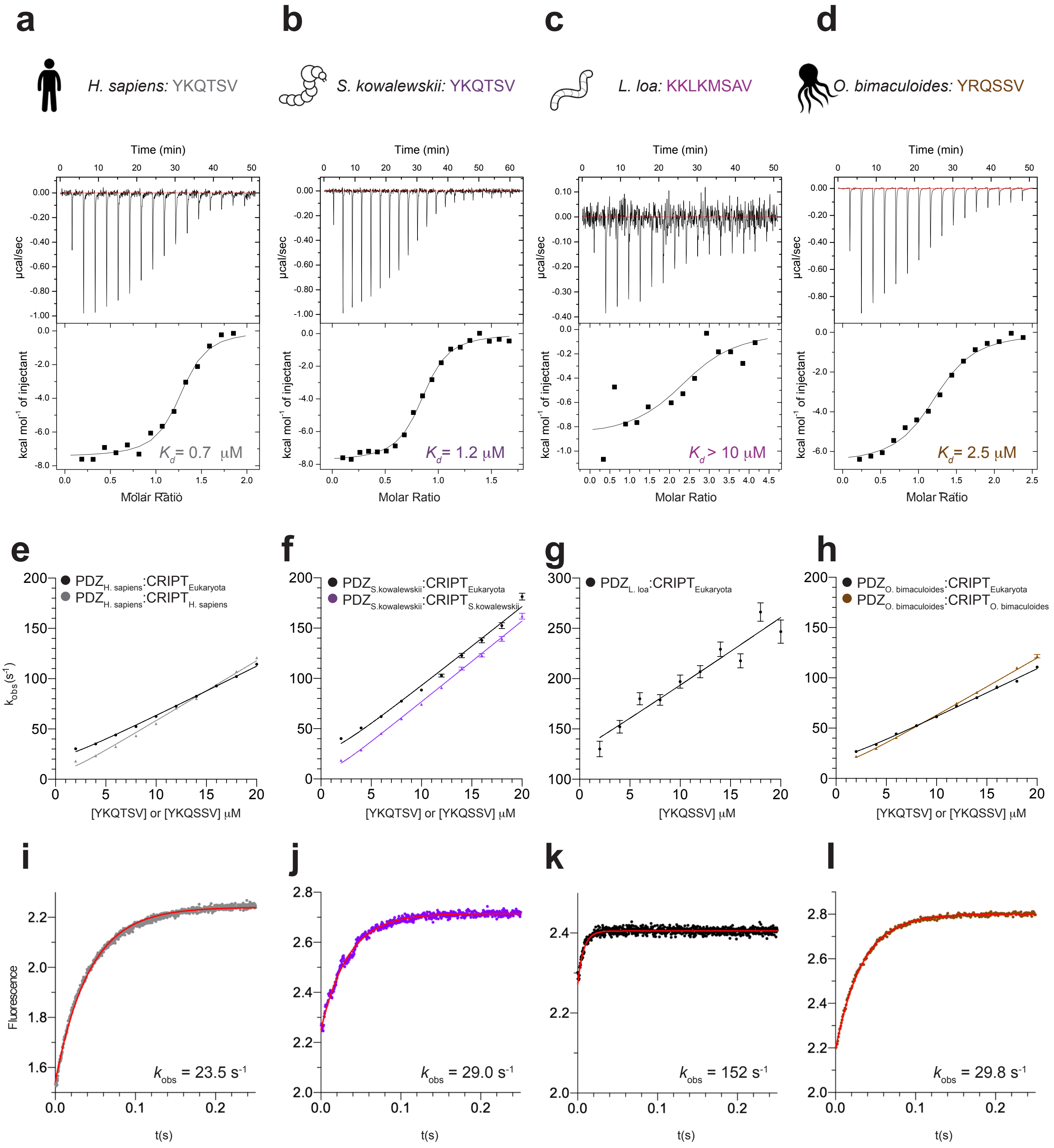

a
b
c
d
H. sapiens: YKQTSV
S. kowalewskii: YKQTSV
L. loa: KKLKMSAV
O. bimaculoides: YRQSSV
Time (min)
µcal/sec
kcal mol⁻¹ of injectant
Molar Ratio
Kd= 0.7 µM
Kd= 1.2 µM
Kd > 10 µM
Kd= 2.5 µM
e
f
g
h
PDZH. sapiens:CRIPTEukaryota
PDZH. sapiens:CRIPTH. sapiens
PDZS.kowalewskii:CRIPTEukaryota
PDZS.kowalewskii:CRIPTS.kowalewskii
PDZL. loa:CRIPTEukaryota
PDZO. bimaculoides:CRIPTEukaryota
PDZO. bimaculoides:CRIPTO. bimaculoides
kobs (s⁻¹)
[YKQTSV] or [YKQSSV] µM
[YKQSSV] µM
i
j
k
l
Fluorescence
t(s)
kobs = 23.5 s⁻¹
kobs = 29.0 s⁻¹
kobs = 152 s⁻¹
kobs = 29.8 s⁻¹

FIGURE 4

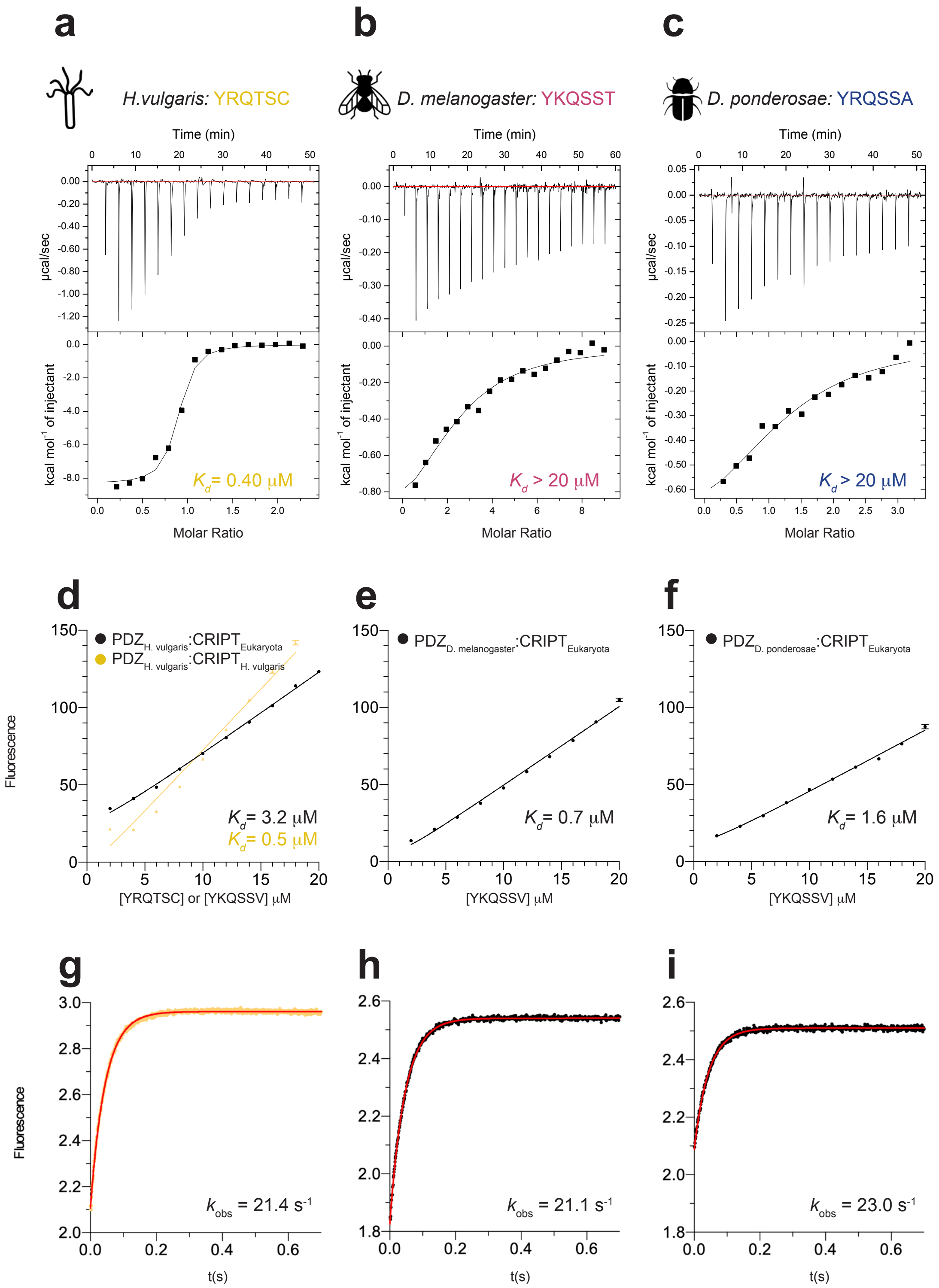

a
H.vulgaris: YRQTSC
Time (min)
µcal/sec
kcal mol⁻¹ of injectant
Kd= 0.40 µM
Molar Ratio
b
D. melanogaster: YKQSST
Time (min)
µcal/sec
kcal mol⁻¹ of injectant
Kd > 20 µM
Molar Ratio
c
D. ponderosae: YRQSSA
Time (min)
µcal/sec
kcal mol⁻¹ of injectant
Kd > 20 µM
Molar Ratio
d
PDZ H. vulgaris:CRIPT Eukaryota
PDZ H. vulgaris:CRIPT H. vulgaris
Fluorescence
Kd= 3.2 µM
Kd= 0.5 µM
[YRQTSC] or [YKQSSV] µM
e
PDZ D. melanogaster:CRIPT Eukaryota
Kd= 0.7 µM
[YKQSSV] µM
f
PDZ D. ponderosae:CRIPT Eukaryota
Kd= 1.6 µM
[YKQSSV] µM
g
Fluorescence
kobs = 21.4 s⁻¹
t(s)
h
kobs = 21.1 s⁻¹
t(s)
i
kobs = 23.0 s⁻¹
t(s)

FIGURE 5

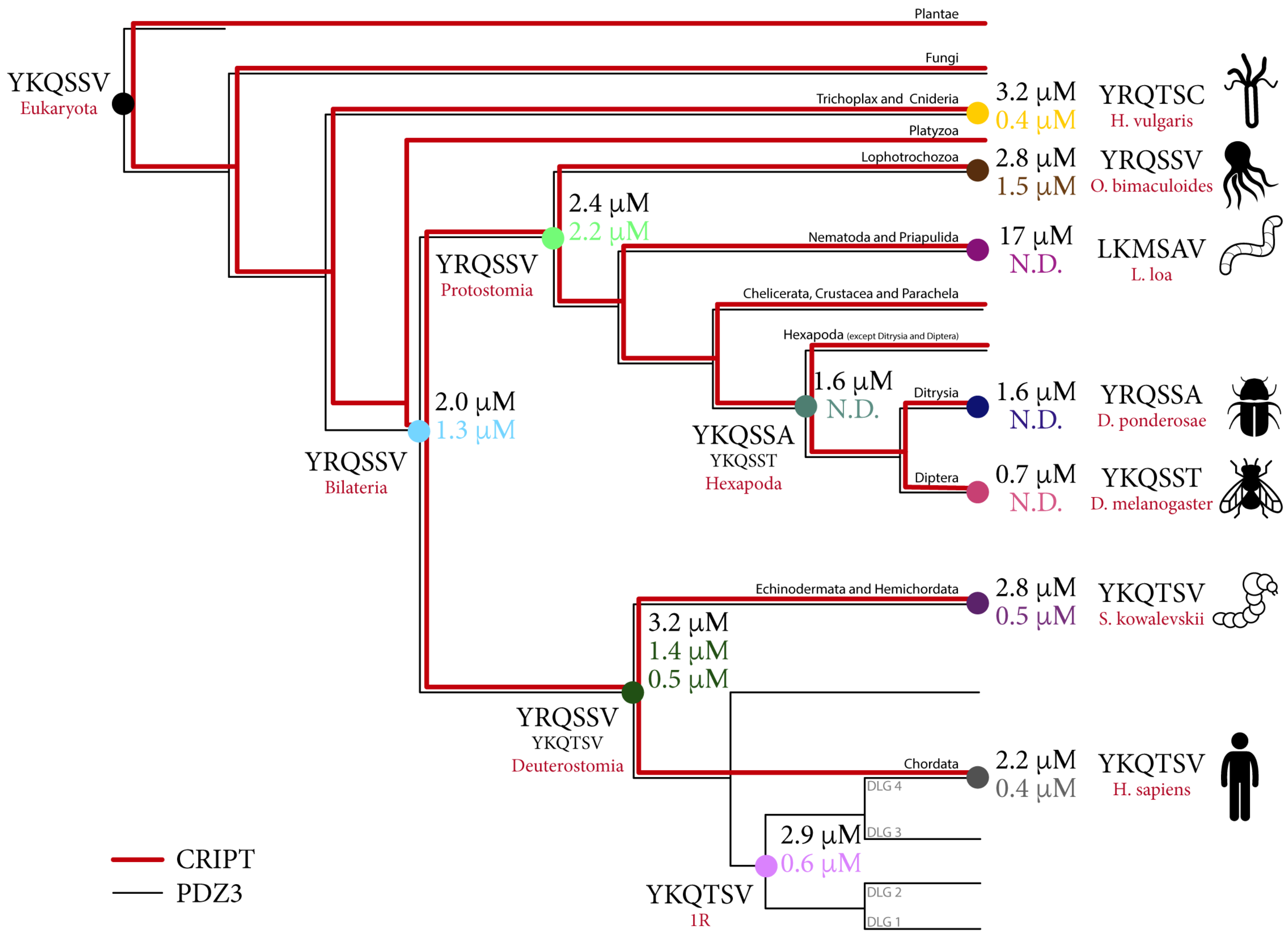

Plantae
Fungi
Trichoplax and Cnideria
Platyzoa
Lophotrochozoa
Nematoda and Priapulida
Chelicerata, Crustacea and Parachela
Hexapoda (except Ditrysia and Diptera)
Ditrysia
Diptera
Echinodermata and Hemichordata
Chordata
DLG 4
DLG 3
DLG 2
DLG 1
YKQSSV
Eukaryota
YRQSSV
Bilateria
2.0 μM
1.3 μM
YRQSSV
Protostomia
2.4 μM
2.2 μM
YKQSSA
YKQSST
Hexapoda
1.6 μM
N.D.
YRQSSV
YKQTSV
Deuterostomia
3.2 μM
1.4 μM
0.5 μM
YKQTSV
1R
2.9 μM
0.6 μM
3.2 μM
0.4 μM
YRQTSC
H. vulgaris
2.8 μM
1.5 μM
YRQSSV
O. bimaculoides
17 μM
N.D.
LKMSAV
L. loa
1.6 μM
N.D.
YRQSSA
D. ponderosae
0.7 μM
N.D.
YKQSST
D. melanogaster
2.8 μM
0.5 μM
YKQTSV
S. kowalevskii
2.2 μM
0.4 μM
YKQTSV
H. sapiens
CRIPT
PDZ3

FIGURE 6

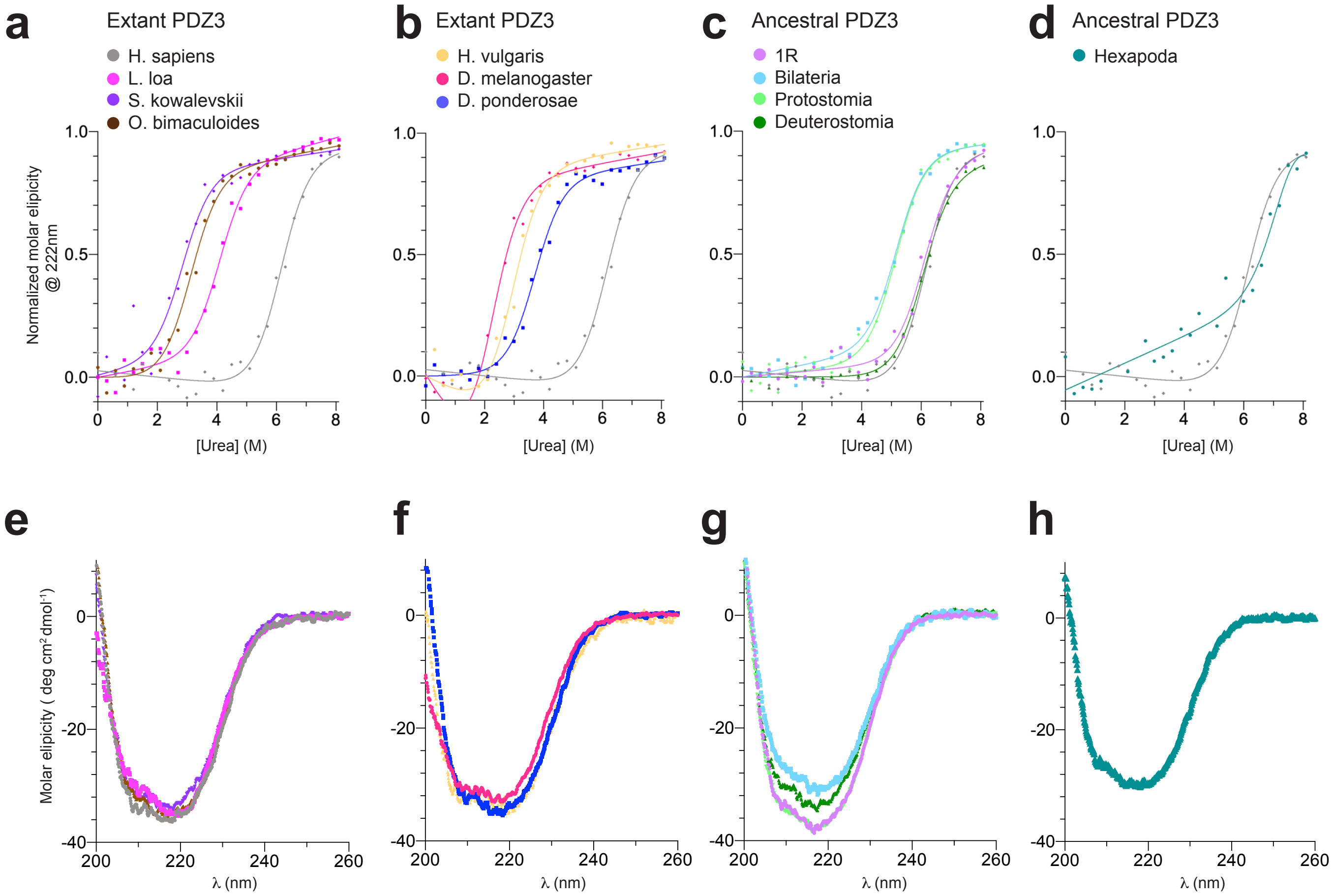

a
Extant PDZ3
● H. sapiens
● L. loa
● S. kowalevskii
● O. bimaculoides
b
Extant PDZ3
● H. vulgaris
● D. melanogaster
● D. ponderosae
c
Ancestral PDZ3
● 1R
● Bilateria
● Protostomia
● Deuterostomia
d
Ancestral PDZ3
● Hexapoda
Normalized molar elipicity @ 222nm
1.0
0.5
0.0
0
2
4
6
8
[Urea] (M)
e
f
g
h
Molar elipicity (deg cm² dmol⁻¹)
0
-20
-40
200
220
240
260
λ (nm)

FIGURE 7

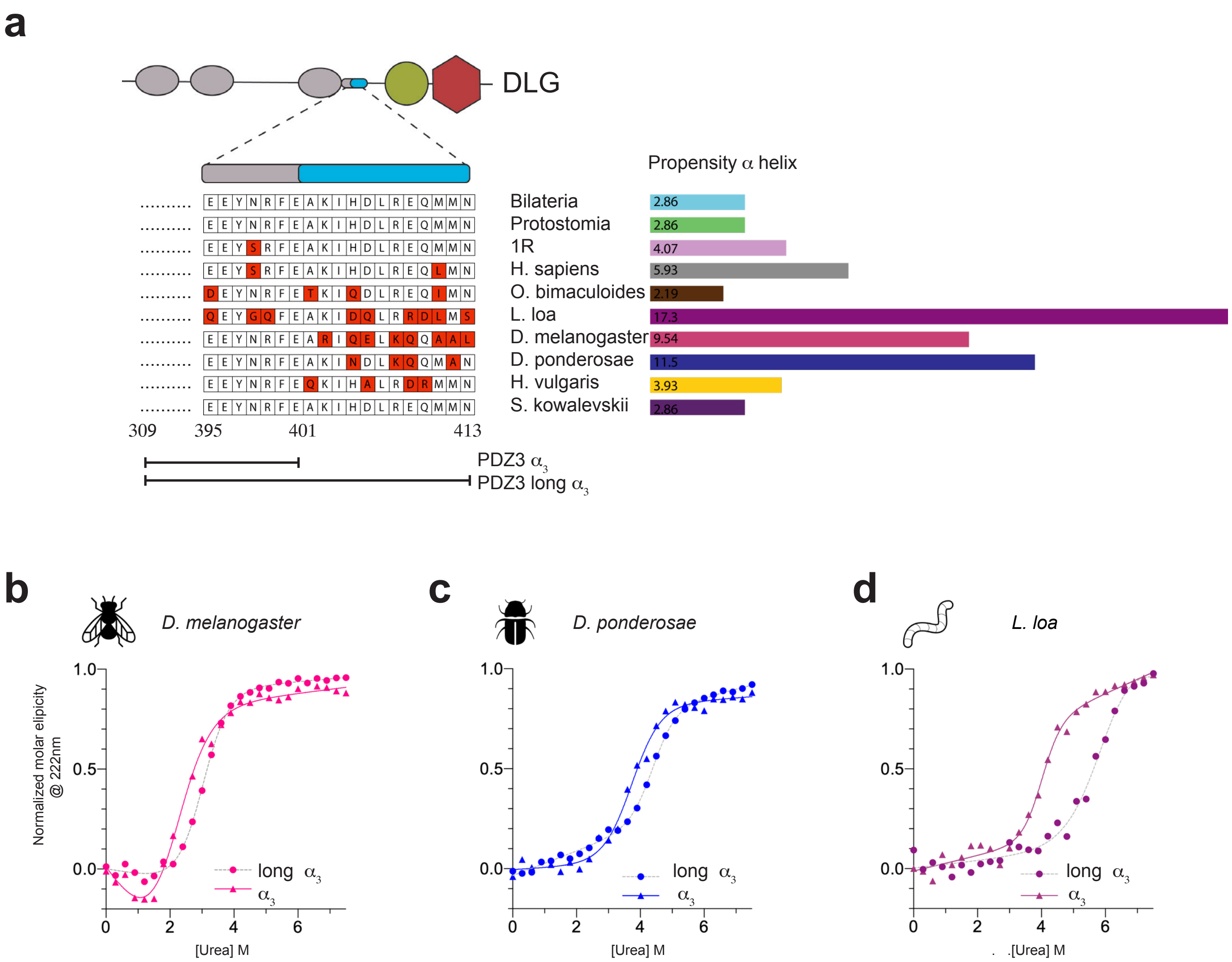

a
DLG
Propensity α helix
.... EEYNRFEAKIHDLREQMMN Bilateria 2.86
.... EEYNRFEAKIHDLREQMMN Protostomia 2.86
.... EEYSRFEAKIHDLREQMMN 1R 4.07
.... EEYSRFEAKIHDLREQLMN H. sapiens 5.93
.... DEYNRFETKIQDLREQIMN O. bimaculoides 2.19
.... QEYGQFEAKIDQLRRDLMS L. loa 17.3
.... EEYNRFEARIQELKQQAAL D. melanogaster 9.54
.... EEYNRFEAKINDLKQQMAN D. ponderosae 11.5
.... EEYNRFEQKIHALRDRMMN H. vulgaris 3.93
.... EEYNRFEAKIHDLREQMMN S. kowalevskii 2.86
309
395
401
413
PDZ3 α3
PDZ3 long α3
b
D. melanogaster
c
D. ponderosae
d
L. loa
Normalized molar elipicity @ 222nm
1.0
0.5
0.0
0
2
4
6
[Urea] M
long α3
α3

Table 1

Binding constants for extant and reconstructed ancestral PDZ3:CRIPT interactions from kinetic experiments. Rate and equilibrium constants were determined for the PDZ3:CRIPT interaction for extant and ancestral PDZ3 domains binding its respective native CRIPT (YKQTSV, YRQSSV or YRQTSC). In addition, binding experiments were performed for all PDZ3 domains to the ancestral Eukaryota CRIPT (YKQSSV) and for all CRIPTs to *H. sapiens* PDZ3. Native interactions are highlighted in bold. Stopped-flow experiments were used to determine the association rate constant (k_{on}) from binding experiments and the dissociation rate constant (k_{off}) from displacement experiments. K_d values were calculated from the ratio of k_{off} and k_{on}. All data were obtained at 10°C in 50 mM sodium phosphate, pH 7.45, 21 mM KCl (I = 150).

CRIPT	PDZ3	k_{on} (µM^{-1} s^{-1})	k_{off} (s^{-1})	K_d (µM)
YKQSSV	H. sapiens	5.3 ± 0.04	11.6 ± 0.01	2.2 ± 0.02
	1R	5.3 ± 0.1	15.3 ± 0.3	2.9 ± 0.12
	Deuterostomia	7.4 ± 0.2	23.6 ± 0.03	3.2 ± 0.07
	Hexapoda	7.3 ± 0.1	11.5 ± 0.1	1.6 ± 0.04
	Protostomia	6.5 ± 0.1	15.5 ± 1.3	2.4 ± 0.23
	Bilateria	5.9 ± 0.1	11.9 ± 0.3	2.0 ± 0.07
	S. kowalevskii	7.6 ± 0.2	20.8 ± 0.3	2.8 ± 0.10
	D. melanogaster	5.4 ± 0.1	3.6 ± 0.01	0.7 ± 0.01
	D. ponderosae	4.3 ± 0.1	6.9 ± 0.03	1.6 ± 0.04
	H.vulgaris	5.7 ± 0.05	18.0 ± 0.2	3.2 ± 0.07
	O. bimaculoides	4.9 ± 0.1	13.6 ± 0.2	2.8 ± 0.08
	L. loa	7.0 ± 0.7	117 ± 3.1	17 ± 2.20
	L. loa P335A	8.4 ± 0.2	70.6 ± 1.0	8.4 ± 0.32
	L. loa Q399R	8.2 ± 0.1	38.0 ± 0.1	4.6 ± 0.07
	L. loa α_3	8.8 ± 0.1	14.4 ± 1.0	1.6 ± 0.13
YKQTSV	**H. sapiens**	**6.3 ± 0.1**	**2.5 ± 0.01**	**0.4 ± 0.01**
	1R	**6.2 ± 0.1**	**3.5 ± 0.02**	**0.6 ± 0.01**
	Deuterostomia	9.3 ± 0.2	5.0 ± 0.03	0.5 ± 0.01 *
	S. kowalevskii	**8.1 ± 0.1**	**4.2 ± 0.04**	**0.5 ± 0.01**
YRQSSV	**Deuterostomia**	**9.4 ± 0.2**	**13.4 ± 0.1**	**1.4 ± 0.03**
	Protostomia	**4.6 ± 0.3**	**10.3 ± 0.1**	**2.2 ± 0.16**
	Bilateria	**5.1 ± 0.2**	**6.8 ± 0.1**	**1.3 ± 0.07**
	O. bimaculoides	**5.8 ± 0.1**	**8.5 ± 0.1**	**1.5 ± 0.02**
	H. sapiens	5.5 ± 0.1	6.8 ± 0.02	1.2 ± 0.03
YRQTSC	**H.vulgaris**	**10.7 ± 0.3**	**4.7 ± 0.2**	**0.4 ± 0.03**
	H. sapiens	10.3 ± 0.2	3.3 ± 0.07	0.3 ± 0.01

Table 2
Binding constants for extant and reconstructed ancestral PDZ3:CRIPT interactions monitored by ITC. Thermodynamic parameters for native ML or AltAll(*) PDZ:CRIPT interactions were determined at 25°C in 50 mM sodium phosphate, pH 7.45, 21 mM KCl (I = 150). A longer *L. loa* CRIPT peptide was used (8 residues) due to low solubility of the hexamer.

CRIPT	PDZ3	$\Delta H(^{kcal}/_{mol})$	$T\Delta S(^{kcal}/_{mol})$	K_d (μM)	
YRQSSA	D. ponderosae	-0.8 ± 0.2	5.4	31 ± 11	
	Hexapoda	-2.3 ± 2.0	3.5	56 ± 25	*
GKKMMSTKNYRQSSA	D. ponderosae α_3	-1.0 ± 0.02	5.5	15 ± 4	
YRQSST	D. melanogaster	-1.0 ± 0.2	5.3	27 ± 20	
	Hexapoda	-1.1 ± 0.1	5.3	21 ± 4	
YRQSpST	D. melanogaster α_3	-1.6 ± 0.01	4.4	40 ± 9	
GKKIMNTKNYKQSST	D. melanogaster α_3	-0.5 ± 0.02	5.6	34 ± 13	
KKLKMSAV	L. loa	-1.1 ± 0.2	5.7	12 ± 7	
GKKMADTKKLKMSAV	L. loa α_3	-0.8 ± 0.02	6.3	6.8 ± 2	
YKQTSV	H. sapiens	-7.5 ± 0.2	0.9	0.7 ± 0.1	
	1R	-7.6 ± 0.2	1.1	0.5 ± 0.1	
	Deuterostomia	-8.8 ± 0.1	-0.6	0.9 ± 0.1	*
	S. kowalevskii	-7.8 ± 0.1	0.3	1.2 ± 0.1	
YRQSSV	Deuterostomia	-6.8 ± 0.1	0.9	2.1 ± 0.2	
	Protostomia	-6.4 ± 0.2	1.3	2.4 ± 0.3	
	Bilateria	-6.8 ± 0.2	1.1	1.6 ± 0.3	
	O. bimaculoides	-6.6 ± 0.2	1.0	2.5 ± 0.4	
YRQTSC	H. vulgaris	-8.3 ± 0.2	0.5	0.4 ± 0.1	

Table 3
Global stability of PDZ3 domains. Global stability of PDZ3 domains was determined by urea denaturation experiments monitored by circular dichroism at 222 nm (see Fig. 6). Experimental data were fitted to a two state model for denaturation to obtain the concentration of urea where 50% of the protein is denatured ($[\text{Urea}]_{50\%}$), the cooperativity of the unfolding ($m_{D\text{-}N}$ value) and the stability ($\Delta G_{D\text{-}N}$) as the product of $[\text{Urea}]_{50\%}$.and $m_{D\text{-}N}$.

PDZ3	$[\text{Urea}]_{50\%}$[ǀ] (M)	$\Delta G_{D\text{-}N}$[ǀ] (kcal mol^{-1})	$[\text{Urea}]_{50\%}$[†] (M)	$\Delta G_{D\text{-}N}$[†] (kcal mol^{-1})	$m_{D\text{-}N}$[†] (kcal mol^{-1})
H. sapiens	6.2 ± 0.4	7.6 ± 1.3	6.5 ± 0.4	6.3 ± 1.3	1.0 ± 0.2
S. kowalevskii	2.7 ± 0.1	3.3 ± 0.9	2.7 ± 0.1	4.3 ± 0.9	1.6 ± 0.3
D. melanogaster	2.1 ± 0.1	2.6 ± 0.3	2.1 ± 0.1	2.4 ± 0.3	1.2 ± 0.1
D. melanogaster α_3	3.0 ± 0.1	4.0 ± 0.4[#]	3.0 ± 0.04	4.4 ± 0.3	1.5 ± 0.1
D. ponderosae	3.7 ± 0.1	4.5 ± 0.8	3.7 ± 0.1	5.1 ± 0.8	1.4 ± 0.2
D. ponderosae α_3	4.5 ± 0.1	6.0 ± 0.6[#]	4.4 ± 0.05	7.2 ± 0.7	1.6 ± 0.2
L. loa	4.1 ± 0.1	5.0 ± 1.4	4.0 ± 0.1	7.0 ± 1.4	1.7 ± 0.3
L. loa α_3	5.7 ± 0.1	7.7 ± 0.7[#]	5.8 ± 0.2	6.6 ± 1.4	1.1 ± 0.2
O. bimaculoides	3.1 ± 0.1	3.8 ± 0.8	3.2 ± 0.1	4.6 ± 0.8	1.5 ± 0.3
H. vulgaris	2.9 ± 0.1	3.6 ± 0.8	2.9 ± 0.1	3.9 ± 0.8	1.3 ± 0.3
Bilateria	5.2 ± 0.2	6.4 ± 0.6	5.4 ± 0.2	4.1 ± 0.6	0.8 ± 0.1
Bilateria AltAll	5.8 ± 0.04	7.1 ± 0.3	5.8 ± 0.04	6.8 ± 0.3	1.2 ± 0.05
Protostomia	5.1 ± 0.1	6.3 ± 1.2	5.1 ± 0.1	6.4 ± 1.2	1.2 ± 0.2
Protostomia AltAll	5.3 ± 0.2	6.5 ± 1.4	5.3 ± 0.2	5.9 ± 1.4	1.1 ± 0.3
Deuterostomia	6.0 ± 0.1	7.4 ± 0.6	6.1 ± 0.1	7.1 ± 0.6	1.2 ± 0.1
Deuterostomia AltAll	4.7 ± 0.1	5.8 ± 0.9	4.7 ± 0.1	5.1 ± 0.9	1.1 ± 0.2
Hexapoda	7.0 ± 0.7	10 ± 4.0[*]	6.7 ± 0.4	17 ± 13	2.5 ± 2.0
Hexapoda AltAll	5.2 ± 0.3	7.4 ± 2.8[*]	5.3 ± 0.4	5.8 ± 3.0	1.1 ± 0.5
1R	6.1 ± 0.2	7.5 ± 2.4	5.9 ± 0.2	9.4 ± 2.3	1.6 ± 0.4
1R AltAll		"	6.6 ± 0.5	13 ± 7.2	1.9 ± 1.1

[ǀ] The $m_{D\text{-}N}$ value was shared among the datasets in the curve fitting; $m_{D\text{-}N}$ = 1.23 kcal mol^{-1}.
Exceptions : [*] Hexapoda and Hexapoda AltAll $m_{D\text{-}N}$ = 1.43 kcal mol^{-1}
[#] PDZ3 α_3 *D.melanogaster*, *L. loa* and *D. ponderosae* $m_{D\text{-}N}$ = 1.35 kcal mol^{-1}
" 1R AltAll was only fitted with a free $m_{D\text{-}N}$ value
The free fitting of *L. loa* PDZ3 with the long α_3 helix suffers from a short and sloping denatured baseline, which likely leads to underestimation of the $m_{D\text{-}N}$ value and thus $\Delta G_{D\text{-}N}$.

[†] Free fitting of both $[\text{Urea}]_{50\%}$ and $m_{D\text{-}N}$.