

Supplementary Table ST2 Examples of domain families showing high tethering number. Tethering number of a domain family is defined as the number of unique family associations made by the domain family in the chosen dataset of multi-domain proteins.

S.No.	Domain family (Interpro ID)	Tethering number	Brief description of the functions of the domain family
1	IPR003593	450	AAA ATPase, a large family of proteins. Members belong to P-loop NTPases superfamily, and are involved in binding to ATP and bringing about energy dependent unfolding of macromolecules.
2	IPR017442	444	Protein kinase domain, involved in transferring the γ phosphate from ATP to one or more amino acids
3	IPR003961	315	Fibronectin, are glycoproteins which mediate binding to variety of substrates such as heparin, collagen, DNA, fibronectin receptors on cell surfaces.
4	IPR019781	296	WD domain, 40 amino acid repeats is involved in variety of functions, they exist as repeats and act as protein-protein interaction modules.
5	IPR019734	277	Tetratricopeptide repeats, contains a minimum of 34 amino acids and occur as repeats. They self-associate via "knobs and holes" mechanism and are largely involved in mediating protein-protein interactions
6	IPR001650	274	DNA/RNA helicase, C-terminal region is the integral part of variety of helicases and helicases related family, utilizing ATP to unwind DNA
7	IPR002110	273	Ankyrin repeats, they exist as tandem repeats. They act as the most common protein interaction motif and are involved in diverse functions
8	IPR001841	268	Zinc finger Ring types, members belonging to this family are specialized type of zinc finger of 40-60 amino acids and are probably involved in mediating protein-protein interactions. There also exist two variants namely C3HC4-type and C3H2C3-type. RING domain has been implicated in many biological processes E3 ubiquitin ligase activity is intrinsic to the RING domain of c-Cbl.
9	IPR014001	262	DEAD like helicase N-terminal, it is found in proteins such as DEAD and DEAH helicases
10	IPR000477	254	Reverse transcriptase, members include RNA dependent DNA polymerase and retroelements
11	IPR001584	242	Integrase catalytic core, members act as endonuclease. This domain is found in Transposase and reverse transcriptase genes.
12	IPR000504	235	RNA recognition motif, RNP-1 consensus motif is present in RRM proteins that have a variety of RNA binding preference and function. This includes hnRNPs, component of small nuclear ribonuclear protein and proteins regulating RNA stability and translation.

13	IPR001611	234	Leucine-rich repeats, members of this family are involved in mediating protein interactions
14	IPR007087	222	Zinc finger C2H2 type , members of this family are involved in binding to DNA, RNA and proteins.
15	IPR003594	203	Histidine kinase- DNA gyrase B- phytochrome-like ATPases domain. Found in several ATP binding proteins like Histidine kinases and topoisomerases
16	IPR002885	201	Pentatricopeptide repeats (PPR) are tandem repeats of a degenerate 35 amino acid motif. They may have some role to play in post-transcriptional process and are thought to be sequence specific RNA binding proteins.
17	IPR001789	201	Receives signals from the sensor partner in bacterial two component system. Found in the N-terminal region of the DNA binding effector.
18	IPR000182	195	GCN5-related N-acetyltransferase , histone acetyltransferase. Involved in regulation of cell growth and development. Particularly important in transcription and DNA repair.
19	IPR018249	190	EF hand 2 , members of the family involved in calcium binding.
20	IPR000421	190	Coagulation factor 5/8 type , the C-terminal domain. The domain is found in extracellular and membrane proteins, carbohydrate binding proteins and phosphatidyl serine binding proteins.