

Supporting Information

Diversity and structural-functional insights of alpha-solenoid proteins

Paula Nazarena Arrías^{1,2†}, Zarifa Osmanli^{1†}, Estefanía Peralta³, Patricio Manuel Chinestrada⁴,
Alexander Miguel Monzon⁵, Silvio C.E. Tosatto^{1,6*}

† These authors contributed equally to this work.

* To whom correspondence should be addressed: silvio.tosatto@unipd.it

¹ Department of Biomedical Sciences, University of Padova, via U. Bassi 58/b, 35121 Padova, Italy

² Department of Protein Science, KTH Royal Institute of Technology, Stockholm SE-10691, Sweden

³ Laboratorio de Investigación y Desarrollo de Bioactivos (LIDeB), Departamento de Ciencias Biológicas, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, B1900ADU La Plata, Buenos Aires, Argentina.

⁴ Laboratorio de Farmacología Molecular, Universidad Nacional de Quilmes, Roque Saenz Peña 352, 1876 Bernal, Buenos Aires, Argentina

⁵ Department of Information Engineering, University of Padova, via Giovanni Gradenigo 6/B, 35131 Padova, Italy

⁶ Institute of Biomembranes, Bioenergetics and Molecular Biotechnologies, National Research Council (CNR-IBIOM), Bari, Italy

Keywords: alpha-solenoid subfolds, repeat evolution, protein sequence logo, repeat structure, amino acid conservation

Abbreviation	Name	Interpro repeat	GO terms associated with the Interpro repeat	Interpro homologous superfamily	Pfam	Pfam clan	CATH ID	Structural unit	Representative structure (PDB)
					PF12796				
					PF13606			α/α /big loop	
					PF13857			or	
					PF13637			α/α /beta-hair	
ANK	Ankyrin	IPR002110	GO:0005515	IPR036770	PF00023	CL0465	1.25.40.20	pin	1N11_A
					PF13174				
					PF13181				
					PF13176				
TPR	Tetratricopeptide	IPR01973	GO:0005515	IPR011990	PF00515	CL0020	1.25.40.10	α/α	4Y6W_A
					IPR00035				
HEAT	HEAT	7	GO:0005515	IPR016024	PF02985	CL0020	1.25.10.10	α/α	4I5N_D
					IPR00131				
PUF	Pumilio	3	GO:0003723	IPR016024	PF22493	CL0020	1.25.10.10	$\alpha/\alpha/\alpha$	3V71_A
					PF12854				
					PF13041				
					PF13812				
PPR	opeptide	5	N/A	IPR011990	PF01535	CL0020	1.25.40.10	α/α	4M57_A
					IPR00022				
ARM	Armadillo	5	GO:0005515	IPR016024	PF00514	CL0020	1.25.10.10	$\alpha/\alpha/\alpha$	4DJS_A
					TAL				
					effector repeat				
TAL	repeat	2	N/A	N/A	PF03377	CL0020	N/A	α/α	3UGM_A
					Sel1-like				
Sel1	repeats	7	N/A	IPR011990	PF08238	CL0020	1.25.40.10	α/α	5B26_A

Table S1 – Alpha-solenoid subfolds. Database identifiers, representative structures, GO terms correspond to protein and DNA/RNA binding (UniProtKB accessed on 17 June 2024).