

Additional file 9. Predicted model scores and ligands for representative EPDR sequences.

Protein	Clade/ profile	C-Score (model)	Predicted ligands	C-Score (ligand)
<i>D. rerio</i> 134034	1/1	-1.20	Phenylpyruvic acid	0.07
<i>M. musculus</i> ENSMUSP00000002885	1/1	-0.78	Tetraethylene glycol monoethyl ether	0.05
			Tetraethylene glycol	0.37
			Peptide	0.06
			Phenylpyruvic acid	0.06
<i>C. owczarzaki</i> 30864	1/1	-0.64	Tetraethylene glycol monoethyl ether	0.04
			Tetraethylene glycol	0.21
			Tetraethylene glycol monoethyl ether	0.07
<i>S. kowalevskii</i> 91225509	2/2	-2.85	Tetraethylene glycol monoethyl ether	0.10
<i>C. globosum</i> HO349164	2/2	-3.03	Phenylpyruvic acid	0.10
			Tetraethylene glycol	0.23
			Tetraethylene glycol monoethyl ether	0.07
<i>N. gruberi</i> EFC42264	2/3	-2.44	Phenylpyruvic acid	0.06
			Tetraethylene glycol	0.15
			Tetraethylene glycol monoethyl ether	0.15
<i>A. californica</i> c13118	2/3	-1.74	Tetraethylene glycol	0.13
			Tetraethylene glycol monoethyl ether	0.07

Details of 3D modelling of EPDR sequences in I-TASSER. C-score (model) is the confidence score of the model on a scale of -5 to 2, where higher values signify higher confidence. C-score (ligand) denotes the confidence of the ligand prediction on a scale of 0 to 1, with higher values again signifying higher confidence.