

Supplementary material

Article

New Three-Finger Protein from Starfish *Asteria rubens* Shares Structure and Pharmacology of Human Brain Neuromodulator Lynx2

Alexander S. Paramonov¹, Mikhail A. Shulepko¹, Alexey M. Makhonin^{1,2}, Maxim L. Bychkov¹, Dmitrii S. Kulbatskii¹, Andrey M. Chernikov^{1,3}, Mikhail Yu. Myshkin¹, Sergey V. Shabelnikov⁴, Zakhar O. Shenkarev^{1,5}, Mikhail P. Kirpichnikov^{1,3} and Ekaterina N. Lyukmanova^{1,3,5,*}

¹ Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Miklukho-Maklaya str. 16/10, 119997, Moscow, Russia; apar@nmr.ru (A.S.P.); mikhailshulepko@gmail.com (M.A.S.); ammakhonin@edu.hse.ru (A.M.M.); maksim.bychkov@gmail.com (M.L.B.); d.kulbatskiy@gmail.com (D.S.K.); chernikov.andrei.m@gmail.com (A.M.C.); mikhail.myshkin@phystech.edu (M.Y.M.); zakhar-shenkarev@yandex.ru (Z.O.S.); kirpichnikov@inbox.ru (M.P.K.); ekaterina-lyukmanova@yandex.ru (E.N.L.)

² AI centre, National Research University Higher School of Economics, Myasnitskaya str. 20, 101000, Moscow, Russia

³ Interdisciplinary Scientific and Educational School of Moscow University «Molecular Technologies of the Living Systems and Synthetic Biology», Faculty of Biology, Lomonosov Moscow State University, Leninskie Gory, 119234, Moscow, Russia.

⁴ Institute of Cytology of the Russian Academy of Sciences, Tikhoretsky prospect 4, 194064, St. Petersburg, Russia; buddasvami@gmail.com (S.V.S.)

⁵ Moscow Institute of Physics and Technology (State University), Institutskiy per. 9, 141701, Dolgoprudny, Moscow Region, Russia.

* Correspondence: ekaterina-lyukmanova@yandex.ru

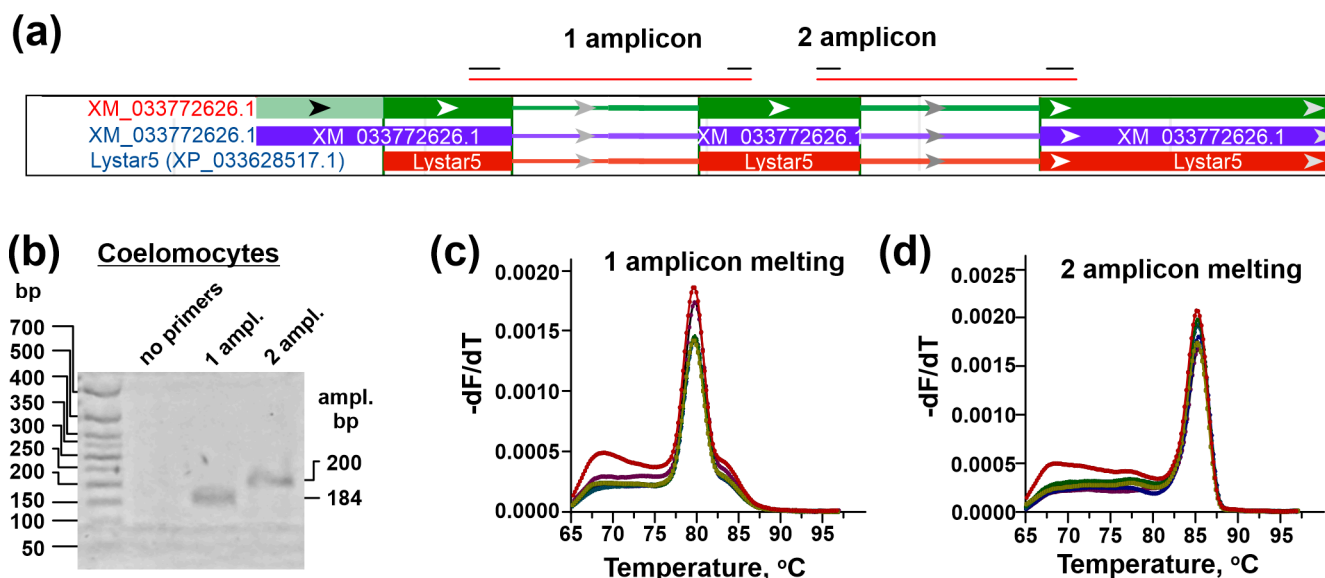


Figure S1. The primer design and analysis of Lystar5 expression in *A. rubens*. (a) The primer design scheme, the amplicons 1 and 2 are shown by red lines. (b) Representative electrophoretic analysis of the product obtained after the PCR reaction with negative control (no primers in the reaction) and in the presence of the primer pair 1 and 2 with cDNA extracted from coelomocytes. The product length is shown on the right. (c,d) the melting curves of amplicons 1 (c) and 2 (d) obtained during real-time PCR.

signal peptide				LU-domain	
10	20	30	40	50	60
MRIKMVFKFR	RNSEGIFCIL	MTLMIILMAV	CFRTASSLQC	FTCEAEDTNE	NCNIKEAPVL
70	80	90	100	110	120
KTCTSRQDRC	LTQVIYSTER	GKLRIKECT	TEDGCTAATT	QLGKRYFCDK	SRPAWGCVEC
130	140	150			
CDTDRCNENG	VAVTRASAAV	VLGALSSVVV	ALHVL		
GPI site			removed in mature form		

Figure S2. Full sequence of Lystar5 from *A. rubens* (XP_033628517.1/LyAr1). Signal peptide predicted by SignalP [1] and the glycosylphosphatidylinositol (GPI) anchoring site predicted by PredGPI [2] are shown by red and blue, respectively.

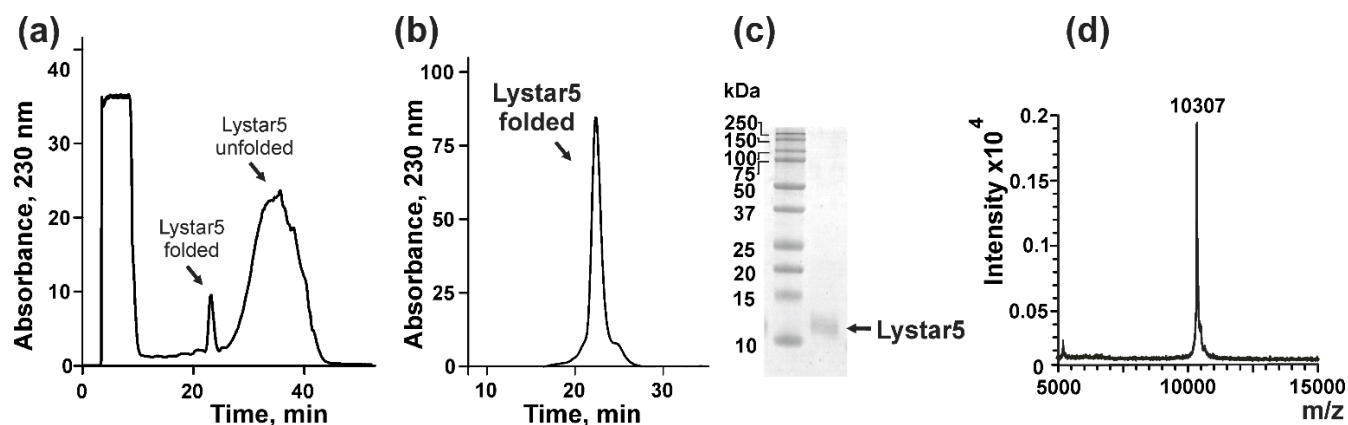


Figure S3. Characterization of the recombinant Lystar5 protein. (a) Representative chromatogram of Lystar5 HPLC purification. Fractions corresponding to the folded and unfolded Lystar5 are shown. (b) Analytical HPLC of folded Lystar5. (c) SDS-PAGE analysis of folded Lystar5. (d) Mass-spectrometry analysis of folded Lystar5.

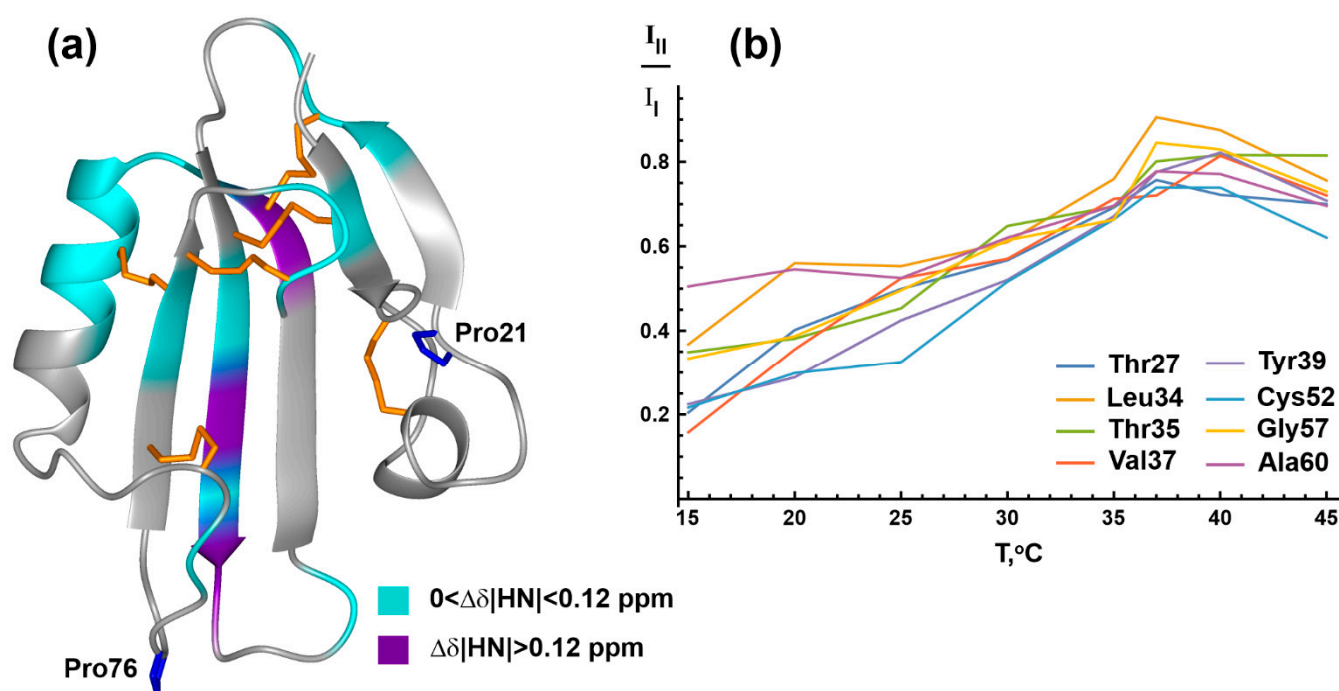


Figure S4. Data on two structural forms of Lystar5. (a) Ribbon representation of 3D structure of Lystar5 obtained using AlphaFold2. Colors show residues for which there was a difference in the chemical shifts of amide groups between the two forms ($\Delta\delta|HN| = \sqrt{(\Delta\delta_{HN})^2 + (\Delta\delta_{NH}/5)^2}$). Residues with the greatest difference ($\Delta\delta|HN| > 0.12\text{ppm}$) are marked in purple. Sidechains of prolines are shown in blue. Disulfides are shown in gold. (b) Temperature dependence of ratio of signal intensities in the ^{15}N -HSQC spectra for the form II and form I for some residues.

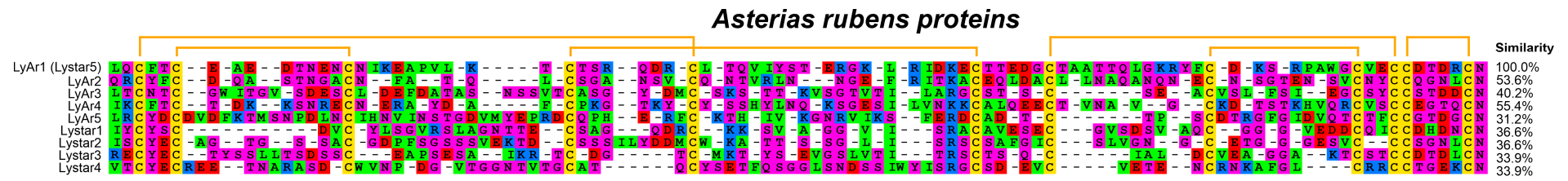


Figure S5. Multiple sequence alignment of TFPs from *Asterias rubens* found by BLAST search (LyAr1-5) and identified early by proteomics (Lystar1-4) [3].

Table S1. List of known Ly6/uPAR proteins which were used for BLAST search of TFP in starfishes.

No.	Name	Description	Organism	Access code
Snake toxins				
1	aBTX	alpha-bungarotoxin	<i>Bungarus multicinctus</i>	P60615
2	alpha cobratoxin	alpha-cobratoxin	<i>Naja kaouthia</i>	P01391
3	LSTIII	alpha-elapitoxin-Ls2a	<i>Laticauda semifasciata</i>	P01379
4	long neurotoxin1	alpha-elapitoxin-Nno2a	<i>Naja oxiana</i>	P01382
5	bCTX	beta-cardiotoxin CTX27	<i>Ophiophagus hannah</i>	Q69CK0
6	bucandin	aucandin	<i>Bungarus candidus</i>	P81782
7	calciseptine	calciseptin	<i>Dendroaspis polylepis polylepis</i>	P22947
8	candoxin	candoxin	<i>Bungarus candidus</i>	P81783
9	cobrotoxin	cobrotoxin	<i>Naja atra</i>	P60770
10	cytotoxin1	cytotoxin 1	<i>Naja kaouthia</i>	P60305
11	cytotoxin2	cytotoxin 2	<i>Naja kaouthia</i>	P01445
12	cytotoxin3	cytotoxin 3	<i>Naja kaouthia</i>	P01446
13	dendroaspin	dendroaspin	<i>Dendroaspis jamesoni kaimosae</i>	P28375
14	denmotoxin	denmotoxin	<i>Boiga dendrophila</i>	Q06ZW0
15	erabutoxin a	erabutoxin a	<i>Laticauda semifasciata</i>	P60775
16	erabutoxin b	erabutoxin b	<i>Laticauda semifasciata</i>	Q90VW1
17	fasciculin-1	Fasciculin-1	<i>Dendroaspis angusticeps</i>	P0C1Y9
18	fasciculin-2	Fasciculin-2	<i>Dendroaspis angusticeps</i>	P0C1Z0
19	haditoxin	Haditoxin	<i>Ophiophagus hannah</i>	A8N286
20	irditoxin	irditoxin subunit A	<i>Boiga irregularis</i>	A0S864
21	kappaBTX	kappa-bungarotoxin	<i>Bungarus multicinctus</i>	P01398
22	mambalgin1	mambalgin-1	<i>Dendroaspis polylepis polylepis</i>	P0DKR6
23	mambalgin2	mambalgin-2	<i>Dendroaspis polylepis polylepis</i>	P0DKS3
24	mambalgin3	mambalgin-3	<i>Dendroaspis angusticeps</i>	C0HJB0
25	micrurotoxin-1	micrurotoxin 1	<i>Micrurus mipartitus</i>	C0HJR1
26	micrurotoxin2	micrurotoxin 2	<i>Micrurus mipartitus</i>	C0HJR2
27	MT1	Muscarinic toxin 1	<i>Dendroaspis angusticeps</i>	P81030
28	MT2	Muscarinic toxin 2	<i>Dendroaspis angusticeps</i>	P18328
29	MT3	Muscarinic toxin 3	<i>Dendroaspis angusticeps</i>	P81031
30	MT4	Muscarinic toxin 4	<i>Dendroaspis angusticeps</i>	Q9PSN1
31	MT7	Muscarinic toxin 7	<i>Dendroaspis angusticeps</i>	Q8QGR0
32	alpha short neurotoxin 1	Short neurotoxin 1	<i>Naja pallida</i>	P01426
33	short neurotoxin 1	Short neurotoxin 1	<i>Naja oxiana</i>	P01427
34	S2C4	Synergistic-type venom protein S2C4	<i>Dendroaspis jamesoni kaimosae</i>	P01407
35	3FTX	Three-finger toxin MALT0070C	<i>Micrurus altirostris</i>	F5CPE6
36	FS2	Toxin FS-2	<i>Dendroaspis polylepis polylepis</i>	P01414
37	LC	Toxin Lc b	<i>Laticauda colubrina</i>	P0C8R8
38	S4C8	Toxin S4C8	<i>Dendroaspis jamesoni kaimosae</i>	P25683
39	WTX	Tryptophan-containing weak neurotoxin	<i>Naja kaouthia</i>	P82935
Mammalia				

No.	Name	Description	Organism	Access code
40	CD59	CD59 glycoprotein	<i>Oryctolagus cuniculus</i>	O77541
41	CD59	CD59 glycoprotein	<i>Papio sp.</i>	Q28785
42	CD59	CD59 glycoprotein	<i>Callithrix sp.</i>	P46657
43	CD59	CD59 glycoprotein	<i>Rattus norvegicus</i>	P27274
44	CD59	CD59 glycoprotein	<i>Chlorocebus aethiops</i>	Q28216
45	CD59	CD59 glycoprotein	<i>Sus scrofa</i>	O62680
46	CD59	CD59 glycoprotein	<i>Macaca fascicularis</i>	Q8SQ46
47	CD59	CD59 glycoprotein	<i>Homo sapiens</i>	P13987
48	CD59	CD59 glycoprotein	<i>Saimiri sciureus</i>	P47777
49	CD59	CD59 glycoprotein	<i>Aotus trivirgatus</i>	P51447
50	CD59	CD59 glycoprotein	<i>Pongo abelii</i>	Q5R510
51	CD59	CD59A glycoprotein	<i>Mus musculus</i>	O55186
52	CD59B	CD59B glycoprotein	<i>Mus musculus</i>	P58019
53	GPIHBP1	Glycosylphosphatidylinositol-anchored high density lipoprotein-binding protein 1	<i>Homo sapiens</i>	Q8IV16
54	GML	Glycosyl-phosphatidylinositol-anchored molecule-like protein	<i>Homo sapiens</i>	Q99445
55	LYNX1	Ly-6/neurotoxin-like protein 1	<i>Macaca mulatta</i>	P0DP62
56	LYNX1	Ly-6/neurotoxin-like protein 1	<i>Bos taurus</i>	Q1RMQ4
57	LYNX1	Ly-6/neurotoxin-like protein 1	<i>Homo sapiens</i>	P0DP58
58	LYNX1	Ly-6/neurotoxin-like protein 1	<i>Pan troglodytes</i>	Q5IS42
59	LYNX1	Ly-6/neurotoxin-like protein 1	<i>Saimiri boliviensis boliviensis</i>	Q5IS87
60	LYNX1	Ly-6/neurotoxin-like protein 1	<i>Mus musculus</i>	P0DP60
61	LYPD1	Ly6/PLAUR domain-containing protein 1	<i>Rattus norvegicus</i>	Q66H42
62	LYPD1	Ly6/PLAUR domain-containing protein 1	<i>Homo sapiens</i>	Q8N2G4
63	LYPD1	Ly6/PLAUR domain-containing protein 1	<i>Mus musculus</i>	Q8BLC3
64	LYPD2	Ly6/PLAUR domain-containing protein 2	<i>Homo sapiens</i>	Q6UXB3
65	LYPD2	Ly6/PLAUR domain-containing protein 2	<i>Mus musculus</i>	Q9DD23
66	LYPD6	Ly6/PLAUR domain-containing protein 6	<i>Homo sapiens</i>	Q86Y78
67	LYPD6	Ly6/PLAUR domain-containing protein 6	<i>Rattus norvegicus</i>	D3ZTT2
68	LYPD6	Ly6/PLAUR domain-containing protein 6	<i>Mus musculus</i>	Q8BPP5
69	LYPD6B	Ly6/PLAUR domain-containing protein 6B	<i>Mus musculus</i>	Q9D7F2
70	LYPD6B	Ly6/PLAUR domain-containing protein 6B	<i>Homo sapiens</i>	Q8NI32
71	LY6G5C	Lymphocyte antigen 6 complex locus protein G5c	<i>Rattus norvegicus</i>	Q8CHN2
72	LY6G5C	Lymphocyte antigen 6 complex locus protein G5c	<i>Macaca mulatta</i>	Q863H0
73	LY6G5C	Lymphocyte antigen 6 complex locus protein G5c	<i>Mus musculus</i>	Q8K1T5

No.	Name	Description	Organism	Access code
74	LY6G5C	Lymphocyte antigen 6 complex locus protein G5c	<i>Canis lupus familiaris</i>	Q9XSV5
75	LY6G5C	Lymphocyte antigen 6 complex locus protein G5c	<i>Homo sapiens</i>	Q5SRR4
76	LY6G6C	Lymphocyte antigen 6 complex locus protein G6c	<i>Homo sapiens</i>	O95867
77	LY6G6C	Lymphocyte antigen 6 complex locus protein G6c	<i>Mus musculus</i>	Q9Z1Q4
78	LY6G6C	Lymphocyte antigen 6 complex locus protein G6c	<i>Bos taurus</i>	A0JNL5
79	LY6G6D	Lymphocyte antigen 6 complex locus protein G6d	<i>Homo sapiens</i>	O95868
80	LY6G6D	Lymphocyte antigen 6 complex locus protein G6d	<i>Rattus norvegicus</i>	Q6MG58
81	LY6G6D	Lymphocyte antigen 6 complex locus protein G6d	<i>Mus musculus</i>	Q9Z1Q3
82	LY6A	Lymphocyte antigen 6A-2/6E-1	<i>Mus musculus</i>	P05533
83	LY6B	Lymphocyte antigen 6B	<i>Rattus norvegicus</i>	Q63317
84	LY6C1	Lymphocyte antigen 6C1	<i>Mus musculus</i>	P0CW02
85	LY6C2	Lymphocyte antigen 6C2	<i>Mus musculus</i>	P0CW03
86	LY6D	Lymphocyte antigen 6D	<i>Homo sapiens</i>	Q14210
87	LY6D	Lymphocyte antigen 6D	<i>Mus musculus</i>	P35459
88	LY6D	Lymphocyte antigen 6D	<i>Bos taurus</i>	Q148C3
89	LY6E	Lymphocyte antigen 6E	<i>Homo sapiens</i>	Q16553
90	LY6E	Lymphocyte antigen 6E	<i>Mus musculus</i>	Q64253
91	LY6F	Lymphocyte antigen 6F	<i>Mus musculus</i>	P35460
92	LY6G	Lymphocyte antigen 6G	<i>Mus musculus</i>	P35461
93	LY6G6E	Lymphocyte antigen 6G6e	<i>Mus musculus</i>	Q8K1T6
94	LY6H	Lymphocyte antigen 6H	<i>Macaca fascicularis</i>	Q4R5M8
95	LY6H	Lymphocyte antigen 6H	<i>Bos taurus</i>	A0JNB3
96	LY6H	Lymphocyte antigen 6H	<i>Homo sapiens</i>	O94772
97	LY6H	Lymphocyte antigen 6H	<i>Mus musculus</i>	Q9WUC3
98	LY6I	Lymphocyte antigen 6I	<i>Mus musculus</i>	Q9WU67
99	LY6K	Lymphocyte antigen 6K	<i>Homo sapiens</i>	Q17RY6
100	LY6K	Lymphocyte antigen 6K	<i>Mus musculus</i>	Q9CWP4
101	LY6L	Lymphocyte antigen 6L	<i>Homo sapiens</i>	H3BQJ8
102	LY6L	Lymphocyte antigen 6L	<i>Mus musculus</i>	H3BJG9
103	PATE1	Prostate and testis expressed protein 1	<i>Homo sapiens</i>	Q8WXA2
104	PAT13	Prostate and testis expressed protein 13	<i>Mus musculus</i>	D3YWX3
105	PAT14	Prostate and testis expressed protein 14	<i>Rattus norvegicus</i>	Q9QXN2
106	PAT14	Prostate and testis expressed protein 14	<i>Mus musculus</i>	Q3UN54
107	PATE2	Prostate and testis expressed protein 2	<i>Mus musculus</i>	Q3UW31
108	PATE2	Prostate and testis expressed protein 2	<i>Homo sapiens</i>	Q6UY27

No.	Name	Description	Organism	Access code
109	PATE3	Prostate and testis expressed protein 3	<i>Homo sapiens</i>	B3GLJ2
110	PATE3	Prostate and testis expressed protein 3	<i>Mus musculus</i>	B3GLJ3
111	PATE4	Prostate and testis expressed protein 4	<i>Mus musculus</i>	Q09098
112	PATE4	Prostate and testis expressed protein 4	<i>Homo sapiens</i>	P0C8F1
113	PSCA	Prostate stem cell antigen	<i>Homo sapiens</i>	O43653
114	PSCA	Prostate stem cell antigen	<i>Mus musculus</i>	P57096
115	PIP1	Protein PIP-1	<i>Sus scrofa</i>	P83106
116	SLURP2	Secreted Ly-6/uPAR domain-containing protein 2	<i>Homo sapiens</i>	P0DP57
117	SLURP2	Secreted Ly-6/uPAR domain-containing protein 2	<i>Macaca mulatta</i>	P0DP61
118	SLURP2	Secreted Ly-6/uPAR domain-containing protein 2	<i>Mus musculus</i>	P0DP59
119	SLURP1	Secreted Ly-6/uPAR-related protein 1	<i>Homo sapiens</i>	P55000
120	SLURP1	Secreted Ly-6/uPAR-related protein 1	<i>Mus musculus</i>	Q9Z0K7
121	SOLD1	Secreted protein of Ly-6 domain 1	<i>Bos taurus</i>	P83107
122	SACA4	Sperm acrosome membrane-associated protein 4	<i>Mus musculus</i>	Q80ZQ0
123	SACA4	Sperm acrosome membrane-associated protein 4	<i>Bos taurus</i>	Q32PB3
124	SAMP14	Sperm acrosome membrane-associated protein 4	<i>Homo sapiens</i>	Q8TDM5
125	UP1	Urinary protein 1	<i>Rattus norvegicus</i>	P81827
126	UP2	Urinary protein 2	<i>Rattus norvegicus</i>	P81828
127	UP3	Urinary protein 3	<i>Rattus norvegicus</i>	P83121
Trematoda				
128	SMP105220	DIF_5	<i>Schistosoma mansoni</i>	B8Y6H3
Viruses				
129	CD59	Surface glycoprotein CD59 homolog	<i>Saimiriine herpesvirus 2 (strain 11)</i>	Q00996
Insecta				
130	QVR	Protein quiver	<i>Drosophila melanogaster</i>	B5A5T4
131	LYNX1	Protein quiver	<i>Nilaparvata lugens</i>	C5IWP3
Aves				
132	LY6E	Lymphocyte antigen 6E	<i>Gallus gallus</i>	Q90986
133	PSCA	UPAR/Ly6 domain-containing protein	<i>Meleagris gallopavo</i>	G1NJY7
Amphibia				
134	WFP	Freeze-responsive liver protein Li16	<i>Lithobates sylvaticus</i>	Q9DG71
135	TFP1	MAC-inhibitory protein	<i>Xenopus laevis</i>	A0A3S8RK06

No.	Name	Description	Organism	Access code
136	PMFG	Plethodontid modulating factor isoform G	<i>Plethodon shermani</i>	Q0GAX2
137	PROD1	Prod 1	<i>Notophthalmus viridescens</i>	A8D0E6
138	PROD1	Prod1	<i>Batrachuperus londongensis</i>	A0A0F6QEW0
139	PROD1	Prod1	<i>Cynops orientalis</i>	A0A0F6QFB4
140	PROD1 long form	Prod1 long form	<i>Aneides lugubris</i>	A0A0F6QDT3
141	TFP2	Tfp2	<i>Xenopus laevis</i>	A0A3S8RJZ9
142	TFP3	Tfp3	<i>Xenopus laevis</i>	A0A1L8F8M0
143	TFP4	Tfp4	<i>Xenopus laevis</i>	A0A3S5HJM2
144	TFP5	Tfp5	<i>Xenopus laevis</i>	A0A1L8G0B7
145	TFP6	Tfp6	<i>Xenopus laevis</i>	A0A3Q8S4U8
146	LYPD6	Ly6/PLAUR domain-containing protein 6	<i>Xenopus laevis</i>	A0A1L8ENX6
Actinopterygii				
147	CD59	CD59	<i>Danio rerio</i>	A0A0B5JW41
148	LY2.3	Cocaine-and amphetamine-regulated transcript protein-like 1	<i>Danio rerio</i>	A0A2R8RTX1
149	LYPD6	Ly6/PLAUR domain-containing protein 6	<i>Danio rerio</i>	Q66IA6
150	LY2.2	MAC-inhibitory protein	<i>Danio rerio</i>	A0A0G2L893
151	protein bouncer	protein bouncer	<i>Oryzias latipes</i>	H2LID1
152	protein bouncer	protein bouncer	<i>Danio rerio</i>	P0DPQ9
153	LY2.1	Three-finger protein 5	<i>Danio rerio</i>	D2CRG0

Table S2. List of putative three-finger proteins which were identified in the genomes of starfishes

No.	Name	Organism	Access code
1	LyAr1 (Lystar5)	<i>Asterias rubens</i>	XP_033628517.1
2	LyAr2	<i>Asterias rubens</i>	XP_033635706.1
3	LyAr3	<i>Asterias rubens</i>	XP_033638913.1
4	LyAr4	<i>Asterias rubens</i>	XP_033640057.1
5	LyAr5	<i>Asterias rubens</i>	XP_033644501.1
6	LyAp1	<i>Acanthaster planci</i>	XP_022094712.1
7	LyAp2	<i>Acanthaster planci</i>	XP_022094807.1
8	LyAp3	<i>Acanthaster planci</i>	XP_022100632.1
9	LyAp4	<i>Acanthaster planci</i>	XP_022109332.1
10	LyAp5	<i>Acanthaster planci</i>	XP_022111363.1
11	LyAp6	<i>Acanthaster planci</i>	XP_022086630.1

Table S3. Pairwise similarity between some TFPs from starfishes and previously known TFPs. Proteins and organisms are listed in the table

	Lyx2 Homo sapiens	LyAr1 Asterias rubens	LyAr2 Asterias rubens	LyAp1 Acanthaster planci	LyAp2 Acanthaster planci	LyAp5 Acanthaster planci
Lyx2 Homo sapiens	100%	50.5%	55.3%	50.5%	58.8%	55.8%
LyAr1 Asterias rubens	-	100.0%	42.2%	85.6%	46.7%	45.1%
LyAr2 Asterias rubens	-	-	100.0%	40.0%	82.9%	54.8%
LyAp1 Acanthaster planci	-	-	-	100.0%	43.3%	41.8%
LyAp2 Acanthaster planci	-	-	-	-	100.0%	58.3%
LyAp5 Acanthaster planci	-	-	-	-	-	100.0%

Table S4. Pairwise similarity between some TFPs from starfishes and previously known TFPs. Proteins and organisms are listed in the table

	LYPD6 Homo sapiens	LYPD6B Homo sapiens	LyAp4 Acanthaster planci	LyAr4 Asterias rubens
LYPD6 Homo sapiens	100%	61.0%	65.9%	61.2%
LYPD6B Homo sapiens	-	100%	52.9%	55.3%
LyAp4 Acanthaster planci	-	-	100%	95.3%
LyAr4 Asterias rubens	-	-	-	100%

Table S5. Pairwise similarity between some TFPs from starfishes and previously known TFPs. Proteins and organisms are listed in the table

	LYNX1 N. lugens	Mambalgin 1 D. polylepis	PROD1 N. viridescens	LyAr3 Asterias rubens	LyAp3 Acanthaster planci	LyAp6 Acanthaster planci
LYNX1 Nilaparvata_lugens	100%	29.8%	31.2%	38.4%	44.9%	35.6%
mambalgin1 Dendroaspis polylepis	-	100%	30.3%	39.3%	26.5%	22.1%
PROD1 Notophthalmus viridescens	-	-	100%	25.8%	22.1%	26.6%
LyAr3 Asterias rubens	-	-	-	100%	63.1%	44.3%
LyAp3 Acanthaster planci	-	-	-	-	100%	39.2%
LyAp6 Acanthaster planci	-	-	-	-	-	100%

Table S6. Chemical shifts of the signals of backbone atoms of Lystar5 in H₂O solution, 800MHz, pH 7.0, 37°C.

No.	Res.	Form I					Form II				
		H ^N	N	C ^α	C	H ^α	H ^N	N	C ^α	C	H ^α
0	MET	-	-	-	-	-					
1	LEU	-	-	52.3	171.6	-					
2	GLN	8.06	121.4	51.5	172.7	5.33					
3	CYS	8.42	118.3	49.0	172.1	4.89	8.33	118.5	48.9	-	4.92
4	PHE	8.97	122.1	56.9	173.3	4.84					
5	THR	8.90	113.1	57.6	169.9	4.78					
6	CYS	8.49	114.6	53.6	169.1	4.98					
7	GLU	9.22	122.1	52.3	171.0	5.15					
8	ALA	9.12	128.1	50.3	173.9	3.67					
9	GLU	8.57	119.1	53.3	174.5	4.63					
10	ASP	8.53	115.9	50.2	173.5	5.02					
11	THR	7.41	111.2	56.7	171.3	-					
12	ASN	-	-	53.3	174.8	-					
13	GLU	8.58	117.6	56.1	174.6	4.07					
14	ASN	7.04	114.9	50.6	172.7	-					
15	CYS	7.70	114.5	54.3	171.4	4.71					
16	ASN	7.21	118.8	49.4	171.3	5.44					
17	ILE	6.38	120.0	58.5	173.5	4.07					
18	LYS	8.58	127.0	53.6	172.7	4.54					
19	GLU	7.92	118.1	52.2	173.4	4.61					
20	ALA	-	-	-	-	-					
21	PRO		-	60.9	172.6	-					
22	VAL	7.70	124.9	62.3	174.5	4.26					
23	LEU	7.58	121.9	52.1	172.6	4.45					
24	LYS	8.61	128.6	52.5	172.5	4.61					
25	THR	8.37	118.9	60.5	171.6	4.47	8.45	119.3	60.5	171.7	4.46
26	CYS	9.27	128.3	50.0	172.8	4.73	9.17	128.1	50.0	172.7	4.77
27	THR	8.45	109.8	59.3	174.0	4.53	8.50	110.0	59.3	174.1	4.52
28	SER	-	-	-	-	-					
29	ARG	-	-	53.2	173.3	-					
30	GLN	7.41	119.1	54.1	171.4	4.10					
31	ASP	8.04	121.2	50.2	174.5	4.62					
32	ARG	8.41	121.5	50.6	171.2	4.98					
33	CYS	8.88	116.1	49.4	171.2	5.46					
34	LEU	9.08	126.1	52.6	174.2	5.33	9.16	126.1	52.4	174.3	5.29
35	THR	9.05	123.0	60.2	169.6	5.07	8.96	122.8	60.3	169.8	5.11
36	GLN	9.80	130.0	51.8	170.8	5.40	9.64	129.2	51.8	170.9	5.36
37	VAL	8.75	125.2	57.9	171.5	5.14	8.83	125.7	57.8	171.4	5.10
38	ILE	8.56	123.9	56.9	171.9	4.40	8.53	124.1	57.4	171.7	4.36
39	TYR	8.34	127.3	54.5	170.8	4.32	8.43	128.3	54.3	170.9	4.37
40	SER	6.36	117.3	53.2	172.2	4.76	6.69	117.7	53.4	-	-
41	THR	8.36	121.0	62.5	173.3	3.98					
42	GLU	8.49	121.3	56.6	174.9	-					
43	ARG	7.46	115.9	53.8	174.7	-					
44	GLY	8.01	107.8	43.5	170.8	3.78/	8.03	108.1	43.5	-	3.78/

No.	Res.	Form I					Form II				
		H ^N	N	C ^α	C	H ^α	H ^N	N	C ^α	C	H ^α
						3.98					3.98
45	LYS	7.01	114.8	51.9	172.3	4.72					
46	LEU	8.68	123.7	52.1	173.3	4.65					
47	ARG	8.52	120.9	53.4	171.7	4.64					
48	ILE	9.38	130.0	57.8	171.0	4.98					
49	ASP	8.95	123.5	50.8	171.9	5.29					
50	LYS	8.98	119.6	52.9	174.1	5.60					
51	GLU	8.40	117.3	53.1	172.9	4.83	8.40	118.1	53.0	173.1	4.83
52	CYS	8.56	118.4	55.2	172.0	5.23	8.71	118.4	55.3	171.9	5.20
53	THR	9.43	117.6	58.0	169.6	4.75	9.39	117.3	58.1	169.7	4.71
54	THR	8.71	110.3	56.1	173.3	5.33	8.66	110.0	56.1	173.4	5.33
55	GLU	9.56	123.8	58.7	176.0	-	9.51	123.9	58.7	176.0	-
56	ASP	8.62	117.6	54.5	176.4	4.46	8.64	117.6	54.4	176.4	4.46
						3.87/					3.85/4
57	GLY	8.29	111.6	44.0	174.8	4.06	8.37	111.9	44.0	174.8	.06
58	CYS	8.58	123.3	59.4	175.0	4.20	8.59	123.4	59.4	175.0	4.20
59	THR	8.44	120.5	63.9	174.2	4.00	8.37	120.6	63.9	174.1	4.01
60	ALA	8.53	125.7	52.8	177.3	4.18	8.57	125.7	52.9	177.2	4.20
61	ALA	8.31	120.1	52.6	178.2	4.20	8.34	120.2	52.6	178.2	4.20
62	THR	8.24	116.2	64.7	174.9	3.86					
63	THR	8.49	119.5	63.5	174.3	4.01					
64	GLN	8.48	120.2	56.4	177.0	4.07					
65	LEU	8.34	121.6	55.2	177.0	4.24					
						3.90/					
66	GLY	7.94	104.1	43.1	172.3	4.58					
67	LYS	7.21	117.3	55.0	174.4	4.20					
68	ARG	7.74	116.3	53.9	172.3	4.27					
69	TYR	8.45	116.1	54.3	172.7	4.91					
70	PHE	9.63	120.6	54.3	174.8	4.98					
71	CYS	9.48	125.9	51.5	170.6	-					
72	ASP	6.89	123.1	50.5	172.8	4.62					
73	LYS	8.26	121.9	54.9	173.4	2.61	8.23	121.6	54.9	173.4	2.52
74	SER	8.22	114.4	56.6	172.5	4.43					
75	ARG	7.51	122.9	52.1	172.6	4.66					
76	PRO		-	61.9	173.2	-					
77	ALA	7.33	116.1	48.7	173.5	5.16					
78	TRP	8.60	118.7	52.8	170.7	5.47					
						3.15/					
79	GLY	8.00	105.4	44.4	169.9	5.34					
80	CYS	9.46	123.1	52.6	170.0	5.44	9.45	122.8	52.6	170.0	5.45
81	VAL	9.44	123.9	58.9	172.5	4.96	9.44	123.9	58.9	172.5	
82	GLU	9.34	127.2	52.1	172.2	4.81	9.30	127.3	52.1	172.2	4.79
83	CYS	8.86	120.4	51.4	170.2	5.87	8.88	120.6	51.4	170.2	5.84
84	CYS	9.17	118.6	52.3	171.4	5.26					
85	ASP	8.49	119.0	51.2	173.0	5.00					
86	THR	7.59	111.0	57.4	170.6	4.71			57.3	170.5	
87	ASP	8.09	118.4	53.1	175.1	-	8.11	118.7	53.0	175.1	
88	ARG	9.34	113.3	54.3	173.2	3.07	9.23	113.1	54.3	173.4	

Table S8. Pairwise similarity calculated for TFPs from *A. rubens*, identified by BLAST search and previously identified TFPs from *A. rubens* (Lystar1-4) [S3].

	XP_033628517.1 (LyAr1)	XP_033635706.1 (LyAr2)	XP_033638913.1 (LyAr3)	XP_033640057.1 (LyAr4)	XP_033644501.1 (LyAr5)	XP_033645811.1 (LYSTAR1)	XP_033638895.1 (LYSTAR2)	XP_033638835.1 (LYSTAR3)	XP_033634394.1 (LYSTAR4)
XP_033628517.1 (LyAr1)	100.0%	45.6%	42.7%	47.3%	45.7%	41.5%	40.2%	42.6%	43.6%
XP_033635706.1 (LyAr2)	-	100.0%	43.0%	51.7%	45.1%	53.6%	49.5%	43.3%	44.4%
XP_033638913.1 (LyAr3)	-	-	100.0%	38.7%	44.1%	47.3%	47.8%	55.8%	53.8%
XP_033640057.1 (LyAr4)	-	-	-	100.0%	41.0%	43.0%	42.6%	41.8%	46.4%
XP_033644501.1 (LyAr5)	-	-	-	-	100.0%	47.0%	47.5%	46.3%	46.5%
XP_033645811.1 (LYSTAR1)	-	-	-	-	-	100.0%	47.1%	40.7%	44.8%
XP_033638895.1 (LYSTAR2)	-	-	-	-	-	-	100.0%	46.7%	45.9%
XP_033638835.1 (LYSTAR3)	-	-	-	-	-	-	-	100.0%	45.1%
XP_033634394.1 (LYSTAR4)	-	-	-	-	-	-	-	-	100.0%

Table S9. Primers used for qPCR experiments.

Gene	Forward primer	Reverse primer	Amplicon length
<i>β-ACTIN</i>	AAC CTT CTT GCA GCT CCT CCG	TCT GAC CCA TAC CCA CCA TCA	199
<i>CHRNA4</i>	TCG TCC TCT ACA ACA AGT GAG	GGT CCA GGA GCC GAA TTT CA	93
<i>CHRNA7</i>	AAT GAC TCG CAA CCA CTC ACC	TCT GTC CAA GAC ATT TGC AGC C	116
<i>ACEE</i>	GGG GAG AAA GCG ATT CCA GAA G	ATC CCT CGC TGA ACT ACA CCA C	230
<i>SYN1</i>	TGC CAA TGG TGG ATT CTC CG	CAG CCC AAT GAC CAA ACT GC	148
<i>SYP</i>	TGC CAA CAA GAC GGA GAG TG	GAA TTC AGC CGA CGA GGA GT	151
<i>DLG4</i>	CTA CCA AGA TGA AGA CAC GCC	TCT GTT CCA TTC ACC TGC AAC	184

References

- [1] Teufel, F.; Almagro Armenteros, J.J.; Johansen, A.R.; Gíslason, M.H.; Pihl, S.I.; Tsirigos, K.D.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 6.0 Predicts All Five Types of Signal Peptides Using Protein Language Models. *Nat Biotechnol* 2022, 40, 1023–1025, doi:10.1038/s41587-021-01156-3.
- [2]. Pierleoni, A.; Martelli, P.L.; Casadio, R. PredGPI: A GPI-Anchor Predictor. *BMC Bioinformatics* **2008**, 9, 392, doi:10.1186/1471-2105-9-392.
- [3] Shabelnikov, S.V.; Bobkov, D.E.; Sharlaimova, N.S.; Petukhova, O.A. Injury Affects Coelomic Fluid Proteome of the Common Starfish, *Asterias Rubens*. *J Exp Biol* **2019**, 222, jeb198556, doi:10.1242/jeb.198556.