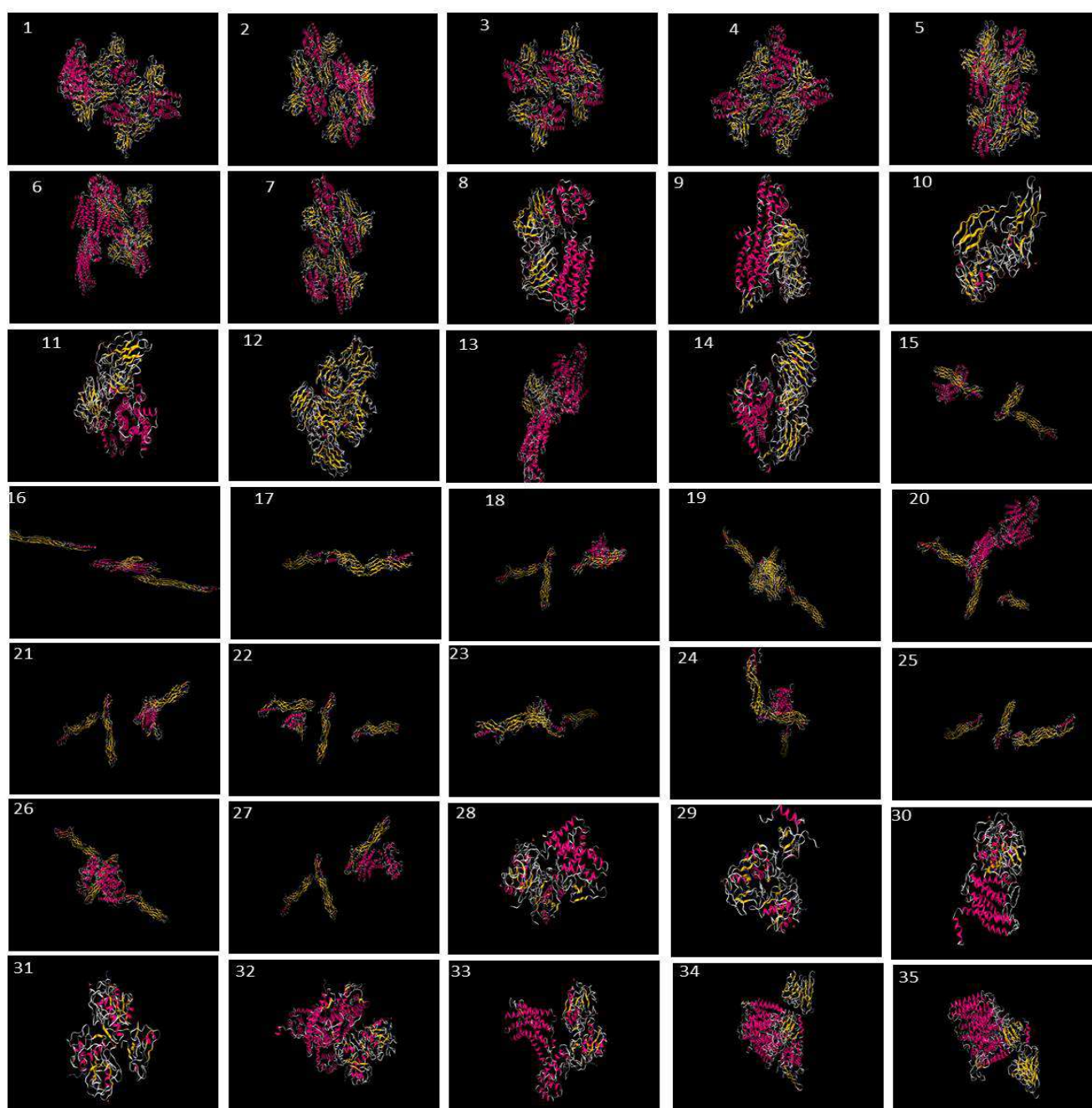
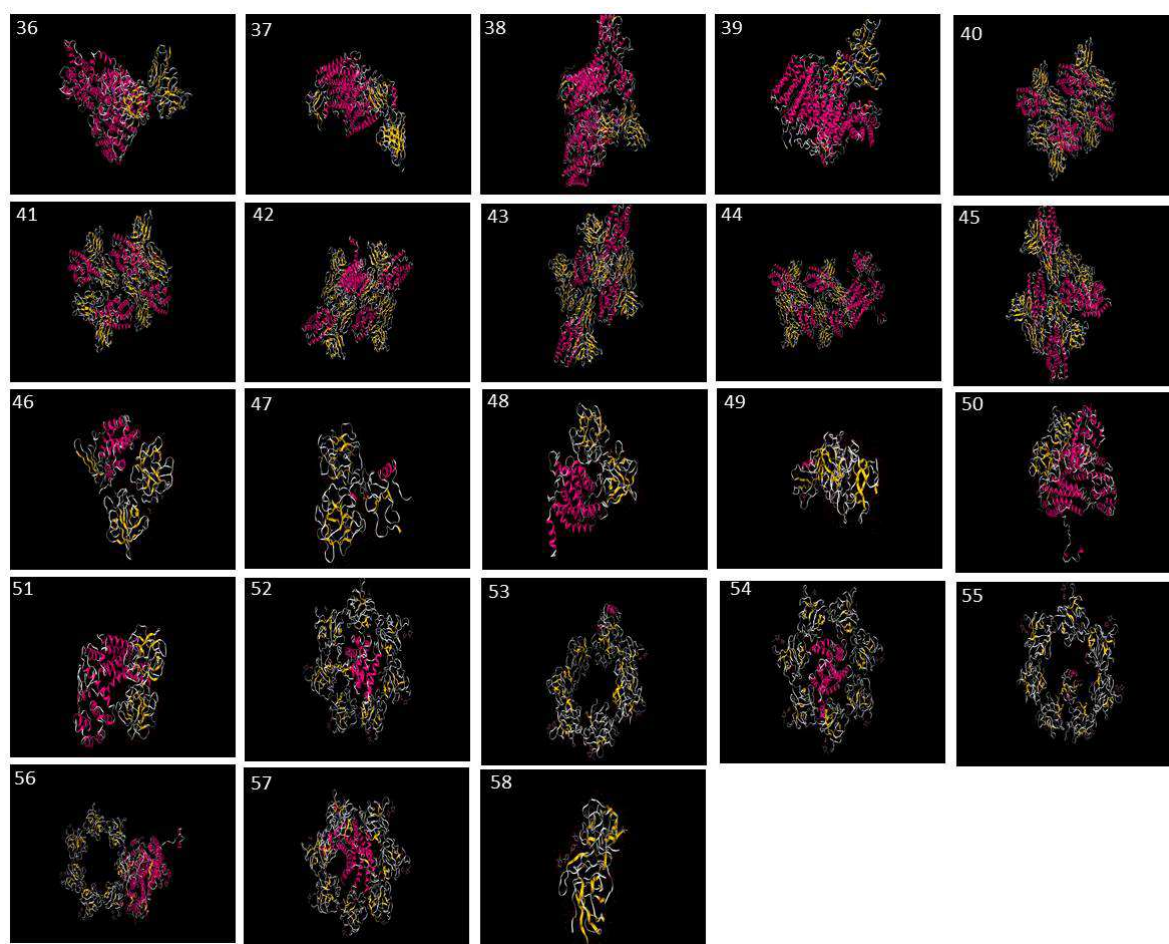


In silico Studies of Parasporin Proteins: Structural and Functional Insights and Proposed Cancer Cell Killing Mechanism for Parasporin 5 and 6

Nasima Aktar^{a,b}, Muhammad Manjurul Karim^a, Shakila Nargis Khan^a, Mustafizur Rahman^c, Anowara Begum^a, Md. Mozammel Hoq^{a*}

Supplementary Information





S. Figure 1: Ligand-receptor domain binding orientation of different parasporin protein. Endotoxin_N (4ARX) domain of class 1 and 6 acts as ligand to bind with 2RH1, 5GLI, 1EDN, 3RZE, 3TYF, 4DAJ and 4XNV receptor molecule of HeLa cells (1-7). Moreover, class 1 interacts with 1NKG domain (8-14); Class 5 interacts with HeLa cells through ETX-MTX2 (1UYJ) ligand (15-21); Class 2 parasporin interacts with 2MPM, 2LNL, 3ODU, 4JL7, 4NY9 and 5X33 receptor of HL-60 cells through ETX-MTX2 (1UYJ) domain (22-27); Aerolysin (2PRE) domain (28-33) and Duf916 (1QLE) domain (34-39); 4ARX, 4ION and 4OWL ligand of Class 3 interacts with 2MPM, 2LNL, 3ODU, 4JL7, 4NY9 and 5X33 receptor of HL-60 cells (40-57); Class 4 ETX-MTX2 ligand molecule (1UYJ) and binds with 1XXZ receptor of CACO-2 cells (58).

Supplementary Tables

S. Table 1: Physico-chemical parameters of different parasporin proteins

PS proteins	No. of A.A	Cellular localization	Certainty	Solubility	PI Site	Potential PI Site	Score (P-value)	pI	Aliphatic Index (AI)	Instability Index	GRAVY	Signal peptide (D-value)	Transmembrane helices	TMpred Score
PS1Aa1	723	Bacterial cytoplasm	0.326	Soluble	None	708	-44.49 (1.766537e-01)	5.73	79.88	39.23 Stable	-0.296	0.167	149-168 (I to O)	1343
													377-397 (I to O)	1671
													405-423 (I to O)	334
													499-517 (I to O)	811
													616-633 (I to O)	530
													647-666 (I to O)	295
													141-161 (O to I)	1098
													377-397 (O to I)	1119
													403-423 (O to I)	98
													500-519 (O to I)	914
616-633 (O to I)	421													
PS1Aa2	742	Bacterial membrane	0.113	Soluble	None	725	-23.16 (2.990767e-02)	5.35	79.42	39.55 Stable	-0.294	0.167	168-187 (I to O)	1411
													396-416 (I to O)	1671
													422-441 (I to O)	328
													518-536 (I to O)	763
													635-653 (I to O)	54
													662-685 (I to O)	295
													157-183 (O to I)	1216
													396-416 (O to I)	1119
													422-442 (O to I)	59
519-538 (O to I)	912													
PS1Aa3	723	Bacterial cytoplasm	0.326	Soluble	None	708	-44.49 (1.766537e-01)	5.73	79.88	39.23 Stable	-0.296	0.167	149-168 (I to O)	1343
													377-397 (I to O)	1671
													403-423 (I to O)	334
													499-517 (I to O)	811
													616-633 (I to O)	530
													643-666 (I to O)	295
													141-164 (O to I)	1098
													377-397 (O to I)	1119
													403-423 (O to I)	98
													500-519 (O to I)	914
616-633 (O to I)	421													

PS1Aa4	723	Bacterial cytoplasm	0.326	Soluble	None	708	-44.75 (1.799218e-01)	5.73	79.34	38.95 Stable	-0.301	0.167	149-168 (I to O)	1343
													377-397 (I to O)	1671
													403-423 (I to O)	334
													499-517 (I to O)	811
													616-633 (I to O)	530
													643-666 (I to O)	295
													141-164 (O to I)	1098
													377-397 (O to I)	1119
													403-423 (O to I)	98
													500-519 (O to I)	914
616-633 (O to I)	421													
PS1Aa5	723	Bacterial cytoplasm	0.326	Soluble	None	708	-44.49 (1.766537e-01)	5.79	79.75	39.61 Stable	-0.298	0.167	149-168 (I to O)	1343
													377-397 (I to O)	1671
													403-423 (I to O)	334
													499-517 (I to O)	811
													616-633 (I to O)	530
													643-666 (I to O)	295
													141-164 (O to I)	1098
													377-397 (O to I)	1119
													403-423 (O to I)	98
													500-519 (O to I)	914
616-633 (O to I)	421													
PS1Aa6	754	Bacterial cytoplasm	0.421	Soluble	None	737	-43.21 (1.614720e-01)	6.01	81.25	37.81 Stable	-0.290	0.154	180-199 (I to O)	1334
													408-428 (I to O)	1671
													434-453 (I to O)	328
													530-548 (I to O)	726
													647-665 (I to O)	16
													674-697 (I to O)	295
													180-198 (O to I)	1115
													408-428 (O to I)	1119
													434-454 (O to I)	59
													531-550 (O to I)	889
PS1Ab1	726	Bacterial cytoplasm	0.281	Soluble	None	701	-34.29 (8.129167e-02)	5.96	81.03	39.65 Stable	-0.324	0.143	180-198 (I to O)	1334
													377-397 (I to O)	1671
													403-423 (I to O)	334
													499-517 (I to O)	811
													149-167 (O to I)	1115
													377-397 (O to I)	1119
													500-519 (O to I)	914

													688-706 (O to I)	171
PS1Ab2	726	Bacterial cytoplasm	0.294	Soluble	None	701	-34.29 (8.129167e-02)	5.77	81.71	39.87 Stable	-0.324	0.141	180-198 (I to O)	1334
													377-397 (I to O)	1671
													403-423 (I to O)	334
													499-517 (I to O)	811
													149-167 (O to I)	1115
													377-397 (O to I)	1119
													403-423 (O to I)	98
													500-519 (O to I)	914
PS1Ac1	777	Bacterial membrane	0.111	Soluble	None	762	-38.00 (1.094560e-01)	5.83	82.11	38.95 Stable	-0.323	0.166	168-187 (I to O)	1347
													396-416 (I to O)	1671
													422-442 (I to O)	334
													518-536 (I to O)	811
													759-777 (I to O)	183
													160-183 (O to I)	1231
													396-416 (O to I)	1119
													422-442 (O to I)	98
													519-538 (O to I)	914
													746-771 (O to I)	237
761-777 (O to I)	234													
PS1Ac2	777	Bacterial membrane	0.111	Soluble	None	762	-38.00 (1.094560e-01)	5.75	81.98	39.42 Stable	-0.327	0.178	168-187 (I to O)	1347
													396-416 (I to O)	1671
													422-442 (I to O)	334
													518-536 (I to O)	811
													759-777 (I to O)	183
													160-183 (O to I)	1231
													396-416 (O to I)	1119
													422-442 (O to I)	98
													519-538 (O to I)	914
													746-771 (O to I)	237
761-777 (O to I)	234													
PS1Ad1	758	Bacterial membrane	0.113	Soluble	None	741	-29.34 (5.320892e-02)	5.23	82.64	42.31 Unstable	-0.279	0.146	180-199 (I to O)	1411
													408-428 (I to O)	1671
													434-454 (I to O)	334
													530-548 (I to O)	726
													726-749 (I to O)	81
													169-195 (O to I)	1216
													408-428 (O to I)	1119
													434-454 (O to I)	98

													531-550 (O to I)	889
													727-749 (O to I)	442
PS2Aa1	338	Bacterial cytoplasm	0.567	Soluble	None	316	-80.18 (8.001545e-01)	5.35	63.11	34.26 Stable	-0.433	0.161	146-163 (I to O)	356
													145-170 (O to I)	282
PS2Aa2	338	Bacterial cytoplasm	0.531	Soluble	None	316	-80.03 (7.981347e-01)	5.21	62.54	34.09 Stable	-0.441	0.164	146-163 (I to O)	356
													145-170 (O to I)	282
PS2Ab1	304	Bacterial cytoplasm	0.320	Soluble	None	287	-34.04 (7.962966e-02)	5.12	75	39.99 Stable	-0.258	0.126	254-276 (I to O)	109
PS3Aa1	825	Bacterial membrane	0.183	Soluble	None	802	-66.14 (5.584945e-01)	6.18	76.22	30.98 Stable	-0.496	0.183	75-99 (I to O)	1655
													303-323 (I to O)	366
													448-471 (I to O)	157
													685-719 (I to O)	157
													792-814 (I to O)	119
													79-99 (O to I)	1951
													451-471 (O to I)	132
													564-579 (O to I)	126
													685-711(O to I)	101
													795-812 (O to I)	626
PS3Ab1	829	Bacterial membrane	0.183	Soluble	None	806	-66.14 (5.584945e-01)	6.19	75.16	31.41 Stable	-0.508	0.167	75-99 (I to O)	1655
													303-323 (I to O)	366
													448-466 (I to O)	158
													690-723 (I to O)	95
													796-818 (I to O)	119
													79-99 (O to I)	1951
													451-471 (O to I)	114
													689-715 (O to I)	220
													799-816 (O to I)	626
PS4Aa1	275	Bacterial cytoplasm	0.286	Soluble	None	261	-40.05 (1.280721e-01)	6.09	81.16	29.03 Stable	-0.171	0.201	17-37 (I to O)	204
													169-197 (I to O)	120
													180-202 (O to I)	475
PS5Aa1	305	Bacterial cytoplasm	0.602	Soluble	None	276	-59.46 (4.238438e-01)	5.99	72.85	35.57 Stable	-0.439	0.126	158-176 (I to O)	392
													158-176 (O to I)	468
PS6Aa1	753	Bacterial Cytoplasm	0.450	Soluble	None	736	-52.36 (2.919484e-01)	5.75	94.79	33.39 Stable	-0.226	0.205	96-116 (I to O)	240
													133-151 (I to O)	1453
													258-279 (I to O)	887
													478-497 (I to O)	116
													96-116 (O to I)	542
													142-167 (O to I)	1571

PS5Aa1	0.61	27	27	27	27	2 – 298	0.90	90.5	8.0	0.2	1.3
PS6Aa1	0.45	15	15	15	15	138 – 746	0.72	74.6	12.5	1.5	11.4

S. Table 3: Ramachandran plot analysis of the predicted structures of parasporin proteins.

PS protein	Ramachandran plot			
	Most favoured regions [A,B,L]	Additional allowed regions [a,b,l,p]	Generously allowed regions [~a,~b,~l,~p]	Disallowed regions
PS1Aa1	77.2	10.4	1.0	11.3
PS1Aa2	77.2	9.6	0.7	12.5
PS1Aa3	77.5	10.2	1.0	11.3
PS1Aa4	78.0	9.5	1.2	11.3
PS1Aa5	77.8	9.8	1.0	11.3
PS1Aa6	76.7	9.0	0.8	13.4
PS1Ab1	78.4	8.9	1.5	11.3
PS1Ab2	59.0	9.4	1.2	30.5
PS1Ac1	61.6	7.4	1.1	29.8
PS1Ac2	60.9	8.2	1.2	29.7
PS1Ad1	76.8	8.8	0.9	13.5
PS2Aa1	80.3	4.1	0.2	15.4
PS2Aa2	90.6	4.7	0.0	4.7
PS2Ab1	85.3	4.0	0.6	10.1
PS3Aa1	76.4	8.5	0.9	14.2
PS3Ab1	75.8	9.3	1.0	13.9
PS4Aa1	90.6	4.7	0.0	4.7
PS5Aa1	90.5	8.0	0.2	1.3
PS6Aa1	74.6	12.5	1.5	11.4

S. Table 4: Domain Information of parasporin proteins

Protein (PS)	Interpro						Pfam					
	Domain	IPR	Location	Database Used	E-value	Biological Function	Family	Domain	Clan	Location	Database	E-value

PS1Aa1	Delta endotoxin, N-terminal	IPR005639	135 – 394	GENE3 D	9.7E-52	GO:0009405 pathogenesis	Endotoxin_N	delta endotoxin, N-terminal domain	n/a	147–348	Pfam	5.2e-13
			168 – 332	Pfam	5.2E-13							
			85 – 336 & 367 – 392	SUPER FAMILY	5.36E-49 & 5.36E-49							
PS1Aa1	Galactose-binding domain-like	IPR008979	625 – 722	GENE3 D	2.9E-8	GO:0009405 pathogenesis	Endotoxin_C	delta endotoxin	CL0202	588–723	Pfam	3.2e-08
			581 – 722	SUPER FAMILY	2.92E-14							
	Delta endotoxin, C-terminal	IPR005638	590 – 722	Pfam	3.2E-8							
PS1Aa2	Delta endotoxin, N-terminal	IPR005639	155 – 413	GENE3 D	2.4E-51	GO:0009405 pathogenesis	Endotoxin_N	delta endotoxin, N-terminal domain	n/a	165–412	Pfam	8.1e-13
			189 – 352	Pfam	8.0E-13							
			104 – 355 & 386 – 411	SUPER FAMILY	6.15E-50 & 6.15E-50							
PS1Aa2	Galactose-binding domain-like	IPR008979	644 – 740	GENE3 D	1.6E-7	GO:0009405 pathogenesis	Endotoxin_C	delta endotoxin	CL0202	607–742	Pfam	9.2e-09
			600 – 741	SUPER FAMILY	9.72E-15							

	Delta endotoxin, C-terminal	IPR005638	609 – 741	Pfam	9.2E-9							
PS1 Aa3	Delta endotoxin, N-terminal	IPR005639	135 – 394	GENE3 D	9.7E-52	GO:0009405 pathogenesis	Endotoxin_N	delta endotoxin, N-terminal domain	n/a	147–348	Pfam	5.2e-13
			168 – 332	Pfam	5.2E-13							
			85 – 336 & 367 – 392	SUPER FAMIL Y	5.36E-49 & 5.36E-49							
	Galactose-binding domain-like	IPR008979	625 – 722	GENE3 D	2.9E-8		Endotoxin_C	delta endotoxin	CL0202	588–723	Pfam	3.2e-08
			581 – 722	SUPER FAMIL Y	2.92E-14							
	Delta endotoxin, C-terminal	IPR005638	590 – 722	Pfam	3.2E-8							
PS1Aa4	Delta endotoxin, N-terminal	IPR005639	135 – 394	GENE3 D	9.7E-52	GO:0009405 pathogenesis	Endotoxin_N	delta endotoxin, N-terminal domain	n/a	147–348	Pfam	5.2e-13
			168 – 332	Pfam	5.2E-13							
			85 – 336 & 367 – 392	SUPER FAMIL Y	5.36E-49 & 5.36E-49							
	Galactose-binding	IPR008979	625 – 722	GENE3 D	3.3E-8		Endotoxin_C	delta endotoxin	CL0202	588–723	Pfam	3.2e-08

	domain-like	9	581 – 722	SUPER FAMILY	3.16E-14			n										
	Delta endotoxin, C-terminal	IPR005638	590 – 722	Pfam	2.7E-8													
PS1Aa5	Delta endotoxin, N-terminal	IPR005639	135 – 394	GENE3 D	1.0E-51	GO:0009405 pathogenesis	Endotoxin_N	delta endotoxin, N-terminal domain	n/a	147–350	Pfam	5e-13						
			168 – 332	Pfam	5.0E-13													
			85 – 336 & 367 – 392	SUPER FAMILY	5.76E-49 & 5.76E-49													
	Galactose-binding domain-like	IPR008979	625 – 722	GENE3 D	2.9E-8								Endotoxin_C	delta endotoxin	CL0202	588–723	Pfam	3.2e-08
	581 – 722	SUPER FAMILY	2.92E-14															
	Delta endotoxin, C-terminal	IPR005638	590 – 722	Pfam	3.2E-8													
PS1Aa6	Delta endotoxin, N-terminal	IPR005639	166 – 425	GENE3 D	8.3E-52	GO:0009405 pathogenesis	Endotoxin_N	delta endotoxin, N-terminal domain	n/a	178–424	Pfam	7.6e-13						
			197 – 364	Pfam	7.5E-13													
			116 – 367 & 398 – 423	SUPER FAMILY	1.16E-48 & 1.16E-48													

	Galactose-binding domain-like	IPR008979	655 – 752	GENE3D	4.2E-8		Endotoxin_C	delta endotoxin	CL0202	619–754	Pfam	3.5e-09
			612 – 753	SUPERFAMILY	5.22E-15							
	Delta endotoxin, C-terminal	IPR005638	621 – 753	Pfam	3.5E-9							
PS1Ab1	Delta endotoxin, N-terminal	IPR005639	135 – 394	GENE3D	6.5E-52	GO:0009405 pathogenesis	Endotoxin_N	delta endotoxin, N-terminal domain	n/a	147–393	Pfam	1.4e-13
			166 – 333	Pfam	1.4E-13							
			85 – 336 & 367 – 392	SUPERFAMILY	8.24E-49 & 8.24E-49							
PS1Ab1	Galactose-binding domain-like	IPR008979	579 – 725	GENE3D	2.5E-7		Endotoxin_C	delta endotoxin	CL0202	588–726	Pfam	4.4e-07
			581 – 725	SUPERFAMILY	2.02E-8							
	Delta endotoxin, C-terminal	IPR005638	590 – 725	Pfam	4.4E-7							
PS1Ab2	Delta endotoxin, N-terminal	IPR005639	135 – 394	GENE3D	9.1E-52	GO:0009405 pathogenesis	Endotoxin_N	delta endotoxin, N-terminal	n/a	147–393	Pfam	2.4e-13
			166 – 333	Pfam	2.4E-13							

			85 – 336 & 367 – 392	SUPER FAMIL Y	6.02E-49 & 6.02E- 49			domain				
	Galactose- binding domain- like	IPR0 0897 9	579 – 725	GENE3 D	2.6E-8		Endoto xin_C	delta endotoxi n	CL0 202	588– 726	Pfam	3.8e-08
			581 – 725	SUPER FAMIL Y	1.59E-8							
	Delta endotoxin, C-terminal	IPR0 0563 8	590 – 725	Pfam	3.8E-8							
PS1Ac1	Delta endotoxin, N-terminal	IPR0 0563 9	155 – 413	GENE3 D	2.5E-51	GO:00094 05 pathogene sis	Endoto xin_N	delta endotoxi n, N- terminal domain	n/a	147– 393	Pfam	2.4e-13
			189 – 352	Pfam	5.7E-14							
			104 – 355 & 386 – 411	SUPER FAMIL Y	4.97E-49 & 4.97E- 49							
	Galactose- binding domain- like	IPR0 0897 9	683 – 776	GENE3 D	3.2E-4							
	Delta endotoxin, C-terminal	IPR0 0563 8	683 – 776	Pfam	1.8E-6							
PS1Ac2	Delta endotoxin, N-terminal		155 – 413	GENE3 D	2.5E-51	GO:00094 05 pathogene sis	Endoto xin_N	delta endotoxi n, N- terminal	n/a	167– 367	Pfam	9.3e-14
			189 – 351	Pfam	9.2E-14							

			104 – 355 & 386 – 411	SUPER FAMIL Y	4.06E-49 & 4.06E-49			domain				
	Galactose-binding domain-like		683 – 776	GENE3 D	3.2E-4		Endotoxin_C	delta endotoxin	CL0 202	679– 777	Pfam	1.8e-06
	Delta endotoxin, C-terminal		683 – 776	GENE3 D	1.8E-6							
PS1Ad1	Delta endotoxin, N-terminal	IPR0 0563 9	167 – 425	GENE3 D	2.4E-51	GO:00094 05 pathogenesis	Endotoxin_N	delta endotoxin, N-terminal domain	n/a	179– 379	Pfam	1.2e-12
			201 – 363	Pfam	1.2E-12						Pfam	
			116 – 367 & 398 – 423	SUPER FAMIL Y	6.28E-50 & 6.28E-50						Pfam	
	Galactose-binding domain-like	IPR0 0897 9	611 – 757	GENE3 D	2.9E-9		Endotoxin_C	delta endotoxin	CL0 202	619– 758	Pfam	1.3e-08
	612 – 757		SUPER FAMIL Y	3.4E-9								
	Delta endotoxin, C-terminal	IPR0 0563 8	621 – 757	Pfam	1.3E-8							
PS2Aa1	Aerolysin-like toxin, beta complex domain	IPR0	130 – 267	GENE3 D	1.1E-17	-	ETX_MTX2	Clostridium epsilon toxin ETX/Bacillus	CL0 345	74– 304	Pfam	3e-07

								mosquit ocidal toxin MTX2				
PS2Aa2	Aerolysin- like toxin, beta complex domain	IPR0 2330 7	130 – 263	GENE3 D	8.8E-18	-	ETX_M TX2	Clostridi um epsilon toxin ETX/Bac illus mosquit ocidal toxin MTX2	CL0	90– 303	Pfam	8.3e-07
PS2Ab1	Aerolysin- like toxin, beta complex domain	IPR0 2330 7	106 – 243	GENE3 D	2.1E-16	-	ETX_M TX2	Clostridi um epsilon toxin ETX/Bac illus mosquit ocidal toxin MTX2	CL0	62– 207	Pfam	1.1e-05
PS3Aa1	Delta endotoxin, N-terminal	IPR0	34 – 257 & 293 – 319	GENE3 D	2.3E-83 & 2.3E- 83	Biological Process: GO:00069 52 defense response GO:00094 05 pathogene sis Molecular	Endoto xin_N	delta endotoxi n, N- terminal domain	N/a	81– 289	Pfam	7.1e-22
			100 – 279	Pfam	7.1E-22							
			34 – 257 & 293 – 319	SUPER FAMIL Y	1.27E-83 & 1.27E- 83							

	Delta endotoxin, central domain	IPR0	322 – 530	SUPER FAMILY	1.46E-33	Function: GO:0005102 receptor binding	Endotoxin_M	delta endotoxin	N/a	327–530	Pfam	3e-11
			333 – 421	GENE3 D	1.8E-8							
	Delta endotoxin, central domain, subgroup 1	IPR0	327 – 530	Pfam	3.0E-11		Endotoxin_C	delta endotoxin	CL0202	540–675	Pfam	1.5e-37
	Galactose-binding domain-like	IPR0	513 – 678	GENE3 D	8.9E-50		RicinB_lectin_2	Ricin-type beta-trefoil lectin domain-like	CL0066	716–812	Pfam	1.1e-14
			532 – 675	SUPER FAMILY	3.68E-35							
	Delta endotoxin, C-terminal	IPR0	540 – 674	Pfam	1.5E-37							
	Ricin B, lectin domain	IPR0	684 – 824	SUPER FAMILY	2.7E-32							
			714 – 825	PROSITE profiles	16.435							
			695 – 825	SMART	2.5E-14							
			727 – 824	Pfam	1.8E-18							

PS3Ab1	Delta endotoxin, N-terminal	IPR0	34 – 257 & 293 – 319	GENE3 D	1.9E-92 & 1.9E-92	Biological Process: GO:0006952 defense response GO:0009405 pathogenesis Molecular Function: GO:0005102 receptor binding	Endotoxin_N	delta endotoxin, N-terminal domain	N/a	81–289	Pfam	7.1e-22	
			100 – 279	Pfam	7.1E-22								
			34 – 257 & 293 – 319	SUPER FAMILY	6.67E-83 & 6.67E-83								
		Delta endotoxin, central domain	IPR0	322 – 530	SUPER FAMILY	1.57E-33		Endotoxin_M	delta endotoxin	N/a	327–530	Pfam	9.5e-11
		333 – 421		GENE3 D	2.0E-8								
		Delta endotoxin, central domain, subgroup 1	IPR0	327 – 530	Pfam	9.5E-11		Endotoxin_C	delta endotoxin	CL0202	540–681	Pfam	7.2e-37
		Galactose-binding domain-like	IPR0	513 – 685	GENE3 D	1.2E-51		RicinB_lectin_2	Ricin-type beta-trefoil lectin domain-like	CL0066	715–816	Pfam	1.6e-15
				532 – 681	SUPER FAMILY	4.89E-37							
		Delta endotoxin, C-terminal	IPR0	540 – 681	Pfam	7.2E-37							

	Ricin B, lectin domain	IPRO	689 – 828	SUPER FAMILY	9.31E-32							
			718 – 829	PROSITE profiles	16.181							
			699 – 829	SMART	7.9E-14							
			731 – 828	Pfam	3.9E-19							
PS4Aa1	Aerolysin-like toxin, beta complex domain	IPRO	39 – 186	GENE3 D	1.2E-24	-	ETX_MTX2	Clostridium epsilon toxin ETX/Bacillus mosquitocidal toxin MTX2	CL0 345	33– 252	Pfam	1.4e-22
PS5Aa1	Aerolysin-like toxin, beta complex domain	IPRO	37 – 196	GENE3 D	2.2E-25		ETX_MTX2	Clostridium epsilon toxin ETX/Bacillus mosquitocidal toxin MTX2	CL0	13– 276	Pfam	5.4e-27

PS6Aa1	Delta endotoxin, N-terminal	IPR0	128 – 355	Gene3 D	3.3E-33	GO:0009405 pathogenesis	Endotoxin_N	delta endotoxin, N-terminal domain	N/a	131–356	Pfam	4.2e-19
--------	-----------------------------	------	-----------	---------	---------	-------------------------	-------------	------------------------------------	-----	---------	------	---------

S. Table 5: Parasporin protein and their probable ligand identification

PS Proteins	Domain	Ligand (PDB)	Ligand present (PDB)	Ligand present (RaptorX)
PS1Aa1	Endotoxin	4ARX	GOL, 13D	BR, NGA, 13D, SO ₄
PS1Aa2	Endotoxin	4ARX	GOL, 13D	BR, NGA, 13D, SO ₄
PS1Aa3	Endotoxin	4ARX	GOL, 13D	BR, NGA, 13D, SO ₄ , DOC, ZN, C2E, CL
PS1Aa4	Endotoxin	4ARX	GOL, 13D	BR, NGA, 13D, SO ₄ , CA
PS1Aa5	Endotoxin	4ARX	GOL, 13D	BR, NGA, 13D, SO ₄ , GUN, ZEA, NA, NEU
PS1Aa6	Endotoxin	4ARX	GOL, 13D	BR, NGA, 13D, SO ₄ ,
PS1Ab1	Endotoxin	4ARX	GOL, 13D	BR, NGA, 13D, SO ₄ ,
PS1Ab2	Endotoxin	4ARX	GOL, 13D	BR, NGA, 13D, SO ₄ ,
PS1Ac1	Endotoxin CBM-like	4ARX 1NKG	GOL, 13D SO ₄ , CA	BR, NGA, 13D, SO ₄ ,
PS1Ac2	Endotoxin CBM-like	4ARX 1NKG	GOL, 13D SO ₄ , CA	BR, NGA, 13D, SO ₄ ,
PS1Ad1	Endotoxin	4ARX	GOL, 13D	BR, NGA, SO ₄ ,
PS2Aa1	ETX_MTX2 Aerolysin DUF916	1UYJ 2PRE 1QLE	U1 SO ₄ , E64 HEA, MN, PC1, CA, CUA, CU	PO ₄ , BOG, EDO, U1, GOL, CL
PS2Aa2	ETX_MTX2 Aerolysin DUF916	1UYJ 2PRE 1QLE	U1 SO ₄ , E64 HEA, MN, PC1, CA, CUA, CU	PO ₄ , BOG, EDO, U1, GOL, CL

PS2Ab1	ETX_MTX2 Aerolysin DUF916	1UYJ 2PRE 1QLE	U1 SO4, E64 HEA, MN, PC1, CA, CUA, CU	PO ₄ , BOG, EDO, U1, GOL, CL
PS3Aa1	Endotoxin_N <u>RicinB lectin 2</u> <u>Ricin B lectin</u>	4ARX 4ION 4OWL	GOL, 13D GOL GOL, GAL, NAG	BR, NGA, 13D, SO ₄ , GOL, GAL
PS3Ab1	Endotoxin_N <u>RicinB lectin 2</u> <u>Ricin B lectin</u>	4ARX 4ION 4OWL	GOL, 13D GOL GOL, GAL, NAG	BR, NGA, 13D, SO ₄ , GOL, GAL
PS4Aa1	ETX_MTX2	1UYJ	U1	PO ₄ , BOG, EDO, LU, GOL
PS5Aa1	ETX_MTX2	1UYJ	U1	PO ₄ , BOG, EDO, U1, CL
PS6Aa1	Endotoxin_N	4ARX	GOL, 13D	BR, NGA, 13D, SO ₄

S. Table 6: Parasporin class showing higher toxicity cancer cell lines, their receptors with PDB ID and probable ligands with which they tend to bind/interact

Parasporin protein	Cell Line	Receptors	PDB ID of Receptors	Binding site for residue	Ligand with whom they Bind
PS 1, 5 and 6	Hela	1. Adrenergic β 2	2RH1	2RH1_A_BC4_2_414	CLR
				2RH1_A_AC2_6_402	SO4
				2RH1_A_BC2_5_412	CLR
				2RH1_A_AC8_11_408	CAU
				2RH1_A_AC9_4_409	BU1
				2RH1_A_BC1_3_411	ACM
				2RH1_A_AC3_4_403	SO4
				2RH1_A_BC5_2_415	PLM
				2RH1_A_BC3_2_413	CLR
				2RH1_A_AC5_6_405	SO4

				2RH1_A_BC6_3_416	12P
				2RH1_A_AC1_9_401	MAL
				2RH1_A_AC7_3_407	SO4
				2RH1_A_AC6_3_406	SO4
				2RH1_A_AC4_5_404	SO4
		2. Endothelin Type- B	5GLI	5GLI_A_AC1_5_1201	SO4
				5GLI_A_AC8_7_1208	OLC
				5GLI_A_AC7_4_1207	OLA
				5GLI_A_AC6_4_1206	OLC
				5GLI_A_AC2_2_1202	SO4
				5GLI_A_AC5_2_1205	SO4
				5GLI_A_AC4_5_1204	SO4
				5GLI_A_AC3_4_1203	SO4
		Endothelin-1	1EDN	N/A	
		3. Histamine H1	3RZE	3RZE_A_AC2_12_1201	D7V
				3RZE_A_AC1_9_1200	5EH
				3RZE_A_AC3_6_1202	PO4
				3RZE_A_AC6_2_1205	OLC
				3RZE_A_AC4_2_1203	PO4
				3RZE_A_AC5_4_1204	PO4
		4. Lysophospholipid	3TYF	3TYF_A_AC2_2_302	GOL
				3TYF_A_AC3_2_303	GOL
				3TYF_A_AC1_2_301	GOL
				3TYF_A_AC9_6_1000	GOL

		5. Muscarinic	4DAJ	4DAJ_A_AC1_12_2000	0HK
				4DAJ_A_AC5_2_2004	PO4
				4DAJ_A_AC3_3_2002	PO4
				4DAJ_A_AC4_1_2003	PO4
				4DAJ_A_AC2_2_2001	PO4
		6. Purinoceptor	4XNV	4XNV_A_AC9_4_1109	OLC
				4XNV_A_AD2_5_1111	OLC
				4XNV_A_AC1_10_1101	BUR
				4XNV_A_AC5_7_1105	Y01
				4XNV_A_AC4_5_1104	Y01
				4XNV_A_AD1_6_1110	OLC
				4XNV_A_AC8_3_1108	OLC
				4XNV_A_AC7_5_1107	OLC
				4XNV_A_AC2_4_1102	CLR
				4XNV_A_AD3_5_1112	OLC
				4XNV_A_AD4_6_1113	OLC
				4XNV_A_AC3_4_1103	Y01
				4XNV_A_AD6_4_1115	ZN
				PS 2, 3 & 4	HL-60
4NY9_A_AC1_9_501	2Q4				
Chemokine CCR3	2MPM	N/A			
Chemokine CXCR1	2LNL	N/A			
Chemokine CXCR2	4JL7	N/A			
Chemokine CXCR4	3ODU	N/A			

		Leukotriene LTB4	5X33	N/A	
PS 4	CACO-2	Somatostatin	1XXZ	N/A	