A simple approach for predicting protein-protein interactions

Mamoon Rashid, Sumathy Ramasamy, Gajendra P.S. Raghava [§]

Bioinformatics Centre, Institute of Microbial Technology, Sector-39A, Chandigarh, India

Supplementary Material

Table S1: Performance comparison of amino acid (AA) and dipeptide (DP) based SVM methods for PPI prediction

Method	**Threshold**	***E. coli***					***S. cerevisiae***					***H. pylori***				
		Sen	Spe	PPV	Acc	MCC	Sen	Spe	PPV	Acc	MCC	Sen	Spe	PPV	Acc	MCC
AA	1.0	56.5	100	100	96.8	0.74	26.3	97.1	90.0	61.7	0.33	32.4	98.6	95.7	65.5	0.41
	0.8	70.9	100	100	97.9	0.83	35.1	95.0	87.5	65.0	0.38	45.2	97.6	95.0	71.4	0.50
	0.6	81.1	100	100	98.6	0.89	44.2	91.2	83.3	67.7	0.40	55.9	95.4	92.4	75.7	0.59
	0.4	88.7	100	100	99.2	0.94	52.7	85.5	78.4	69.1	0.41	66.0	91.0	88.0	78.5	0.59
	0.2	93.0	100	100	99.5	0.96	60.9	78.6	74.0	69.8	0.40	75.3	86.8	85.1	81.1	0.63
	0.0	95.7	100	99.9	99.7	0.98	69.1	70.2	69.9	69.7	0.39	82.5	80.0	80.5	81.3	0.63
	-0.2	97.9	100	99.9	99.8	0.99	76.2	60.0	65.6	68.1	0.37	88.4	71.1	75.4	79.8	0.60
	-0.4	99.0	100	99.7	99.9	0.99	82.4	48.6	61.6	65.5	0.33	93.2	60.0	70.0	76.6	0.56
	-0.6	99.4	99.9	99.4	99.9	0.99	87.8	37.1	58.2	62.4	0.29	95.8	47.6	64.6	71.7	0.50
	-0.8	99.6	99.3	92.1	99.4	0.96	92.0	26.4	55.6	59.2	0.25	97.9	30.7	58.5	64.3	0.39
	-1.0	100	80.3	28.4	81.7	0.48	95.2	17.1	53.4	56.1	0.20	99.1	17.8	54.7	58.4	0.30
DP	1.0	79.0	100	100	98.5	0.88	28.2	98.5	57.8	63.3	0.38	30.0	99.5	98.2	64.7	0.41
	0.8	88.4	100	100	99.2	0.94	38.8	97.2	61.4	68.0	0.44	48.6	98.8	97.5	73.7	0.55
	0.6	93.2	100	100	99.5	0.96	48.9	94.8	65.0	71.8	0.49	63.0	97.1	95.6	80.0	0.64
	0.4	96.2	100	100	99.7	0.98	57.8	91.6	68.5	74.7	0.53	74.7	94.2	92.8	84.4	0.70
	0.2	98.1	100	100	99.9	0.99	66.0	86.0	71.7	76.0	0.53	82.4	91.0	90.2	86.7	0.74
	0.0	99.2	100	100	99.9	1.0	72.8	78.6	74.3	75.7	0.52	88.5	85.2	85.7	86.8	0.74
	-0.2	99.4	100	100	100	1.0	79.1	68.2	76.5	73.6	0.48	92.9	77.1	80.2	85.0	0.71
	-0.4	99.6	100	100	100	1.0	85.0	55.0	78.5	70.0	0.42	96.5	66.1	74.0	81.3	0.66
	-0.6	99.7	100	100	100	1.0	90.0	40.4	80.2	65.2	0.35	98.3	50.8	66.6	74.5	0.56
	-0.8	99.7	100	100	100	1.0	94.1	25.6	81.3	59.9	0.27	99.3	29.1	58.3	64.2	0.40
	-1.0	99.7	86.4	36.5	87.4	0.56	96.9	13.5	81.3	55.2	0.19	99.7	12.1	53.1	55.9	0.24

Sen, Spec, Acc and MCC are sensitivity, specificity, accuracy and Matthews Correlation Coefficient, respectively. AA and DP are amino acid and dipeptide composition, respectively.

Table S2: Performance comparison of SVM based methods for PPI prediction

Method	**Threshold**	***E. coli***					***S. cerevisiae***					***H. pylori***				
		Sen	Spe	PPV	Acc	MCC	Sen	Spe	PPV	Acc	MCC	Sen	Spe	PPV	Acc	MCC
SA	1.0	44.4	100	100	96.0	0.65	22.0	99.0	95.8	60.5	0.33	21.9	99.8	99.1	60.8	0.35
	0.8	59.5	100	100	97.1	0.76	31.7	97.9	93.7	64.8	0.39	38.1	99.5	98.6	68.8	0.47
	0.6	72.6	100	100	98.0	0.84	42.1	95.9	91.0	69.0	0.45	53.1	98.1	96.6	75.6	0.57
	0.4	83.2	100	100	98.8	0.91	52.2	92.2	87.0	72.2	0.48	65.8	96.0	94.3	80.9	0.65
	0.2	90.3	100	100	99.3	0.95	61.5	85.7	81.2	73.6	0.49	77.5	92.7	91.4	85.1	0.71
	0.0	94.5	100	100	99.6	0.97	70.8	76.6	75.1	73.7	0.47	85.3	86.8	86.6	86.0	0.72
	-0.2	96.7	100	100	99.8	0.98	78.9	63.6	68.4	71.2	0.43	91.9	78.7	81.2	85.3	0.71
	-0.4	98.2	100	100	99.9	0.99	86.1	46.8	61.8	66.5	0.36	96.2	65.4	73.6	80.8	0.65
	-0.6	99.2	100	99.7	99.9	0.99	91.9	30.1	56.8	61.0	0.28	98.8	39.9	62.2	69.3	0.48
	-0.8	100	99.7	96.6	99.7	0.98	95.6	16.6	53.4	56.1	0.20	99.9	15.4	54.1	57.6	0.28
	-1.0	100	58.4	15.8	61.4	0.30	98.0	7.8	51.5	52.9	0.14	100	4.7	51.2	52.3	0.15
PA	1.0	61.6	100	100	97.2	0.77	25.1	97.2	90.0	61.2	0.32	32.6	98.1	94.6	65.4	0.41
	0.8	75.6	100	100	98.2	0.86	34.5	94.9	87.1	64.7	0.37	47.0	97.1	94.1	72.0	0.51
	0.6	84.6	100	100	98.9	0.91	43.7	91.5	83.8	67.6	0.40	58.6	94.5	91.4	76.5	0.57
	0.4	91.0	100	100	99.3	0.95	52.7	86.5	79.6	69.6	0.42	67.9	91.2	88.5	79.5	0.61
	0.2	94.3	100	99.9	99.6	0.97	60.6	80.0	75.2	70.3	0.41	75.7	85.5	83.9	80.6	0.62
	0.0	96.5	100	99.9	99.7	0.98	68.0	72.2	71.0	70.1	0.40	82.5	78.5	79.3	80.5	0.61
	-0.2	98.1	100	99.9	99.9	0.99	74.7	62.9	66.8	68.8	0.38	88.6	70.3	74.9	79.5	0.60
	-0.4	99.3	100	99.8	99.9	0.99	81.5	51.4	62.6	66.4	0.34	93.3	59.9	69.9	76.6	0.56
	-0.6	99.5	100	99.4	99.9	0.99	87.1	39.7	59.1	63.4	0.30	96.1	48.5	65.1	72.3	0.51
	-0.8	99.5	99.5	93.4	99.5	0.96	91.5	27.9	55.9	59.7	0.25	97.6	34.8	59.9	66.2	0.42
	-1.0	100	84.6	33.7	85.7	0.53	95.0	17.6	53.5	56.3	0.20	98.8	20.7	55.5	59.7	0.31

Sen, Spe, Acc and MCC are sensitivity, specificity, accuracy and Matthews Correlation Coefficient, respectively. SA, split amino acid composition; PA, pseudo-amino acid composition.

Table S3: The list of top ten amino acid residue selected on the basis of ACB values computed from pair wise interaction data.

S. No	*E. coli*				*S. cerevisiae*				*H. pylori*			
	ACB+		**ACB-**		**ACB+**		**ACB-**		**ACB+**		**ACB-**	
	Value	Feature	Value	Feature	Value	Feature	Value	Feature	Value	Feature	Value	feature
1	0.052	V1	0.052	R1	0.064	Q2	0.024	V2	0.097	D1	0.028	F1
2	0.044	V2	0.054	E2	0.058	Q1	0.031	Y2	0.076	H1	0.031	G1
3	0.043	A2	0.059	W1	0.053	E1	0.039	H1	0.057	C1	0.040	I1
4	0.042	M1	0.079	R2	0.045	E2	0.040	H2	0.051	E1	0.044	P1
5	0.038	G2	0.081	D1	0.039	D1	0.051	F1	0.049	D2	0.051	W1
6	0.037	A1	0.096	Y1	0.032	N2	0.054	W1	0.042	N2	0.057	M1
7	0.036	P2	0.097	C2	0.031	N1	0.055	F2	0.040	E2	0.065	F2
8	0.034	L1	0.113	H1	0.027	D2	0.058	W2	0.031	T2	0.084	C2
9	0.034	M2	0.149	H2	0.008	T2	0.078	C2	0.026	N1	0.097	M2
10	0.031	I1	0.150	C1	0.005	M1	0.097	C1	0.022	K1	0.133	W2

ACB is average compositional bias; plus and minus sign of ACB shows dominancy of that feature towards interacting and non-interacting pairs, respectively. 1 and 2 refer to protein1 and protein2 of pair wise interaction data.

Table S4: The list of amino acid residues with ACB values computed from pair wise interaction data.

S. No	*E. coli*				*S. cerevisiae*				*H. pylori*			
	ACB+		**ACB-**		**ACB+**		**ACB-**		**ACB+**		**ACB-**	
	Value	Feature	Value	Feature	Value	Feature	Value	Feature	Value	Feature	Value	feature
1	0.048	V	0.008	S	0.061	Q	0.00	S	0.073	D	0.003	A
2	0.040	A	0.024	N	0.049	E	0.00	L	0.045	E	0.004	R
3	0.038	M	0.033	E	0.033	D	0.005	I	0.034	N	0.008	S
4	0.034	G	0.039	K	0.032	N	0.010	A	0.028	H	0.013	C
5	0.028	L	0.042	W	0.005	T	0.013	P	0.014	K	0.014	L
6	0.020	I	0.057	Y	0.004	K	0.017	G	0.006	T	0.016	G
7	0.020	P	0.060	D	0.001	M	0.018	R	0.006	Y	0.018	P
8	0.013	T	0.065	R			0.023	V	0.004	Q	0.018	V
9	0.010	F	0.124	C			0.027	Y			0.023	I
10	0.007	Q	0.131	H			0.039	H			0.047	F
11							0.053	F			0.077	M
12							0.056	W			0.092	W
13							0.087	C				

ACB is average compositional bias; plus and minus sign of ACB shows dominancy of that feature towards interacting and non-interacting pairs, respectively. Each value is the average of two values coming from similar amino acid of two proteins of a pair.

Table S5: The list of top 20 dipeptides selected on the basis of ACB values computed from pair wise interaction data.

S. No	*E. coli*				*S. cerevisiae*				*H. pylori*			
	ACB+		**ACB-**		**ACB+**		**ACB-**		**ACB+**		**ACB-**	
	Value	Feature	Value	Feature	Value	Feature	Value	Feature	Value	Feature	Value	feature
1	0.264	M1F1	0.290	H1Y1	0.152	Q2Q2	0.146	Y2H2	0.609	H1H1	0.262	W2R2
2	0.204	P1N1	0.296	K2H2	0.139	Q1Q1	0.146	C2P2	0.425	C1C1	0.272	W2Y2
3	0.190	F1F1	0.302	R2H2	0.125	N2N2	0.149	C2V2	0.369	G1C1	0.275	M1S1
4	0.182	M1M1	0.306	H2R2	0.113	T2Q2	0.150	P1W1	0.260	Y1T1	0.277	H2W2
5	0.182	F1Q1	0.307	C1S1	0.111	E1Q1	0.151	F2H2	0.260	D1D1	0.282	C2W2
6	0.177	I1F1	0.307	C1E1	0.111	M2E2	0.154	F1W1	0.258	C1R1	0.286	M2C2
7	0.177	F1M1	0.308	C2S2	0.109	Q1D1	0.158	F1F1	0.256	A1C1	0.304	P1M1
8	0.175	M1I1	0.310	C1Y1	0.106	M1E1	0.162	F1C1	0.253	P1R1	0.314	F1H1
9	0.167	P2Y2	0.311	N2C2	0.105	Q1T1	0.164	Y2C2	0.248	T1V1	0.343	P1C1
10	0.160	W1N1	0.315	W1M1	0.104	Q1E1	0.174	C2C2	0.247	G1H1	0.352	W1S1
11	0.157	V1W1	0.316	E2C2	0.104	N1N1	0.176	F2F2	0.209	D1I1	0.352	S1C1
12	0.157	W1G1	0.321	W1Y1	0.100	F2M2	0.184	W1G1	0.208	T2W2	0.368	W2P2
13	0.157	Q1E1	0.334	W1H1	0.098	E1L1	0.187	C1R1	0.204	C1W1	0.374	M1W1
14	0.155	V2M2	0.338	E1C1	0.096	D1Q1	0.205	C1P1	0.201	C1H1	0.421	I2W2
15	0.151	T1S1	0.340	H2C2	0.094	Q2H2	0.207	A1C1	0.199	C1G1	0.424	T1W1
16	0.151	S1A1	0.341	D1C1	0.092	Q2E2	0.212	P2C2	0.195	Q1C1	0.444	Y1W1
17	0.150	I1L1	0.342	C2R2	0.092	L2E2	0.221	W1P1	0.192	K1F1	0.472	W1C1
18	0.146	V1V1	0.344	S1C1	0.092	G2Q2	0.223	W2G2	0.189	M1Y1	0.473	W2C2
19	0.144	A1S1	0.368	P2C2	0.091	N1E1	0.265	C1C1	0.189	C1E1	0.491	C2R2
20	0.140	W1T1	0.383	C1Q1	0.089	L2Q2	0.297	M2C2	0.187	D1G1	0.589	Q1W1

ACB is average compositional bias; plus and minus sign of ACB shows dominancy of that feature towards interacting and non-interacting pairs, respectively. 1 and 2 refer to protein1 and protein2 of pair wise interaction data.

Table S6: The list of dipeptide with ACB values computed from pair wise interaction data.

S. No	*E. coli*				*S. cerevisiae*				*H. pylori*			
	ACB+		**ACB-**		**ACB+**		**ACB-**		**ACB+**		**ACB-**	
	Value	Feature	Value	Feature	Value	Feature	Value	Feature	Value	Feature	Value	feature
1	0.167	PN	0.224	CC	0.145	QQ	0.117	TC	0.239	CC	0.177	WV
2	0.159	MF	0.228	MC	0.114	NN	0.118	CS	0.218	HH	0.178	NW
3	0.145	MI	0.231	NC	0.109	ME	0.118	CM	0.174	DD	0.178	MK
4	0.142	VM	0.236	WM	0.098	QE	0.122	PW	0.167	WD	0.184	FM
5	0.141	MM	0.236	KH	0.095	QD	0.122	FW	0.152	AC	0.187	MM
6	0.132	WT	0.237	DH	0.093	EL	0.122	GC	0.147	DQ	0.188	IC
7	0.131	IF	0.238	CE	0.091	EQ	0.128	YH	0.146	TV	0.191	PM
8	0.127	YM	0.243	HI	0.089	TQ	0.134	YC	0.137	CG	0.202	FW
9	0.120	VV	0.256	HY	0.088	LQ	0.143	CG	0.134	DI	0.216	MC
10	0.120	MY	0.262	RH	0.087	FM	0.143	CV	0.130	ED	0.224	IW
11	0.118	FF	0.269	HN	0.085	LE	0.144	FH	0.128	DN	0.226	FH
12	0.118	GL	0.279	YH	0.082	NE	0.166	CR	0.124	PD	0.232	SC
13	0.117	SM	0.284	CR	0.080	QS	0.167	FF	0.124	VE	0.240	PC
14	0.117	TQ	0.285	PC	0.079	DQ	0.167	WP	0.112	DG	0.250	WF
15	0.113	VG	0.297	SC	0.079	EF	0.175	CP	0.111	YT	0.256	MW
16	0.109	IG	0.297	DC	0.073	QN	0.176	AC	0.111	GC	0.262	YW
17	0.107	FM	0.307	CS	0.073	ET	0.179	PC	0.110	CH	0.282	WS
18	0.105	VW	0.310	HC	0.072	DK	0.203	WG	0.110	NY	0.283	WP
19	0.104	WG	0.327	EC	0.069	DN	0.215	MC	0.106	FE	0.318	QW
20	0.104	PM	0.335	CQ	0.069	QT	0.219	CC	0.105	QC	0.472	WC

ACB is average compositional bias; plus and minus sign of ACB shows dominancy of that feature towards interacting and non-interacting pairs, respectively. Each value is the average of two values coming from similar amino acid of two proteins of a pair.

Table S7: Performance comparison of SVM based methods for PPI prediction

Method	**Threshold**	***E. coli***					***S. cerevisiae***					***H. pylori***				
		Sen	Spe	PPV	Acc	MCC	Sen	Spe	PPV	Acc	MCC	Sen	Spe	PPV	Acc	MCC
BM	1.0	69.9	100	100	97.8	0.83	22.1	94.8	81.1	58.5	0.25	34.0	95.6	88.6	64.8	0.38
	0.8	75.9	100	100	98.3	0.86	32.4	90.7	77.8	61.6	0.29	45.5	92.1	85.2	68.8	0.43
	0.6	79.9	100	100	98.5	0.89	41.5	85.7	74.3	63.6	0.30	56.0	89.1	83.7	72.6	0.48
	0.4	83.7	100	100	98.8	0.90	49.7	79.5	70.8	64.6	0.31	64.3	85.3	81.3	74.8	0.51
	0.2	87.0	100	100	99.1	0.93	57.7	72.8	68.0	65.2	0.31	71.3	79.4	77.6	75.3	0.51
	0.0	90.0	100	99.7	99.3	0.94	64.9	65.8	65.5	65.3	0.31	77.1	73.5	74.4	75.3	0.51
	-0.2	92.4	99.9	98.8	99.4	0.95	71.7	57.9	63.0	64.8	0.30	82.4	65.3	70.4	73.8	0.48
	-0.4	94.1	99.8	96.8	99.3	0.95	77.8	49.0	60.4	63.4	0.28	87.4	57.5	67.3	72.5	0.47
	-0.6	95.0	99.5	93.2	99.1	0.94	83.5	39.4	57.9	61.4	0.26	91.2	46.6	63.1	68.9	0.42
	-0.8	96.0	98.3	81.8	98.2	0.88	88.5	28.9	55.5	58.7	0.22	94.7	34.5	59.1	64.6	0.37
	-1.0	97.2	92.8	51.4	93.1	0.68	93.1	18.7	53.4	55.9	0.18	96.5	23.2	55.7	59.8	0.29
BT	1.0	56.4	100	100	96.8	0.74	28.7	98.3	94.3	63.5	0.38	28.5	99.5	98.3	64.0	0.40
	0.8	72.1	100	100	98.0	0.84	38.5	96.8	92.3	67.6	0.43	46.1	98.6	97.1	72.4	0.53
	0.6	83.9	100	100	98.8	0.91	48.5	94.1	89.2	71.3	0.48	63.0	97.0	95.4	80.0	0.64
	0.4	91.4	100	100	99.4	0.95	57.5	90.8	86.2	74.2	0.51	74.3	94.4	93.0	84.3	0.70
	0.2	94.6	100	100	99.6	0.97	65.6	85.2	81.6	75.4	0.52	81.8	90.4	89.5	86.1	0.73
	0.0	97.2	100	100	99.8	0.98	72.8	78.1	76.9	75.5	0.51	87.7	84.6	85.1	86.2	0.72
	-0.2	98.5	100	100	99.9	0.99	78.9	67.6	70.9	73.3	0.47	92.2	76.5	79.7	84.3	0.70
	-0.4	99.3	100	100	99.9	1.0	84.6	55.0	65.3	69.8	0.42	95.5	64.5	72.9	80.0	0.63
	-0.6	99.6	100	100	100	1.0	89.7	40.6	60.2	65.2	0.35	98.0	49.0	65.8	73.5	0.54
	-0.8	99.9	99.9	99.0	99.9	0.99	93.6	26.2	55.9	59.9	0.27	98.8	29.1	58.2	63.9	0.39
	-1.0	100	68.0	19.6	70.3	0.37	96.2	14.4	52.9	55.3	0.19	99.7	12.3	53.2	56.0	0.25

Where Sen, Spec, Acc and MCC are sensitivity, specificity, accuracy and Matthews Correlation Coefficient, respectively. BM, Biochemical monopeptide composition; BT, Biochemical tripeptide composition

Table S8: The list of top 20 biochemical tripeptides selected on the basis of ACB values computed from pair wise interaction data.

S. No	*E. coli*				*S. cerevisiae*				*H. pylori*			
	ACB+		**ACB-**		**ACB+**		**ACB-**		**ACB+**		**ACB-**	
	Value	Feature	Value	Feature	Value	Feature	Value	Feature	Value	Feature	Value	Feature
1	0.193	B1B1B1	0.188	O2J2G2	0.123	X2X2X2	0.094	O1J1J1	0.272	U1J1U1	0.132	B1G1G1
2	0.171	B1B1J1	0.192	J1U1J1	0.121	U1B1U1	0.097	J2O2J2	0.209	O1O1U1	0.133	O1X1O1
3	0.146	B1J1B1	0.193	U2O2O2	0.118	U1G1U1	0.097	J1X1J1	0.196	O1O1O1	0.134	G1G1J1
4	0.143	B2B2B2	0.197	O1B1U1	0.114	U2G2U2	0.097	J2G2B2	0.186	O1U1U1	0.135	B1B1B1
5	0.133	J2X2J2	0.197	J1U1O1	0.105	U1X1O1	0.099	G2J2J2	0.181	U1U1J1	0.137	G1B1G1
6	0.127	B1G1B1	0.201	O2J2O2	0.102	U1J1U1	0.099	J1G1B1	0.169	U1G1G1	0.140	J2J2B2
7	0.108	U1U1X1	0.204	J1O1X1	0.100	U1X1U1	0.104	B2J2J2	0.167	U1J1O1	0.141	J2G2B2
8	0.108	B1X1B1	0.204	J1G1U1	0.095	B2U2O2	0.109	J1B1J1	0.167	U1B1J1	0.144	O2J2J2
9	0.106	B2G2B2	0.207	X2O2O2	0.093	B1X1U1	0.109	G1G1J1	0.165	J1U1X1	0.146	G1X1X1
10	0.105	J1G1B1	0.208	X1O1O1	0.090	B1U1X1	0.114	B1J1J1	0.165	U1B1U1	0.155	J2X2J2
11	0.104	B1B1G1	0.231	O2O2X2	0.090	U1O1U1	0.117	G1J1J1	0.152	U1U1O1	0.159	J1J1G1
12	0.102	X2X2X2	0.240	J1O1O1	0.087	U2B2U2	0.117	J2G2J2	0.151	U1U1U1	0.166	J1B1B1
13	0.102	G1O1G1	0.242	O1J1O1	0.086	X1U1U1	0.118	J1J1G1	0.149	J1G1U1	0.179	J2J2O2
14	0.102	G1B1B1	0.243	O1J1B1	0.083	X1X1X1	0.128	J2J2B2	0.139	U1O1U1	0.180	J2G2J2
15	0.101	B2G2G2	0.244	G1U1J1	0.082	U2O2U2	0.132	X2J2J2	0.138	J2U2U2	0.181	B2J2J2
16	0.100	B2B2J2	0.245	U1J1U1	0.082	U1U1J1	0.149	J2B2J2	0.135	G1U1B1	0.185	J2B2B2
17	0.096	G1G1G1	0.264	O2O2J2	0.081	B2U2U2	0.155	J1J1B1	0.127	J1U1O1	0.194	B2J2B2
18	0.096	G1B1J1	0.282	U1O1J1	0.081	B1U1U1	0.158	J2J2G2	0.126	J2U2O2	0.205	B2B2J2
19	0.096	G2B2G2	0.289	O1O1J1	0.079	U2J2B2	0.176	J1J1J1	0.121	O1B1U1	0.208	J2O2J2
20	0.095	J1B1J1	0.321	U1J1O1	0.077	U2U2X2	0.228	J2J2J2	0.121	U2O2U2	0.261	J1B1J1

ACB is average compositional bias; plus and minus sign of ACB shows dominancy of that feature towards interacting and non-interacting pairs, respectively. 1 and 2 refer to protein1 and protein2 of pair wise interaction data.

Table S9: The list of biochemical tripeptide with ACB values computed from pair wise interaction data.

S. No	*E. coli*				*S. cerevisiae*				*H. pylori*			
	ACB+		**ACB-**		**ACB+**		**ACB-**		**ACB+**		**ACB-**	
	Value	Feature	Value	Value	Feature	Value	Value	Feature	Value	Value	Feature	Value
1	0.168	BBB	0.137	JBU	0.116	UZU	0.074	JZZ	0.196	UJU	0.087	BZB
2	0.136	BBJ	0.140	JOU	0.104	UBU	0.074	JBB	0.144	UUJ	0.087	ZZJ
3	0.117	BZB	0.142	ZUJ	0.103	XXX	0.075	BJB	0.130	UOU	0.088	JZB
4	0.104	BJB	0.144	XUO	0.086	UOU	0.075	ZJZ	0.127	JUU	0.089	JXJ
5	0.094	BBZ	0.145	UUJ	0.083	BXU	0.077	JBZ	0.126	JUO	0.092	BJJ
6	0.092	BXB	0.152	OOZ	0.083	UXO	0.080	JOJ	0.119	OUU	0.093	ZJB
7	0.086	ZBB	0.153	JOX	0.081	BUX	0.081	JZO	0.119	UZZ	0.094	ZBZ
8	0.083	JBB	0.154	UOO	0.081	UXU	0.083	JXJ	0.114	UBU	0.096	JOB
9	0.082	ZBJ	0.158	OJZ	0.081	BUU	0.085	ZJB	0.111	UUU	0.096	OXO
10	0.080	BZZ	0.172	OUJ	0.076	UJB	0.093	OJJ	0.105	BUU	0.097	BZJ
11	0.079	ZBZ	0.172	OJB	0.075	UXX	0.098	JZB	0.103	JUX	0.098	JJB
12	0.076	ZZZ	0.177	JZU	0.074	UJU	0.100	ZZJ	0.100	UUO	0.100	ZJZ
13	0.070	XXX	0.188	UJU	0.074	BUO	0.100	XJJ	0.096	UUZ	0.104	BBB
14	0.069	JZB	0.200	OOX	0.074	UUX	0.104	JZJ	0.089	OOU	0.115	OJJ
15	0.069	BZJ	0.201	JOO	0.069	XBU	0.108	ZJJ	0.087	XUO	0.125	BBJ
16	0.067	ZOZ	0.207	XOO	0.068	UXB	0.109	BJJ	0.087	UUB	0.130	JJZ
17	0.065	XZX	0.210	UOJ	0.068	XUU	0.129	JBJ	0.087	ZUB	0.140	BJB
18	0.065	ZXJ	0.222	OJO	0.067	UUJ	0.138	JJZ	0.085	UOB	0.148	JZJ
19	0.064	XBB	0.245	UJO	0.063	XUB	0.141	JJB	0.084	OUO	0.175	JBB
20	0.061	JBJ	0.276	OOJ	0.062	UBX	0.202	JJJ	0.084	UBJ	0.184	JBJ

ACB is average compositional bias; plus and minus sign of ACB shows dominancy of that feature towards interacting and non-interacting pairs, respectively. Each value is the average of two values coming from similar amino acid of two proteins of a pair.

Table S10: Performance comparison of SVM based methods for PPI prediction

Method	**Threshold**	***S. cerevisiae***					***H. pylori***				
		Sen	Spe	PPV	Acc	MCC	Sen	Spe	PPV	Acc	MCC
PSSM	1.0	--	--	--	--	--	25.7	99.2	96.9	62.4	0.37
	0.8	--	--	--	--	--	38.7	98.4	95.9	68.5	0.46
	0.6	--	--	--	--	--	53.2	95.8	92.7	74.5	0.54
	0.4	--	--	--	--	--	64.7	92.7	89.9	78.7	0.60
	0.2	--	--	--	--	--	74.7	88.4	86.6	81.6	0.64
	0.0	--	--	--	--	--	83.0	82.0	82.1	82.5	0.65
	-0.2	--	--	--	--	--	89.6	71.4	75.8	80.5	0.62
	-0.4	--	--	--	--	--	93.8	58.2	69.2	76.0	0.56
	-0.6	--	--	--	--	--	97.1	41.7	62.5	69.4	0.46
	-0.8	--	--	--	--	--	99.0	23.7	56.5	61.4	0.34
	-1.0	--	--	--	--	--	99.7	12.3	53.2	56.0	0.25
HB	1.0	21.2	99.1	95.9	60.2	0.32	28.6	99.5	98.3	64.1	0.40
	0.8	31.5	98.1	94.4	64.8	0.40	46.1	99.0	98.0	72.6	0.53
	0.6	42.1	96.3	92.0	69.2	0.46	62.9	97.1	95.6	80.0	0.64
	0.4	52.4	93.0	88.3	72.7	0.50	74.9	94.6	93.3	84.7	0.71
	0.2	62.4	87.4	83.3	74.9	0.52	83.1	91.1	90.3	87.1	0.74
	0.0	71.2	79.1	77.3	75.2	0.51	88.2	85.7	86.1	87.0	0.74
	-0.2	79.5	65.6	69.8	72.5	0.46	92.5	77.6	80.5	85.0	0.71
	-0.4	87.1	46.2	61.8	66.6	0.36	95.9	66.0	73.8	80.9	0.65
	-0.6	93.2	26.5	55.9	59.9	0.27	98.3	51.0	66.7	74.6	0.56
	-0.8	97.2	11.7	52.4	54.4	0.17	99.3	29.2	58.4	64.3	0.40
	-1.0	98.9	4.3	50.8	51.6	0.10	99.7	12.3	53.2	56.0	0.25

Where Sen, Spec, Acc and MCC are sensitivity, specificity, accuracy and Matthews Correlation Coefficient, respectively. PSSM, Position Specific Scoring Matrix composition; HB, Hybrid feature where dipeptide and biochemical tripeptide compositions are joined together. '--'means that the SVM models were not trained for Yeast data set using PSSM.