Supplementary Materials:

Part S1. The influence of different k_+ values on predictions

The dataset we used to train the SVM classifiers was quite unbalanced: the ratio of binders versus non-binders was 1:20. When using penalty parameters, the objective function of SVM becomes:

$$\min_{\substack{w,b,\xi \\ w_i,b,\xi}} \frac{1}{2} w^T w + C_+ \sum_{y_i=1} \xi_i + C_- \sum_{y_i=-1} \xi_i \\
\text{subject to} \quad y_i (w_T \phi(x_i) + b) \ge 1 - \xi_i, \\
\xi_i \ge 0, i = 1, ..., l.$$
(1)

where $C_{+}=k_{+}C$ and $C_{-}=k_{-}C$ are the penalty parameters of error terms for the positive and the negative class, respectively; k_{+} and k_{-} are the weight parameters for the positive and the negative class, respectively; C > 0 is a constant that represents the upper boundary of the penalty parameters.

In this study, we set a higher weight to the binder class (k_+) while keeping k_- to 1. Because different values of k_+ and k_- would affect the classification sensitivity and specificity of the SVM, we plotted a receiver operating characteristic (ROC) curve by varying k_+ (Figure S3). The area under the curve in the ROC plot represents of the performance of the model and our model achieved a value of 0.948. To determine the value of k_+ that would achieve balanced sensitivity (*SE*), prediction accuracy for binders (Q_+) and Matthews correlation coefficient (*C*), we analyzed how these classification criteria changed with the value of k_+ (Figure S4 and S5 in the supplementary materials). The value of k_+ is decided by the user: higher value of k_+ gives higher *SE* but lower Q_+ and *C*; lower of k_+ gives lower *SE* but higher Q_+ and *C*. Figure S4 and S5 suggest that $k_+ = 4$ may be a balanced value.

Table S1. The number of binders and non-binders included in the data set for training the MIEC-SVM classifier

	Binder	Non-binder
Abl_human	31	620
Boi1_yeast	25	500
c-Src_human	61	1220
Fyn_human	27	540
Grb2_mouse	19	380
Itk_human	5	100
Lsb3_yeast	25	500
Lyn_human	28	560
Myo3_yeast	6	120
Myo5_yeast	52	1040
Nbp2_human	25	500
P85a_human	29	580
Rvs167_yeast	19	380
Sla1_yeast	30	600
Spta2_human	20	400
Yes_human	29	580
Yha2_yeast	40	800
Ysc84_yeast	20	400
Total	491	9820

10010 02. 11	the dominion of the interaction parts for the rist of the domain for encounting willes
Peptide residue	SH3 residue
Ala_6	Gly13, Asp14, Asn15, Thr16, Asn31, His32, Asn33, Trp36, Glu38, Trp47
Pro ₋₅	Gly13, Asp14, Asn15, Thr16, Asn31, His32, Asn33, Gly34, Trp36, Glu38, Trp47
Ser_4	Ser12, Gly13, Asp14, Asn15, Thr16, His32, Asn33, 00, Gly34, Trp36, Trp47
Tyr ₋₃	Phe9, Val10, Ala11, Ser12, Gly13, Asp14, Asn15, Thr16, Glu35, Trp36, Trp47
Ser_2	Ser12, Asn33, 00, 00, Gly34, Glu35, Trp36, Trp47
Pro_{-1}	Gly34, Glu35, Trp36, Pro49, Ser50, Asn51, Tyr52
Pro ₀	Phe9, Trp36, Pro49, Asn51, Tyr52
Pro ₁	Tyr7, Asp8, Asn51, Tyr52
Pro ₂	Leu6, Tyr7, Asn51, Tyr52
Pro ₃	Tyr7, Asp8, Asn51, Tyr52

Table S2. The definition of the interaction pairs for the Abl SH3 domain for calculating MIECs

Model	MIECs*	Kernel function	SE_{train}	SP_{train}	SE_{test}	SP_{test}	Q_+	Q_{-}	С
			(%)	(%)	(%)	(%)	(%)	(%)	
1	$\Delta E_{vdw}, \Delta E_{ele}$	Linear	87.5	88.2	76.9	87.4	23.4	98.7	0.377
		Polynomial	65.0	85.4	62.4	85.2	17.4	97.8	0.269
		RBF	80.4	88.8	74.9	88.3	24.3	98.6	0.381
		Sigmoid	71.3	86.9	67.3	86.7	20.3	98.1	0.316
2	$\Delta E_{vdw}, \Delta G_{polar}$	Linear	89.9	88.3	79.5	87.6	24.3	98.8	0.394
		Polynomial	69.8	87.4	66.2	87.3	20.7	98.1	0.317
		RBF	71.1	88.1	67.5	87.8	21.8	98.2	0.333
		Sigmoid	66.3	85.4	63.8	85.4	17.9	97.9	0.279
3	$\Delta E_{vdw}, \Delta E_{ele}, \Delta G_{GB}$	Linear	89.2	89.1	77.5	88.2	24.8	98.7	0.393
		Polynomial	48.6	87.5	46.5	87.3	15.5	97.0	0.206
		RBF	79.1	88.6	74.1	88.2	24.0	98.5	0.374
		Sigmoid	72.4	86.8	69.0	86.5	20.4	98.2	0.322

Table S3.The performance of the SVM classifiers using various kernel functions and various combinations of domain-peptide MIECs

* ΔE_{vdw} , ΔE_{ele} , and ΔG_{GB} are van der Waals, electrostatic and polar contribution to desolvation, respectively. $\Delta G_{polar} = \Delta E_{ele} + \Delta G_{GB}$

No.	Domains	p-vc	ılue [*]
		MM/GBSA	MM/PBSA
1	Abl_human	2.77e ⁻⁸	8.15e ⁻¹²
2	Boi1_yeast	$1.75e^{-6}$	0.0050
3	c-Src_human	$1.27e^{-9}$	7.71e ⁻⁸
4	Fyn_human	8.16e ⁻⁵	$1.95e^{-4}$
5	Grb2_mouse	1.09e ⁻⁹	8.16e ⁻⁴
6	Lsb3_yeast	1.52e ⁻⁸	3.12e ⁻⁵
7	Lyn_human	$4.67e^{-22}$	0.0023
8	Myo5_yeast	0.55	5.37e ⁻⁴
9	Nbp2_human	5.76e ⁻¹¹	3.96e ⁻¹⁰
10	P85a_human	6.32e ⁻²³	$1.24e^{-9}$
11	Rvs167_yeast	8.36e ⁻¹³	$1.03e^{-12}$
12	Sla1_yeast	7.29e ⁻⁸	3.92e ⁻⁵
13	Spta2_human	5.82e ⁻⁴	0.010
14	Yes_human	5.56e ⁻¹²	9.04e ⁻⁷
15	Yha2_yeast	$4.24e^{-22}$	7.63e ⁻¹⁹
16	Ysc84_yeast	6.73e ⁻¹⁰	2.40e ⁻⁵

Table S4. Distinct distributions of the binding free energies calculated by MM/GBSA and MM/PBSA for binders and non-binders.

*P-value was calculated by Student's *t*-test. Two domains, Itk and Myo3, were not included in the calculation because they only had a small number of known binders (5 and 6, respectively).

No.	Domains	SE	n+	N+	SP	n-	N-	\overline{Q}_+	<i>Q</i> _
_		(%)			(%)			(%)	(%)
1	Abl_human	90.3	28	31	75.3	467	620	15.5	99.4
2	Boi1_yeast	56.0	14	25	83.4	417	500	14.4	97.4
3	c-Src_human	44.3	27	61	87.0	1062	1220	14.6	96.9
4	Fyn_human	92.6	25	27	79.9	383	540	137	99.5
5	Grb2_mouse	63.2	12	19	79.7	303	380	13.5	97.7
6	Lsb3_yeast	76.0	19	25	78.8	473	600	13.0	98.7
7	Lyn_human	92.9	26	28	86.6	485	560	25.7	99.6
8	Myo5_yeast	82.7	43	52	21.1	219	1040	4.98	96.1
9	Nbp2_human	96.0	24	25	74.6	373	500	15.9	99.7
10	P85a_human	93.1	27	29	88.6	514	580	29.0	99.6
11	Rvs167_yeast	15.8	3	19	93.2	354	380	10.3	95.7
12	Sla1_yeast	96.7	29	30	79.0	474	600	18.7	99.8
13	Spta2_human	46.7	10	20	87.7	367	400	23.3	97.3
14	Yes_human	82.8	24	29	81.0	470	580	17.9	98.9
15	Yha2_yeast	75.0	30	40	84.9	679	800	19.9	98.5
16	Ysc84_yeast	90.0	18	20	80.3	321	400	18.6	99.4
	Total	74.8	359	480	75.9	7361	9700	13.3	98.4

Table S5. The classification accuracies of SVMs trained on the total binding free energies calculated by MM/GBSA.

Table S6. Comparison of MIEC-SVM and SH3-hunter on predicting the interactions between twenty peptides and four SH3 domains. These 20 peptides were not included in the training set of MIEC-SVM. It is not clear to us whether they were included in the training set of SH3-hunter.

No.	Peptides		Src			Yes			Abl			Grb2	
		Exp	MIEC	SH3-									
			-SVM	hunter									
1	LASRPLPLLP	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	×	Х	\checkmark	×	Х	×
2	ISQRALPPLP	\checkmark	\checkmark	\checkmark	\checkmark	\times	\checkmark	×	\times	\checkmark	×	\times	\times
3	ITMRPLPALP	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\times	\times	\checkmark	\checkmark	\checkmark	\checkmark
4	RSGRPLPPIP	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\times	×	\checkmark	\times	\checkmark	\times
5	PPWWAPPPIP	\times	\times	\checkmark	\times	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	\times	\checkmark	\times
6	APTYPPPPPP	×	\times	\times	×	\times	\checkmark	\checkmark	\checkmark	\checkmark	×	\checkmark	\times
7	LTPQSKPPLP	×	\times	\times									
8	LGEFSKPPIP	×	\times	\times	×	\times							
9	SSAPQRPPLP	×	\times	\times									
10	VVPLGRPEIP	×	\times	\times									
11	MPPPVPPRPP	\checkmark	\times	\times	×	\times	\checkmark	×	\times	\checkmark	×	\times	\checkmark
12	VPPLVAPRPP	×	\times	\times									
13	GQPAGDPDPP	×	\times	\times	×	\times							
14	ATSEGLPILP	×	\times	\times									
15	KWDSLLPALP	\checkmark	\times	\times	\checkmark	\times	\times	×	\times	\times	\checkmark	\times	\times
16	YWDMPLPRLP	\checkmark	\checkmark	\times	\checkmark	\checkmark	\times						
17	YYQRPLPPLP	\checkmark	\checkmark	\times	\checkmark	\checkmark	\times	×	\times	\times	\checkmark	\checkmark	\times
18	YFSRALPGLP	\checkmark	\times	\checkmark	\checkmark	\times	\checkmark	×	\times	\checkmark	\checkmark	\times	\times
19	SLWDPLPPIP	\checkmark	\checkmark	\times	\checkmark	\checkmark	\times	×	\times	\times	\checkmark	\checkmark	\times
20	DPYDALPETP	\checkmark	×	×	\checkmark	×	×	\times	×	×	\checkmark	×	X
	Accuracy		16	13		15	13		20	14		14	11

007.	The 210 peptides test	ed in the pep	inde unitaly experi
No.	Peptide	Prediction	Experiments
1	PPKFSPPPPP	\checkmark	
2	PPHWAPPAPP	\checkmark	
3	PPTWTPPKPP	×	
4	KPTYPPPPPP	\checkmark	
5	GPRWSPPPVP	\checkmark	
6	GPRFPVPPVP	\checkmark	
7	APKKPAPPVP	\checkmark	×
8	APTMPPPLPP	\checkmark	
9	PPPYPPPVP	\checkmark	
10	PPPYPPPDIP	\checkmark	
11	AAAMQKPSLP	×	×
12	AAMFQAPKTP	×	×
13	AFSFPHPINP	×	×
14	AGLMLTPTGP	×	×
15	AIWYTLPILP	×	×
16	AKRYVVPGHP	×	×
17	ALSMFAPLLP	×	×
18	ALWWFIPESP	×	×
19	ASQMLRPFAP	×	×
20	CEAFLKPWAP	×	×
21	DDFWPNPKFP	×	×
22	DDIMMSPHSP	×	×
23	DEAFKNPTKP	×	×
24	DGMMLGPKYP	×	×
25	DNGWIHPLDP	×	×
26	DNYYGTPKPP	×	×
27	DQGYPDPPNP	\checkmark	×
28	DVFMGPPGSP	×	×
29	ECLFPSPTRP	×	×
30	EDDFLEPATP	×	×
31	EGNFRTPMLP	×	×
32	EILWSAPLGP	×	×
33	EWKYLKPRTP	×	×
34	FALMGSPKPP	×	×
35	FPNWTYPVGP	×	×
36	FSHYPQPSPP	×	×
37	FSIFDSPSNP	×	×
38	GAGYPPPTMP	×	×
39	GGRFKRPTTP	×	×
40	GGRFNLPHAP	×	×
41	GLPYSHPPQP	×	×
42	GVVMPSPVKP	×	×

Table S7. The 210	peptides tested in the	peptide array ex-	periment ^a

43	ILEYLHPRLP	×	×
44	IMQYTGPMLP	×	×
45	IMRYLAPEGP	×	×
46	ISSFSPPEKP	×	×
47	IVQWEEPVEP	×	×
48	KCGFPLPGVP	×	×
49	KPKFTTPEYP	×	×
50	KRMFPLPEVP	×	×
51	KSPFGVPGMP	×	×
52	KSQFRLPFKP	×	×
53	KSQYLQPKQP	×	×
54	LAAYAAPGYP	×	×
55	LAEYEMPIQP	×	×
56	LEDYKKPLPP	×	×
57	LEWMQNPEAP	×	×
58	LLNYIAPGEP	×	×
59	LSPFMIPLFP	×	×
60	MARWNQPQPP	×	\checkmark
61	MSFWLIPSRP	×	×
62	MTEMNPPTQP	×	×
63	NARFKRPVLP	×	×
64	PDPFKAPSRP	×	×
65	PGPYGLPGFP	×	×
66	PKKFHVPGLP	×	×
67	PLRWGPPEAP	×	×
68	PPPWAPPCSP	×	×
69	PPWMQPPPPP	\checkmark	\checkmark
70	QSIYGSPLSP	×	×
71	QSNYSYPQVP	×	×
72	RDSYPGPEDP	×	×
73	RFAFDRPGLP	×	\checkmark
74	RPAFGGPAIP	×	×
75	RWNFSPPEFP	×	×
76	SAPMPEPGAP	×	×
77	SEGWIEPSYP	×	×
78	SGHYSVPKLP	×	×
79	SLNFSSPDPP	×	×
80	SMPFAPPTLP	×	×
81	SNQFVGPIPP	×	×
82	SSLFYSPSSP	×	×
83	STKYNGPPFP	×	×
84	TALFTHPEGP	×	×
84 85	TALFTHPEGP TFDMNRPLLP	× ×	× ×

87	TPGFQNPQLP	×	×
88	TPTWESPARP	×	×
89	TRLYYTPEDP	×	×
90	TTTFDKPTVP	×	×
91	TWEFTQPGLP	×	×
92	TWVYYLPLLP	×	×
93	VGEWYKPDRP	×	×
94	VTAWQQPQPP	×	×
95	VVSMTPPHSP	×	×
96	WDNMVGPGKP	×	×
97	WPMFSAPSSP	×	×
98	YGDYTLPDVP	×	×
99	YITFIGPSWP	×	×
100	YLKYKDPQSP	×	×
101	YRMFISPLYP	×	×
102	AASLALPPQP	×	×
103	AAVEAYPEIP	×	×
104	ADIPKSPTKP	×	×
105	AFHSITPAPP	×	×
106	ALDQSMPVTP	×	×
107	ANFGLFPELP	×	×
108	ANTGGAPLNP	×	×
109	APCIKIPAAP	×	×
110	APRQQRPQQP	×	\checkmark
111	ATPVSGPTTP	×	×
112	AVPLIFPERP	×	×
113	CEDLPQPESP	×	×
114	CGPKPPPGFP	×	×
115	CIVKLVPSKP	×	×
116	CMIASTPPAP	×	×
117	DGLRFVPSLP	×	×
118	DRVHSFPTQP	×	×
119	DTFGDEPNNP	×	×
120	EIGKVPPPIP	×	×
121	EPAPRSPVPP	×	×
122	EREEGAPETP	×	×
123	ESLDDAPVAP	×	×
124	EVERNVPDPP	×	×
125	FCPCDTPYIP	×	×
126	FRGQGCPSTP	×	×
127	GALCSNPSCP	×	×
128	GEEQRPPETP	×	×
129	GIAVAQPILP	×	×
130	GNMGVVPPGP	×	×

131	GPSLPGPFSP	×	×
132	GPSPGGPQPP	×	×
133	GPSSLGPSNP	×	×
134	HSAGVIPIKP	×	×
135	HYGTMDPNIP	×	×
136	IIEETEPTVP	×	×
137	ISCDSSPVLP	×	×
138	KARSGPPTIP	×	×
139	KERIKQPPSP	×	×
140	KGDAGPPGIP	×	×
141	KVALGIPNLP	×	×
142	LADLGIPVMP	×	×
143	LAKQVDPYIP	×	×
144	LELQRLPERP	×	×
145	LFRISLPVAP	×	×
146	LFRLGPPKPP	×	×
147	LHLGDLPAEP	×	×
148	MVAEEAPPPP	×	×
149	NEPPPPPPP	×	×
150	NGVLVRPRDP	×	×
151	NKCPAGPSGP	×	×
152	NMDVTFPSMP	×	×
153	PAAHGTPGAP	×	×
154	PANGEYPLNP	×	×
155	PDGSEIPLPP	×	×
156	PDWPLPPDWP	×	×
157	PFVHPKPPPP	×	×
158	PGTKGFPGSP	×	×
159	PKHSPPPPTP	\checkmark	×
160	PLRDPHPTPP	×	×
161	PPHQAIPLLP	×	×
162	PQLPLTPSTP	×	×
163	PSRGWAPPGP	×	\checkmark
164	PSSLMSPSTP	×	×
165	QPPPPPPQGP	×	×
166	QRLGPGPALP	×	×
167	RAALNLPLLP	×	×
168	RDFPGPPHAP	×	×
169	RKAVYLPGVP	×	×
170	RLSPTLPLLP	×	×
171	RNCLLRPGSP	×	×
172	RQLKLSPQVP	×	×
173	RQRLITPSPP	×	×
174	RTPPSPPGCP	×	×

175	RTPTTPPVFP	×	×
176	SERRDAPPPP	×	×
177	SETSPIPKTP	×	×
178	SHPSPAPVLP	×	×
179	SIDSGPPPLP	×	×
180	SKRAFEPRTP	×	×
181	SMVNSLPTFP	×	×
182	SQLVEFPLGP	×	×
183	SRDAASPDKP	×	×
184	SRHGLSPATP	×	×
185	SSDSLGPPRP	×	×
186	SSDSMFPYIP	×	×
187	SSLILPPKTP	×	×
188	SSSQLVPWKP	×	×
189	SSTKSKPGSP	×	×
190	STDPPKPPLP	×	×
191	STPCGEPNAP	×	×
192	SWPDDVPKIP	×	×
193	TGKKQVPLNP	×	×
194	TIEPVGPSDP	×	×
195	TKKRPAPRAP	×	×
196	TLWKAKPDEP	×	×
197	TSDGHCPLHP	×	×
198	TSFKIVPIVP	×	×
199	TTTTTTPDKP	×	×
200	VAPSSLPPPP	×	×
201	VDGDFIPDDP	×	×
202	VDPKYVPVKP	×	×
203	VEKVIYPGLP	×	×
204	VENHPNPAAP	×	×
205	VEPAAVPGEP	×	×
206	VEPNTVPHTP	×	×
207	VHQSTIPSNP	×	×
208	VSLGWEPVRP	×	×
209	WPPICDPPQP	×	×
210	YVRRRRPHKP	×	×

^{*a*} $\sqrt{}$: binder; \times : non-binder

	10	20	30		40	ř.	50	_60	70	G	80	-
YB65_SCHPO S Y V	KALYAYT	AQ - S DM EL	S IQEGDI	QVTNRN-			$\cdots \cdot AG \cdots N$	G W - S E G I L M	Ň	- G V T -	GQFPA	NYVTDV
MYOC_DICDI Q Q Y	IALYEYD	0 A M - Q P D E L	TFKENDV	INLIK-K-			VD A	DW-WQGELV	/R	- T K Q I	GML P S	NYVQQI
MYOB_DICDI P T A	KALYDYD	DAS - STD EL	SFKEGDI	IF I VQ - K -			DN G	GW-TQGELF	(8	-GQK-	GWAPT	NYLQYN
MYSB_ACACA PQV	KALYDYD	AQ - TGD EL	TFKEGDT	I I V HQ - K -			DP A	G <mark>W</mark> - WE <mark>G</mark> ELN		- G K R -	GWVP A	NYVQDI
VAV_HUMAN G T A	KAR YDF C	AR - DRS EL	SLKEGDI	IKILNKK -			••••GQ••Q	GW-WRGEIY	(- G R V -	GWF P A	NYVEED
LASP1_CAEEL F A V	KAIYDYA	AA - DKD E I	SFLEGDI	IVNCE - K -			1 D D	GW-MTGTVQ	R T	- LQW -	GML P A	NYVQPH
ABP1_SACEX P WA	TAE YDYE	AG - EDN EL	TFAENDK	IINIE-F-			VD D	DW-WLGELE	T T	- GQK -	GLFPS	NYVVLG
PLCG2_HUMAN R T V	KAL YDYK	AK - R S D E L	SFCRGAL	IHNVS - K -			E P G	G W - WKG DYC	iT	- R I Q -	QYFPS	NYVEDI
PLCG1_BOVIN CAV	KALFDYK	AQ - RED EL	TFTKSAI	IQNVE - K -			••••QE••G	GW-WRGDYC		-GKKQ	L WF P S	NYVEEM
DRK_DROME ML V	QAL YDF V	PQ-ESGEL	DFRRGDV	ITVTD-R-			S D E	NW-WNGEI-		- GNRK	GIFPA	T Y V T P Y
SEM5_CAEEL KF V	Q A L F D F N	PQ-ESGEL	AFKRGDV	ITL IN - K -			D D P	NW-WEGQLN	*	- N R R -	GIFPS	NYVCPY
GRB2_CHICK TYV	QALFDFD	PQ - EEG EL	GFRRGDF	IQVLD-N-			S D P	NW - WKGACI		- GQT -	GMF P R	NYVTPV
HCLS1_HUMAN I S A	VAL YDYQ	GE - GSD EL	SFDPDDV	ITDIE - M -			VD E	GW-WRGRCH	1	-GHF -	GLFPA	NYVKLL
MYSC_ACACA EQA	RALYDFA	AE - NP D EL	TFNEGAV	VTVIN-K-			S N P	DW-WEGEL?	4	- GQR -	GVFPA	SYVELI
CDC25_CANAL DT V	I AL YDF P	GT-QSSHL	PLNLGDT	IHVLS - K -			S A T	GW-WDGVVM	IGNSGE	- L Q R -	GWF P H	NYVRSV
YKA7_CAEEL T Y A	TAIYDYN	ISN-EAGDL	NFAVGSQ	IMVTA - R -			VN E	EW-LEGECH		-GRS -	GIFPS	QFVDCP
NCF2_HUMAN S Q V	EALFSYE	AT - Q P E DL	EFQEGDI	ILVLS - K -		• • • •	VN E	EW-LEGECH	(- G K V -	GIFPK	VFVEDC
TXK_HUMAN IQV	KALYDFL	. P R - E P C N L	ALRRAEE	YLILE - K -	• • • • •	• • • •	Y N P	HW - WKARDF	RL	-GNE -	GLIPS	NYVTEN
TEC_HUMAN E I V	VAMYDEQ	AA - EGH DL	RLERGQE	YLILE - K -			ND V	HW-WRARD	(Y	-GNE -	GYIPS	NYVTGK
BTK_HUMAN KKV	VALYDYM	IPM - NAN DL	QLRKGDE	YFILE-E-			S N L	PW-WRARD	CN	- GQE -	GYIPS	NYVTEA
BTKL_DROME KL V	VALYPEK	AI-EGGDL	SLEKNAE	YEVID-D-			SQ E	HW-WKVKD/	\L	-GNV-	GYIPS	NYVKPK
SRK1_SPOLAPIY	VGKYDYD	DSR-TDDDL	SFKKGDL	MY115-T-		• • • •	DEG	DW - WF ARS F	CDTA	-GKE -	GYIPS	NYVAEY
SIK_HYDATTIF	VALYDYE	AR-ISEDL	SFKKGER	LQTIN-T-			G	DW - WYARSI	. 1 1	-NSE -	GYIPS	TYVAPE
FGR_HUMAN T L F	TALYDYE	AR - TED DL	TEIKGEK	HILN-N-			TEG	DW - WEARSI		-GKT -	GCTPS	NYVAPV
DLK_CHICKKLV	VALEDYE	AV NOR	OVINCEN	OVIR C			su	WI ADEL	V T	-GQE -	C VVD	NEVADV
BLK_MOUSE KFV	VALIDIA	AV-NDRDL	OVLKGEK	LQVLR-S-				DW-WLARSI	. v 1	-GRE -	GLIP	NEVAPV
SRC64_DROME R V V	VALTDIK	SK-DESDL	S F MKG DR	NEVID-D-			DE E	DW - WR V VNI		- RQE -	GUIPL	NEVALE
CRKL_HUMAN E TV	KILIDFP ALVDPP	GR - DAE DL	FERGET	DVIIE-K-			DEND.	QW-WSARNI	D	-GRV-	GLUDY	PTVERL
ARI DROME OI F	VALNDEO	AG GEN OL	SI VYGEO	VPILS VN			VS C	EW CEAUSI		GNV	GWVDS	NUTRI
SPEL DUCTI	VALIDIQ	AS ODT DI	SEEEVEI	VALLS IN-			ED E	EW I EVVE	NK D	NEE	OL VBC	NAVERO
DPK DROME	LAKHDES	ATADDALEL	SERVICI	KIL NME -			DD	NW-VPAFLI		GKE	GLIPS	NYLEMK
SEMS CAFEL	VAEHDEO	AG - SPD EL	SEKRONT	KVINKD.			ED	HW. VKAELI		GNE	GEIPS	NVIRMT
CSK CHICKTEC	LAKNNEH	GT - AFO DI	PESKODV	TIVAVT -			KDP	NW-YKAKNE	v	GRE .	GLIPA	NYVOKR
NCKI HUMAN VVV	VAKEDYV	A0 -0E0 EL	DIKKNER	WLLD-D-				SW-WRVRNS	M	- NKT -	GEVES	NYVERK
SPTCA DROME + + + + E C V	VALVDYT	EK-SPREV	SMKKGDV	TLLN-S-			NN K	DW-WKVEV		- DRO -	GEVPA	AVIKKI
SPTAL HUMAN OR V	MALYDEO	AR-SPREV	TMKKGDV	TLLS - S -			K	DW-WKVEA		- DHO -	GIVPA	VYVRRL
SCD2 SCHPO K V I	RALNDYT	AR - KAT EV	SEAKGDE	HVIGRE -			· · · · ND · · K	AW-YEVCNI		-GTR -	GEVEV	SHEEEL
BEMI YEASTKVI	KAKYSYO	AO . TSK FL	SEMEGEE	YVSG - D -			EK	DW-YKASNE	S T	-GKE -	GVVPK	TYPEVE
SLAT YEASTKKV	RAIYDYE	OV-ONADEEL	TEHENDY	DVFD-D-			KD A	DW-LLVKST	V S	-NEF -	GFIPG	NYVEPE
RASAL BOVIN R R V	RAILPYT	KVPDTD EI	SELKGDM	FIVHN-E-			L E D	WMWVTNLE	ат	- DEO -	GLIVE	DLVEEV
YKA7 CAEEL P YG	LAKEDYA	PT-OSDEM	GLRIGDT	VLISK-K-			VD A	EW-FYGENO	N 0	-RTF -	GIVPS	SYLDIK
STE6 SCHPO F O T	TAISDYE	NSSNPS FL	KFSAGDT	IIVIE-V-			L E D	GW-CDGICS		- E K R -	GWF P T	SCIDSS
NCF1 HUMAN Q T Y	RAIADYE	KT-SGSEM	ALSTODY	VEVVE-K-				GW-WFCQM	(- AKR -	GWIPA	SFLEPL
NCF1_HUMAN E P Y	VAIKAYT	AV - EGD EV	SLLEGEA	VEVIH-K-			L L D	GW-WVIRKI)	- DVT -	GYFPS	MYLQKS
SCD2_SCHPO L F G	I VQ F DF A	AE - RPD EL	EAKAGEA	1111A - R -			S N H	EW-LVAKP	GRL	-GGP -	GLIPL	SFIQLR
BEMI_YEAST L Y A	I VL YDF K	AE - KAD EL	TTYVGEN	LFICA-H-			HN C	EW-FIAKP	GRL	-GGP -	GLVPV	GFVSII
SLAI_YEAST KRG	I VQ Y DF M	1AE - SQD EL	TIKSGDK	VYILDDK-			KS K	DW-WMCQLV	/D S	- G K S -	GLVPA	QFIEPV
NCF2_HUMAN E A H	RVLFGFV	PE-TKEEL	Q VMPGNI	VFVLK-K-			G N D	NW - AT VMF M	4	- GQK -	GLVPC	NYLEPV
BOI2_YEAST P MY	IAINEYF	KR - ME D EL	DMKPGDK	KVITDD -			E E Y K D	G W - YFGRNI	. R T	- NEE -	GLYPV	VFTQK1
BZZ1_YEASTGKN	KVL YAYV	QK - DDD EI	TITPGDK	ISLVARD -			TG S	G <mark>W - TKINNI</mark>	DT T	- G E T -	GLVPT	TYIRIS
BUD14_YEAST DKL	YALYAFN	GH-DSSHC	QLGQDEP	CILLN-D-		• • • •	Q D A	YW-WLVKR	T D	- G K I -	GFAPA	EILETF
HCK_HUMAN E D I I V	VALYDYE	AI-HHEDL	SFQKGDQ	NVVLE-E-		• • • •	G	EW-WKARSI	A T	- RKE -	GYIPS	NYVARV
BBC1_YEASTFKV	VAQFPYK	SD-YEDDL	NFEKDQE	ITVTS - V -	• • • • •	• • • •	E D A	EW-YFGEYC	DSNGD	- VIE -	GIFPK	SFVAVQ
ABLI_HUMAN ····NLF	VALYDEV	AS -GDN TL	STIKGEK	KVLGYN-			HN G	EW-CEAQTH	N · · · · · ·	- GQ	UWVPS	NYTTPV
BIO1_YEASTPLY	IAVNQYS	KR - MED EL	NMKPGDK	IKVITDD-			GEYND	GW-YYGRNI	. R T	- KEE -	GLYPA	VFTKRI
EVN HUMAN CTOUT	VALUDTE	AP TED	SPUKOPH	COLLN C			G	WE ARSI	T T	GET.	GVIDO	NYVAPS
FYN_HUMAN GIGVILF	VALYDYE	AR-TEDDL	SPHKGEK	FQILN-S-			SEG	DW - WEARSI	.1 1	-GET -	GEIDE	NYVAPV
GRB2_MOUSEMEA	IAKIDEK	TN DBO FI	ALBODEE	VILD C			Q	NW-IKAEL		CHE.	OVADE	NTIEMK
	VALVBYD	CI HPD DI	SEVVOEN	ITLLD-S-			se1	WE WE AVEL	I T	VVE	GEIDE	STLVERS -
MYO3 VEAST K DP K F	FAANDEP	OSOSSS	PLKKGDL	VELSE-D.			EP	GW-SLAKII	DesGerre	SKE	GWVPT	AVMTRV
MYOS VEAST	FAANDEP	GSGSPS	PLKKGDV	IVITR - F -			EP	GW-SLAKLI	DesGere	SKE	GWVPT	AVMEPH
NBP2 HUMAN	VALVDEE	PE-NDN-FL	RLAEGDI	VEISV-K-			HG	GW-I VAENE	S. G.	SKT	GLVPE	FEVSTIOP
P85A HUMAN - AEG VOY	RALYDYK	KE-REE- DI	DLHLGDI	TVNKGSL	VALGE	S DGO	EARPEE	GW-LNGVNE	ТТ	GER.	GDEPG	TYVEYIGP
RVS167 YEAST	TALYDYO	A0 - AAG DI	SEPAGAV	IEIVO-R-			···· TPDVN	EW-WTGRY		- 600 -	OVERG	NYVOLN
SLAI YEAST	RAVYAYE	PO - TPE FI	ALOEDDI	YLLO-K-			···· SD. 10	DW-WTVKER	VI-GSDS	EEPV.	GLVPS	TYLEFAPY
SPTA2 HUMAN FI V	LALYDYO	EK-SPR	TMKKGDU	TLLN-S-			K	DW-WKVEVN		- DRO -	GEVPA	AYVKKL
YES HUMAN	VALYDYE	AR - TTE DI	SEKKGER	FOIIN-N-			TEG	DW - WEARS	A T	-GKN-	GYLPS	NYVAPA
LSB3 YEAST - GSSPKA	VALYSEA	GE-ESGDI	PERKGDV	ITILK-K-			S DSON	DW-WTGRV		-GRE -	GIEPA	NYVELV
YHA2 YEAST R R V	RALYDLT	TN-EPDEL	SFRKGDV	ITVLE-0-			VY R	DW-WKGALF		- G NM -	GIFPL	NYVTPI
YSC84 YEAST ATPTA	VALYNEA	GE-QPGDL	AFKKGDV	ITILK-K-			S DSON	DW-WTGRTM		-GKE -	GIFPA	NYVRVS
-												
Conservation		dia addi 🦰					_					
	3859491	32-23399	3812542	52553 - 4 -			13 0	2 - 428313		-342 -	+ 6 6 + 3	599311

Figure S1. The multiple sequence alignment of the SH3 domains generated using MUSCLE. The sequences of the eighteen SH3 domains studied in this work are shown at the bottom of the multiple sequence alignment. The alignment is colored based on the sequence conservation (conservation larger than 25%) using the ClustalX coloring scheme. The conservation score reflecting the conservation of physico-chemical properties in the alignment is shown at the bottom of the figure. The figure was generated by Jalview.



Figure S2. The spatial distributions of residues at the twenty-five important positions of the Bio1 SH3 domains (two different orientations). The SH3 domain is shown using the strand model; the peptide is shown using the stick model; the C_{α} carbons of the important residues are shown using the CPK model. The twelve conserved residues are colored in red and the others in green.



Figure S3. The ROC curve for the unified MIEC-SVM model.



Figure S4. Sensitivity (*SE*), specificity (*SP*), accuracy of the binder class (Q_+) and Matthews correlation coefficient (*C*) of the MIEC-SVM versus the weight parameter (k_+) for the binder class in cross validations.



Figure S5. Sensitivity (*SE*), specificity (*SP*), accuracy of the binder class (Q_+) and Matthews correlation coefficient (*C*) versus the weight parameter (k_+) for the binder class in leave-one-SH3-out cross validations.























Figure S6. The distributions of the binding free energies of binders and non-binders for the 16 SH3 domains calculated by MM/GBSA (left) and MM/PBSA (right). Red and green bars represent non-binders and binders, respectively.