

Supplementary information:

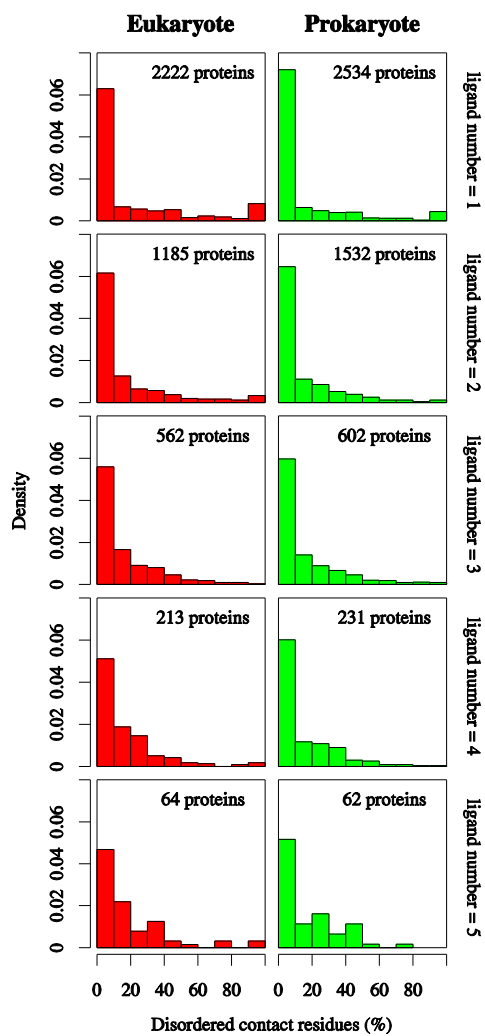


Figure S1. Distribution of disordered contact residues of proteins with 1, 2, 3, 4 and 5 bound ligands in eukaryotes and prokaryotes.

Figure S2. Association of contact residue disorder with GO annotations in ordered single domain proteins and ordered multiple domain proteins. Legend “Single” is ordered single domain proteins and legend “Mutiple” is orderd multiple domain proteins. (A) Eukaryote proteins with one ligand. (B) Prokaryote proteins with one ligand. (C) Eukaryote proteins with more ligands. (D) Prokaryote proteins with more ligands.

Figure S2A. Eukaryote proteins with one ligand.

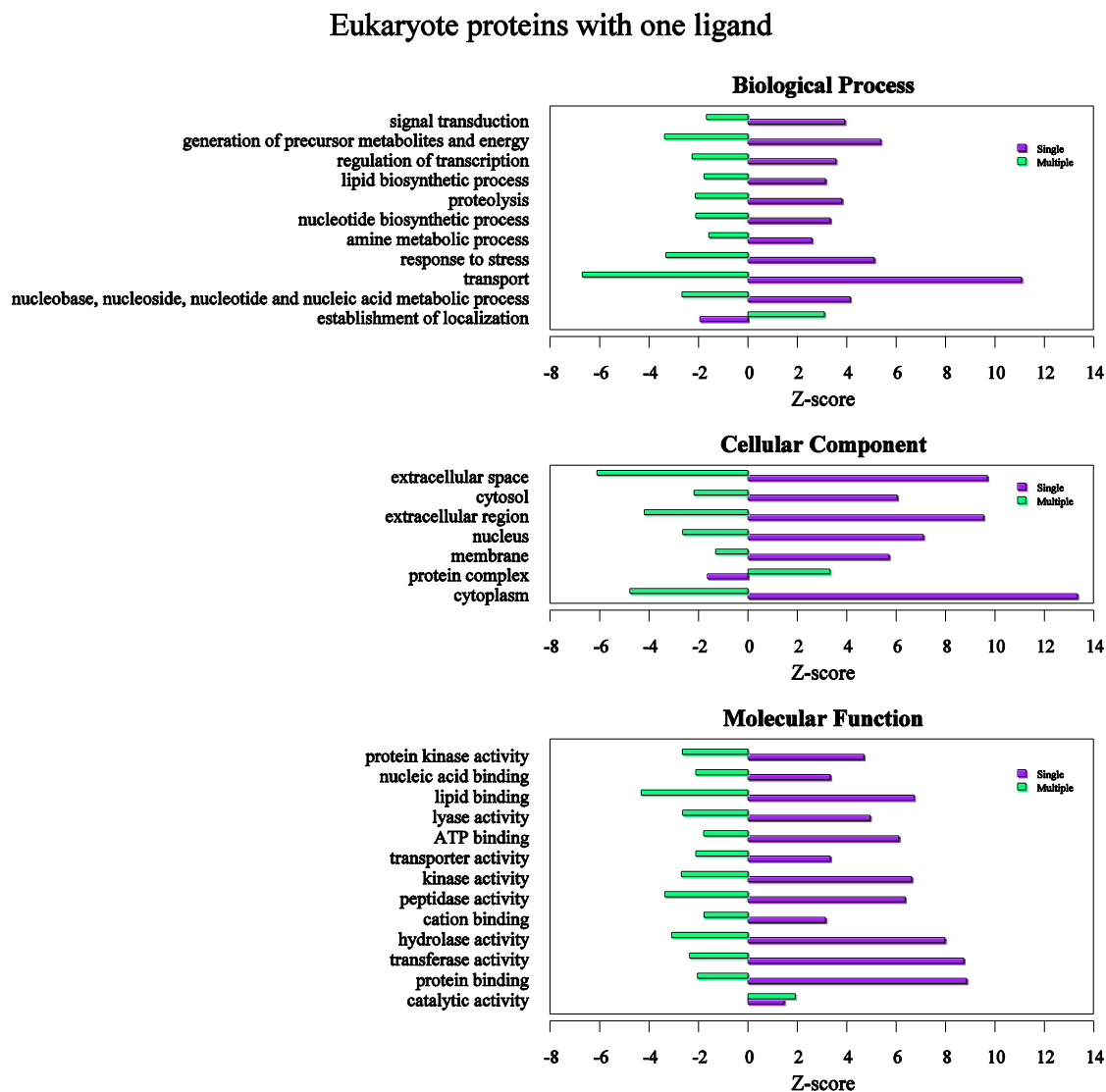


Figure S2B. Prokaryote proteins with one ligand.

Prokaryote proteins with one ligand

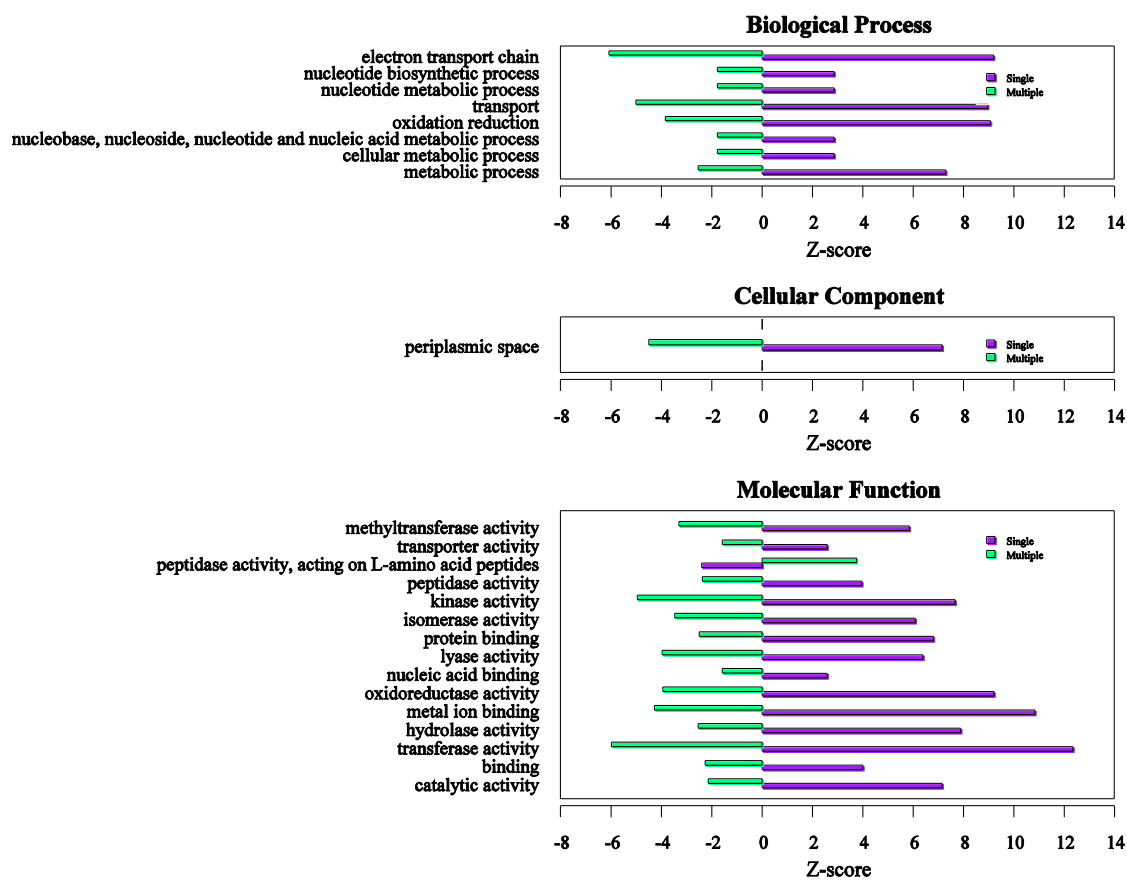


Figure S2C. Eukaryote proteins with more ligands.

Eukaryote proteins with more ligands

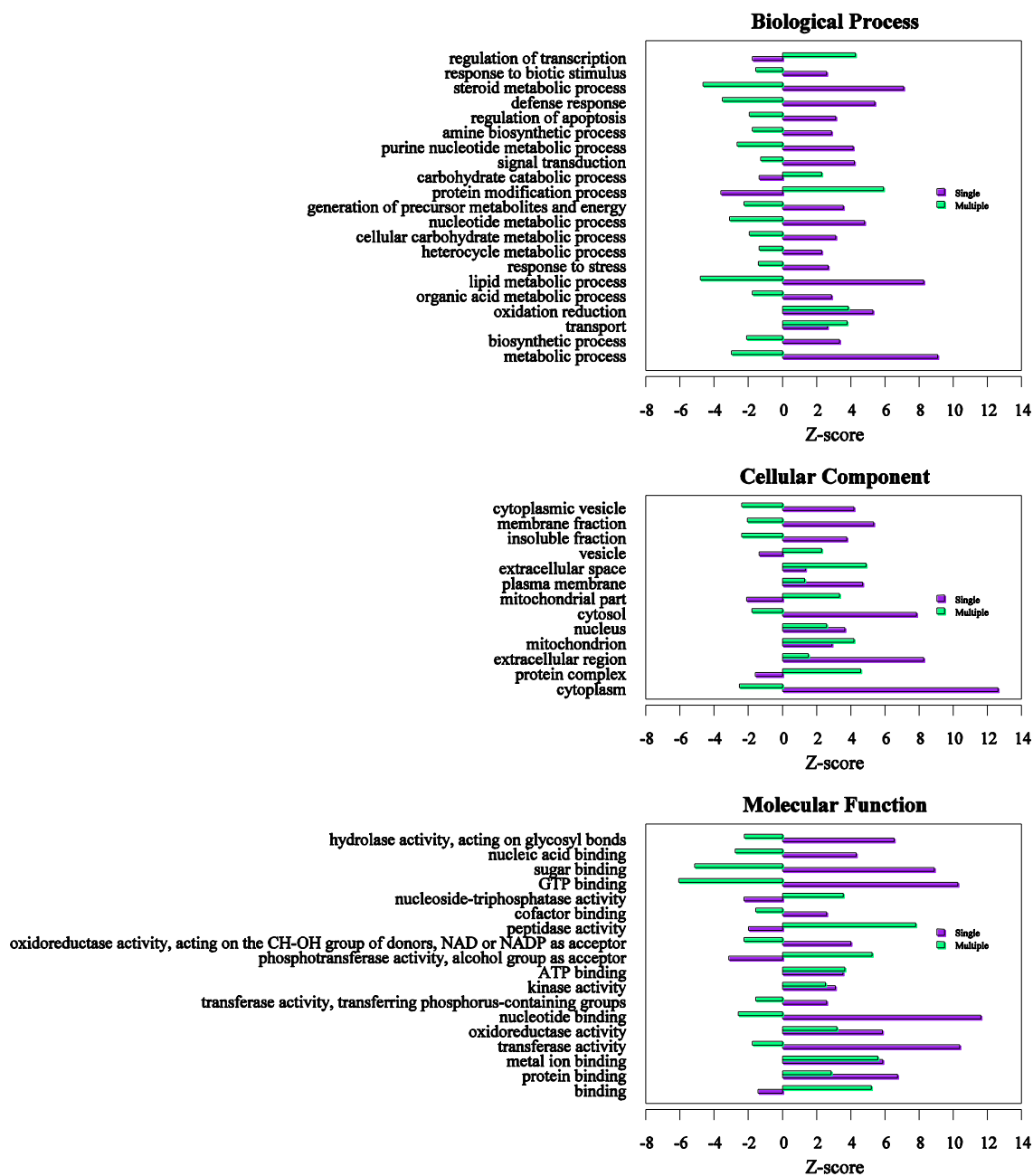
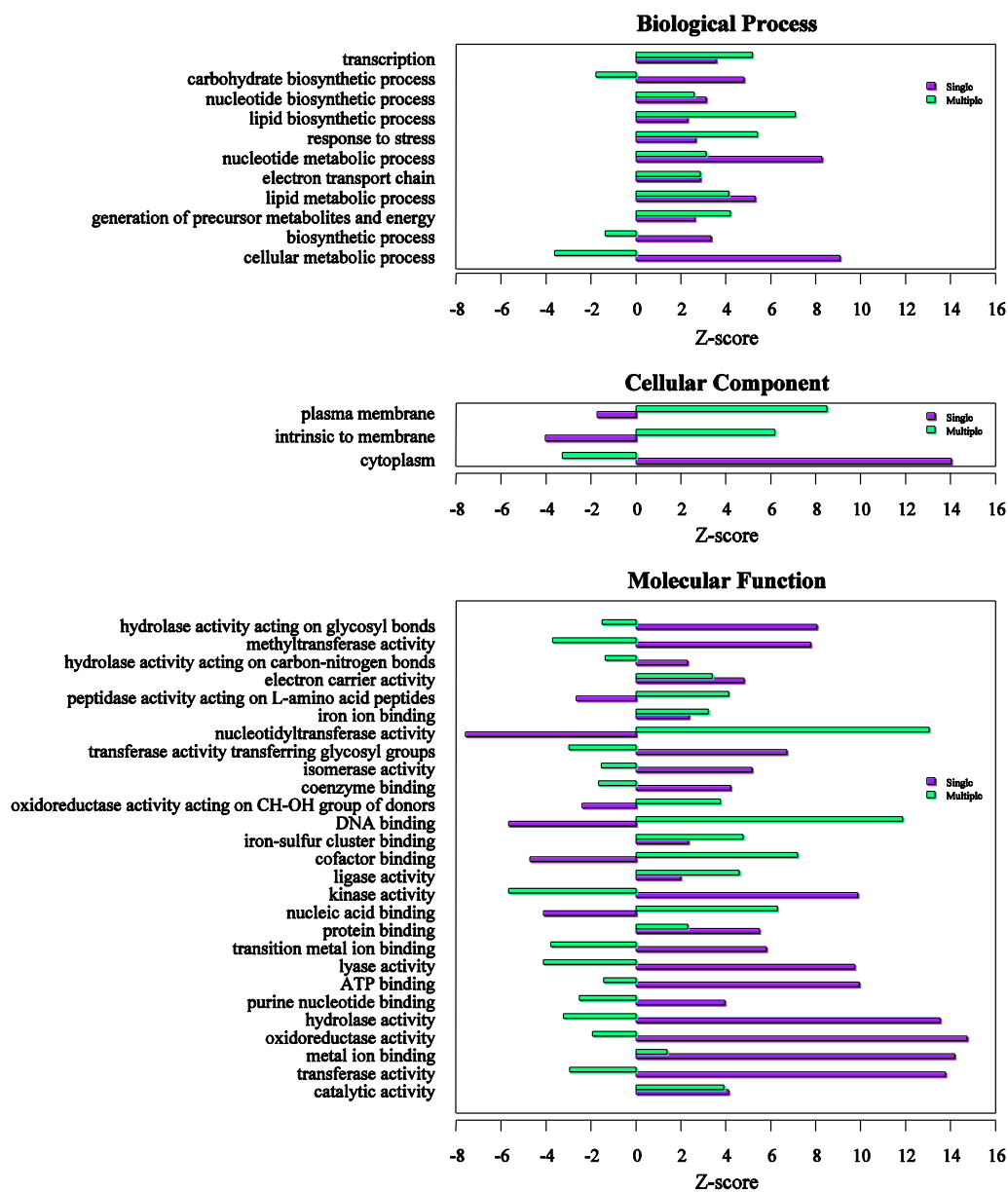
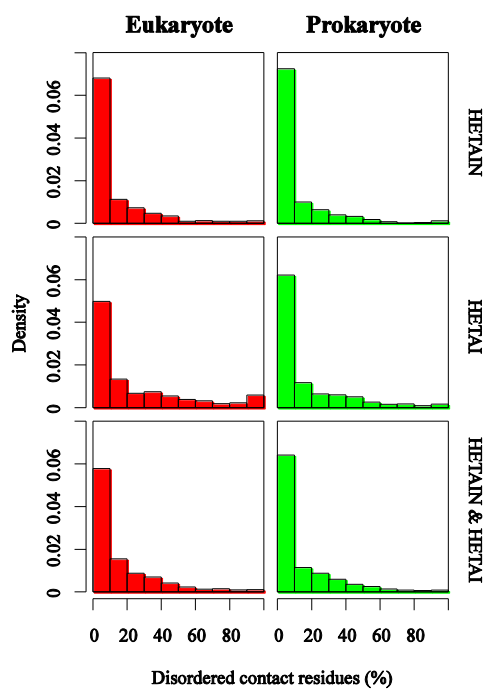


Figure S2D. Prokaryote proteins with more ligands.

Prokaryote proteins with more ligands





FigureS3. Distribution of disordered contact residues of proteins binding HETAIN, HETAI and HETAIN&HETAI in eukaryotes and prokaryotes.



Figure S4. The top 58 superfamily domains with more than 20 proteins.

Table S1A. GO annotations from eukaryotic proteins with one ligand that are significantly over- or under-represented in disordered contact residues (P -value < 0.2). S_zscore and M_zscore are Z-score values on ordered single domain and ordered multiple domain. S_pvalue and M_pvalue are P -values on ordered single domain and ordered multiple domains.

GO	Name	Namespace	S_zscore	M_zscore	S_pvalue	M_pvalue
0051234	establishment of localization	biological process	-1.949	3.103	5.135e-02	1.914e-03
0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	biological process	4.133	-2.661	3.584e-05	7.794e-03
0006810	transport	biological process	11.069	-6.710	1.768e-28	1.941e-11
0006950	response to stress	biological process	5.104	-3.325	3.319e-07	8.846e-04
0009308	amine metabolic process	biological process	2.583	-1.583	9.785e-03	1.133e-01
0009165	nucleotide biosynthetic process	biological process	3.334	-2.109	8.560e-04	3.492e-02
0006508	proteolysis	biological process	3.812	-2.129	1.378e-04	3.327e-02
0008610	lipid biosynthetic process	biological process	3.139	-1.785	1.696e-03	7.429e-02
0045449	regulation of transcription	biological process	3.550	-2.259	3.849e-04	2.387e-02
0006091	generation of precursor metabolites and energy	biological process	5.375	-3.375	7.658e-08	7.382e-04
0007165	signal transduction	biological process	3.913	-1.677	9.111e-05	9.353e-02
0005737	cytoplasm	cellular component	13.337	-4.783	1.406e-40	1.723e-06
0043234	protein complex	cellular component	-1.634	3.317	1.023e-01	9.099e-04
0016020	membrane	cellular component	5.705	-1.308	1.166e-08	1.910e-01
0005634	nucleus	cellular component	7.092	-2.638	1.323e-12	8.330e-03
0005576	extracellular region	cellular component	9.538	-4.199	1.462e-21	2.681e-05
0005829	cytosol	cellular component	6.037	-2.185	1.572e-09	2.886e-02
0005615	extracellular space	cellular component	9.700	-6.118	3.006e-22	9.480e-10
0003824	catalytic activity	molecular function	1.456	1.918	1.454e-01	5.507e-02
0005515	protein binding	molecular function	8.851	-2.045	8.645e-19	4.090e-02
0016740	transferase activity	molecular function	8.742	-2.365	2.284e-18	1.802e-02
0016787	hydrolase activity	molecular function	7.978	-3.079	1.489e-15	2.078e-03
0043169	cation binding	molecular function	3.139	-1.785	1.696e-03	7.429e-02
0008233	peptidase activity	molecular function	6.361	-3.361	2.003e-10	7.763e-04
0016301	kinase activity	molecular function	6.625	-2.688	3.468e-11	7.184e-03
0005215	transporter activity	molecular function	3.334	-2.109	8.560e-04	3.492e-02
0005524	ATP binding	molecular function	6.114	-1.794	9.700e-10	7.285e-02
0016829	lyase activity	molecular function	4.944	-2.634	7.671e-07	8.435e-03
0008289	lipid binding	molecular function	6.728	-4.313	1.718e-11	1.611e-05
0003676	nucleic acid binding	molecular function	3.334	-2.109	8.560e-04	3.492e-02
0004672	protein kinase activity	molecular function	4.695	-2.654	2.668e-06	7.963e-03

Table S1B. GO annotations from prokaryotic proteins with one ligand that are significantly over- or under-represented in disordered contact residues (P -value < 0.2). S_zscore and M_zscore are Z-score values on ordered single domain and ordered multiple domain. S_pvalue and M_pvalue are P -values on ordered single domain and ordered multiple domains.

GO	Name	Namespace	S_zscore	M_zscore	S_pvalue	M_pvalue
0008152	metabolic process	biological process	7.291	-2.547	3.08e-13	1.09e-02
0044237	cellular metabolic process	biological process	2.855	-1.774	4.31e-03	7.60e-02
0006139	nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	biological process	2.855	-1.774	4.31e-03	7.60e-02
0055114	oxidation reduction	biological process	9.07	-3.841	1.20e-19	1.22e-04
0006810	transport	biological process	8.954	-5.017	3.44e-19	5.26e-07
0009117	nucleotide metabolic process	biological process	2.855	-1.774	4.31e-03	7.60e-02
0009165	nucleotide biosynthetic process	biological process	2.855	-1.774	4.31e-03	7.60e-02
0022900	electron transport chain	biological process	9.193	-6.084	3.81e-20	1.17e-09
0042597	periplasmic space	cellular component	7.15	-4.504	8.69e-13	6.67e-06
0003824	catalytic activity	molecular function	7.16	-2.143	8.10e-13	3.21e-02
0005488	binding	molecular function	3.993	-2.261	6.52e-05	2.37e-02
0016740	transferase activity	molecular function	12.348	-5.984	4.98e-35	2.17e-09
0016787	hydrolase activity	molecular function	7.891	-2.537	3.00e-15	1.12e-02
0046872	metal ion binding	molecular function	10.84	-4.283	2.22e-27	1.85e-05
0016491	oxidoreductase activity	molecular function	9.205	-3.945	3.42e-20	7.98e-05
0003676	nucleic acid binding	molecular function	2.583	-1.583	9.79e-03	1.13e-01
0016829	lyase activity	molecular function	6.383	-3.968	1.74e-10	7.25e-05
0005515	protein binding	molecular function	6.8	-2.499	1.04e-11	1.24e-02
0016853	isomerase activity	molecular function	6.089	-3.475	1.14e-09	5.11e-04
0016301	kinase activity	molecular function	7.662	-4.954	1.82e-14	7.25e-07
0008233	peptidase activity	molecular function	3.953	-2.372	7.72e-05	1.77e-02
0070011	peptidase activity, acting on L-amino acids	molecular function	-2.4	3.754	1.64e-02	1.74e-04
0005215	transporter activity	molecular function	2.583	-1.583	9.79e-03	1.13e-01
0008168	methyltransferase activity	molecular function	5.851	-3.301	4.89e-09	9.62e-04

Table S1C. GO annotations from eukaryotic proteins with more ligands that are significantly over- or under-represented in disordered contact residues (P -value < 0.2). S_zscore and M_zscore are Z -score values on ordered single domain and ordered multiple domain. S_pvalue and M_pvalue are P -values on ordered single domain and ordered multiple domains.

GO	Name	Namespace	S_zscore	M_zscore	S_pvalue	M_pvalue
0008152	metabolic process	biological process	9.076	-3.007	1.13e-19	2.64e-03
0009058	biosynthetic process	biological process	3.334	-2.109	8.56e-04	3.49e-02
0006810	transport	biological process	2.62	3.771	8.78e-03	1.63e-04
0055114	oxidation reduction	biological process	5.288	3.841	1.24e-07	1.22e-04
0006082	organic acid metabolic process	biological process	2.855	-1.774	4.31e-03	7.60e-02
0006629	lipid metabolic process	biological process	8.261	-4.821	1.45e-16	1.43e-06
0006950	response to stress	biological process	2.654	-1.429	7.96e-03	1.53e-01
0046483	heterocycle metabolic process	biological process	2.282	-1.369	2.25e-02	1.71e-01
0044262	cellular carbohydrate metabolic process	biological process	3.103	-1.949	1.91e-03	5.14e-02
0009117	nucleotide metabolic process	biological process	4.802	-3.119	1.57e-06	1.82e-03
0006091	generation of precursor metabolites and energy	biological process	3.55	-2.259	3.85e-04	2.39e-02
0006464	protein modification process	biological process	-3.63	5.903	2.84e-04	3.58e-09
0016052	carbohydrate catabolic process	biological process	-1.369	2.282	1.71e-01	2.25e-02
0007165	signal transduction	biological process	4.186	-1.299	2.84e-05	1.94e-01
0006163	purine nucleotide metabolic process	biological process	4.133	-2.661	3.58e-05	7.79e-03
0009309	amine biosynthetic process	biological process	2.855	-1.774	4.31e-03	7.60e-02
0042981	regulation of apoptosis	biological process	3.103	-1.949	1.91e-03	5.14e-02
0006952	defense response	biological process	5.39	-3.519	7.06e-08	4.33e-04
0008202	steroid metabolic process	biological process	7.073	-4.658	1.51e-12	3.19e-06
0009607	response to biotic stimulus	biological process	2.583	-1.583	9.79e-03	1.13e-01
0045449	regulation of transcription	biological process	-1.779	4.262	7.53e-02	2.03e-05
0005737	cytoplasm	cellular component	12.608	-2.541	1.91e-36	1.10e-02
0043234	protein complex	cellular component	-1.598	4.57	1.10e-01	4.87e-06
0005576	extracellular region	cellular component	8.269	1.512	1.35e-16	1.31e-01
0005739	mitochondrion	cellular component	2.891	4.169	3.84e-03	3.06e-05
0005634	nucleus	cellular component	3.635	2.585	2.78e-04	9.74e-03
0005829	cytosol	cellular component	7.848	-1.807	4.21e-15	7.08e-02
0044429	mitochondrial part	cellular component	-2.109	3.334	3.49e-02	8.56e-04
0005886	plasma membrane	cellular component	4.695	1.293	2.67e-06	1.96e-01
0005615	extracellular space	cellular component	1.328	4.878	1.84e-01	1.07e-06
0031982	vesicle	cellular component	-1.369	2.282	1.71e-01	2.25e-02
0005626	insoluble fraction	cellular component	3.754	-2.4	1.74e-04	1.64e-02
0005624	membrane fraction	cellular component	5.333	-2.067	9.67e-08	3.88e-02
0031410	cytoplasmic vesicle	cellular component	4.168	-2.388	3.08e-05	1.69e-02
0005488	binding	molecular function	-1.439	5.203	1.50e-01	1.96e-07
0005515	protein binding	molecular function	6.726	2.84	1.75e-11	4.51e-03
0046872	metal ion binding	molecular function	5.853	5.563	4.82e-09	2.65e-08
0016740	transferase activity	molecular function	10.36	-1.777	3.78e-25	7.56e-02
0016491	oxidoreductase activity	molecular function	5.833	3.177	5.46e-09	1.49e-03
0000166	nucleotide binding	molecular function	11.622	-2.594	3.20e-31	9.49e-03
0016772	transferase activity, transferring phosphorus-containing groups	molecular function	2.583	-1.583	9.79e-03	1.13e-01

0016301	kinase activity	molecular function	3.09	2.51	2.00e-03	1.21e-02
0005524	ATP binding	molecular function	3.531	3.646	4.13e-04	2.67e-04
0016773	phosphotransferase activity, alcohol group as acceptor	molecular function	-3.163	5.244	1.56e-03	1.57e-07
0016616	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	molecular function	3.993	-2.261	6.52e-05	2.37e-02
0008233	peptidase activity	molecular function	-2.001	7.789	4.54e-02	6.74e-15
0048037	cofactor binding	molecular function	2.583	-1.583	9.79e-03	1.13e-01
0017111	nucleoside-triphosphatase activity	molecular function	-2.259	3.55	2.39e-02	3.85e-04
0005525	GTP binding	molecular function	10.259	-6.076	1.08e-24	1.23e-09
0005529	sugar binding	molecular function	8.886	-5.145	6.31e-19	2.68e-07
0003676	nucleic acid binding	molecular function	4.31	-2.782	1.63e-05	5.40e-03
0016798	hydrolase activity, acting on glycosyl bonds	molecular function	6.52	-2.238	7.03e-11	2.52e-02

Table S1D. GO annotations from prokaryotic proteins with more ligand that are significantly over- or under-represented in disordered contact residues (P -value < 0.2). S_zscore and M_zscore are Z -score values on ordered single domain and ordered multiple domain. S_pvalue and M_pvalue are P -values on ordered single domain and ordered multiple domains.

GO	Name	Namespace	S_zscore	M_zscore	S_pvalue	M_pvalue
0044237	cellular metabolic process	biological process	5.169	-2.896	2.350e-07	3.776e-03
0009058	biosynthetic process	biological process	6.276	-3.447	3.483e-10	5.665e-04
0006091	generation of precursor metabolites and energy	biological process	2.583	-1.583	9.785e-03	1.133e-01
0006629	lipid metabolic process	biological process	-1.429	2.654	1.530e-01	7.963e-03
0022900	electron transport chain	biological process	9.739	-1.533	2.052e-22	1.252e-01
0009117	nucleotide metabolic process	biological process	4.644	-3.011	3.420e-06	2.605e-03
0006950	response to stress	biological process	3.993	-2.261	6.515e-05	2.374e-02
0008610	lipid biosynthetic process	biological process	5.803	-2.945	6.516e-09	3.227e-03
0009165	nucleotide biosynthetic process	biological process	7.178	-4.729	7.055e-13	2.258e-06
0016051	carbohydrate biosynthetic process	biological process	-1.369	2.282	1.709e-01	2.248e-02
0006350	transcription	biological process	-5.102	10.378	3.361e-07	3.132e-25
0005737	cytoplasm	cellular component	14.022	-3.282	1.149e-44	1.030e-03
0031224	intrinsic to membrane	cellular component	-4.046	6.168	5.205e-05	6.936e-10
0005886	plasma membrane	cellular component	-1.739	8.496	8.208e-02	1.955e-17
0003824	catalytic activity	molecular function	4.103	3.897	4.075e-05	9.719e-05
0016740	transferase activity	molecular function	13.767	-2.951	4.006e-43	3.171e-03
0046872	metal ion binding	molecular function	14.184	1.372	1.144e-45	1.701e-01
0016491	oxidoreductase activity	molecular function	14.735	-1.949	3.822e-49	5.134e-02
0016787	hydrolase activity	molecular function	13.525	-3.230	1.111e-41	1.240e-03
0017076	purine nucleotide binding	molecular function	3.948	-2.534	7.880e-05	1.128e-02
0005524	ATP binding	molecular function	9.937	-1.442	2.870e-23	1.493e-01
0016829	lyase activity	molecular function	9.721	-4.124	2.444e-22	3.727e-05
0046914	transition metal ion binding	molecular function	5.792	-3.792	6.969e-09	1.496e-04
0005515	protein binding	molecular function	5.478	2.312	4.300e-08	2.080e-02
0003676	nucleic acid binding	molecular function	-4.128	6.288	3.665e-05	3.219e-10
0016301	kinase activity	molecular function	9.861	-5.677	6.166e-23	1.368e-08
0016874	ligase activity	molecular function	1.979	4.592	4.783e-02	4.381e-06
0048037	cofactor binding	molecular function	-4.729	7.178	2.258e-06	7.055e-13
0051536	iron-sulfur cluster binding	molecular function	2.327	4.757	1.999e-02	1.962e-06
0003677	DNA binding	molecular function	-5.679	11.857	1.356e-08	1.985e-32
0016614	oxidoreductase activity, acting on CH-OH group of donors	molecular function	-2.400	3.754	1.638e-02	1.738e-04
0050662	coenzyme binding	molecular function	4.216	-1.667	2.481e-05	9.552e-02
0016853	isomerase activity	molecular function	5.154	-1.549	2.549e-07	1.215e-01
0016757	transferase activity, transferring glycosyl groups	molecular function	6.695	-2.998	2.158e-11	2.717e-03
0016779	nucleotidyltransferase activity	molecular function	-7.604	13.051	2.872e-14	6.298e-39
0005506	iron ion binding	molecular function	2.347	3.222	1.890e-02	1.272e-03
0070011	peptidase activity, acting on L-amino acid peptides	molecular function	-2.661	4.133	7.794e-03	3.584e-05
0009055	electron carrier activity	molecular function	4.797	3.389	1.611e-06	7.014e-04
0016810	hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds	molecular function	2.282	-1.369	2.248e-02	1.709e-01
0008168	methyltransferase activity	molecular function	7.775	-3.713	7.528e-15	2.046e-04

0016798	hydrolase activity, acting on glycosyl bonds	molecular function	8.050	-1.518	8.270e-16	1.290e-01
---------	---	--------------------	-------	--------	-----------	-----------

Table S2. Frequency of amino acids on the whole proteins, the contact residues and the hub contact residues

Residues	Frequency in protein residues	Frequency in contact residues	Frequency in hub contact residues
I	5.886	4.133	2.183
V	7.149	4.565	2.217
L	9.247	6.279	3.596
F	4.181	4.905	4.412
C	1.431	2.986	3.297
M	2.070	2.078	1.608
A	8.242	5.416	2.861
G	7.554	9.225	5.262
T	5.362	5.580	6.216
S	5.688	6.162	5.262
W	1.373	2.349	3.309
Y	3.506	5.108	6.491
P	4.639	2.895	1.321
H	2.517	6.263	13.373
N	4.157	5.061	5.641
Q	3.564	2.965	2.803
D	5.899	6.620	11.776
E	6.697	5.221	7.261
K	5.760	5.007	4.412
R	5.072	7.186	6.698

Table S3. Scop domains significantly enriched in proteins with HETAIN, HETAI and HETAIN&HETAI. (*P*-value < 0.2)

Scop domain	Name	HN_zscore	HI_zscore	HT&HI_zscore	HN_pvalue	HI_pvalue	HT&HI_pvalue
46458	Globin-like	5.978	-7.048	-2.396	2.254E-09	1.810E-12	1.658E-02
48508	Nuclear receptor ligand-binding domain	5.604	-4.463	-3.456	2.093E-08	8.075E-06	5.473E-04
47616	GST C-terminal domain-like	3.116	-3.886	-1.435	1.836E-03	1.020E-04	1.512E-01
46626	Cytochrome c	4.015	-4.513	-1.670	5.938E-05	6.399E-06	9.491E-02
47473	EF-hand	-4.078	4.944	-2.995	4.552E-05	7.671E-07	2.744E-03
48726	Immunoglobulin	-1.830	-3.607	-7.338	6.725E-02	8.587E-03	8.587E-03
49899	Concanavalin A-like lectins/glucanases	-5.185	2.924	-4.448	2.160E-07	3.453E-03	8.674E-06
81296	E set domains	-3.019	1.339	-2.086	2.532E-03	1.807E-01	3.701E-02
49503	Cupredoxins	-4.189	3.370	-1.669	2.801E-05	7.512E-04	9.505E-02
51011	Glycosyl hydrolase domain	-4.034	2.098	-2.743	5.475E-05	3.592E-02	6.081E-03
50814	Lipocalins	3.466	-4.236	-1.435	5.289E-04	2.277E-05	1.512E-01
50129	GroES-like	-2.362	-2.762	3.243	1.818E-02	5.741E-03	1.184E-03
51735	NAD(P)-binding Rossmann-fold domains	3.689	-7.865	-1.879	2.250E-04	3.678E-15	6.020E-02
53383	PLP-dependent transferases	4.514	-5.245	-2.562	6.376E-06	1.559E-07	1.041E-02
53335	S-adenosyl-L-methionine-dependent methyltransferases	-2.405	-2.912	-3.165	1.616E-02	3.596E-03	1.552E-03
52833	Thioredoxin-like	2.472	-2.528	-2.806	1.343E-02	1.148E-02	5.023E-03
53067	Actin-like ATPase domain	-3.028	-2.007	1.395	2.463E-03	4.473E-02	1.631E-01
51569	Aldolase	-2.485	1.365	-1.435	1.294E-02	1.722E-01	1.512E-01
53187	Zn-dependent exopeptidases	-3.653	3.418	-1.689	2.588E-04	6.316E-04	9.118E-02
52518	Thiamin diphosphate-binding fold (THDP-binding)	-3.163	-3.563	4.844	1.564E-03	3.668E-04	1.273E-06
51395	FMN-linked oxidoreductases	4.384	-3.703	-2.426	1.165E-05	2.131E-04	1.526E-02
54452	MHC antigen-recognition domain	-2.550	-3.543	-4.868	1.078E-02	3.954E-04	1.129E-06
56436	C-type lectin-like	-3.536	3.466	-2.135	4.067E-04	5.289E-04	3.273E-02
56059	Glutathione synthetase ATP-binding domain-like	-1.575	-1.575	1.405	1.153E-01	1.153E-01	1.601E-01
55424	FAD/NAD-linked reductases, dimerisation (C-terminal) domain	4.944	-3.519	-3.074	7.641E-07	4.333E-04	2.116E-03