

# Supporting Information for

## PASSer: Prediction of Allosteric Sites Server

Hao Tian,<sup>†</sup> Xi Jiang,<sup>‡</sup> and Peng Tao<sup>\*,†</sup>

*<sup>†</sup>Department of Chemistry, Center for Research Computing, Center for Drug Discovery, Design, and Delivery (CD4), Southern Methodist University, Dallas, Texas, United States of America*

*<sup>‡</sup>Department of Statistical Science, Southern Methodist University, Dallas, Texas, United States of America*

E-mail: [ptao@smu.edu](mailto:ptao@smu.edu)

**Table S1.** Summarization of 90 selected proteins in alphabetical order of PDB ID.

No.	Protein name	PDB ID
1	GLUTAMINE PHOSPHORIBOSYLPYROPHOSPHATE AMIDOTRANSFERASE	1AO0
2	PROTEIN (CARBAMOYL-PHOSPHATE SYNTHASE)	1CE8
3	PROTEIN (GLYCEROL-3-PHOSPHATE CYTIDYLYLTRANSFERASE)	1COZ
4	VITAMIN D NUCLEAR RECEPTOR	1DB1
5	INDUCIBLE NITRIC OXIDE SYNTHASE	1DD7
6	LAC REPRESSOR	1EFA
7	CYTOCHROME P450ERYF	1EGY
8	PANTOTHENATE KINASE	1ESM
9	FK506-BINDING PROTEIN	1FAP
10	FRUCTOSE-1,6-BISPHOSPHATASE	1FTA
11	ANNEXIN V	1HAK
12	ANTHRANILATE SYNTHASE	1I7S
13	groEL protein	1KP8
14	L-LACTATE DEHYDROGENASE	1LDN
15	L-LACTATE DEHYDROGENASE (T- AND R- STATE TETRAMER COMPLEX)	1LTH
16	Ribonucleoside-diphosphate reductase 2 alpha chain	1PEQ
17	PHOSPHOFRUCTOKINASE	1PFK
18	NAD-dependent malic enzyme, mitochondrial	1PJ3
19	ACETYLCHOLINESTERASE	1QTI
20	Parathion hydrolase	1QW7
21	Integrin alpha-L	1RD4
22	DIHYDROFOLATE REDUCTASE	1RX2
23	Dual specificity mitogen-activated protein kinase kinase 1	1S9J
24	Caspase-7	1SHJ
25	Protein-tyrosine phosphatase, non-receptor type 1	1T49
26	Aspartate carbamoyltransferase catalytic chain	1TUG
27	glucokinase isoform 2	1V4S
28	STALKED-CELL DIFFERENTIATION CONTROLLING PROTEIN	1W25
29	ACETYL-COENZYME A CARBOXYLASE	1W96

Continued on next page

**Table S1 – continued from previous page**

No.	Protein name	PDB ID
30	Kinesin-like protein KIF11	1X88
31	ribonucleotide reductase, B12-dependent	1XJE
32	Retinoic acid receptor RXR-alpha	1XLS
33	Probable uracil phosphoribosyltransferase	1XTU
34	Glycogen phosphorylase, muscle form	1Z8D
35	Copper-containing nitrite reductase	1ZDS
36	Glutamate receptor 2	2AL4
37	URIDYLATE KINASE	2BND
38	PYRUVATE DEHYDROGENSAE KINASE ISOENZYME 2	2BU2
39	THREONINE SYNTHASE 1, CHLOROPLASTIC	2C2B
40	Tryptophan synthase alpha chain	2CLH
41	Hemoglobin alpha subunit	2D5Z
42	BirA bifunctional protein	2EWN
43	Epidermal growth factor receptor	2GS7
44	D-alanine-D-alanine ligase	2I80
45	Mevalonate kinase	2OI2
46	isomerase domain of glutamine-fructose-6-phosphate transaminase (isomerizing)	2POC
47	PROTEIN (PURINE REPRESSOR)	2PUC
48	Phenylpyruvate decarboxylase	2Q5O
49	Pyruvate carboxylase protein	2QF7
50	RIBONUCLEOTIDE REDUCTASE R1 PROTEIN	2R1R
51	5'-AMP-ACTIVATED PROTEIN KINASE CATALYTIC SUBUNIT ALPHA-1	2V92
52	PYRUVATE KINASE ISOZYMES R/L	2VGI
53	METHIONYL-TRNA SYNTHETASE	2X1L
54	MYOSIN-2 HEAVY CHAIN	2XO8
55	2-C-METHYL-D-ERYTHRITOL 4-PHOSPHATE CYTIDYLYLTRANSFERASE, CHLOROPLASTIC	2YC3
56	cAMP and cAMP-inhibited cGMP 3',5'-cyclic phosphodiesterase 10A	2ZMF
57	POL polyprotein	3AO1
58	UDP-N-acetylglucosamine 2-epimerase	3BEO

Continued on next page

**Table S1 – continued from previous page**

No.	Protein name	PDB ID
59	6-phosphofructokinase isozyme 2	3CQD
60	Isocitrate dehydrogenase kinase/phosphatase	3EPS
61	Proto-oncogene tyrosine-protein kinase Src	3F3V
62	Alpha-isopropylmalate synthase	3F6G
63	2-isopropylmalate synthase	3FIG
64	RTX toxin RtxA	3FZY
65	Mitogen-activated protein kinase 14	3GCP
66	Pyruvate kinase isozymes M1/M2	3GR4
67	Tyrosine-protein kinase ABL2	3GVU
68	Casein kinase II subunit alpha	3H30
69	Pyruvate carboxylase	3HO8
70	Protein FimX	3HV8
71	Reverse transcriptase/ribonuclease H	3I0R
72	cAMP-specific 3',5'-cyclic phosphodiesterase 4D	3IAD
73	cAMP-dependent protein kinase catalytic subunit alpha	3IDB
74	Macrophage migration inhibitory factor	3IJG
75	HD domain protein	3IRH
76	DNA-directed RNA polymerase subunit alpha	3IYD
77	Pantothenate kinase 3	3MK6
78	Suppressor of kinetochore protein 1	3MKS
79	ORF 17	3NJQ
80	Mitogen-activated protein kinase 8	3O2M
81	Estrogen receptor	3OS8
82	Toxin B	3PEE
83	UDP-glucose 6-dehydrogenase	3PJG
84	UDP-glucose 6-dehydrogenase	3PTZ
85	NMDA glutamate receptor subunit	3QEL
86	Ubiquitin-conjugating enzyme E2 R1	3RZ3
87	Glutaminase kidney isoform, mitochondrial	3UO9

Continued on next page

**Table S1 – continued from previous page**

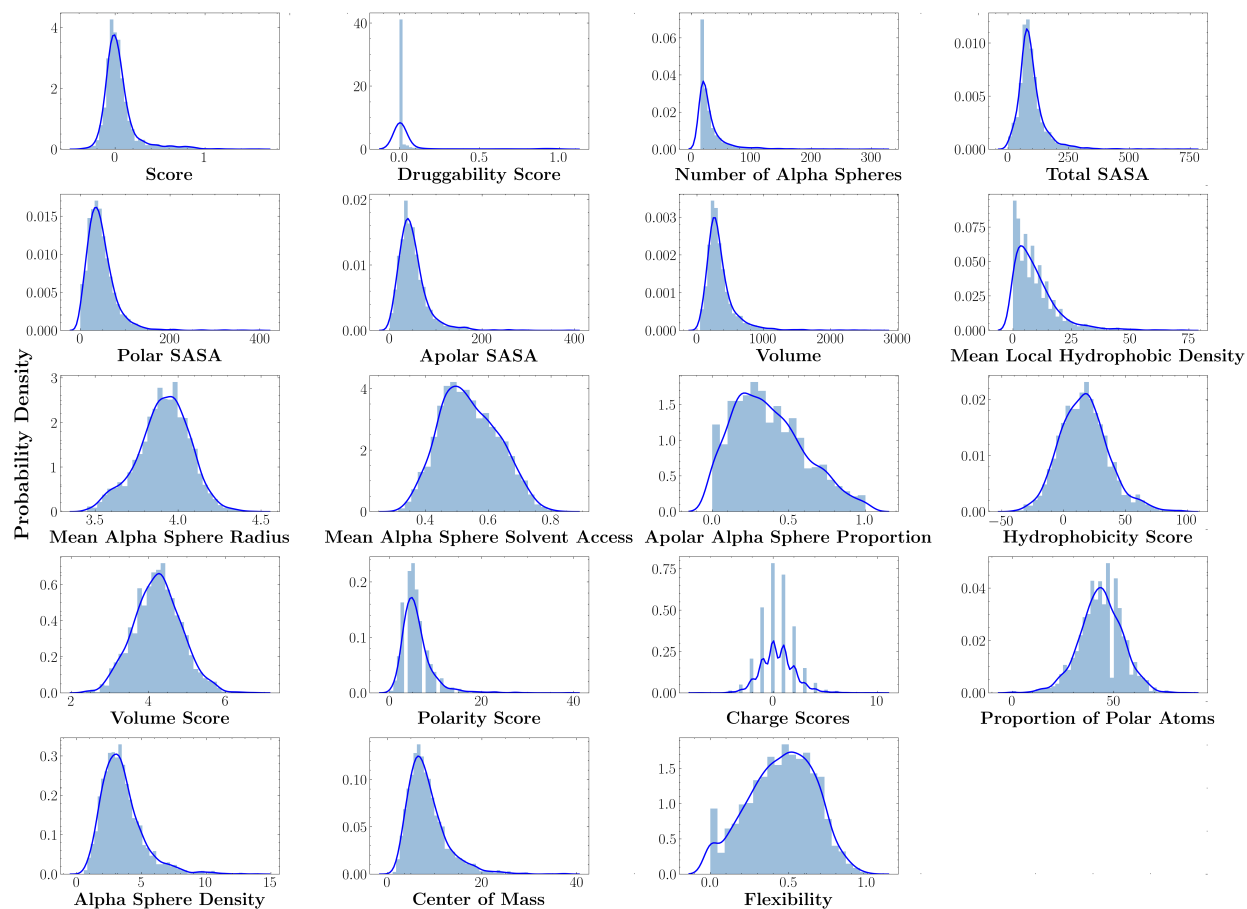
No.	Protein name	PDB ID
88	GLUCOSE-1-PHOSPHATE THYMIDYLYLTRANSFERASE	3ZLK
89	LYSINE ACETYLTRANSFERASE	4AVB
90	CHAPERONE PROTEIN DNAK	4B9Q

**Table S2.** 19 site descriptors calculated by FPocket.

No.	Description
1	Score
2	Druggability score
3	Number of alpha spheres
4	Total solvent-accessible surface area (SASA)
5	Polar SASA
6	Apolar SASA
7	Volume
8	Mean local hydrophobic density
9	Mean alpha sphere radius
10	Mean alpha sphere solvent access
11	Apolar alpha sphere proportion
12	Hydrophobicity score
13	Volume score
14	Polarity score
15	Charge scores
16	Proportion of polar atoms
17	Alpha sphere density
18	Centre of mass - alpha sphere max distance
19	Flexibility

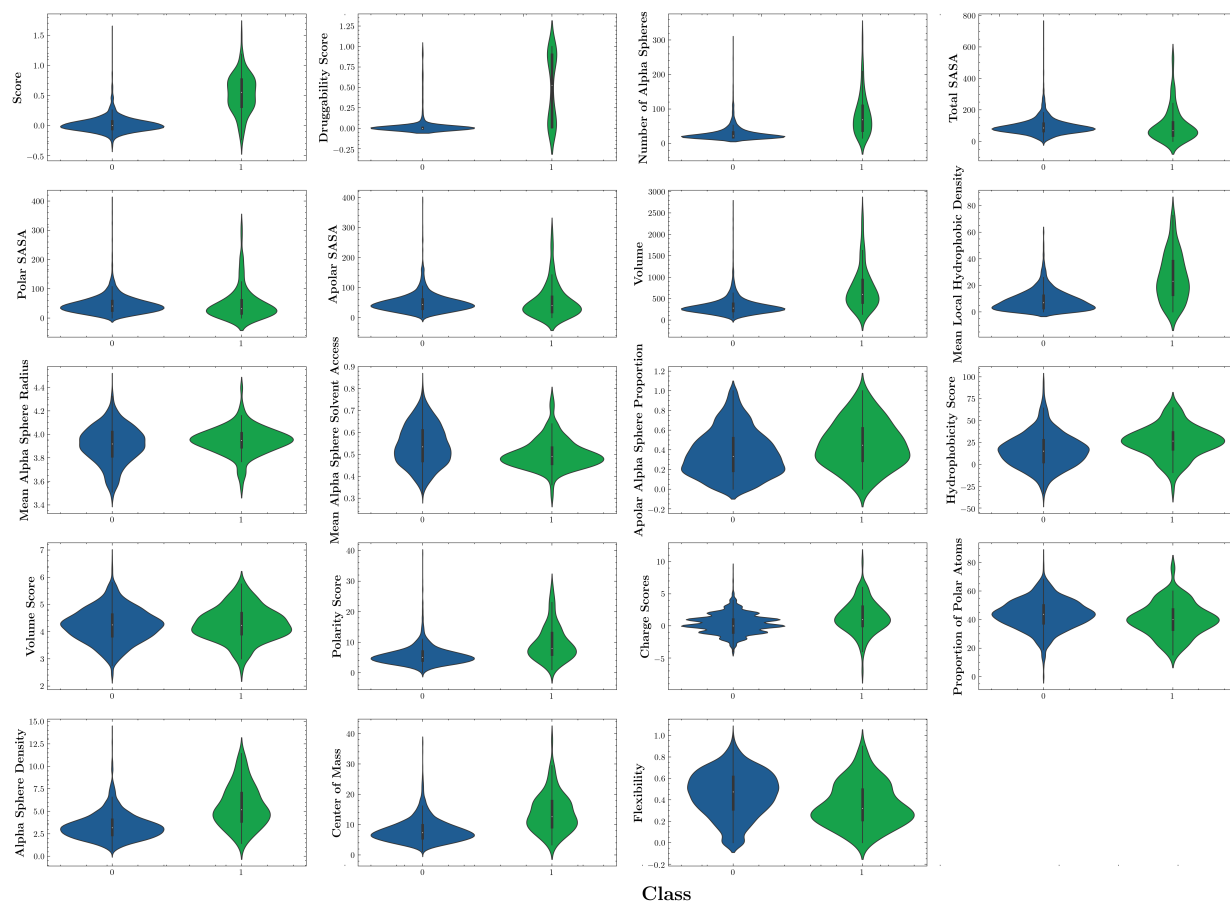
**Table S3.** The correlations between the original features and two additional ranking feature sets.

Feature No.	Inner ranking	Overall ranking
1	-0.7	-0.8
2	-0.5	-0.5
3	-0.6	-0.7
4	-0.8	-0.8
5	-0.8	-0.8
6	-0.8	-0.8
7	-0.7	-0.8
8	-0.8	-0.9
9	-0.9	-1.0
10	-0.9	-1.0
11	-0.9	-1.0
12	-0.9	-1.0
13	-0.9	-1.0
14	-0.8	-0.8
15	-0.9	-0.9
16	-0.9	-1.0
17	-0.9	-0.9
18	-0.9	-0.9
19	-0.6	-1.0

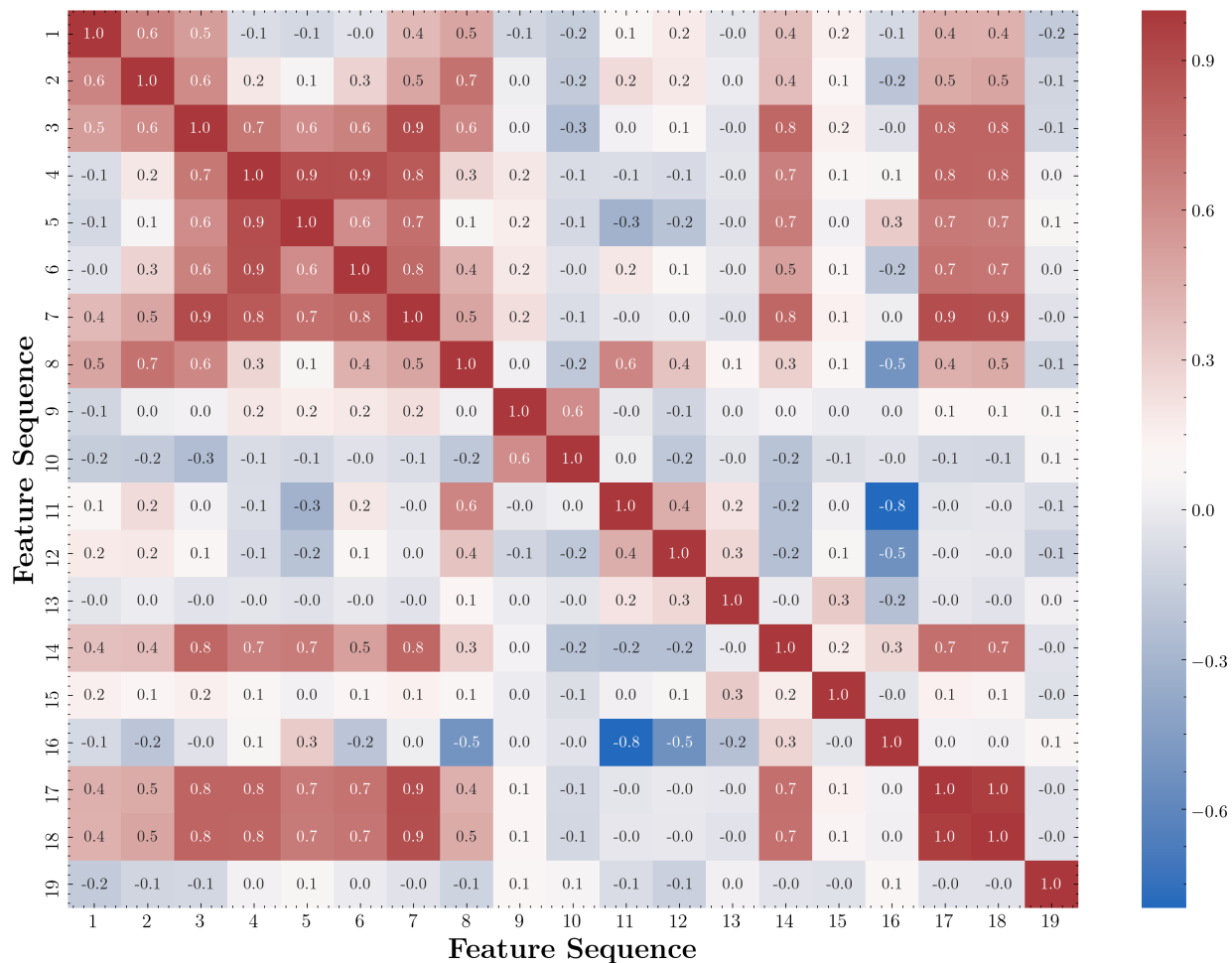


**Figure S1.** Feature value distribution. A histogram plot and a kernel density estimate (blue line) is shown for each feature calculated by FPocket.





**Figure S2.** Violin plots. The distribution of allosteric pockets (class 1) and non-allosteric pockets (class 0) with probability density is plotted for each feature calculated by FPocket.



**Figure S3.** The correlation matrix of 19 features. The numbers in feature sequence correspond to the numbers in Table S2. The correlation matrix is symmetrical along the diagonal. White color represents no correlation between the two related features while red and blue represent positive and negative correlations, respectively. The darker the color, the greater the correlation.