

ARTICLE TYPE

Cite this: DOI: 10.1039/xxxxxxxxxx

Supplementary Information for Na⁺-binding modes involved in thrombin's allosteric response as revealed by molecular dynamics simulations, correlation networks and Markov modeling

Jiajie Xiao,^{a,b} and Freddie R. Salsbury, Jr^{*a}

Received Date

Accepted Date

DOI: 10.1039/xxxxxxxxxx

www.rsc.org/journalname

^a Department of Physics, Wake Forest University, Winston Salem, NC, USA. Fax: +1 (336) 758-6142; Tel: +1 (336) 758-4975; E-mail: salsbufr@wfu.edu

^b Department of Computer Science, Wake Forest University, Winston Salem, NC, USA.

Table S1 Na⁺-binding statistics for Na⁺-residue contacts. ResID is followed by the sequential residue numbering scheme for thrombin in literature^{1,2}. The statistics were calculated over all Na⁺ ions and runs. The present table was sorted by the fourth column to highlight residues contributing kinetically stable Na⁺-binding. The NA values in the table stand for not available, indicating no unbinding event was detected within each 0.1 ns time step.

ResID (sequential numbering scheme)	ResNum (chymotrypsin numbering scheme)	Occupancy	Mean Residence Time (ns)	Standard errors in residence time (ns)	Number of Na ⁺ -binding events
235	ASP189	0.56	31.2	0.2	356
236	ALA190	0.55	20.5	0.2	336
272	LYS224	0.04	9.1	0.3	232
269	ARG221	0.04	6.5	0.2	457
213	THR172	0.02	4.6	0.2	155
210	LYS169	0.02	3.8	0.2	145
268	ASP221A	0.09	3.4	0.1	588
238	GLU192	0.07	2.7	0.1	199
94	ASP63	0.06	2.5	0.0	196
92	GLU61	0.04	2.2	0.1	214
27	GLU14E	0.07	2.1	0.0	432
61	GLU39	0.01	2.1	0.1	139
71	ASP49	0.04	2.0	0.1	262
241	SER195	0.00	1.8	0.2	20
234	GLY188	0.00	1.7	0.2	47
8	ASP1A	0.09	1.6	0.0	396
205	GLU164	0.05	1.5	0.0	228
79	HSD57	0.00	1.3	0.1	24
30	GLU14H	0.15	1.3	0.0	909
230	GLU186B	0.09	1.3	0.0	559
87	ASP60E	0.18	1.3	0.0	856
118	GLU86	0.08	1.3	0.0	689
42	ASP21	0.13	1.3	0.0	899
291	ASP243	0.15	1.3	0.0	926
39	GLU18	0.17	1.3	0.0	910
31	SER14I	0.03	1.3	0.1	363
295	GLU247	0.35	1.3	0.0	1802
158	ASP125	0.03	1.2	0.0	186
290	ILE242	0.00	1.2	0.1	59
270	ASP222	0.12	1.2	0.0	928
130	GLU97A	0.04	1.2	0.0	541
262	SER214	0.00	1.1	0.1	42
211	ASP170	0.14	1.1	0.0	958
34	ASP15	0.28	1.1	0.0	1773
44	GLU23	0.08	1.1	0.0	726
273	TYR225	0.01	1.1	0.1	188
36	ARG17	0.33	1.0	0.0	2150
22	ASP14	0.02	1.0	0.1	376
229	ASP186A	0.12	1.0	0.0	1001
149	ASP116	0.01	1.0	0.0	140
16	GLU8	0.08	1.0	0.0	635
265	GLU217	0.02	1.0	0.0	354
6	GLU0	0.09	0.9	0.0	727
219	ASP178	0.02	0.9	0.0	197

167	GLN131	0.03	0.8	0.0	679
266	GLY219	0.26	0.8	0.0	4471
160	GLU127	0.09	0.8	0.0	782
21	GLU13	0.11	0.8	0.0	1232
33	ILE14K	0.02	0.7	0.0	561
264	GLY216	0.01	0.7	0.0	191
182	GLU146	0.03	0.6	0.0	328
166	LEU130	0.00	0.5	0.1	76
156	LEU123	0.00	0.5	0.1	3
108	GLU77	0.00	0.5	0.0	109
293	PHE245	0.01	0.4	0.0	332
163	ALA129A	0.01	0.4	0.0	147
294	GLY246	0.01	0.4	0.0	320
25	GLU14C	0.01	0.4	0.0	215
207	PRO166	0.00	0.3	0.0	131
161	THR128	0.01	0.3	0.0	159
3	GLY-3	0.01	0.3	0.0	199
271	GLY223	0.01	0.3	0.0	199
212	SER171	0.01	0.3	0.0	191
183	THR147	0.01	0.3	0.0	253
72	ARG50	0.02	0.3	0.0	579
197	GLN156	0.02	0.3	0.0	689
69	ILE47	0.00	0.3	0.1	2
35	GLY16	0.07	0.3	0.0	1653
40	GLY19	0.03	0.3	0.0	694
203	ILE162	0.00	0.3	0.0	7
93	ASN62	0.02	0.3	0.0	401
70	SER48	0.00	0.2	0.0	141
250	PRO204	0.01	0.2	0.0	189
292	GLN244	0.02	0.2	0.0	675
140	LYS107	0.04	0.2	0.0	1169
228	PRO186	0.01	0.2	0.0	431
32	TYR14J	0.01	0.2	0.0	454
4	SER-2	0.01	0.2	0.0	351
170	TYR134	0.01	0.2	0.0	337
102	HSD71	0.00	0.2	0.0	127
1	THR-5	0.01	0.2	0.0	381
7	ALA1B	0.00	0.2	0.0	148
29	LEU14G	0.00	0.2	0.0	101
226	TYR184A	0.00	0.2	0.0	73
253	ASN205	0.01	0.2	0.0	212
84	PRO60B	0.04	0.2	0.0	1020
186	ALA147C	0.00	0.2	0.0	134
91	THR60I	0.02	0.2	0.0	635
267	CYS220	0.06	0.2	0.0	1917
263	TRP215	0.00	0.2	0.0	124
142	LYS109	0.00	0.2	0.0	172
168	ALA132	0.01	0.2	0.0	191
133	ASP100	0.01	0.2	0.0	248
2	PHE-4	0.01	0.2	0.0	199
233	ARG187	0.04	0.2	0.0	1483
19	SER11	0.03	0.2	0.0	934
164	SER129B	0.01	0.2	0.0	317
43	ALA22	0.00	0.2	0.0	205

85	PRO60C	0.02	0.2	0.0	677
278	HSD230	0.00	0.2	0.0	18
187	ASN147D	0.01	0.1	0.0	209
131	ASN98	0.00	0.1	0.0	156
89	ASN60G	0.01	0.1	0.0	380
185	THR147B	0.00	0.1	0.0	157
184	TRP147A	0.01	0.1	0.0	265
231	GLY186C	0.01	0.1	0.0	282
248	LYS202	0.00	0.1	0.0	54
191	GLY147H	0.00	0.1	0.0	40
147	PHE114	0.01	0.1	0.0	210
232	LYS186D	0.01	0.1	0.0	264
5	GLY-1	0.01	0.1	0.0	276
252	ASN204B	0.00	0.1	0.0	98
23	LYS14A	0.02	0.1	0.0	928
74	VAL52	0.00	0.1	0.0	2
274	GLY226	0.02	0.1	0.0	992
41	SER20	0.01	0.1	0.0	551
107	TYR76	0.00	0.1	0.0	173
17	LYS9	0.01	0.1	0.0	405
110	ASN78	0.00	0.1	0.0	117
190	LYS147G	0.00	0.1	0.0	121
192	GLN151	0.00	0.1	0.0	75
171	LYS135	0.00	0.1	0.0	107
251	PHE204A	0.01	0.1	0.0	427
105	THR74	0.00	0.1	0.0	106
124	PRO92	0.02	0.1	0.0	628
127	ASN95	0.01	0.1	0.0	553
283	LYS235	0.00	0.1	0.0	111
194	SER153	0.00	0.1	0.0	114
227	LYS185	0.01	0.1	0.0	227
129	ARG97	0.02	0.1	0.0	832
152	HSD119	0.00	0.1	0.0	14
126	TYR94	0.01	0.1	0.0	463
195	VAL154	0.00	0.1	0.0	173
20	LEU12	0.01	0.1	0.0	609
214	ARG173	0.01	0.1	0.0	502
18	LYS10	0.01	0.1	0.0	627
26	ARG14D	0.02	0.1	0.0	861
261	VAL213	0.00	0.1	0.0	6
169	GLY133	0.00	0.1	0.0	64
104	ARG73	0.00	0.1	0.0	66
180	LEU144	0.00	0.1	0.0	52
56	ARG35	0.00	0.1	0.0	199
217	ILE176	0.00	0.1	0.0	49
287	GLN239	0.00	0.1	0.0	98
24	THR14B	0.01	0.1	0.0	268
60	GLN38	0.01	0.1	0.0	362
196	LEU155	0.00	0.1	0.0	71
125	ARG93	0.02	0.1	0.0	894
120	ILE88	0.00	0.1	0.0	118
181	LYS145	0.00	0.1	0.0	95
143	LYS110	0.00	0.1	0.0	139
188	VAL147E	0.00	0.1	0.0	210

86	TRP60D	0.00	0.1	0.0	193
279	VAL231	0.00	0.1	0.0	9
216	ARG175	0.01	0.1	0.0	364
200	ASN159	0.00	0.1	0.0	49
220	ASN179	0.00	0.1	0.0	69
215	ILE174	0.00	0.1	0.0	91
123	HSD91	0.01	0.1	0.0	382
80	CYS58	0.00	0.1	0.0	10
96	LEU65	0.00	0.1	0.0	32
132	LEU99	0.00	0.1	0.0	85
128	TRP96	0.01	0.1	0.0	445
15	PHE7	0.01	0.1	0.0	342
106	ARG75	0.01	0.1	0.0	299
288	LYS240	0.01	0.1	0.0	387
117	LEU85	0.00	0.1	0.0	11
281	ARG233	0.00	0.1	0.0	148
146	ALA113	0.00	0.1	0.0	81
138	LEU105	0.00	0.1	0.0	46
173	ARG137	0.00	0.1	0.0	72
101	LYS70	0.00	0.1	0.0	24
134	ARG101	0.01	0.1	0.0	365
159	ARG126	0.01	0.1	0.0	233
237	CYS191	0.00	0.1	0.0	37
285	TRP237	0.01	0.1	0.0	286
57	LYS36	0.00	0.1	0.0	188
58	SER36A	0.01	0.1	0.0	343
103	SER72	0.00	0.1	0.0	80
59	PRO37	0.01	0.1	0.0	309
284	LYS236	0.00	0.1	0.0	218
249	SER203	0.00	0.1	0.0	29
218	THR177	0.00	0.1	0.0	73
145	VAL112	0.00	0.1	0.0	44
150	TYR117	0.00	0.1	0.0	109
122	ILE90	0.01	0.1	0.0	276
144	PRO111	0.00	0.1	0.0	108
119	LYS87	0.01	0.1	0.0	273
82	LEU60	0.00	0.1	0.0	228
47	MET26	0.00	0.1	0.0	141
45	ILE24	0.01	0.1	0.0	248
14	LEU6	0.00	0.1	0.0	217
62	LEU40	0.00	0.1	0.0	19
137	ALA104	0.00	0.1	0.0	36
153	PRO120	0.00	0.1	0.0	18
121	TYR89	0.00	0.1	0.0	220
90	PHE60H	0.00	0.1	0.0	74
254	ARG206	0.00	0.1	0.0	169
255	TRP207	0.00	0.1	0.0	19
189	GLY147F	0.00	0.1	0.0	77
55	PHE34	0.00	0.1	0.0	44
276	TYR228	0.00	0.1	0.0	22
136	ILE103	0.00	0.1	0.0	70
111	ILE79	0.00	0.1	0.0	48
280	PHE232	0.00	0.1	0.0	25
199	VAL158	0.00	0.1	0.0	26

282	LEU234	0.00	0.1	0.0	107
135	ASP102	0.00	0.1	0.0	53
206	ARG165	0.00	0.1	0.0	27
109	ARG77A	0.00	0.1	0.0	166
95	LEU64	0.00	0.1	0.0	28
38	VAL17	0.00	0.1	0.0	29
209	CYS168	0.00	0.1	0.0	30
88	LYS60F	0.00	0.1	0.0	63
83	TYR60A	0.00	0.1	0.0	129
13	PRO5	0.00	0.1	0.0	67
113	LYS81	0.00	0.1	0.0	34
193	PRO152	0.00	0.1	0.0	34
98	ARG67	0.00	0.1	0.0	35
289	VAL241	0.00	0.1	0.0	143
148	SER115	0.00	0.1	0.0	42
28	LEU14F	0.00	0.1	0.0	45
165	LEU129C	0.00	0.1	0.0	48
11	LEU3	0.00	0.1	0.0	51
78	ALA56	0.00	0.1	0.0	52
81	LEU59	0.00	0.1	0.0	54
46	GLY25	0.00	0.1	0.0	65
208	VAL167	0.00	0.1	0.0	89
198	VAL157	0.00	0.1	0.0	98
12	ARG4	0.00	0.1	0.0	103
9	CYS1	0.00	0.1	0.0	4
10	GLY2	0.00	0.1	0.0	2
37	ILE16	0.00	0.1	0.0	16
48	SER27	0.00	0.1	0.0	19
49	PRO28	0.00	0.1	0.0	9
50	TRP29	0.00	0.1	0.0	3
51	GLN30	0.00	0.1	0.0	2
53	MET32	0.00	0.1	0.0	3
54	LEU33	0.00	0.1	0.0	4
63	LEU41	0.00	0.1	0.0	12
64	CYS42	0.00	0.1	0.0	2
68	LEU46	0.00	0.1	0.0	2
73	TRP51	0.00	0.1	0.0	30
75	LEU53	0.00	0.1	0.0	6
76	THR54	0.00	0.1	0.0	9
77	ALA55	0.00	0.1	0.0	10
97	VAL66	0.00	0.1	0.0	3
100	GLY69	0.00	0.1	0.0	7
112	GLU80	0.00	0.1	0.0	10
114	ILE82	0.00	0.1	0.0	24
115	SER83	0.00	0.1	0.0	8
116	MET84	0.00	0.1	0.0	26
139	MET106	0.00	0.1	0.0	18
141	LEU108	0.00	0.1	0.0	17
151	ILE118	0.00	0.1	0.0	13
157	PRO124	0.00	0.1	0.0	4
162	ALA129	0.00	0.1	0.0	26
172	GLY136	0.00	0.1	0.0	2
174	VAL138	0.00	0.1	0.0	12
175	THR139	0.00	0.1	0.0	7

176	GLY140	0.00	0.1	0.0	8
177	TRP141	0.00	0.1	0.0	18
178	GLY142	0.00	0.1	0.0	5
179	ASN143	0.00	0.1	0.0	23
201	LEU160	0.00	0.1	0.0	2
202	PRO161	0.00	0.1	0.0	2
204	VAL163	0.00	0.1	0.0	5
221	MET180	0.00	0.1	0.0	23
223	CYS182	0.00	0.1	0.0	1
224	ALA183	0.00	0.1	0.0	1
225	GLY184	0.00	0.1	0.0	2
240	ASP194	0.00	0.1	0.0	3
247	MET201	0.00	0.1	0.0	5
256	TYR208	0.00	0.1	0.0	11
258	MET210	0.00	0.1	0.0	2
260	ILE212	0.00	0.1	0.0	2
275	PHE227	0.00	0.1	0.0	10
277	THR229	0.00	0.1	0.0	18
286	ILE238	0.00	0.1	0.0	52
52	VAL31	0.00	NA	NA	0
65	GLY43	0.00	NA	NA	0
66	ALA44	0.00	NA	NA	0
67	SER45	0.00	NA	NA	0
99	ILE68	0.00	NA	NA	0
154	VAL121	0.00	NA	NA	0
155	CYS122	0.00	NA	NA	0
222	PHE181	0.00	NA	NA	0
239	GLY193	0.00	NA	NA	0
242	GLY196	0.00	NA	NA	0
243	GLY197	0.00	NA	NA	0
244	PRO198	0.00	NA	NA	0
245	PHE199	0.00	NA	NA	0
246	VAL200	0.00	NA	NA	0
257	GLN209	0.00	NA	NA	0
259	GLY211	0.00	NA	NA	0

References

- 1 B. Fuglestad, P. M. Gasper, M. Tonelli, J. A. McCammon, P. R. Markwick and E. A. Komives, *Biophysical Journal*, 2012, **103**, 79–88.
- 2 J. Xiao, R. L. Melvin and F. R. Salsbury, *Physical Chemistry Chemical Physics*, 2017, **19**, 24522–24533.
- 3 F. Noé, H. Wu, J. H. Prinz and N. Plattner, *Journal of Chemical Physics*, 2013, **139**, year.
- 4 J. H. Prinz, H. Wu, M. Sarich, B. Keller, M. Senne, M. Held, J. D. Chodera, C. Schütte and F. Noé, *Journal of Chemical Physics*, 2011, **134**, 1–23.
- 5 R. J. G. B. Campello, D. Moulavi and J. Sander, *Advances in Knowledge Discovery and Data Mining*, 2013, 160–172.
- 6 R. L. Melvin, R. C. Godwin, J. Xiao, W. G. Thompson, K. S. Berenhaut and F. R. Salsbury, *Journal of Chemical Theory and Computation*, 2016, **12**, 6130–6146.
- 7 R. L. Melvin, J. Xiao, R. C. Godwin, K. S. Berenhaut and F. R. Salsbury, *Protein Science*, 2018, **27**, 62–75.

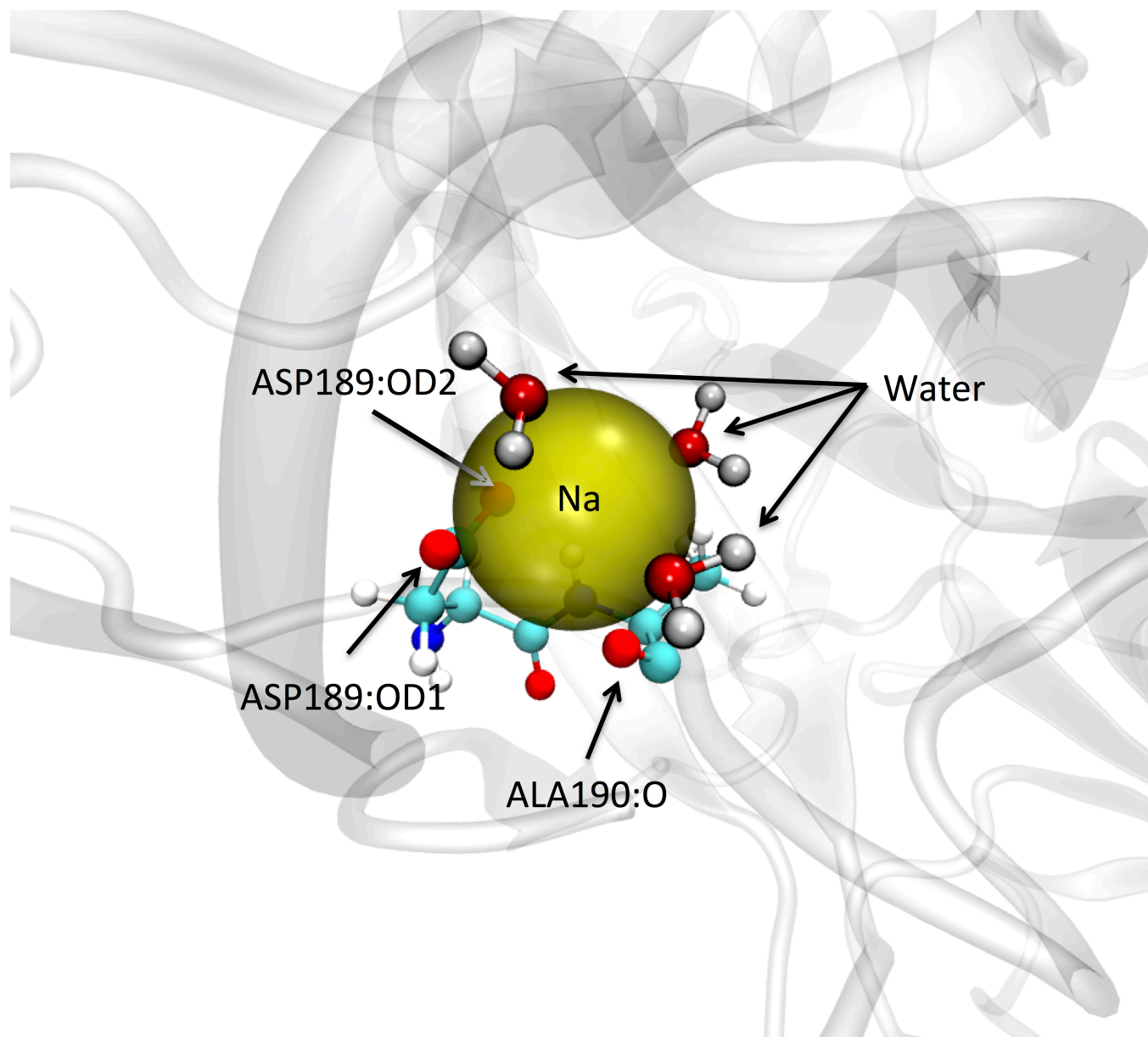


Fig. S1 The bound Na⁺ at the “inner” Na⁺-binding site is coordinated to the side chain oxygen atoms of ASP189, main chain oxygen atom of ALA190, and three surrounding water molecules. The bound Na⁺ is visualized as a yellow sphere using van der Waals representation. Oxygen, hydrogen, carbon and nitrogen atoms are colored in red, white, cyan and blue respectively.

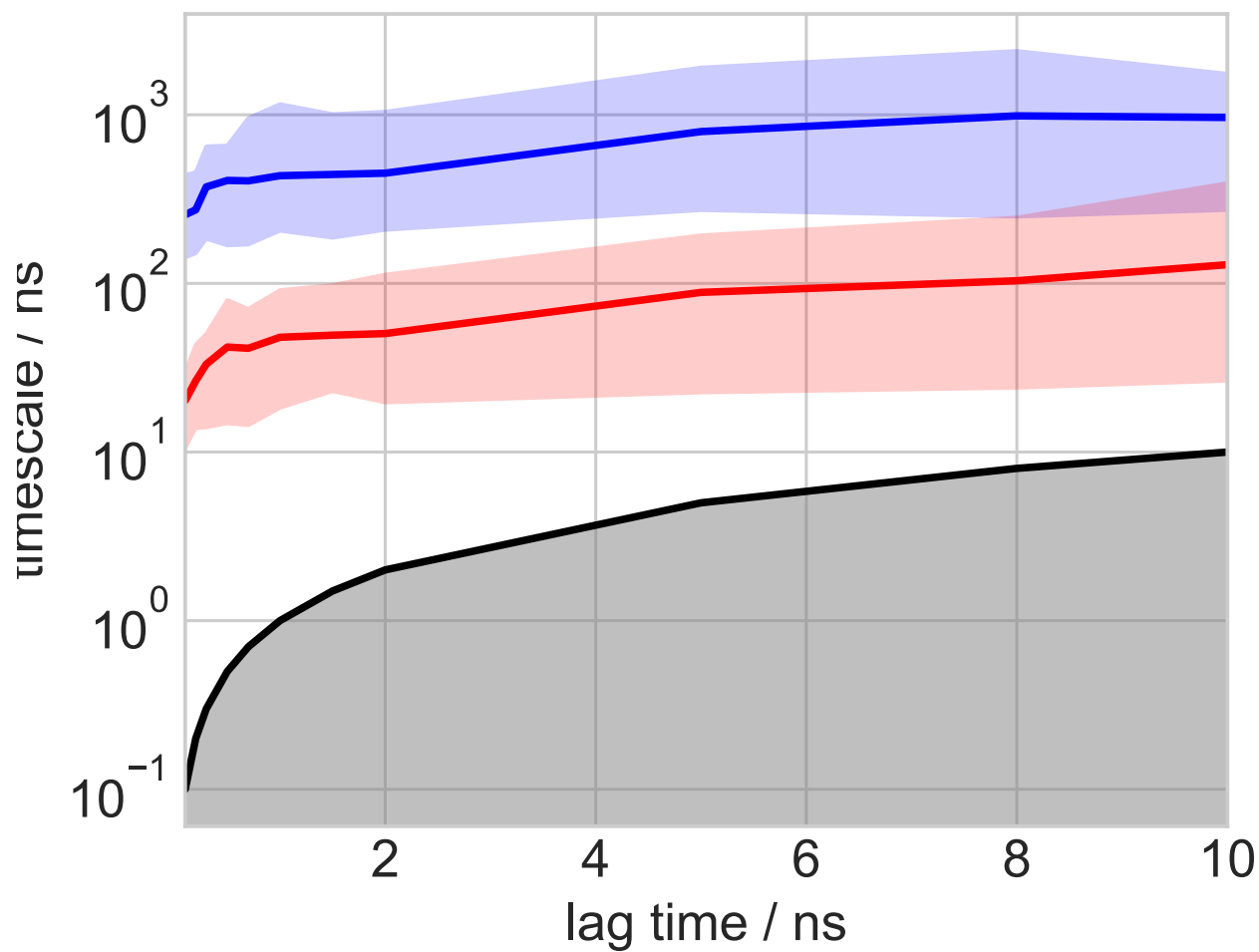


Fig. S2 Implied timescales based on the transition matrix from the reversible Bayesian hidden Markov model estimator³ indicate 1 ns is a proper Markov lag time. The time scales of the underlying processes start to be level with 95% confidence interval after 1 ns. The relaxation time for the two slowest processes is about 435 ns and 47 ns.

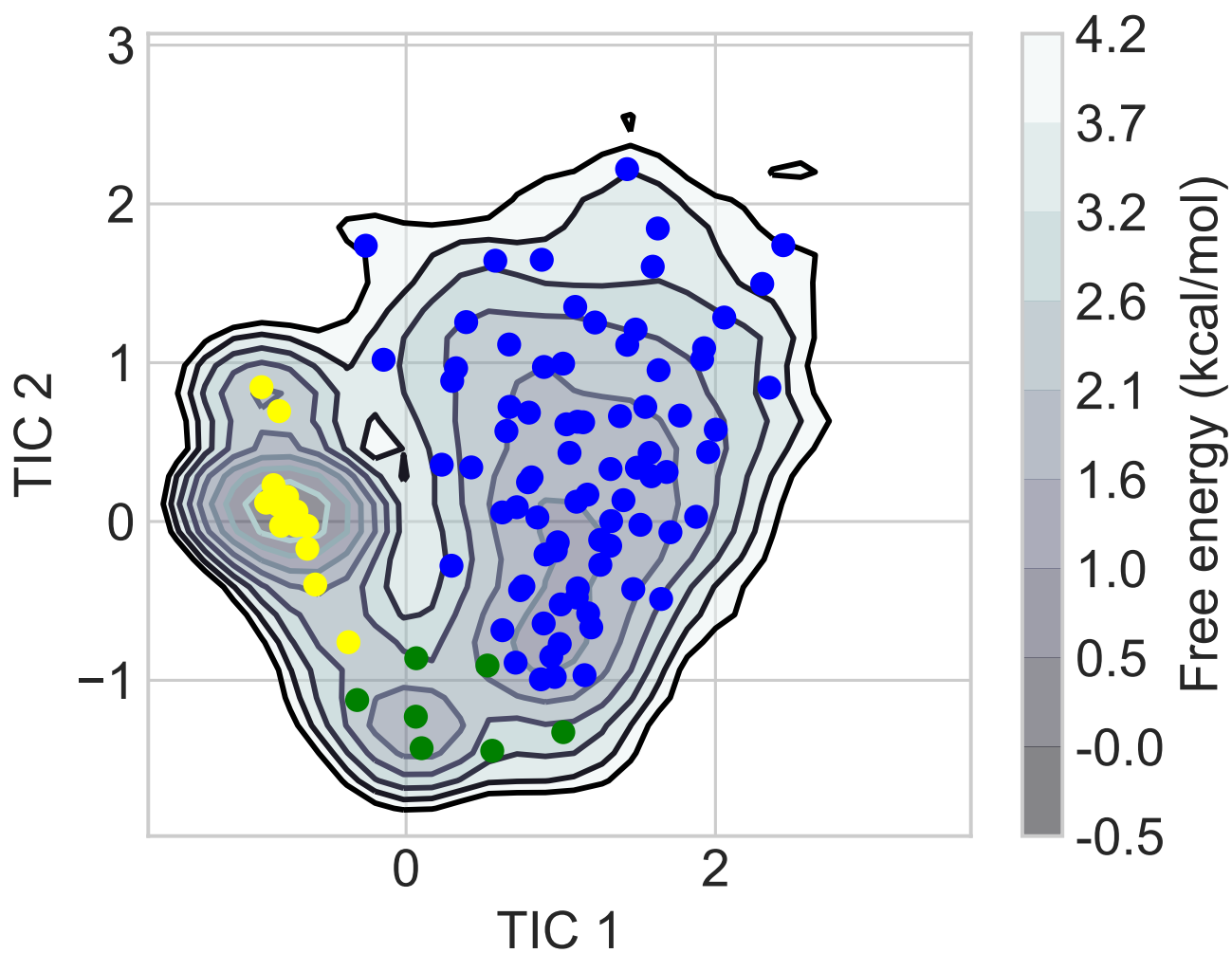


Fig. S3 Assignment of the metastable states in hidden Markov modeling. The green, blue and yellow dots respectively present microstates inside the “outer”-binding, unbound or non-specifically bound elsewhere, and “inner”-binding states.

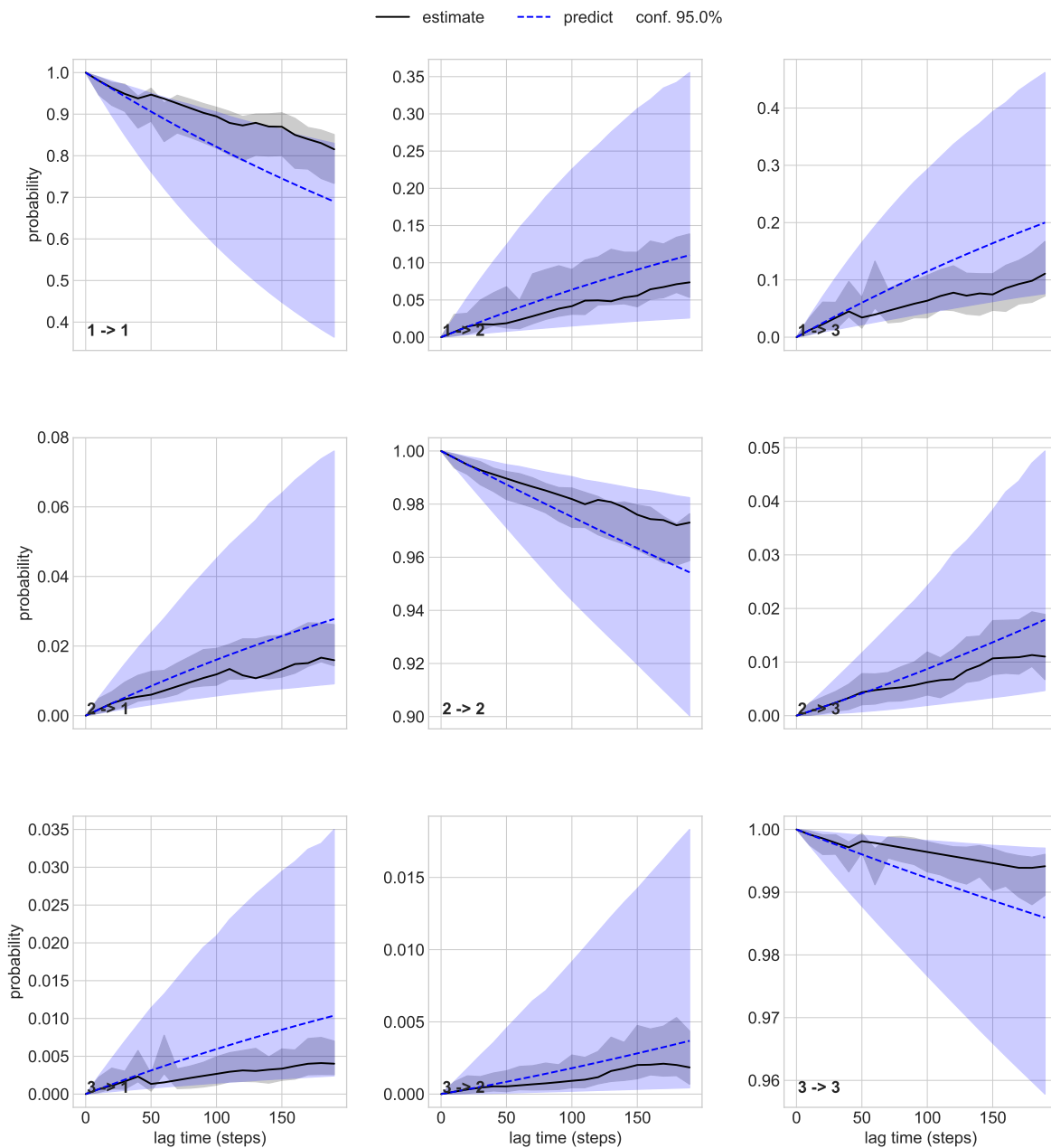


Fig. S4 The resulting HMM with a Markov lag time of 1 ns successfully passes the Chapman-Kolmogorov test⁴ with 95% confidence interval. Each step in the figures represent 0.1 ns.

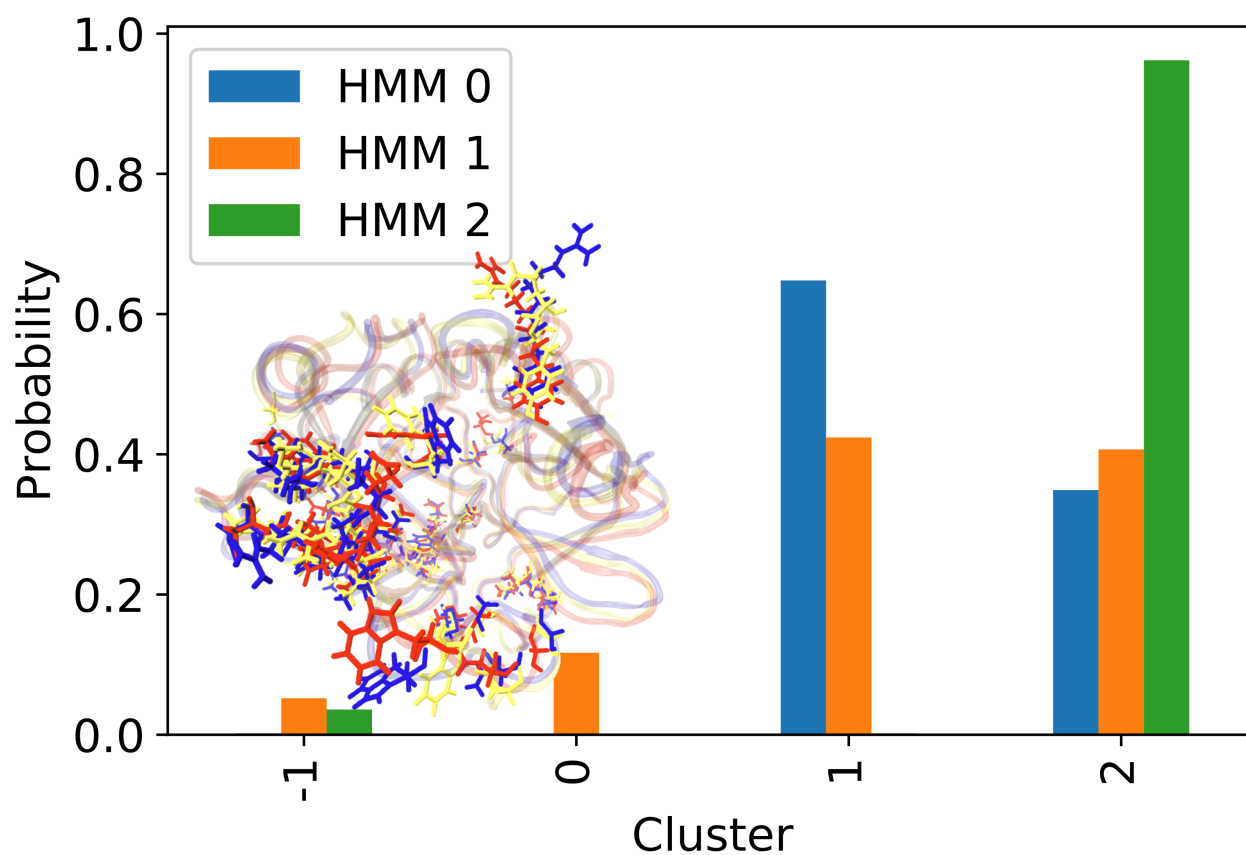


Fig. S5 Distributions of the backbone torsion angles of the residues allosterically responding to Na^+ -binding shown in Figure 3. The median structures of the cluster 0, 1, and 2 are respectively displayed in yellow, blue, and red. The side chains of the residues that allosterically respond to Na^+ -binding (see nodes in Figure 3a&b) show different orientations upon the Na^+ -binding statuses. HMM0, HMM1, and HMM2 stand for the outer Na^+ -bound, Na^+ -unbound, and inner Na^+ -bound modes. Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN) clustering algorithm⁵, which has recently proven sensitive and effective in grouping local conformational changes and highlighting representative features^{6,7}, was adapted here with a minimum cluster size of 10 observations to eliminate the number of clusters representing noise (denoted as cluster -1 in the figure). Other parameters in the clustering algorithm was set by the default values as followed: `min_samples=None`, `metric='euclidean'`, `alpha=1.0`, `p=None`, `algorithm='best'`, `leaf_size=40`, `memory=Memory(location=None)`, `approx_min_span_tree=True`, `gen_min_span_tree=False`, `core_dist_n_jobs=4`, `cluster_selection_method='eom'`, `allow_single_cluster=False`, `prediction_data=False`, `match_reference_implementation=False`. More details of these parameters can be found from the online documentation of HDBSCAN API at <https://hdbscan.readthedocs.io/en/0.8.18/api.html>