Supporting Information

Equally Weighted Multiscale Elastic Network Model and Its Comparison with Traditional and Parameter-Free Models

Weikang Gong, Yang Liu, Yanpeng Zhao, Shihao Wang, Zhongjie Han, Chunhua Li*

Faculty of Environmental and Life Sciences, Beijing University of Technology, Beijing 100124, China.
Beijing International Science and Technology Cooperation Base for Intelligent Physiological Measurement and Clinical Transformation, Beijing University of Technology, Beijing 100124, China.

*All correspondence should be addressed to Chunhua Li (Email: chunhuali@bjut.edu.cn).

The authors state no conflict of interest.

TABLES.

Table S1. Information of molecular dynamics simulations for the six proteins used in this study [a]

Protein Name	PDB ID	Number of residues	Total atoms	Box Type
Fasciculin-1	1FAS	61	15,709	Octahedral (61.711Å, 61.711Å, 61.711Å)
Glutaredoxin-1	1KTE	105	19,490	Octahedral (66.137Å, 66.137Å, 66.137Å)
Chemotaxis protein CheY	1CHN	126	18,022	Octahedral (64.964Å, 64.964Å, 64.964Å)
Focal adhesion kinase 1	1K40	126	37,686	Octahedral (81.900Å, 81.900Å, 81.900Å)
Extracellular globin	1ASH	147	23,810	Octahedral (70.914Å, 70.914Å, 70.914Å)
Allergen Bos d 2	1BJ7	150	19,371	Octahedral (66.426Å, 66.426Å, 66.426Å)

[a] 10 ns MD simulation is conducted using NAMD 2.5 with Charmm 22 force field for each protein.

Table S2. The *B_PCC* values obtained by pfENM with inverse 12th power of inter-residue distance

PDB ID / Model	1FAS	1KTE	1CHN	1K40	1ASH	1BJ7
pfGNM	0.45	0.48	0.08	0.63	0.45	0.12
pfANM	0.64	0.50	0.60	0.73	0.58	0.31

Table S3. Values of overlaps and *RMSIPs* between motional modes from ANMs and the principle components of motions sampled by MD simulations for the six proteins

Model \ Metric	Protein ID	O_1^{max}	O_2^{max}	O_3^{max}	CO_1^{20}	CO_2^{20}	CO_3^{20}	$RMSIP_3^{20}$	$RMSIP_6^{20}$	$RMSIP_{10}^{20}$	$RMSIP_{20}^{20}$
traditional ANM	1FAS	0.39	0.28	0.46	0.73	0.53	0.60	0.62	0.33	0.57	0.54
	1KTE	0.51	0.25	0.32	0.74	0.53	0.65	0.65	0.33	0.60	0.54
	1CHN	0.11	0.13	0.32	0.22	0.33	0.53	0.38	0.25	0.42	0.39
	1K40	0.32	0.31	0.40	0.66	0.58	0.68	0.64	0.33	0.62	0.56
	1ASH	0.24	0.32	0.40	0.48	0.58	0.58	0.55	0.30	0.52	0.53
	1BJ7	0.31	0.35	0.19	0.63	0.68	0.48	0.60	0.32	0.51	0.45
pfANM	1FAS	0.27	0.20	0.21	0.64	0.37	0.43	0.49	0.29	0.49	0.49
	1KTE	0.43	0.37	0.32	0.69	0.55	0.60	0.61	0.32	0.56	0.50
	1CHN	0.15	0.19	0.37	0.33	0.41	0.66	0.49	0.27	0.53	0.49
	1K40	0.49	0.28	0.28	0.72	0.56	0.61	0.64	0.33	0.58	0.52
	1ASH	0.25	0.22	0.35	0.48	0.43	0.50	0.47	0.29	0.42	0.43
	1BJ7	0.46	0.48	0.25	0.74	0.75	0.60	0.70	0.35	0.66	0.59
mANM	1FAS	0.20	0.14	0.18	0.38	0.36	0.33	0.35	0.24	0.38	0.37
	1KTE	0.45	0.26	0.25	0.60	0.38	0.45	0.48	0.28	0.44	0.48
	1CHN	0.11	0.08	0.07	0.17	0.14	0.11	0.14	0.16	0.17	0.17
	1K40	0.29	0.30	0.35	0.58	0.58	0.65	0.61	0.32	0.60	0.55
	1ASH	0.33	0.30	0.28	0.47	0.52	0.57	0.52	0.30	0.50	0.53
	1BJ7	0.06	0.01	0.02	0.07	0.03	0.04	0.05	0.09	0.06	0.07
equally weighted mANM	1FAS	0.33	0.20	0.34	0.64	0.45	0.59	0.57	0.31	0.54	0.51
	1KTE	0.49	0.32	0.32	0.77	0.53	0.68	0.66	0.33	0.63	0.57
	1CHN	0.19	0.31	0.39	0.40	0.54	0.70	0.56	0.31	0.57	0.53
	1K40	0.32	0.27	0.36	0.59	0.56	0.66	0.60	0.32	0.58	0.53
	1ASH	0.25	0.28	0.28	0.47	0.61	0.63	0.57	0.31	0.54	0.55
	1BJ7	0.32	0.35	0.19	0.65	0.68	0.51	0.62	0.32	0.53	0.48

Table S4. *B_PCC* values obtained by mGNM and equally weighted mGNM for B-factor prediction on small-, medium-, and large-size structures

PDB ID	mGNM			equally weighted mGNM		
	η_1	η_2	*B_PCC*	η_1	η_2	*B_PCC*
Small-size structures						
1AIE	6	8	0.70	26	26	0.64
1AKG	26	26	0.37	1	26	0.37
1BX7	1	1	0.75	1	1	0.75
1ETL	2	3	0.68	1	2	0.66
1ETM	2	2	0.43	2	2	0.43
1ETN	7	9	0.08	26	26	0.06
1FF4	1	1	0.69	1	1	0.69
1GK7	3	4	0.84	7	23	0.84
1GVD	2	26	0.75	2	2	0.75
1HJE	1	2	0.81	1	1	0.77
1KYC	26	26	0.79	26	26	0.79
1NOT	1	2	0.75	1	1	0.70
1O06	26	26	0.92	26	26	0.92
1OB4	1	26	0.77	2	4	0.76
1OB7	2	8	0.87	26	26	0.75
1P9I	4	8	0.79	2	5	0.69
1PEF	1	10	0.89	26	26	0.89
1PEN	1	26	0.46	2	3	0.39
1Q9B	1	1	0.82	1	1	0.82
1RJU	4	6	0.52	10	10	0.51
1U06	1	2	0.54	1	1	0.48
1UOY	2	4	0.72	5	5	0.71
1USE	3	21	0.84	3	6	-0.09
1VRZ	1	4	0.90	26	26	0.80
1XY2	1	1	0.64	1	1	0.64
1YJO	5	7	0.43	1	1	0.38
1YZM	1	1	0.94	1	1	0.94
2DSX	1	26	0.32	4	26	0.30
2JKU	2	4	0.81	3	3	0.81
2NLS	1	5	0.63	1	1	0.62
2OL9	1	26	0.91	1	1	0.90
2OLX	1	2	0.91	3	3	0.90
6RXN	1	2	0.59	2	2	0.59
Medium-size structures						
1ABA	5	26	0.71	6	26	0.71

1CYO	1	26	0.77	1	1	0.77
1FK5	2	20	0.60	3	6	0.59
1GXU	4	5	0.80	10	10	0.76
1I71	1	26	0.48	4	26	0.45
1LR7	1	1	0.69	1	1	0.69
1N7E	2	26	0.56	4	26	0.52
1NNX	3	26	0.79	5	26	0.79
1NOA	1	1	0.65	1	1	0.65
1OPD	1	1	0.64	1	1	0.64
1QAU	1	25	0.67	2	4	0.65
1R7J	2	17	0.78	4	5	0.30
1UHA	2	3	0.72	9	10	0.71
1ULR	1	7	0.66	1	2	0.64
1USM	1	1	0.87	1	1	0.87
1V05	1	1	0.69	1	1	0.69
1W2L	5	8	0.72	1	26	0.71
1X3O	2	4	0.70	1	1	0.68
1Z21	1	26	0.68	1	2	0.58
1ZVA	3	5	0.81	12	12	0.77
2BF9	1	26	0.63	1	26	0.62
2BRF	14	15	0.81	1	1	0.81
2CE0	5	6	0.73	1	2	0.72
2E3H	1	26	0.70	26	26	0.67
2EAQ	1	1	0.78	1	1	0.78
2EHS	2	2	0.76	1	1	0.76
2FQ3	1	16	0.74	1	2	0.72
2IP6	1	26	0.62	1	2	0.59
2MCM	1	1	0.88	1	1	0.88
2NUH	6	7	0.87	1	2	0.86
2PKT	2	6	0.39	4	26	-0.18
2PLT	3	4	0.46	4	4	0.45
2QJL	1	2	0.60	1	2	0.60
2RB8	3	4	0.75	9	9	0.72
3BZQ	1	26	0.54	4	26	0.51
5CYT	1	6	0.45	2	5	0.45
Large-size structures						
1AHO	26	26	0.69	4	26	0.69
1ATG	2	11	0.61	2	9	0.61
1BYI	2	3	0.53	1	2	0.52
1CCR	3	4	0.58	7	8	0.57
1E5K	1	2	0.81	1	1	0.80
1EW4	1	2	0.68	1	2	0.67

1IFR	1	16	0.71	1	2	0.69
1NKO	16	16	0.67	6	26	0.68
1NLS	25	26	0.69	7	7	0.68
1O08	1	1	0.40	1	1	0.40
1PMY	1	1	0.74	1	1	0.74
1PZ4	2	2	0.89	1	2	0.89
1QTO	1	26	0.49	4	26	0.47
1RRO	1	1	0.49	1	1	0.49
1UKU	1	1	0.72	1	1	0.72
1V70	1	26	0.60	26	26	0.60
1WBE	1	1	0.67	1	1	0.67
1WHI	1	26	0.57	3	26	0.49
1WPA	3	19	0.58	4	5	0.55
2AGK	4	5	0.71	1	4	0.71
2C71	2	26	0.67	1	2	0.67
2CG7	2	26	0.51	3	3	0.46
2CWS	1	1	0.73	1	1	0.73
2HQK	2	21	0.83	3	19	0.82
2HYK	1	2	0.60	1	2	0.60
2I24	1	8	0.54	1	1	0.51
2IMF	3	4	0.65	2	3	0.65
2PPN	1	2	0.66	2	3	0.65
2R16	1	5	0.60	1	1	0.60
2V9V	1	1	0.60	1	1	0.60
2VIM	1	18	0.41	1	5	0.40
2VPA	3	26	0.75	3	5	0.72
2VYO	1	1	0.80	1	1	0.80
3SEB	1	1	0.89	1	1	0.89
3VUB	25	26	0.66	1	2	0.66

η_1 and η_2 are the two parameters of the kernel function.

SUPPORTING FIGURES.

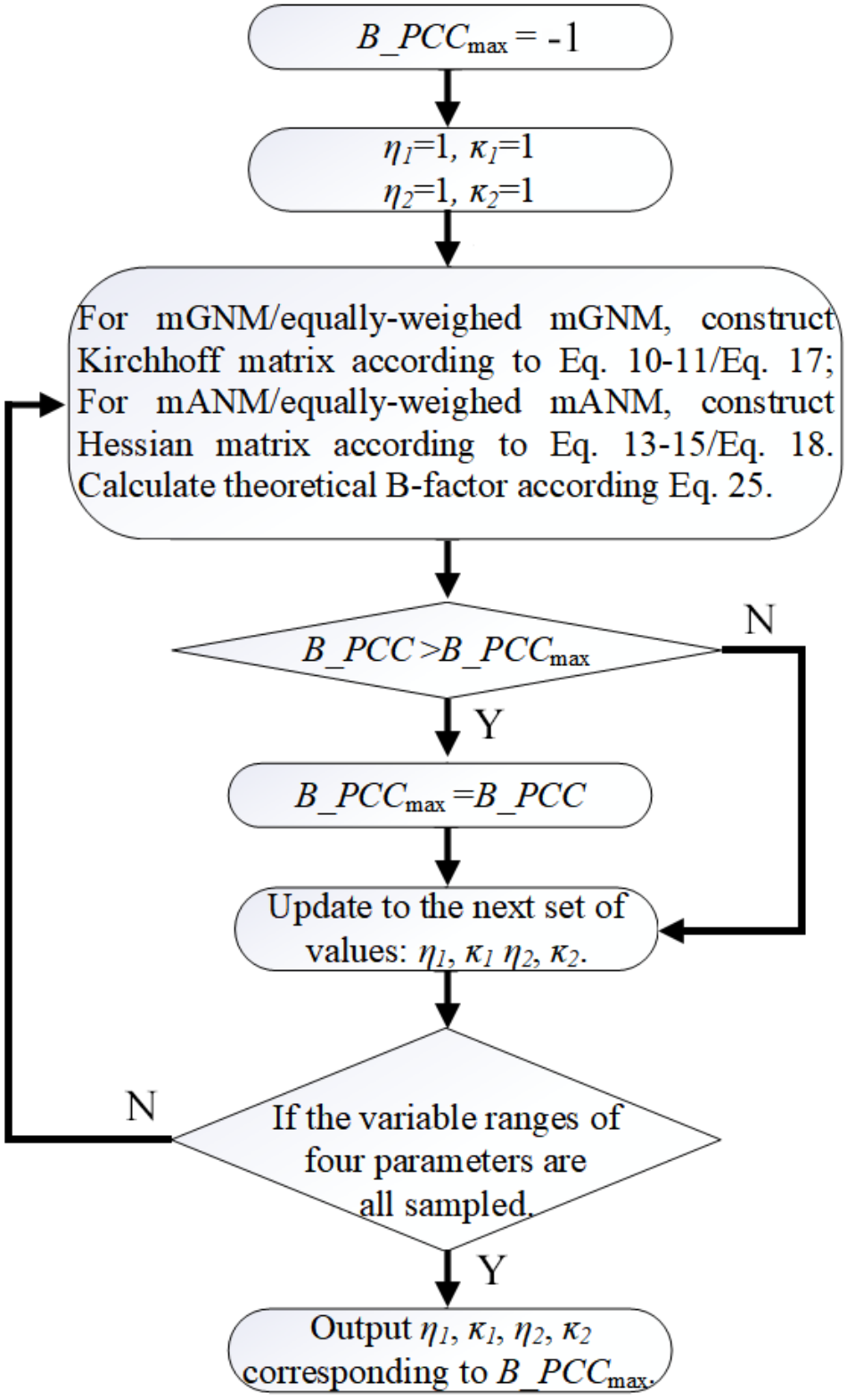

Figure S1. Optimization flowchart of η and κ parameters in mENM and equally weighted mENM models.

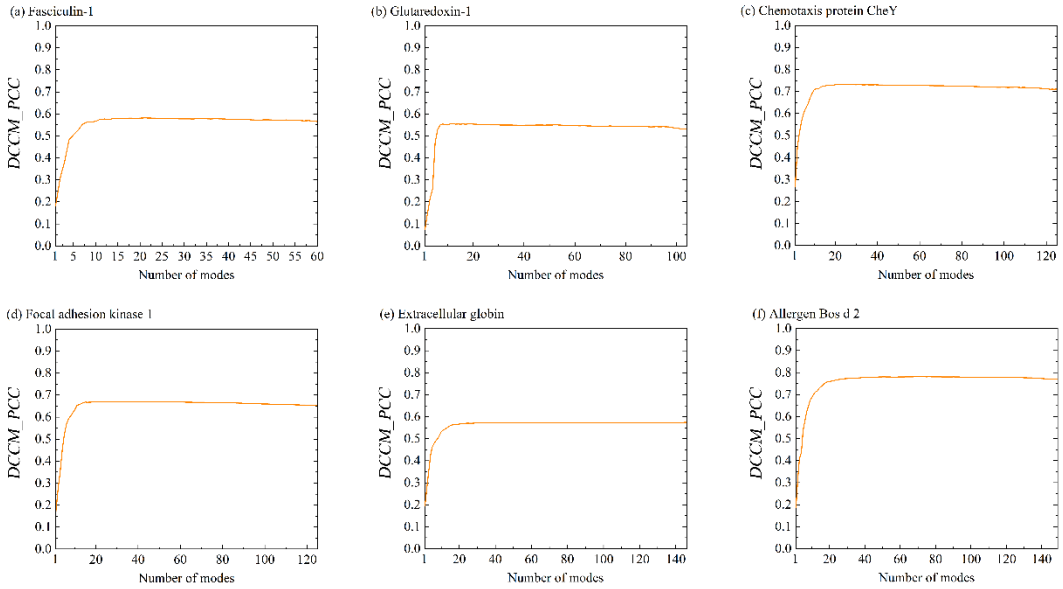

Figure S2. Changes of pearson correlation coefficients between DCCMs (*DCCM_PCC*) obtained from MD ensembles and pfGNM (with a cutoff distance of 7.0 Å) with the number of motional modes considered in pfGNM for proteins 1FAS (a), 1KTE (b), 1CHN (c), 1K40 (d), 1ASH (e) and 1BJ7 (f).

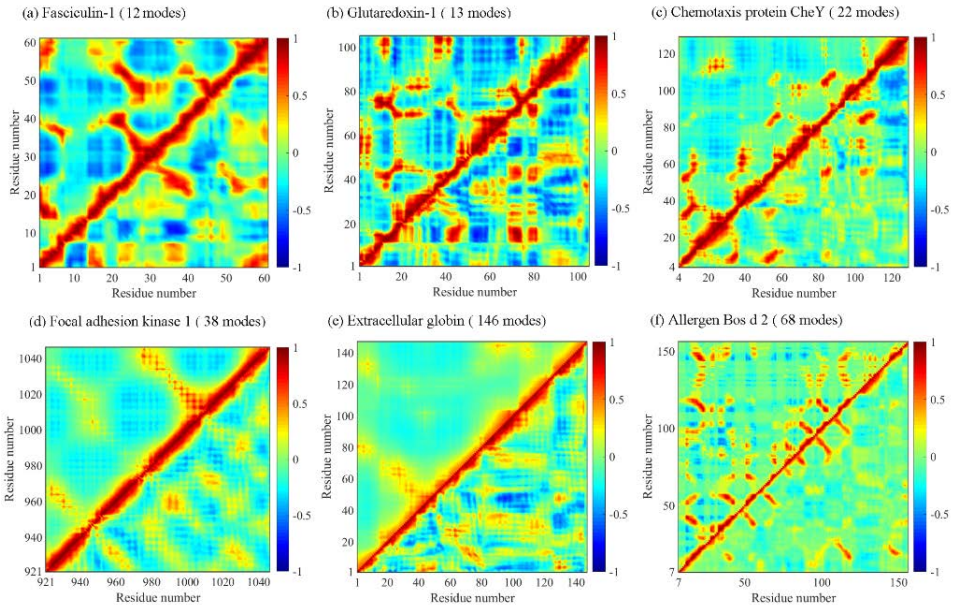

Figure S3. DCCMs obtained from MD ensembles (lower right triangle) and traditional GNM (upper left triangle) at the best *DCCM_PCC* value (corresponding number of motional modes given in parentheses) for proteins 1FAS (a), 1KTE (b), 1CHN (c), 1K40 (d), 1ASH (e) and 1BJ7 (f).

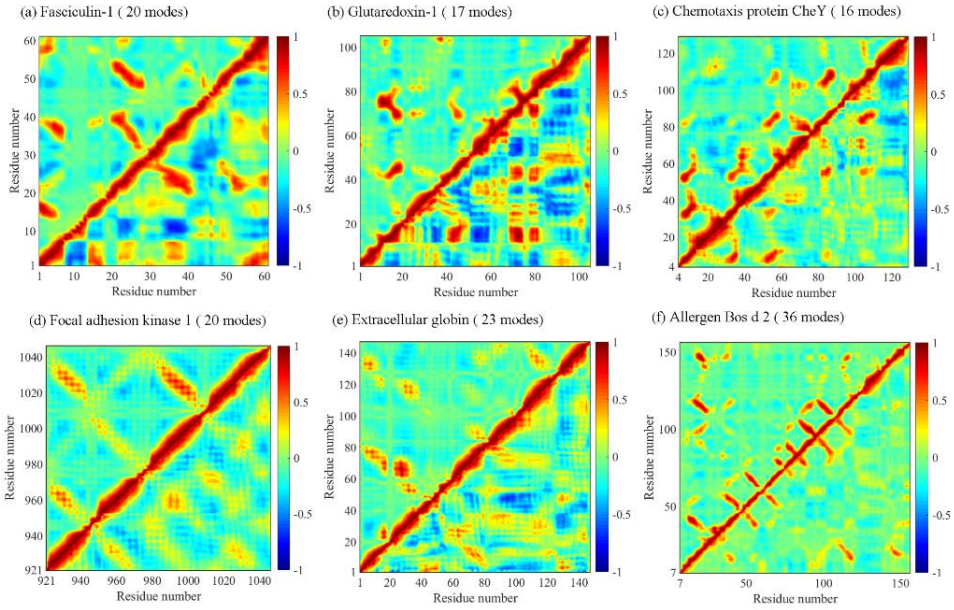

Figure S4. DCCMs obtained from MD ensembles (lower right triangle) and pfGNM (upper left triangle) at the best *DCCM_PCC* value (corresponding number of motional modes given in parentheses) for proteins 1FAS (a), 1KTE (b), 1CHN (c), 1K40 (d), 1ASH (e) and 1BJ7 (f).

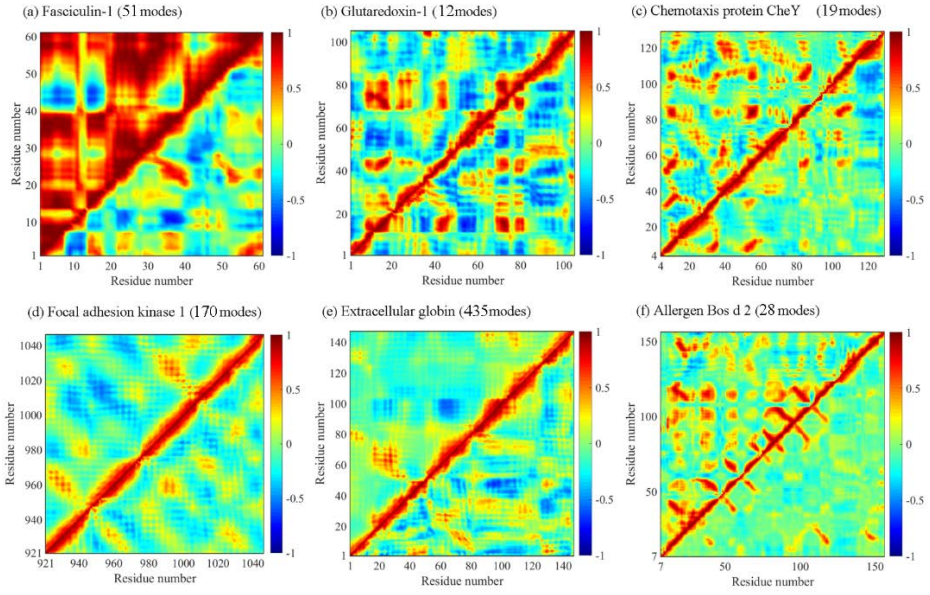

Figure S5. DCCMs obtained from MD ensembles (lower right triangle) and traditional ANM (upper left triangle) at the best *DCCM_PCC* value (corresponding number of motional modes given in parentheses) for proteins 1FAS (a), 1KTE (b), 1CHN (c), 1K40 (d), 1ASH (e) and 1BJ7 (f).

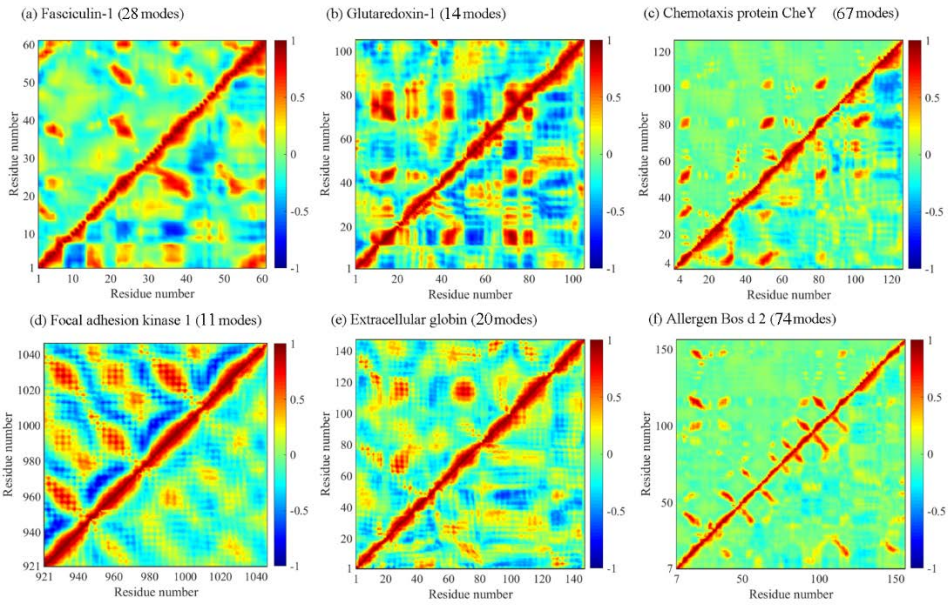

Figure S6. DCCMs obtained from MD ensembles (lower right triangle) and pfANM (upper left triangle) at the best *DCCM_PCC* value (corresponding number of motional modes given in parentheses) for proteins 1FAS (a), 1KTE (b), 1CHN (c), 1K40 (d), 1ASH (e) and 1BJ7 (f).

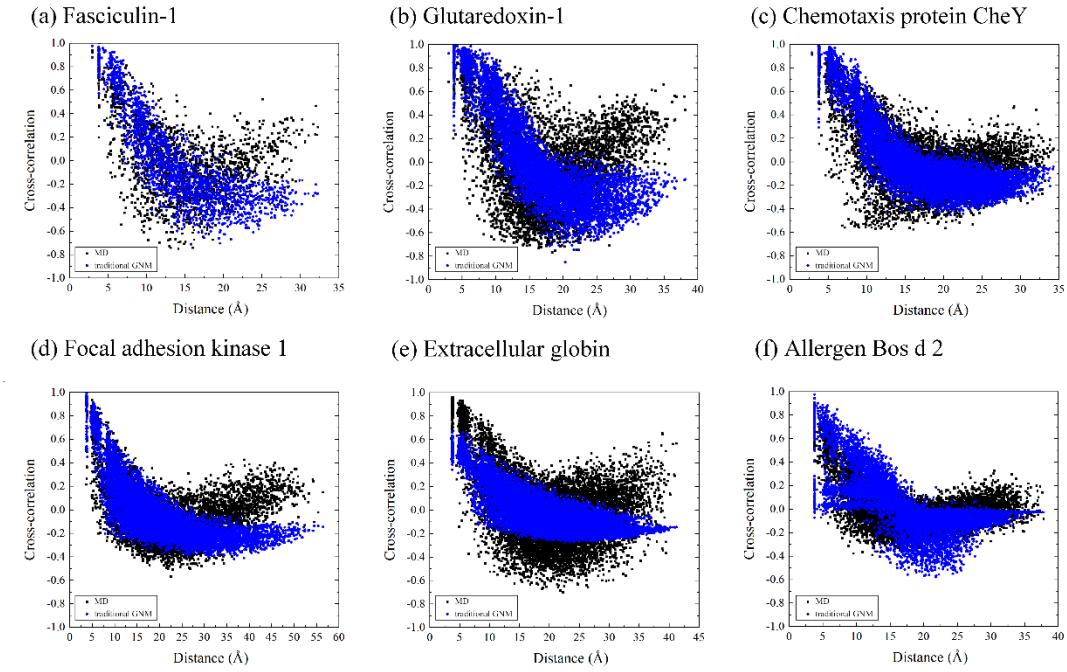

Figure S7. Distributions of dynamical cross-correlations obtained from MD ensembles (black) and traditional GNM (blue) (at the best *DCCM_PCC* value) with respect to the inter-residue distance for proteins 1FAS (a), 1KTE (b), 1CHN (c), 1K40 (d), 1ASH (e) and 1BJ7 (f).

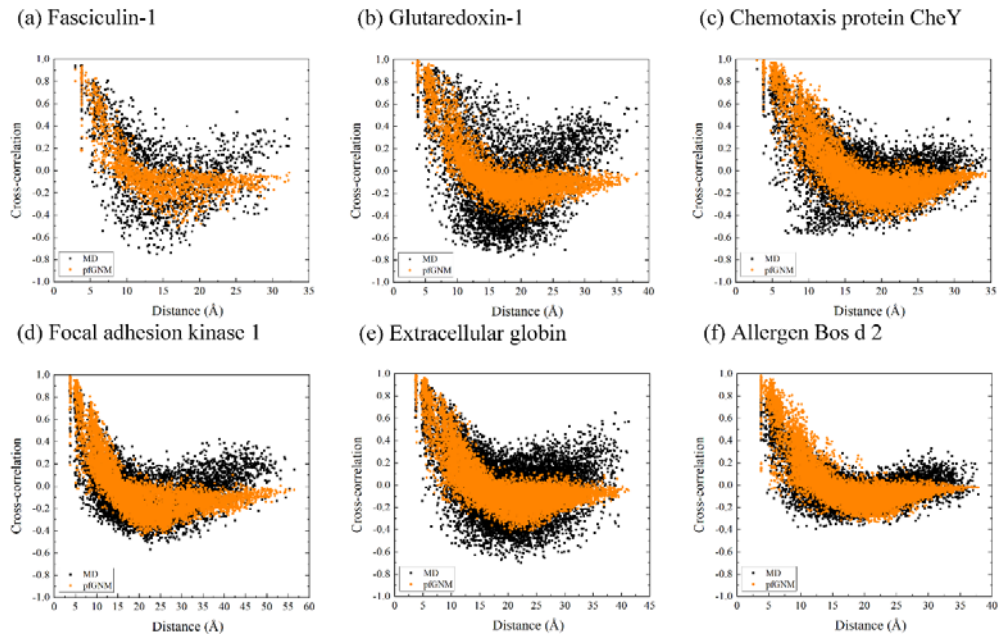

Figure S8. Distributions of dynamical cross-correlations obtained from MD ensembles (black) and pfGNM (orange) (at the best *DCCM_PCC* value) with respect to the inter-residue distance for proteins 1FAS (a), 1KTE (b), 1CHN (c), 1K40 (d), 1ASH (e) and 1BJ7 (f).

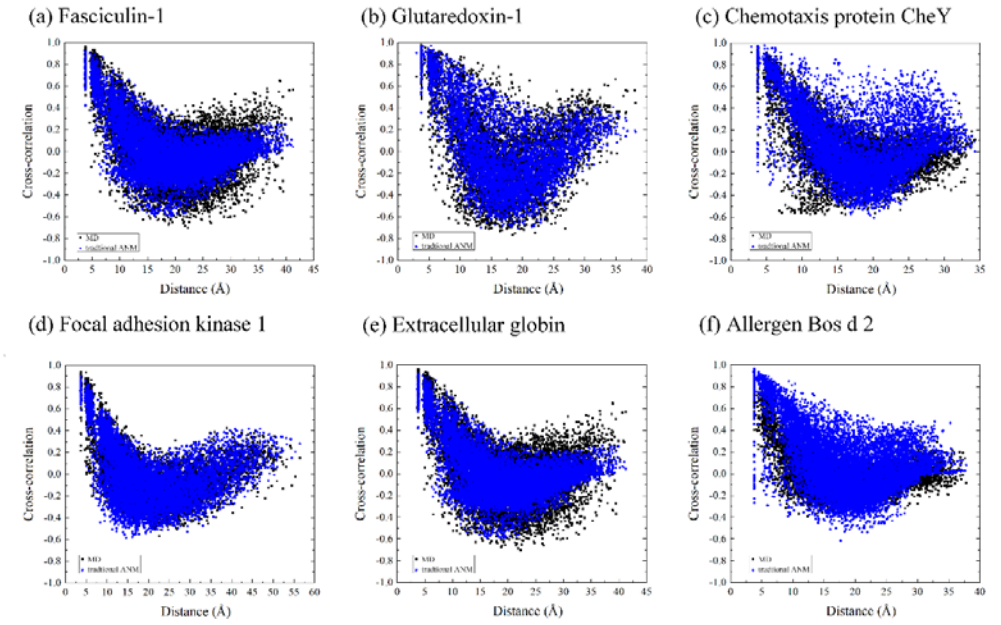

Figure S9. Distributions of dynamical cross-correlations obtained from MD ensembles (black) and traditional ANM (blue) (at the best *DCCM_PCC* value) with respect to the inter-residue distance for proteins 1FAS (a), 1KTE (b), 1CHN (c), 1K40 (d), 1ASH (e) and 1BJ7 (f).

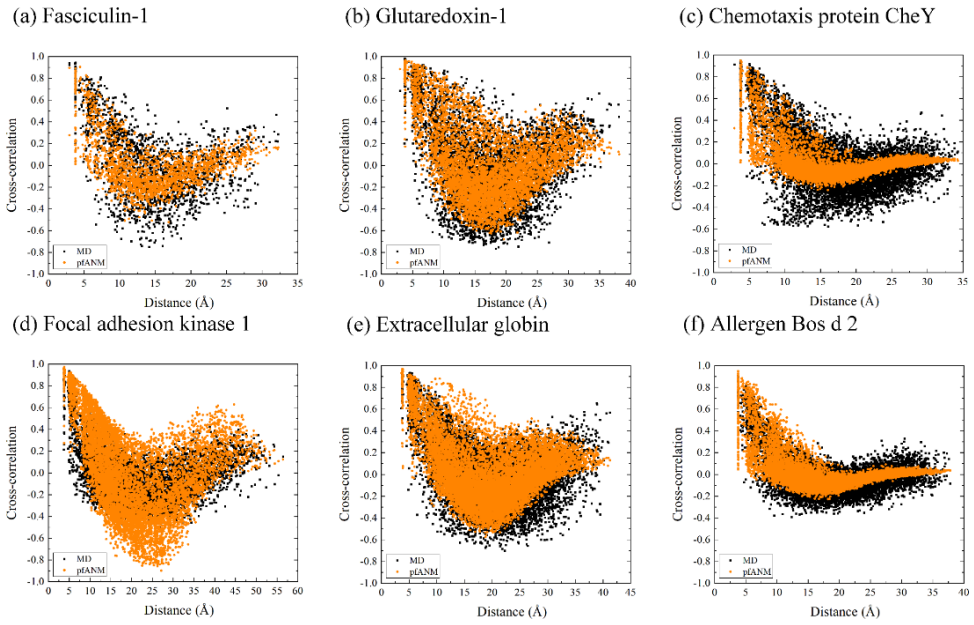

Figure S10. Distributions of dynamical cross-correlations obtained from MD ensembles (black) and pfANM (orange) (at the best *DCCM_PCC* value) with respect to the inter-residue distance for proteins 1FAS (a), 1KTE (b), 1CHN (c), 1K40 (d), 1ASH (e) and 1BJ7 (f).