Supporting Information

A Universal Pattern in the Percolation and Dissipation of Protein Structural Perturbations

Nandakumar Rajasekaran,[1] Ashok Sekhar[2] & Athi N. Naganathan[1*]

[1]Department of Biotechnology, Bhupat & Jyoti Mehta School of Biosciences, Indian Institute of Technology Madras (IITM), Chennai 600036, India.

[2]Departments of Molecular Genetics, Biochemistry and Chemistry, The University of Toronto, Toronto, ON, Canada M5S 1A8.

e-mail: athi@iitm.ac.in

Table 1: Protein Mutations

Sl.No.	Protein	PDB ID	Protein Class	No.of residues	Mutation(s)	Reference(s)
1	Ubiquitin	1UBQ	α+β	76	L43A, L69S	[1,2]
2	Staphylococcal Nuclease	3BDC	β	129	L103A, L125A, V66A, I92A	[3]
3	Engrailed homeodomain	2JWT	α	59	L16A	[4]
4	Titin M10 domain	2Y9R	β	98	H56P	[5]
5	Thymidylate Synthase	1BID	α+β	265	C146S	[6]
6	FAS Domain	2LTB	β	134	R555W	[7]
7	FF Domain	1UZC	α	69	L24A,A39G	[8,9]
8	DHFR	1RE7	α/β	159	G67V	[10]
9	Troponin	2CTN	α	88	L29Q	[11]
10	Human Lysozyme	2NWD	α+β	130	I23A, I56V, I89V	[12]
11	Human superoxide dismutase	1HL5	β	153	G93A, G37R, G85R, A4V, V148G, H46R, V148I	[13]
12	HTRF1	1BA5	α	53	K52C	[14]

Table 2: Protein-Ligand Binding

Sl. No.	Protein	PDB ID	Protein Class	No. of residues	Ligand/Protein	No. of residues	Reference
1	RafRBD	1RRB/ 1GUA	α+β	76	WT Ras	166	[15]
2	RafRBD	1RRB/ 1GUA	α+β	76	D30E/E31K Ras	166	[15]
3	Chymotrypsin Inhibitor - CI2	2CI2	α+β	65	Bovine α-Chymotrypsin	245	[16]
4	Rev1-BRCT domain	2M2I	α/β	94	Proliferating cell nuclear antigen (PCNA)	258	[17]
5	Ubiquitin	2JY6	α+β	76	Ubiquilin 1 UBA	52	[18]
6	Ubiquilin 1 UBA	2JY6	α	52	Ubiquitin	76	[18]
7	Human PD1 (hPD1- Programmed cell death protein 1)	2M2D	β	118	PD- Ligand 1	222	[19]
8	Human PD1 (hPD1- Programmed cell death protein 1)	2M2D	β	118	PD-Ligand 2	202	[19]
9	Ubiquitin (polyUbq)	1UBQ	α+β	76	Polyubiquitin linkage at K6, K11, K27, K29, K33, K48, K63		[20]
10	Pin1 (Pin-CDC)	1I6C	β	39	Cdc25 Peptide	10	[21]
11	Pin1 (Pin-Tau)	1I6C	β	39	Tau Peptide	13	[21]
12	Mcl-1	2JM6	α	162	Maritoclax		[22]
13	Hypoxia Inducible Factor- 2α (HIF)	3F1P	α/β	114	Compund 2 - [$C_{12}H_6ClFN_4O_3$] (Drug Molecule)		[23]
14	Liver Fatty Acid Binding Protein (LFABP)	2JU3	β	127	Mono-olein		[24]
15	Birch Pollen Allergen (Bet)	4QIP	α+β	159	Sodium dodecyl sulphate (SDS)		[25]

Table 3: Phosphorylation

Sl. No.	Protein	PDB ID	Protein Class	No. of residues	Phosphorylation Site	Reference
1	UBL	5TR5	α+β	76	65	[26]
2	PDZ3	1BFE	β	119	92	[27]

Table 4: Goodness of fit measures for the exponential fitting to experimental chemical shift data.*

Mutations

Mutation	RMSE	R-squared	F-statistic	P-value
Ubq L43A , L69S	0.182	0.10	70.1	6.77E-22
SNase L103A , L125A	0.0604	0.20	118	9.69E-37
EnHD L16A	0.202	0.22	74.7	1.18E-16
Titin H56P	0.0649	0.20	46.5	1.33E-14
ThySyn C146S	0.0478	0.26	84.9	5.11E-24
FAS R555W	0.0708	0.26	54.2	1.06E-17
DHFR G67V	0.0678	0.30	47.7	1.16E-16
Troponin L29Q	0.0156	0.45	96.5	1.36E-22
Human Lysozyme I23A, I56V, I89V	0.0338	0.13	169	3.43E-53
SOD G93A, G37R, G85R	0.0502	0.17	98	3.09E-34
SOD A4V, V148G	0.0519	0.14	71.5	7.92E-25

Ligand Binding

Protein-Ligand Pair	RMSE	R-squared	F-statistic	P-value
Raf - WTRas	0.0325	0.39	71.8	5.71E-18
Raf - mutRas	0.0318	0.38	69.4	1.32E-17
CI2 - Chymotrypsin	0.324	0.37	28.7	1.95E-09
RevBRCT - PCNA	0.0296	0.15	34.3	1.54E-11
Uba -Ubq	0.0808	0.41	49.5	2.15E-12
Ubq - Uba	0.11	0.27	40.5	1.84E-12
HPD - L1	0.141	0.11	69.4	4.13E-20
HPD - L2	0.124	0.21	71.5	1.68E-20
PolyUb K6, K11, K27, K29	0.0452	0.16	135	6.07E-42
PolyUb K33, K48, K63	0.0238	0.20	128	2.59E-37
Pin - Tau	0.0696	0.26	20.3	1.81E-06

Phosphorylation

Protein	RMSE	R-squared	F-statistic	P-value
UBL p65	0.109	0.705	167	2.43E-27
PDZ3 p92	0.1	0.57	86.8	3.07E-20

* F-statistic and its corresponding p-value highlight the significance of the R-squared values. F-statistic performs a f-test where the null hypothesis is the zero model (all coefficients in our model is zero), and its magnitude depends on number of degrees of freedom (so we cannot directly compare across proteins). The p-values indicate the significance of the obtained F-values. Low p-value for the F-value, as we obtain for all the systems, means the probability of getting the corresponding R-squared by random chance is negligible.

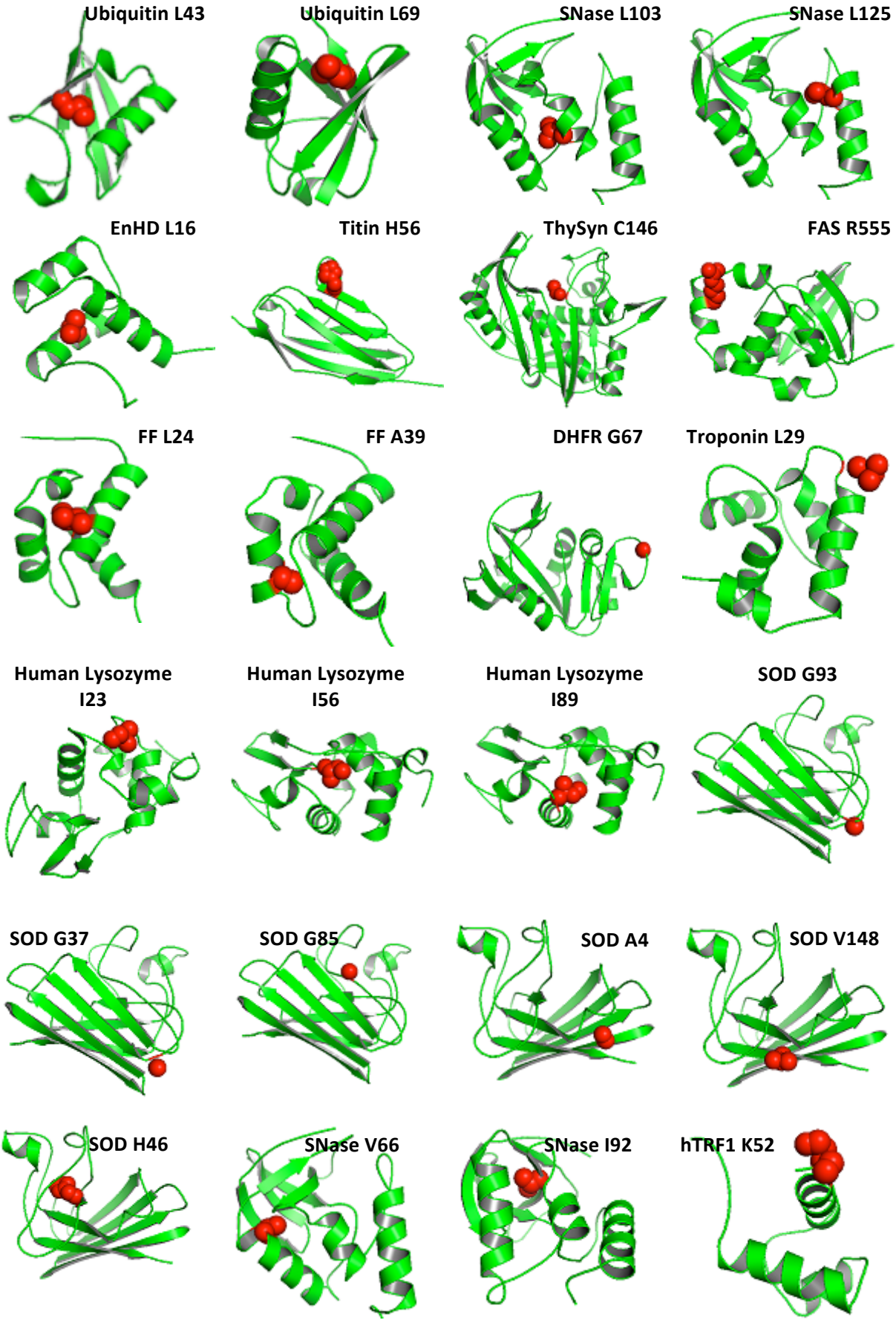

Figure S1 Structures of the proteins employed in Figure 1 of the main text. The mutated residue is shown in red.

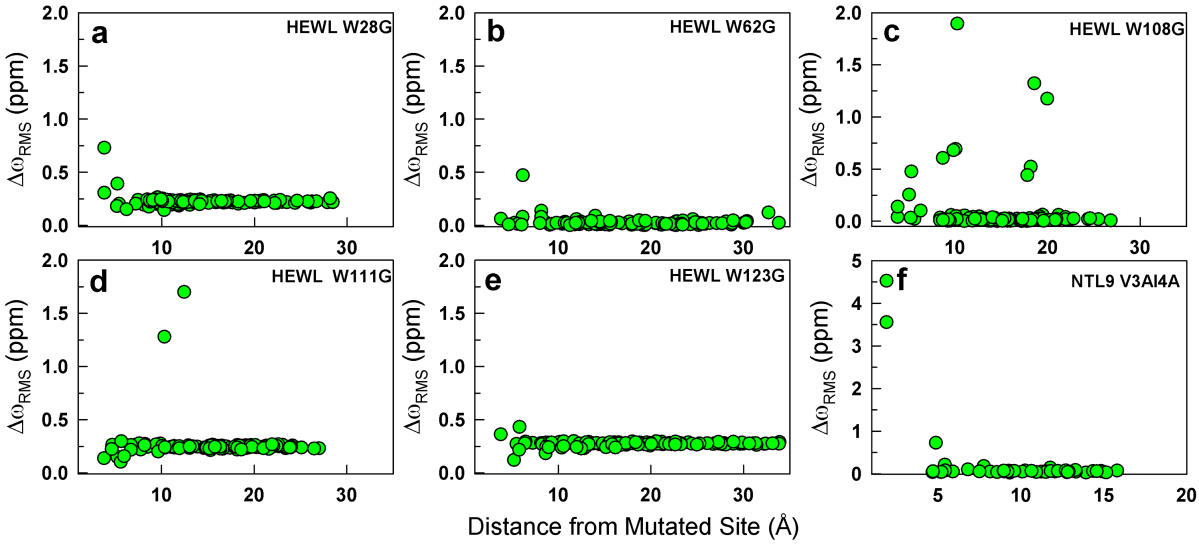

Figure S2 (a-e) Plot of chemical shift differences, $\Delta\omega_{\mathrm{RMS}} = \sqrt{\frac{1}{N}\sum_i \left(\frac{\Delta\varpi_i}{\Delta\varpi_{i,STD}}\right)^2}$ (where $\Delta\varpi_i$ is the shift diefference in p.p.m. units, $\Delta\varpi_{i,STD}$ is a nuclei specific value and *N* is the number of nuclei considered[28]) between denatured states of all alanine wild type (WT^{Ala}) Hen Egg White Lysozyme (HEWL) and the indicated mutants as a function of C_α-C_α distance from the mutated site. The experiments were carried out at 293 K and at pH 2.0. The distances were calculated using the cysteine containing wild type structure (PDB ID: 6LYZ).[29] (f) Plot of chemical shift differences between denatured states of N-terminal domain of ribosomal protein L9 (NTL9) and its double mutant (V3AI4A) against distance from the C_α coordinate center of the mutations (PDB ID: 2HBB). The experiments were carried out at 285 K under solution conditions of 5.5 pH and 8.3 M urea.[30,31]

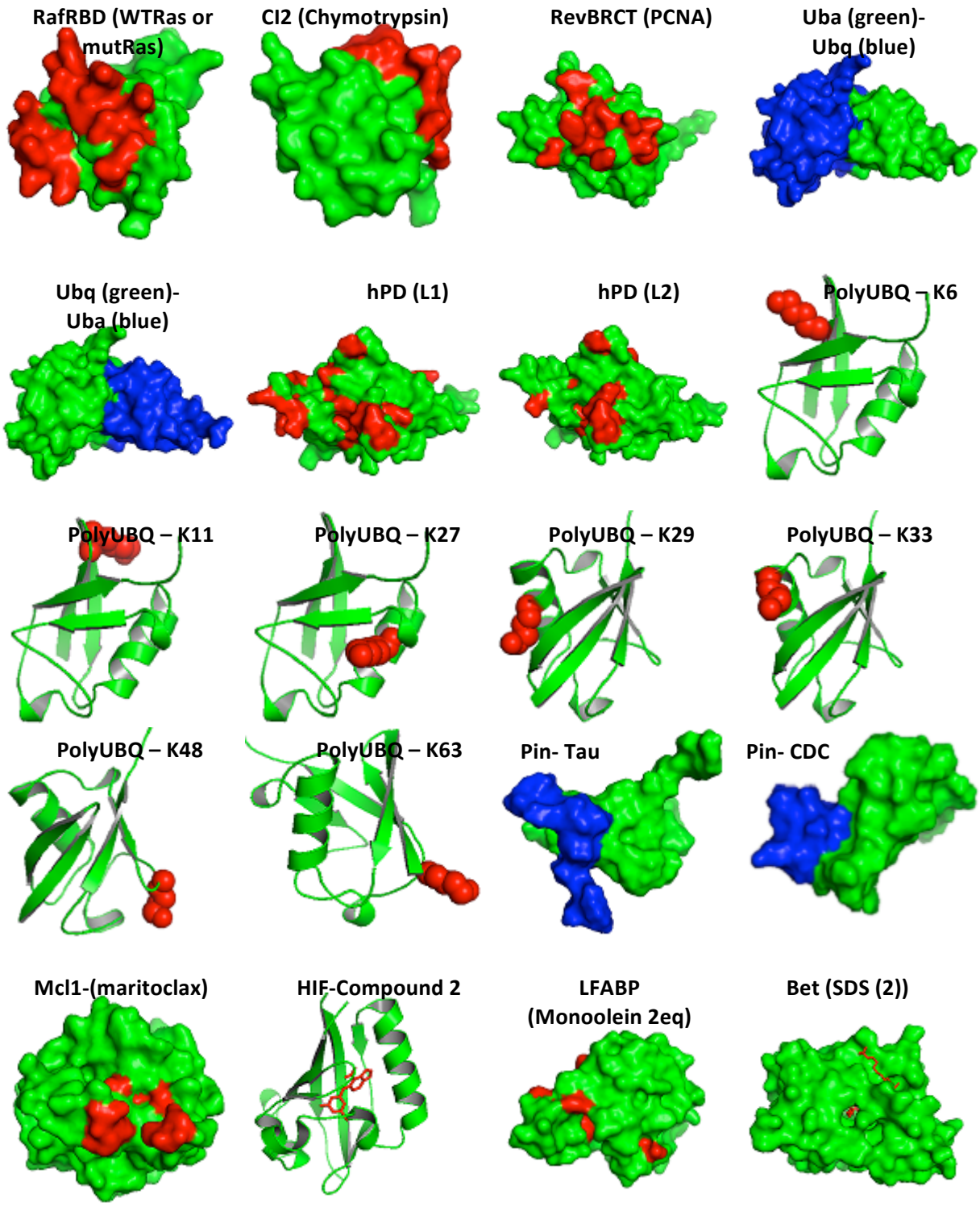

Figure S3 Structures of the proteins used in Figure 2 of the main text. WT protein is shown in green. If the structure of the protein-protein complex is available, the partner protein is shown in blue. If the complex structure is unavailable, the experimentally identified binding residues are colored red. Small molecule ligands are shown as sticks.

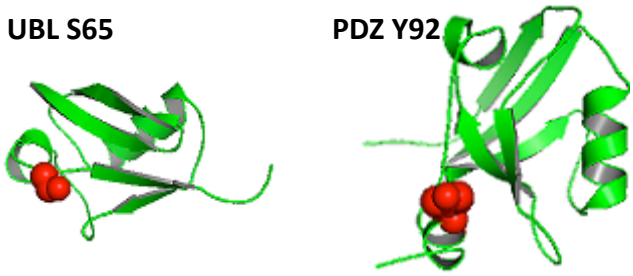

Figure S4 Structures of the proteins used in Figure 3. The phosphorylation site is shown in red.

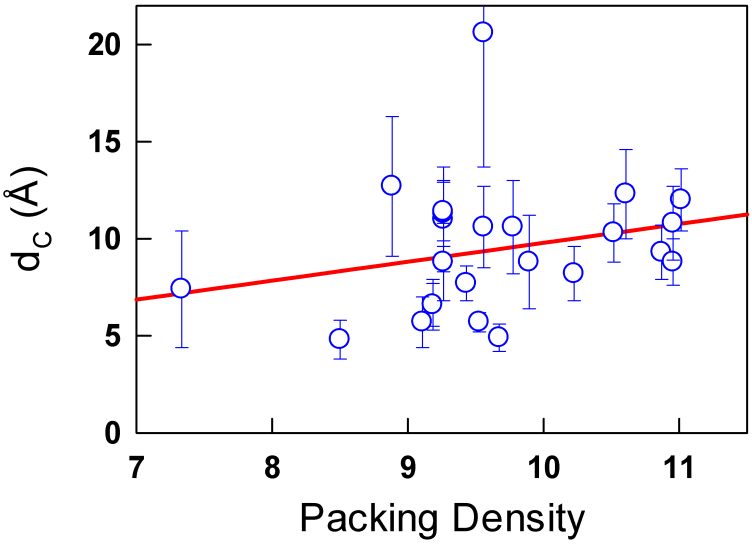

Figure S5 Dependence on the coupling distance (ordinate) on the packing density (abcissa). Packing density is defined as the average number of residue neighbors within a cut-off distance of 6 Å in a given protein.

References

1. Haririnia, A. et al. Mutations in the hydrophobic core of ubiquitin differentially affect its recognition by receptor proteins. *J Mol Biol* **375**, 979-96 (2008).
2. Lee, S.Y. et al. Alanine scan of core positions in ubiquitin reveals links between dynamics, stability, and function. *J Mol Biol* **426**, 1377-89 (2014).
3. Roche, J. et al. Structural, energetic, and dynamic responses of the native state ensemble of staphylococcal nuclease to cavity-creating mutations. *Proteins* **81**, 1069-80 (2013).
4. Mayor, U. et al. The complete folding pathway of a protein from nanoseconds to microseconds. *Nature* **421**, 863-7 (2003).
5. Rudloff, M.W., Woosley, A.N. & Wright, N.T. Biophysical characterization of naturally occurring titin M10 mutations. *Protein Sci* **24**, 946-55 (2015).
6. Sapienza, P.J. & Lee, A.L. Widespread Perturbation of Function, Structure, and Dynamics by a Conservative Single-Atom Substitution in Thymidylate Synthase. *Biochemistry* **55**, 5702-5713 (2016).
7. Underhaug, J. et al. Mutation in transforming growth factor beta induced protein associated with granular corneal dystrophy type 1 reduces the proteolytic susceptibility through local structural stabilization. *Biochim Biophys Acta* **1834**, 2812-22 (2013).
8. Korzhnev, D.M. et al. Nonnative interactions in the FF domain folding pathway from an atomic resolution structure of a sparsely populated intermediate: an NMR relaxation dispersion study. *J Am Chem Soc* **133**, 10974-82 (2011).
9. Vallurupalli, P. et al. Studying "invisible" excited protein states in slow exchange with a major state conformation Hsp70 biases the folding pathways of client proteins. *J Am Chem Soc* **134**, 8148-61 (2012).
10. Narayanan, S.P., Maeno, A., Wada, Y., Tate, S. & Akasaka, K. Sequential backbone resonance assignments of the E. coli dihydrofolate reductase Gly67Val mutant: folate complex. *Biomol NMR Assign* **10**, 125-9 (2016).
11. Robertson, I.M. et al. The structural and functional effects of the familial hypertrophic cardiomyopathy-linked cardiac troponin C mutation, L29Q. *J Mol Cell Cardiol* **87**, 257-69 (2015).
12. Ahn, M. et al. The Significance of the Location of Mutations for the Native-State Dynamics of Human Lysozyme. *Biophys J* **111**, 2358-2367 (2016).
13. Sekhar, A. et al. Probing the free energy landscapes of ALS disease mutants of SOD1 by NMR spectroscopy. *Proc Natl Acad Sci U S A* **113**, E6939-45 (2016).
14. Sekhar, A., Rosenzweig, R., Bouvignies, G. & Kay, L.E. Hsp70 biases the folding pathways of client proteins. *Proc Natl Acad Sci U S A* **113**, E2794-801 (2016).
15. Terada, T. et al. Nuclear magnetic resonance and molecular dynamics studies on the interactions of the Ras-binding domain of Raf-1 with wild-type and mutant Ras proteins. *J Mol Biol* **286**, 219-32 (1999).
16. Whitley, M.J. & Lee, A.L. Exploring the role of structure and dynamics in the function of chymotrypsin inhibitor 2. *Proteins* **79**, 916-24 (2011).
17. Pustovalova, Y., Maciejewski, M.W. & Korzhnev, D.M. NMR mapping of PCNA interaction with translesion synthesis DNA polymerase Rev1 mediated by Rev1-BRCT domain. *J Mol Biol* **425**, 3091-105 (2013).
18. Zhang, D., Raasi, S. & Fushman, D. Affinity makes the difference: nonselective interaction of the UBA domain of Ubiquilin-1 with monomeric ubiquitin and polyubiquitin chains. *J Mol Biol* **377**, 162-80 (2008).
19. Cheng, X. et al. Structure and interactions of the human programmed cell death 1 receptor. *J Biol Chem* **288**, 11771-85 (2013).

20. Castaneda, C.A. et al. Linkage-specific conformational ensembles of non-canonical polyubiquitin chains. *Phys Chem Chem Phys* **18**, 5771-88 (2016).
21. Wintjens, R. et al. 1H NMR study on the binding of Pin1 Trp-Trp domain with phosphothreonine peptides. *J Biol Chem* **276**, 25150-6 (2001).
22. Doi, K. et al. Discovery of marinopyrrole A (maritoclax) as a selective Mcl-1 antagonist that overcomes ABT-737 resistance by binding to and targeting Mcl-1 for proteasomal degradation. *J Biol Chem* **287**, 10224-35 (2012).
23. Scheuermann, T.H. et al. Allosteric inhibition of hypoxia inducible factor-2 with small molecules. *Nat Chem Biol* **9**, 271-6 (2013).
24. Lagakos, W.S. et al. Liver fatty acid-binding protein binds monoacylglycerol in vitro and in mouse liver cytosol. *J Biol Chem* **288**, 19805-15 (2013).
25. Grutsch, S. et al. Ligand binding modulates the structural dynamics and compactness of the major birch pollen allergen. *Biophys J* **107**, 2972-81 (2014).
26. Aguirre, J.D., Dunkerley, K.M., Mercier, P. & Shaw, G.S. Structure of phosphorylated UBL domain and insights into PINK1-orchestrated parkin activation. *Proc Natl Acad Sci U S A* **114**, 298-303 (2017).
27. Zhang, J., Petit, C.M., King, D.S. & Lee, A.L. Phosphorylation of a PDZ domain extension modulates binding affinity and interdomain interactions in postsynaptic density-95 (PSD-95) protein, a membrane-associated guanylate kinase (MAGUK). *J Biol Chem* **286**, 41776-85 (2011).
28. Bouvignies, G. et al. Solution structure of a minor and transiently formed state of a T4 lysozyme mutant. *Nature* **477**, 111-4 (2011).
29. Sziegat, F. et al. Disentangling the coil: modulation of conformational and dynamic properties by site-directed mutation in the non-native state of hen egg white lysozyme. *Biochemistry* **51**, 3361-72 (2012).
30. Meng, W., Luan, B., Lyle, N., Pappu, R.V. & Raleigh, D.P. The denatured state ensemble contains significant local and long-range structure under native conditions: analysis of the N-terminal domain of ribosomal protein L9. *Biochemistry* **52**, 2662-71 (2013).
31. Meng, W. & Raleigh, D.P. Analysis of electrostatic interactions in the denatured state ensemble of the N-terminal domain of L9 under native conditions. *Proteins* **79**, 3500-10 (2011).