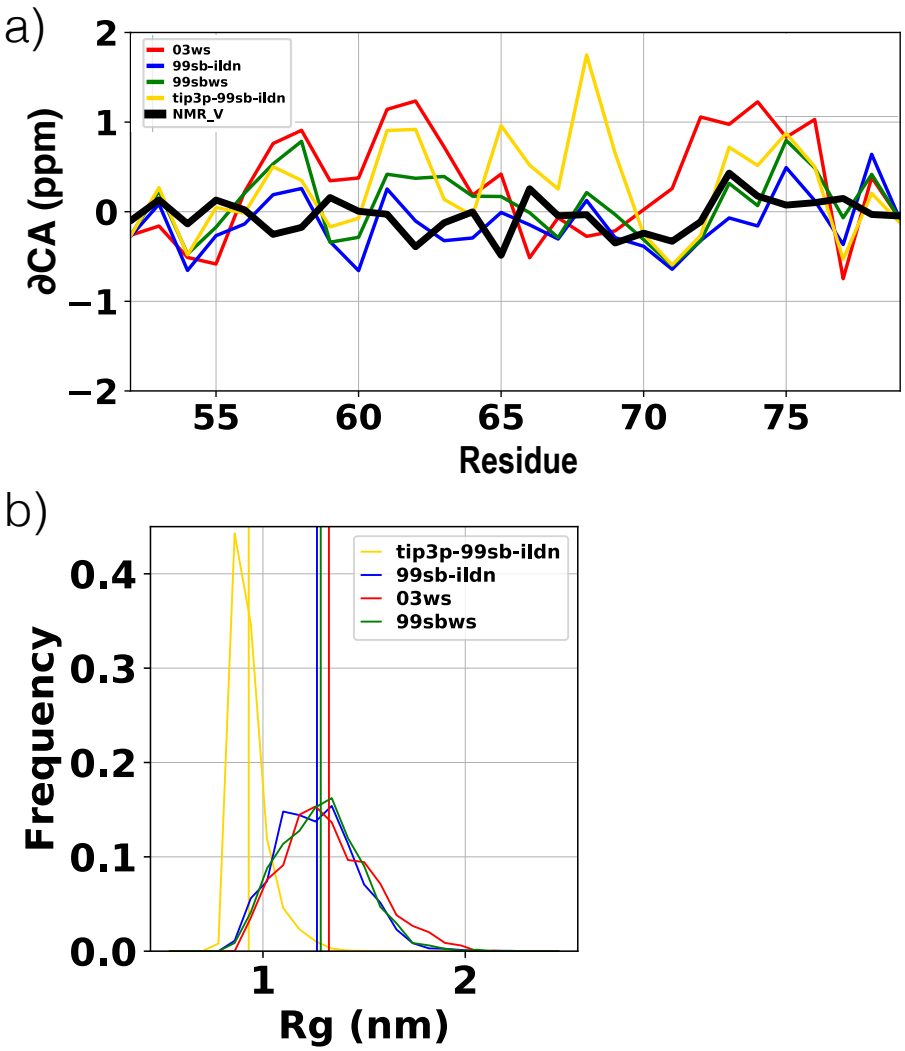
Sequence specificity despite intrinsic disorder: how a disease-associated Val/Met polymorphism shifts tertiary interactions in a long disordered protein

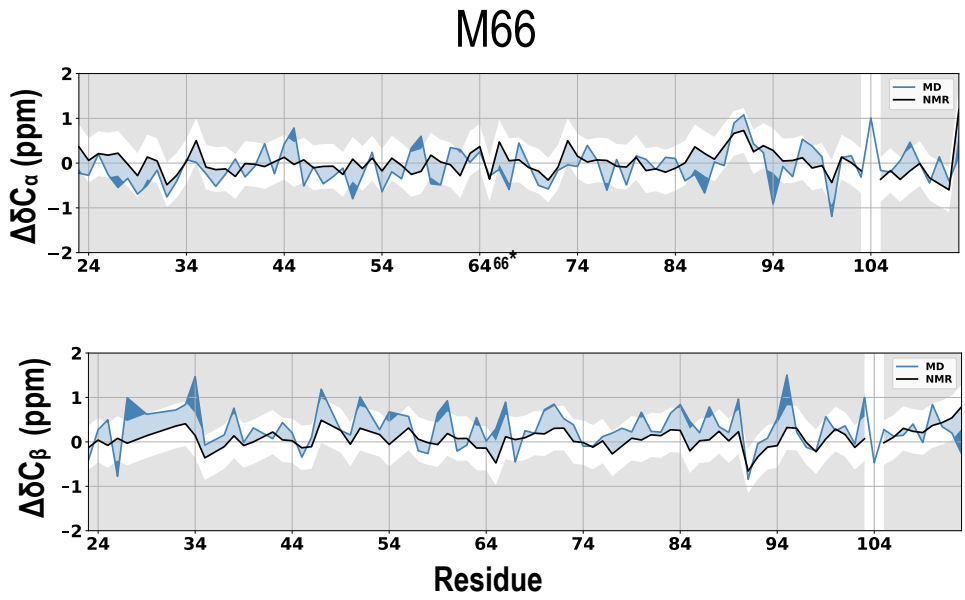
Ruchi Lohia[1], Reza Salari[1], Grace Brannigan[1,2*],

1 Center for Computational and Integrative Biology, Rutgers University, Camden, NJ, USA
2 Department of Physics, Rutgers University, Camden, NJ, USA

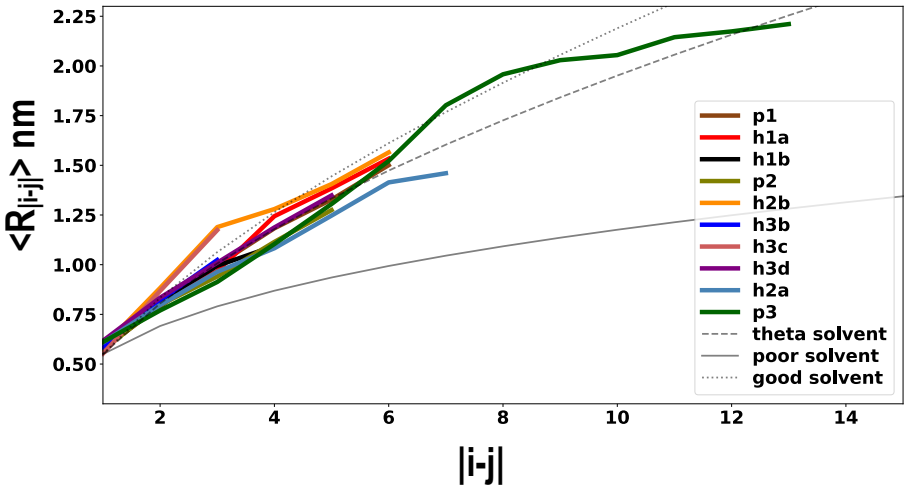
* grace.brannigan@rutgers.edu(GB)



S1 Fig. Force field comparison. We ran 500ns of T-REMD simulations of a 30 residue fragment of the V66 prodomain with several commonly used force field and water model combinations. The first 200ns of data was discarded as equilibration. (a) Comparison of calculated C_α chemical shifts from MD ensembles at 280K for Amber99sb*-ildn-q [1,2] with Tip4p-D [3] (RMSD 0.36 ppm), Amber99sbws [1,4] (RMSD 0.42 ppm), Amberff03sbws [4,5] (RMSD 0.73 ppm), Amber99sb-ildn with Tip3P [6] (RMSD 0.65 ppm), calculated using SPARTA+ [7] and NMR C_α chemical shifts (black line) from Ref. 8 at 280K. (b) R_g distribution for each force field. Tip3P generates very collapsed ensembles and the remaining three force field generates similar R_g distribution. The $\langle R_g \rangle$ is shown with vertical lines.



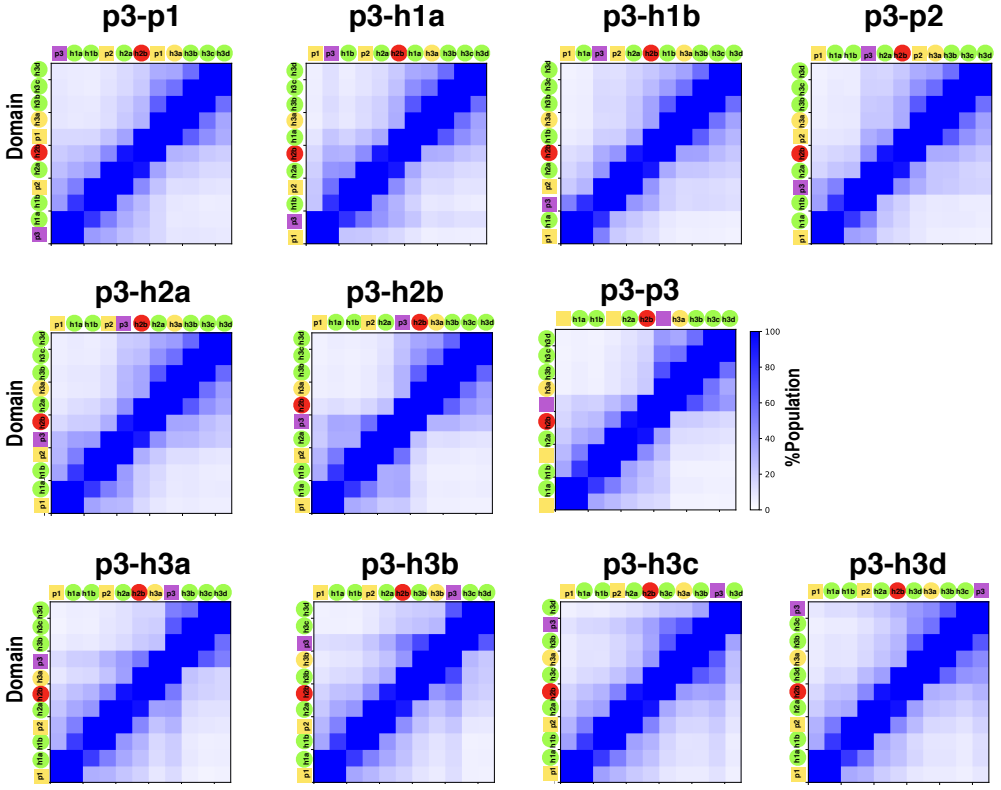
S2 Fig. Comparison of MD and NMR observables. C_α (top) and C_β (bottom) chemical shifts from NMR at 280K (black lines) [8] and MD at 300K (blue lines), calculated using SPARTA+ [7] for M66 sequence. Chemical shifts outside the range of 0.5 ppm from NMR chemical shifts are shaded grey.



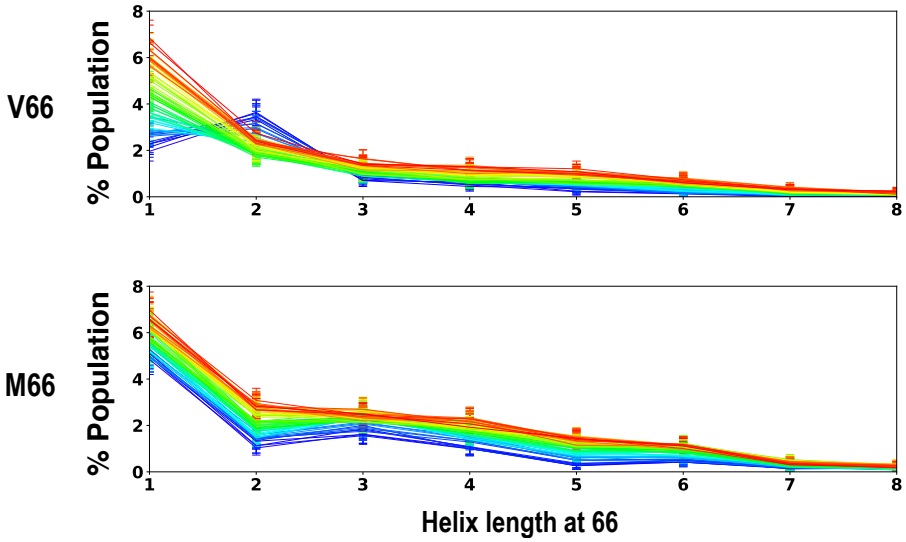
S3 Fig. Scaling behavior of each identified domain. Ensemble averaged interchain distance profiles for each domain in V66 sequence. Theoretical polymer scaling limits are shown with grey lines (prefactor A = 0.55 nm), while Flory exponents for each curve are given in Table S1.

Table S1. Flory exponent (ν) and prefactor (A) for each proregion domain, calculated from MD trajectories. Errors represent fit error of $\langle R_{|i-j|} \rangle$ vs $A|i-j|^{\nu}$, weighted by each point's standard deviation.

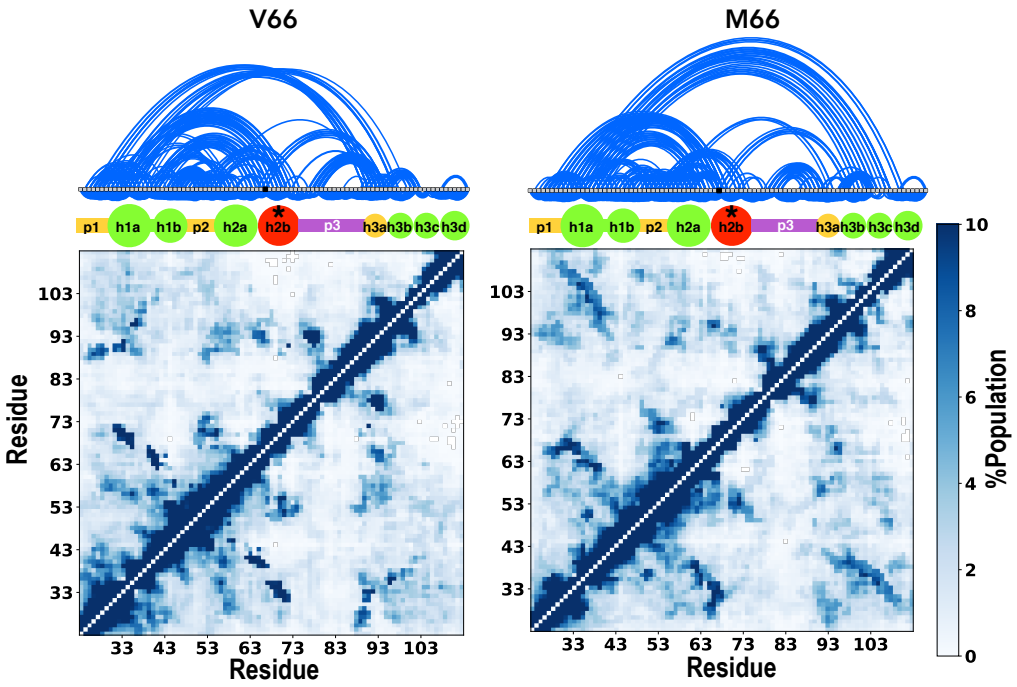
domain	A (V66)	ν (V66)	A (M66)	ν (M66)
p1	0.53 ±0.16	0.58 ±0.24	0.52 ±0.15	0.60 ±0.23
h1a	0.53 ±0.16	0.59 ±0.25	0.53 ±0.16	0.60 ±0.24
h1b	0.61 ±0.24	0.44 ±0.40	0.60 ±0.26	0.44 ±0.41
p2	0.56 ±0.22	0.50 ±0.34	0.59 ±0.18	0.51 ±0.29
h2a	0.56 ±0.17	0.50 ±0.22	0.57 ±0.17	0.48 ±0.23
h2b	0.64 ±0.14	0.51 ±0.18	0.60 ±0.15	0.56 ±0.20
p3	0.49 ±0.09	0.63 ±0.10	0.50 ±0.09	0.62 ±0.10
h3b	0.56 ±0.32	0.55 ±0.62	0.55 ±0.28	0.60 ±0.57
h3c	0.52 ±0.13	0.75 ±0.29	0.52 ±0.14	0.74 ±0.31
h3d	0.58 ±0.18	0.52 ±0.27	0.59 ±0.19	0.49 ±0.28



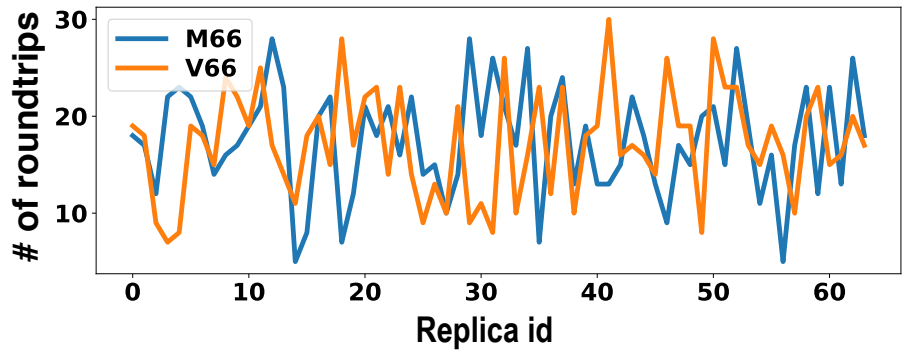
S4 Fig. Effect of perturbing monomer properties on freely-jointed, self-avoiding heteropolymer Contact probability maps from MC simulations, analogous to those in Figure 3 of the main text, in which the monomer with properties of linker domain p3 is swapped with every other monomer on the chain, with the new location represented by the purple square in the graph annotation. As the p3 bead is shifted along the chain, p3 and p1 consistently bound a white "forbidden" region that has little interaction with the rest of the protein.



S5 Fig. Effects of temperature and Val66Met mutation on helix propensity around residue 66. Frequency of helix of a given length at residue 66 in V66 (top) and M66 (bottom) in the temperature range of 300K to 385 K. With the increase in temperature the color transitions from cooler (blue) to hotter (red). It is entropically unfavorable for V66 and its neighboring residue to be simultaneously in the helical region of the Ramachandran map, as indicated by the decreasing helical propensity with increasing temperature. For longer helices, the trend will depend more on the additional side-chains in the helix, and the trend with temperature is reversed, but it remains weaker than the analogous trend for the M66 sequence. Errors represent standard error of a Bernoulli trial with n number of samples, where n is the product of total number unique replicas forming the helix of given length at residue 66 at a given temperature and average number of roundtrips per replica, 17.



S6 Fig. Residue level contacts for the entire prodomain. Contact probability between every residue pair for V66 (left) and M66 (right). A linear network of transient tertiary contacts is also shown (top) for each form. Two residue pairs are in contact if the distance between C_α-C_α atoms between the two residues are 0.8nm or less. The contact networks were build using Cytoscape [9] with a linear representation of residues. Each protein residue comprises a node in the network, with interactions between residues represented as edges. The strength of individual interactions can be interpreted by the thickness of the edge line on the network diagram. If the separation between residues forming the contact is more than 3, its edge is drawn above the node; otherwise, the edge is drawn at the bottom of the node. To focus on significant interactions, interactions showing more than 4% persistence were considered in network visualization.



S7 Fig. Number of round trip completed by each replica. Both V66 and M66 complete a minimum 5 roundtrips for each replica and an average of 17 round-tips per replica over the course of 76.8 μs (1.2 μs $\times$64) of simulations.

References

1. Lindorff-Larsen K, Piana S, Palmo K, Maragakis P, Klepeis JL, Dror RO, et al. Improved side-chain torsion potentials for the Amber ff99SB protein force field. Proteins. 2010;78(8):1950–8. doi:10.1002/prot.22711.

2. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, Simmerling C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins Struct Funct Bioinforma. 2006;65(3):712–725. doi:10.1002/prot.21123.

3. Piana S, Donchev AG, Robustelli P, Shaw DE. Water Dispersion Interactions Strongly Influence Simulated Structural Properties of Disordered Protein States. J Phys Chem B. 2015;119(16):5113–5123. doi:10.1021/jp508971m.

4. Best RB, Zheng W, Mittal J. Balanced Protein–Water Interactions Improve Properties of Disordered Proteins and Non-Specific Protein Association. J Chem Theory Comput. 2014;10(11):5113–5124. doi:10.1021/ct500569b.

5. Best RB, Hummer G. Optimized Molecular Dynamics Force Fields Applied to the Helix−Coil Transition of Polypeptides. J Phys Chem B. 2009;113(26):9004–9015. doi:10.1021/jp901540t.

6. Jorgensen WL. Quantum and statistical mechanical studies of liquids. 10. Transferable intermolecular potential functions for water, alcohols, and ethers. Application to liquid water. J Am Chem Soc. 1981;103(2):335–340. doi:10.1021/ja00392a016.

7. Shen Y, Bax A. SPARTA+: a modest improvement in empirical NMR chemical shift prediction by means of an artificial neural network. J Biomol NMR. 2010;48(1):13–22. doi:10.1007/s10858-010-9433-9.

8. Anastasia A, Deinhardt K, Chao MV, Will NE, Irmady K, Lee FS, et al. Val66Met polymorphism of BDNF alters prodomain structure to induce neuronal growth cone retraction. Nat Commun. 2013;4:2490. doi:10.1038/ncomms3490.

9. Ahlstrom LS, Baker JL, Ehrlich K, Campbell ZT, Patel S, Vorontsov II, et al. Network visualization of conformational sampling during molecular dynamics simulation. J Mol Graph Model. 2013;46:140–9. doi:10.1016/j.jmgm.2013.10.003.