Supporting Information

Detection of Native-State Non-Additivity in Double Mutant Cycles via Hydrogen Exchange

Joshua A. Boyer[‡], Cristina J. Clay[§], K. Scott Luce[], Marshall H. Edgell[],and Andrew L. Lee [*,‡,§]

EX1 vs. EX2 Behavior. In lieu of the dependence of all measurements that follow on the exchange regime of the amides, EX2 behavior is demonstrated here for the V18A/V54A cycle (Figure S1). In the EX2 regime, the direct comparison of each residue's measured $\log(k_{ex})$ at varying pH's will produce a slope of unity [1], with the intercept reflecting the difference in pH's. If the exchange is independent from pH, EX1 behavior is present and the intercept will approach zero. The variants of the V18A/V54A cycle are sufficient to infer EX2 behavior for all eglin c mutants and conditions studied here. This cycle contains the most drastic destabilizations with small but significant changes in both structure and dynamics of the single mutant variants [2-3]. The globally measured destabilization incurred by the double mutant is equivalent to the destabilization of the WT protein induced by the presence of 2.2 M GdnHCl, and therefore should serve as a proxy for this, and less severe buffer conditions.

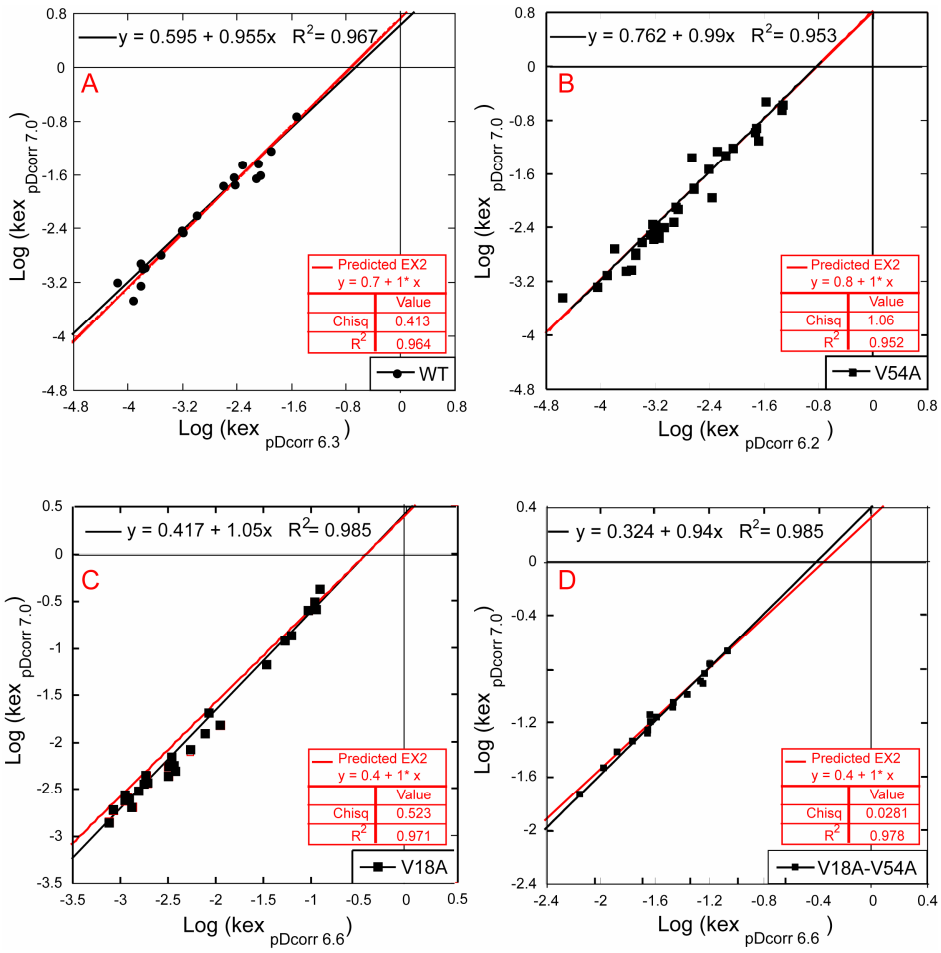

Figure S1. EX1 vs. EX2 behavior in the V18A/V54A cycle variants. Panels (A) WT (B) V54A (C) V18A and (D) V18A/V54A. By comparing the log(k_{ex}) of each reside at varying pH's, a linear correlation is found with a slope approaching one in each of the variants in this cycle and a intercept indicative of the difference in pH. This relationship is indicative of the EX2 exchange scheme. If EX1 behavior were present, the intercept would approach zero, indicating the exchange rates independence from pH.

Local vs. global exchange in the reference, WT. Because of the multiple exchange mechanisms accessible to a labile amide-hydrogen bond, to aid in analysis it is important to first ascertain the means by which each amide exchanges. In order to identify residues participating in a global or a mixed exchange mechanism, we examined the dependence of local free energies on denaturant concentrations[4-6]. It follows that residues exhibiting linear dependencies with slopes approximating the globally determined m-value are in fact exchanging globally. Via this

analysis, residues 16, 21, 35, 54, 56, and 63 globally exchange in the WT reference (see, for example, Figure S2 for residue 35). We note that this laboratory has measured the global stability for WT eglin c, using fluorescence, hundreds of times over the last 6-7 years. For reasons not entirely clear, the fitted stabilities have gradually increased from ~6.0 to ~6.5 kcal/mol. Importantly, however, *changes* in stability upon mutation do not change and WT stability is always remeasured as a reference. Table S1 shows the HX-derived fitted free energies and m-values from the denaturant dependence studies of globally exchanging residues. These global residues have fitted stability values within error of the globally determined fluorescence stability when a proline isomerization correction and a feasible D_2O solvent isotope effect of up to 1.4 kcal/mole are applied (see Table S1)[7].

Residues with significantly lower slope prior to the intersection with the global mechanism denote a mixed mechanism, and residues with no slope exchange via a local breathing/unfolding mechanism (refer to *Figure S2*). As such, residues 25, 37, and 52 have been identified as exchanging through mixed mechanisms and the remaining residues exhibit local exchange.

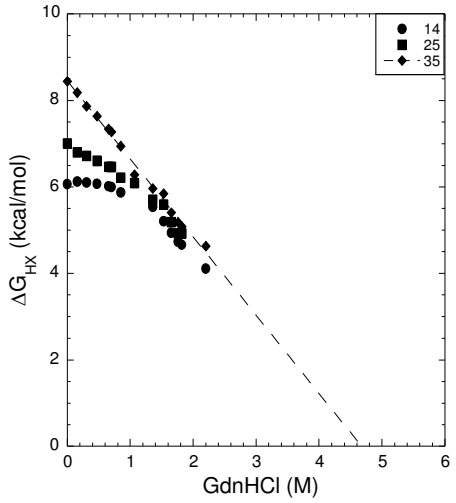

Figure S2. Determination of the amide exchange mechanisms in wild-type eglin c. Residues that exchange only with the unfolding of the protein, exchange via the global mechanism. These residues can be identified by a strict dependence between the local stability (ΔG_{HX}) and chaotrope concentration (GdnHCl) that approximates the m-value determined via traditional stability measurements (dashed line). Here residue Y35 exchanges through the global mechanism. Residue F25 experiences a mixed exchange mechanism, identified by a lesser dependence on co-solute. Lastly, residue V14 exchanges through local-breathing events or a local mechanism that is independent of denaturant concentration.

Table S1: ΔG_{HX} at 25 °C for globally exchanging residues extrapolated to 0 M GdnHCl from HX GdnHCl titration (as shown in Figure S2).

Residue	**ΔG_{HX}**[a]	**m value**[b]	**R^2**
K16	7.86±0.04	-1.46±0.05	0.999
A21	8.64±0.03	-1.74±0.03	0.997
Y35	8.45±0.05	-1.81±0.04	0.993
V54	7.99±0.04	-1.63±0.03	0.995
Y56	7.85±0.09	-1.61±0.07	0.984
V63	7.16±0.05	-1.44±0.04	0.988
Global Stability[c]	6.59±0.17	-1.80±0.01	
+ Proline Corrected[d]	7.22±0.17		
+ Solvent Isotope Effects[e]	8.62±0.17		

[a] All values are in kcal/mol.
[b] All values are in kcal/mol•M GdnHCl.
[c] As measured by fluorescence (see supp. info. text).
[d] Proline isomerization correction, 0.722, calculated as described [7].
[e] Possible solvent isotope effect 1.4 kcal/mol [7].

Local vs. global exchange in mutants. Another mode of analysis for ascribing exchange mechanisms that provides analogous results to GdnHCl dependence is standardization of $\Delta\Delta G_{HX,i}^{mut\text{-}WT}$ [1]. In dividing the local free-energy change ($\Delta\Delta G_{HX,i}^{mut\text{-}WT}$) by the globally-measured, $\Delta\Delta G_U^{mut\text{-}WT}$, residues can be classified by ζ-values (here $\zeta = \Delta\Delta G_{HX,i}/\Delta\Delta G_U$). Using the cutoff of Neira et al., the exchange mechanisms of the variants were determined. Residues with $\zeta > 0.8$ were classified as globally exchanging and $0.6 < \zeta < 0.8$ as possessing a mixed mechanism [1]. The classification of each residue of the variants can be seen in Table S2.

Table S2: Residue exchange mechanism in cycle variants

Resid eg C	WT[a]	V18I[b]	L27I[b]	V18I[b] L27I	V34L[b]	P58Y	V34L P58Y	V18A	V54A	V18A V54A	CI2[c]	Resid CI2[c]
10	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local		5
13	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	8
14	Local	Local	Local	Local	Local	Local	Local	Local	Local	Mixed	Local	9
15	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local		
16	Global	Local	Local	Local	Local	Global	Global	Global	Global	Global	Global	11
17	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local		
18	Local	Global	Local	Global	Local	Local	Local	Local	Local	Local	Local	13
20	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local		
21	Global	Global	Local	Global	Global	Global	Global	Global	Global	Global	Local	16
22	Local	Global	Local	Global	Local	Local	Local	Global	Local	Mixed	Local	17
23	Local	Global	Local	Global	Local	Local	Local	Global	Local	Local	Local	18
24	Local	Local	Local	Local	Local	Local	Local	Mixed	Local	Local		19
25	Mixed	Local	Local	Local	Local	Local	Local	Mixed	Local	Mixed	Global	20
26	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Global	21
27	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	22
29	Local	Local	Local	Local	Local	Local	Local	Mixed	Local	Local	Local	24
32	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	27
33	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	28
34	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local		
35	Global	Local	Local	Local	Local	Local	Global	Global	Global	Global	Global	30
37	Mixed	Local	Local	Local	Local	Global	Global	Mixed	Local	Mixed	Local	32
39	Local	Local	Local	Local	Local	Global	Local	Local	Local	Local		34
51	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local		46
52	Mixed	Local	Local	Local	Local	Local	Mixed	Mixed	Mixed	Global	Global	47
53	Local	Local	Local	Local	Local	Local	Local	Local	Local	Mixed		48
54	Global	Local	Local	Local	Global	Mixed	Global	Global	Global	Global	Global	49
55	Local	Local	Local	Local	Local	Local	Local	Mixed	Local	Mixed	Global	50
56	Global	Local	Global	Local	Local	Global	Global	Global	Global	Global	Global	51
57	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	52
62	Local	Local	Local	Local	Local	Local	Local	Mixed	Local	Mixed	Local	56
63	Global	Local	Local	Local	Global	Global	Global	Global	Global	Global		57
68	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local		62
69	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local	63
70	Local	Local	Local	Local	Local	Local	Local	Local	Local	Local		64

[a]Measured via GdnHCl dependency. [b]$|\Delta\Delta G_U| <$ 0.5 kcal/mol [c]Neira et al., 1997. Stabilities of variants decrease from left-to-right. Variants highlighted in gray are not included in Figure S3. Residues highlighted in yellow are believed to be false negatives generated by shortcomings of the method as $\Delta\Delta G_U$ approaches zero.

As with CI2, a majority of the residues exchange via local exchange and the bulk maintain WT's mechanisms of exchange (Figure S3). As one would expect, the reliability of this method breaks down as $|\Delta\Delta G_U|$ approaches zero (gray in Table S2), in these instances we assume the mechanism of WT is maintained since the overall stability is minimally affected. Also of note, as the stability of the variant decreases (moving left-to-right in Table S2) it appears that the number of residues exchanging through a global or mixed mechanism increases.

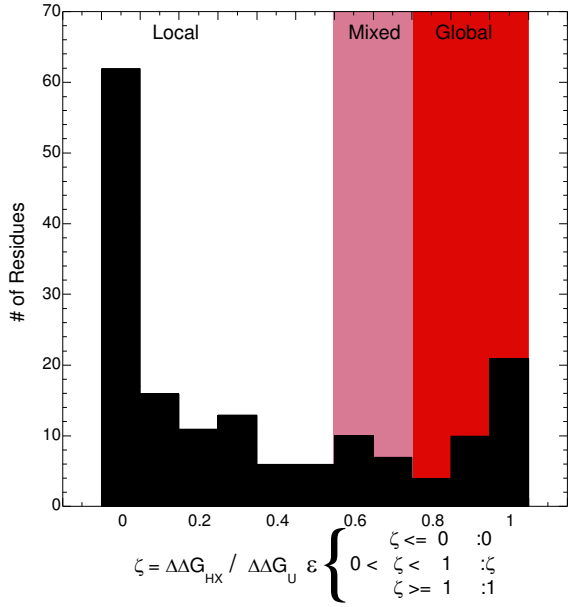

Figure S3 Distribution of residue exchange mechanisms. Each mutant's residues exchange mechanism was determined by standardizing $\Delta\Delta G_{HX}$ by $\Delta\Delta G_U$($\zeta = \Delta\Delta G_{HX}/ \Delta\Delta G_U$). ζ-Values less than zero were redefined as zero, and φ-values greater than one as one. Residues were then categorized by binning each mechanism such that ζ <0.6 delineated local exchangers, 0.6 <= ζ <= 0.8 mixed, and ζ > 0.8 demarcated global exchangers. As suggested (Neira et al., 1997) a majority residues exchange locally. Mutants with $|\Delta\Delta G_U|$< 0.5 kcal/mol were excluded (shown gray in table S2) from this analysis and WT mechanism assumed to be maintained.

Response to mutation via hydrogen exchange. The following description details the individual amides' change in HX stability (i.e. change in free energy upon "unfolding", with positive $\Delta\Delta G_{HX}$ corresponding to slowed exchange in the mutant relative to WT) in relation to WT upon mutation ($\Delta\Delta G_{HX}$) in three separate double-mutant cycles of varying additivity and distance, as measured between C_β to C_β of the mutated residues. The additive cycle is composed of mutations V18I and L27I and yields an essentially zero, globally- determined, coupling of -0.12 ± 0.20 kcal/mol. These positions are ~15 Å apart in the WT structure. The changes in ΔG_{HX} upon these mutations can be seen in Figure S4A. In the V18I mutation, positions 18, 20, 23, 33, 35, 37, 53, 56, 63, and 69 show significant change as they exceed the 0.3 kcal/mol cutoff (see methods) with positions 20 and 56 showing large changes. The L27I mutation shows changes at positions 14, 17, 22, 24, 27, 29, 33, 37, 56, 57, 68, and 69 with large changes apparent at 14 and 56. When compared to WT, the double mutant displays changes at 13,17, 18, 21, 26, 27, 29, 32, 33, 53, 57, and 69. This mutation marginally destabilizes the protein globally by -0.3 kcal/mol, yet nearly all of the significant changes indicate increases in local stability. However, sites shown to exchange through the global mechanism in WT display destabilization equivalent to the globally determined value.

The smaller coupled cycle consists of mutations V34L and P58Y. These positions are separated by ~17 Å and resulted in a non-zero coupling of -0.38 ± 0.13 kcal/mol as determined by classical means and exceeds the cutoff suggested by Chen and Stites[8]. The changes in ΔG_{HX} in variants of the V34L/P58Y cycle are shown in Figure S4B. In the V34L mutation, residues 13, 14, 17, 21, 22, 23, 24, 25, 26, 27, 32, 33, 35, 37, 52, 53, 57, 68, and 69 showed changes greater than the 0.3 kcal/mol cutoff. This mutation marginally destabilizes the protein globally by -0.3 kcal/mol, yet nearly all of the significant changes indicate increases in local stability.

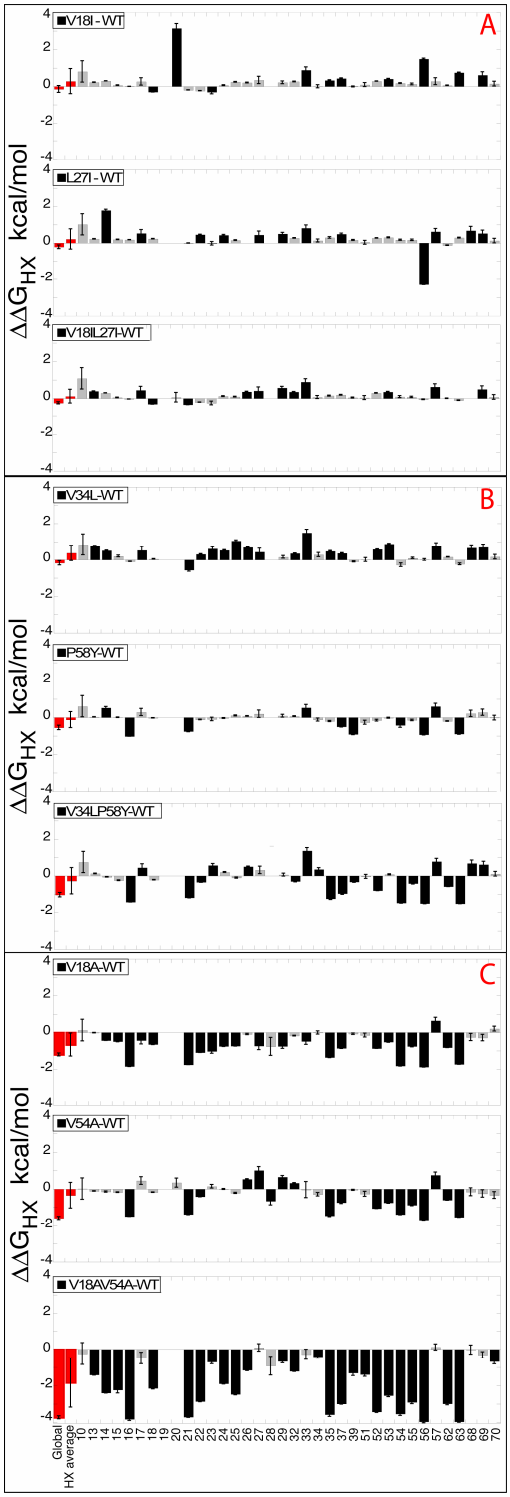

Figure S4 HX-determined changes in local stabilities upon mutation. $\Delta\Delta G_{HX}$ shown as mutant minus WT in the additive V18IL27I (A), the small coupled V34LP58Y (B), and the large coupled V18AV54A (C) cycles. Black bars show significant changes and red bars show globally measured and HX average values.

However, sites shown to exchange through the global mechanism in WT display destabilization equivalent to the globally determined value. In the P58Y variant, residues 14, 16, 21, 33, 37, 39, 54, 56, 57, and 63 display changes in response to the mutation, with residues 16, 21, 39, 56, and 63 revealing destabilization that exceeds that of the globally determined change of -0.67±0.08 kcal/mol. Moreover, the double mutant of this cycle appears to adopt traits from each of the single mutants, as there are increases in local stabilities and decreases that exceed the global destabilization. Residues 16, 17, 21, 22, 23, 26, 32, 33, 34, 35, 37, 39, 52, 54, 55, 56, 57, 62, 63, 68, and 69 all show sizeable changes.

The V18A/V54A cycle displays the largest thermodynamic coupling of -0.95 ± 0.19 kcal/mol. An ~8Å distance separates these residues. Changes in local stabilities for the mutants of this cycle are shown in Figure S4C. V54A has been discussed previously [2]. Similar to P58Y, the globally exchanging residues' response to the V18A mutation exceeds the classically-measured destabilization of -1.21 ± 0.08 kcal/mol. Nearly all amide reporters show significant change in free energies with this mutation, specifically residues 14, 15, 16, 17, 18, 21, 22, 23, 24, 25, 27, 29, 33, 35, 37, 52, 53, 54, 55, 56, 57, 62, and 63. This large- scale response is even more pronounced in the double mutant of this cycle. Residues 13, 14, 15, 16, 18, 21, 22, 23, 24, 25, 26, 29, 32, 34, 35, 37, 39, 51, 52, 53, 54, 55, 56, 62, 63, and 70 all show significant changes upon the dual mutation. Consequently, the variants containing the V18A mutation have some indications of small structural perturbation discussed later.

Site-specifically observed (SSO) coupling values. The site-specific $\Delta\Delta G_{HX}$'s determined above (Figure S4) were assembled into a double mutant cycle at each reporting residue via equation 2 (main text). Just as one would determine a thermodynamic coupling using $\Delta\Delta G_U$'s, the sum of

the single mutants' $\Delta\Delta G_{HX,i}$'s subtracted from the double mutant's $\Delta\Delta G_{HX,i}$ provides a coupling value reported at residue i . These SSO coupling values have been tabulated and can be seen graphically in Figure 2 (main text). In the V18I/L27I cycle there are SSO coupling values that differ significantly from zero (>0.4 kcal/mol), despite the absence of a non-zero coupling measured by traditional fluorescence-monitored gdnHCl titration. Residues 14, 22, 33, 35, 37, 56, 63, and 69 show such couplings. Interestingly, residues 14, 33, and 69, all exchange via the local mechanism in every variant of the cycle. In the small non-additive cycle (V34L/P58Y), residues 13, 14, 15, 22, 25, 32, 33, 35, 37, 39, 52, 53, 54, 56, 57, and 62 show significant coupling values. Again, local-exchanging residues (13, 14, 15, 22, 32, 33, 53, 57, and 62) report a non-zero coupling. Finally, in the large non-additive cycle (V18A/V54A) many residues display significant coupling values. Residues 13, 14, 15, 16, 18, 21, 22, 24, 25, 26, 29, 32, 35, 37, 39, 51, 52, 53, 55, 57, 62, 63, and 70 show non-zero additivity. In the variants of this cycle, several residues change mechanisms in part to the disruptive mutations including V18A. Despite the disturbance, residues 13, 15, 18, 26, 32, 39, 51, 57, and 70 maintain their local exchanging mechanism and report not only non-zero coupling values, but several exceed the values of residues that exchange through global and mixed mechanisms.

Structural comparison across variants. Residual dipolar couplings ($N\text{-}H^N$ RDCs) were measured using non-ionic lipid bicelles[3] or stretched 6% polyacrylamide gels [9] to elucidate the mutations' effects on eglin c backbone structure. Residual dipolar couplings were determined for V34L and V18A/V54A with bicelles, while WT, P58Y, and V34L/P58Y required the use of the gel system due to an apparent interaction between variants containing the P58Y mutation and the lipid bicelles. V54A and V18A RDC analyses have been published previously [10]. In all

variants, alignment was sufficient to generate maximal RDCs > 30 Hz and 14 Hz using bicelles and gels, respectively. Linear relationships were calculated for each mutant compared to WT. Residuals from these fits were added to data collated previously [10] for statistical analyses. The standard deviation of RDC residuals was determined to be 1.9Hz. Using a 95% confidence interval, any residue with a residual > 2 standard deviations from the mean can be considered an outlier. This RDC structural metric has been discussed previously for the V18A and V54A mutations [10]. These data are shown here with the new mutants in Figure S5.

In conjunction with RDC's, chemical shift perturbation was used to search for aberrant backbone structural features. In equation S1, we employ a chemical shift vector($\vec{\delta}$) representation to enable quick and direct comparison between variant and WT chemical shifts via a linear fit to $\vec{\delta}_{WT}$ vs. $\vec{\delta}_{mut}$ and identification of outliers by studentizing residuals as shown in Figure *S6*.

$$\vec{\delta} = \sqrt{{\delta_{1}}_{H}^{2} + (0.1 \times \delta_{15_N})^2} \qquad \text{(eq. S1)}$$

The standard deviation of residuals for all combined chemical shift vectors ($\vec{\delta}$) measured here was 0.08 ppm.

In Figures S5 and S6, note the residues (demarked with asterisks) with residuals beyond the red lines– denoting the 2σ limit. Here, the V18I-L27I cycle displays no evidence for structural change in $\vec{\delta}$ except for the sites at or adjacent to mutation (specifically 18 and 28) and therefore RDC measurements were not pursued. The V34L/P58Y cycle showed minor perturbation in $\vec{\delta}$ about the P58Y mutation in both single and double variants. Whereas within the V34L variant, there appears to be evidence for slight repacking of the core, accommodating the larger side chain. Unsurprisingly, residues 34 and 25 show changes in the RDC (Figure S5) comparison as these sites pack against each other in the WT structure. The site adjacent to the

mutation, Y35, and residues 25, 29, 32, and W10's indole (members of the adjacent, aromatic portion of the core) show small but notable changes in the $\vec{\delta}$ comparison, further supporting slight alteration to accommodate the extra atoms of the leucine moiety. Finally, the double mutant cycle containing the V18A mutation (V18A-V54A) exhibits response (both $\vec{\delta}$ and RDC) at the same residues affected in the single mutant (V18A), yet it seems the presence of the second mutation exacerbates the effects in some reporters (residues 12, 16, 25, 33, 35, 54, 64, 68) and quenches it in others (residues 18, 22, 32, 50, 51, 55, 56, 57, 62, 64, 65).

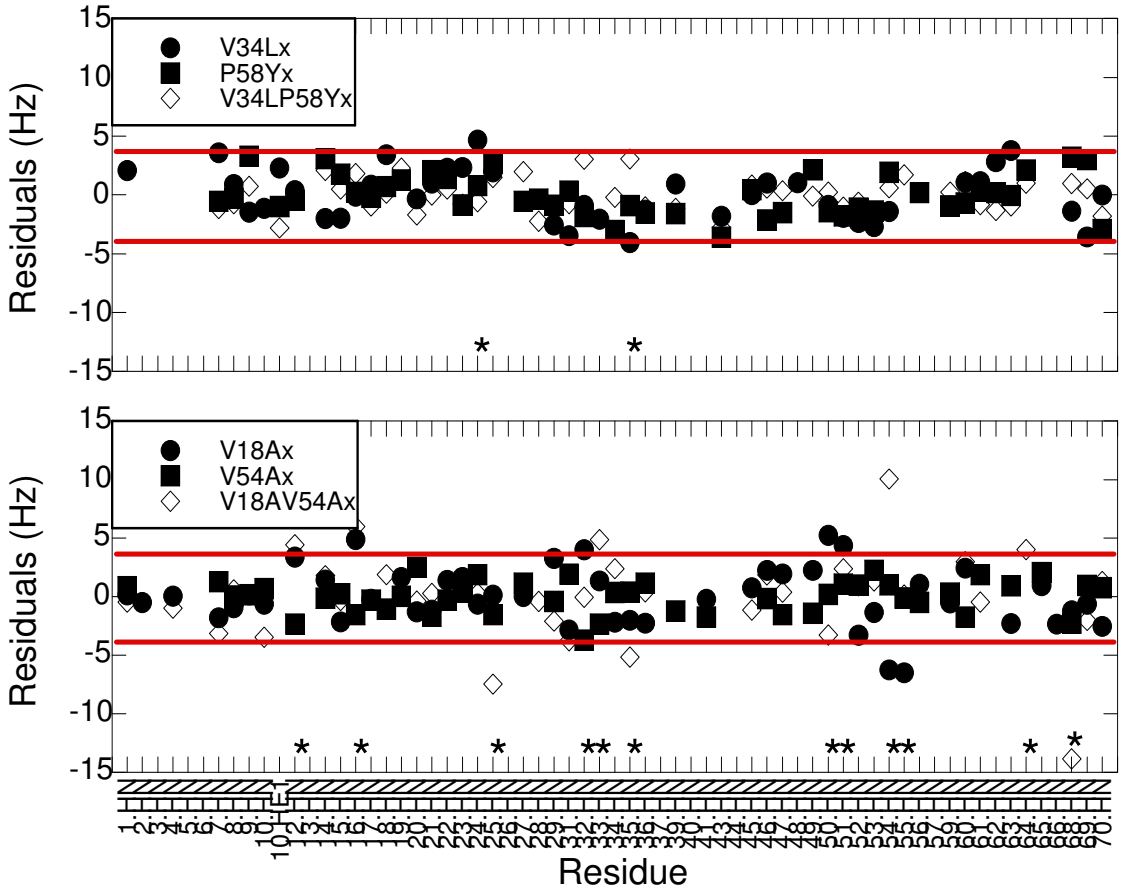

Figure S5 Comparison of residual dipolar couplings (RDCs) between wild-type and mutant proteins. Mutant residues' RDC values were directly compared to the Reference WT values by linear fit. Resulting residuals are shown per residue. The residuals of all mutants were collated with previous data (Clarkson & Lee 2006) yielding a standard deviation of 1.89Hz. Thus with a 95% confidence interval all residuals that exceed 3.8 Hz are deemed outliers. The 3.8 Hz cutoff is shown as a red horizontal line. In the legend, "x" denotes the background F10W mutation (see main text). Asterisks denote outliers in one of the variants.

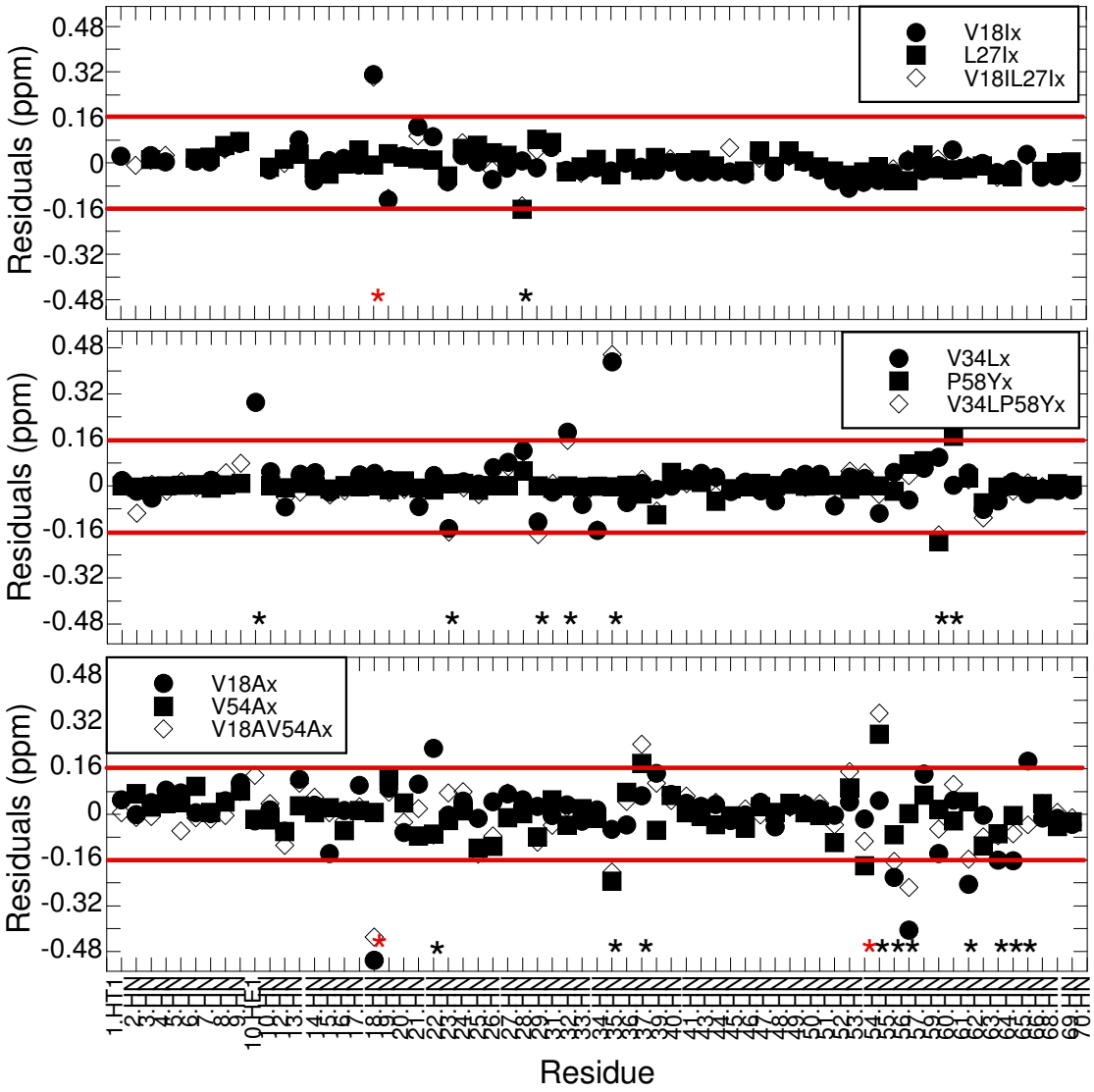

***Figure* S6 Chemical shift perturbation due to mutation.** In order to quickly identify residues with possible aberrant structure each mutant's residues chemical shift vector (eq. S1) was directly compared to that of the reference WT via a linear fit. Residuals of that fit are shown per residue. The residuals of all mutants were collated for statistical analysis yielding a standard deviation of 0.08 ppm. Thus with a 95% confidence interval all residuals that exceed 0.16 ppm are deemed outliers. The 0.16 ppm cutoff is shown as a red horizontal line. Asterisks denote outliers in one of the variants and red asterisks indicate sites of mutation.

References

(1) Neira, J. L.; Itzhaki, L. S.; Otzen, D. E.; Davis, B.; Fersht, A. R. *J Mol Biol* **1997**, *270*, 99.

(2) Boyer, J. A.; Lee, A. L. *Biochemistry* **2008**, *47*, 4876.

(3) Clarkson, M. W.; Lee, A. L. *Biochemistry* **2004**, *43*, 12448.

(4) Maity, H.; Lim, W. K.; Rumbley, J. N.; Englander, S. W. *Protein Sci* **2003**, *12*, 153.

(5) Mayne, L.; Englander, S. W. *Protein Sci* **2000**, *9*, 1873.

(6) Bai, Y. *Chem Rev* **2006**, *106*, 1757.

(7) Huyghues-Despointes, B. M.; Scholtz, J. M.; Pace, C. N. *Nat Struct Biol* **1999**, *6*, 910.

(8) Chen, J.; Stites, W. E. *Biochemistry* **2001**, *40*, 14004.

(9) Whitley, M. J.; Zhang, J.; Lee, A. L. *Biochemistry* **2008**, *47*, 8566.

(10) Clarkson, M. W.; Gilmore, S. A.; Edgell, M. H.; Lee, A. L. *Biochemistry* **2006**, *45*, 7693.