Supplementary Information for:
Experimental evidence for a frustrated energy landscape in a 3-helix bundle protein family

Beth G. Wensley, Sarah Batey, Fleur A. C. Bone, Zheng Ming Chan, Nuala R. Tumelty, Annette Steward, Lee Gyan Kwa, Alessandro Borgia and Jane Clarke

This file contains 3 sections:

- Supplementary Results and Discussion  (p2)
- Supplementary Tables  (p12)
- Supplementary Figures  (p14)

Each section is complete with its own set of References
Supplementary Results and Discussion

Is an internal friction model appropriate for describing the response of the spectrin domains to changes in solvent viscosity, $\eta$?

In our manuscript we have fitted the kinetic data to eq. 1, a modified version of Kramers’ theory which ascribes deviations from the expected $1/\eta$ dependence of the rate constants to the presence of internal friction, $\sigma$:

$$k = \frac{C}{\eta + \sigma} \exp\left(-\frac{\Delta G^\ddagger}{RT}\right)$$ (1).

In this Supplementary Results and Discussion we consider alternative models for explaining solvent viscosity effects on protein dynamics and examine in more depth whether the whether the mechanistic explanation we have offered is reasonable given the magnitude of the internal friction observed.

**Power-law dependence of protein dynamics and solvent viscosity effects.**

An alternative, power-law relationship between rate constants and viscosity data has been proposed:

$$k = A\eta^{-\kappa}$$ (2)

For a simple system, where the rate constant is directly proportional to solvent viscosity, $\kappa = 1$, equivalent to the slope of relative rate constant vs. relative viscosity $= 1$ in the plots shown in **Fig. 2b, Main Text**.

Experimentally, values for $\kappa$ have been determined for a range of peptides and full-sized proteins. In some systems $\kappa$ is close to 1 (0.96 for a short (Gly-Ser)$_l$ peptide, 0.93 for protein L), but in others $\kappa$ is significantly lower (0.55 for cytochrome c and 0.53 for a short $\alpha$-helix). Plots of this kind (**Fig. 1 below**) show that R15 is unremarkable, with $\kappa$ falling close to 1 (mean $\kappa = 0.85 \pm 0.08$). However, R16 and, particularly, R17 fall well outside the range of what has been observed previously for any system ($\kappa = 0.31 \pm 0.06$ and $0.17 \pm 0.09$, respectively), reflecting the
unusually weak dependence of their rate constants for folding and unfolding on solvent viscosity (Table 1).

Table 1. Results from the power-law fit for the spectrin domains

| Determined using:        | R15     | R16     | R17     |
|--------------------------|---------|---------|---------|
| $k_f$ (@ $\Delta G = 1.5$ kcal/mol) | 0.84±0.010 | 0.32±0.11 | 0.10±0.11 |
| $k_u$ (@ $\Delta G = 1.5$ kcal/mol) | 0.94±0.08 | 0.36±0.010 | 0.27±0.12 |
| $k_f = k_u$ (@ $\Delta G = 0$ kcal/mol) | 0.78±0.06 | 0.24±0.07 | 0.13±0.05 |
| mean (±S.D)              | 0.85±0.08 | 0.31±0.06 | 0.17±0.09 |

Figure 1. Power law fits of spectrin kinetic data; R15, black; R16, red; R17, blue.
(a) using $k_f$ values at $\Delta G = 1.5$ kcal mol$^{-1}$. (b) using $k_u$ values at $\Delta G = 1.5$ kcal mol$^{-1}$.
(c) using $k_f = k_u$ values at $\Delta G = 0.0$ kcal mol$^{-1}$.

The power-law fit (eq. 2) requires an explanation for deviation from the $1/\eta$ dependence which does not invoke internal friction.
Hofrichter and co-workers\textsuperscript{7}, intriguingly, see different $\kappa$ values for folding of a small (21-residue) helical peptide ($\kappa \sim 0.6$) vs. a small (16-residue) $\beta$-hairpin ($\kappa \sim 1$). Although both the internal friction and power-law fits for the helical peptide are approximately equally good, they argue that invoking significant internal friction in such a small system is unsatisfactory. They postulate differences in the energy barriers for the elementary steps in helix and turn formation might explain the differences in viscosity dependence. Since all three spectrin domains are alpha helical, and significant helical structure is present in all three transition states, this explanation for the very different $\kappa$ values for the spectrin domains seem unlikely, in particular in these large systems where the elementary steps are unlikely to be rate limiting.

Frauenfelder and co-workers have invoked a slaving model, whereby the protein chain is slaved by the hydrating solvent shell and thus the motions of the protein chain are dominated by $\alpha$-relaxation of the solvent\textsuperscript{3}. They use temperature-dependent folding vs. viscosity data to argue that the slaving model can even explain the large-scale motions in CO-myoglobin and in cytochrome c, which give particularly low values of $\kappa$ ($\sim 0.55$)\textsuperscript{3,8}. This differs from the analysis of Eaton and Hagen and their co-workers who invoke internal friction as the source of deviation from the viscosity dependence, in the fast reorganisations of these largely folded systems\textsuperscript{1,6}.

Is it possible that the large differences in viscosity scaling can be attributed to differences in solvent slaving – to differences in the interactions of the spectrin domains with the solvent? The strength of the comparative study of proteins from the same family is that we are comparing like with like. The proteins in question are exactly the same length and they all have the same native structure. The overall sequence composition is very similar in terms of hydrophobicity, and number of charged and aromatic residues. Furthermore the rate limiting transition states for folding are similar in terms of solvent accessibility. Thus it seems very unlikely that a solvent-slaving model can explain such a vast range in the response of the rate constants of folding and unfolding to solvent viscosity ($\kappa$ values that range from 0.9 – 0.2). The internal friction model offers a structurally more reasonable explanation (see below and Main Text).
Determining the relative magnitude of the landscape roughness.

(i) Evaluation of internal friction.

From eq. 1, at constant $\Delta G^\ddagger$,

$$k \propto \frac{C}{\eta + \sigma}$$  \hspace{1cm} (3)

Thus a plot of $\frac{1}{k}$ vs. solvent viscosity, $\eta$, allows internal friction, $\sigma$, to be evaluated in cP. The results are shown in Fig. 2 and Table 2, below.

**Figure 2.** Fits of spectrin kinetic data used to determine internal friction, $\sigma$. (a) R15, (b) R16, (c) R17. Black, using $k_f$ values at $\Delta G = 1.5$ kcal mol$^{-1}$; Green, using $k_u$ values at $\Delta G = 1.5$ kcal mol$^{-1}$; blue, using $k_f = k_u$ values at $\Delta G = 0.0$ kcal mol$^{-1}$. 
Table 2 Values of internal friction, $\sigma$.

| Determined using: | $\sigma$ (cP) |
|-------------------|---------------|
|                   | R15 | R16 | R17 |
| $k_f (\Delta G = 1.5$ kcal/mol) | 0.25±0.23 | 3.4±1.7 | 18±25 |
| $k_u (\Delta G = 1.5$ kcal/mol) | 0.09±0.14 | 3.6±1.7 | 5±3 |
| $k_f = k_u (\Delta G = 0$ kcal/mol) | 0.42±0.16 | 6.3±2.8 | 13±7 |
| mean (±S.D)       | 0.25±0.16 | 4.4±1.6 | 12.0±6.6 |

Although associated with significant error, the internal friction of the R16 and R17 domains are significantly and consistently higher than those of R15, consistent with the low solvent viscosity dependence of the data. We note that the only systems where such a high value of $\sigma$ has been observed previously is for reconfigurations of almost completely folded haem proteins$^{1,6}$.

We can now use these data to evaluate the differences in landscape roughness.

(ii) Evaluation of the relative roughness of the energy landscape of the spectrin domains.

Zwanzig$^9$ presents an expression for the effect of roughness on a one-dimensional energy landscape:

$$D^* = D \exp \left( -\frac{\epsilon}{RT} \right)^2$$

(4)

where $D$ is the diffusion coefficient characteristic of a smooth landscape with a given shape, $\epsilon$ is the characteristic energy of the roughness and $D^*$ is the effective diffusion coefficient.

This expression can be used to determine the relative roughness of the folding landscapes of the slow folding proteins R16 and R17, compared to R15, by assuming
that the additional roughness, $\Delta \varepsilon$, can be attributed to change of internal friction, $\sigma$, and that then

$$\sigma^* = \sigma \exp \left[ \frac{\Delta \varepsilon}{RT} \right]^2$$

(5).

By associating $\sigma$ with the smooth landscape of R15 and $\sigma^*$ with the rough landscape of R16 or R17, we can calculate $\Delta \varepsilon$ as:

$$\Delta \varepsilon = RT \sqrt{\ln \frac{\sigma^*}{\sigma}}$$

(6)

Using the Ansari et al.\textsuperscript{1} adaptation of Kramers’ equation (eq.1) $\frac{\sigma_{R16}}{\sigma_{R15}}$ can be evaluated. The ratios of internal friction would be:

$$\frac{\sigma_{R16}}{\sigma_{R15}} = \frac{4.4}{0.25} = 17.6$$

(7)

and

$$\frac{\sigma_{R17}}{\sigma_{R15}} = \frac{12.0}{0.25} = 48.0$$

(8)

Thus, using eq. 6, the characteristic roughness in R16 and R17, above the intrinsic roughness of R15, $\Delta \varepsilon$, is $\sim 1.7 \cdot RT$ and $\sim 2.0 \cdot RT$ respectively.

This result suggests that the increase in the ruggedness of the folding landscapes of the slow folding spectrin domains can be accounted for by barriers of the order of 1.7–2 $RT$. This is similar to the roughness observed in single molecule and ensemble measurements of unfolded proteins and small peptides, and theoretical studies (0.5 – 3 $RT$).\textsuperscript{11-14} We propose in our manuscript that non-native mis-docking of preformed helices might be responsible for the roughness in the folding landscape. It is not unreasonable to propose that such transient non-native structure formation would lead
to kinetic traps in the landscape of the order seen in the dynamics of fast folding proteins or of transient contact formation in unfolded states.

**Conclusion**

Analysis of the kinetic data show that the solvent viscosity of the slow folding spectrin domains fall outside that seen in any other protein system. Importantly the same features are seen in folding and unfolding, suggesting this behaviour is a feature of the transition states for folding. Since the spectrin domains have the same size, structure, similarly collapsed transition states and sequence, the most likely origin of this behaviour is increased internal friction in the slow folding proteins. The magnitude of the increased landscape roughness is reasonable for the proposed transient misfolding as the protein diffuses across the energy landscape. What is remarkable is that such relatively small increases in average height of the small potential barriers that describe fluctuations in the free energy landscape for folding can result in such dramatic changes in response to solvent.

† It has been pointed out by the referees of this manuscript that Zwanzig’s quadratic expression results from the assumption of random roughness in the landscape, with a Gaussian distribution of barrier heights, and gives a lower limit of the roughness. In the alternative periodic potential limit, there would be a series of well-defined intermediates as the energy dips up and down. This would result in a higher estimate of roughness \((3-4 \, RT)\). At this limit one would perhaps expect to see populated intermediates which we do not observe in the spectrin domains.
Appendix

In the core-swapped proteins the faster folding is accompanied by a significant decrease in internal friction (and thus in the magnitude of the characteristic landscape roughness, compared to the ‘major parent’ proteins) – see Table 3.

Table 3  Data from fits of core-swapped proteins

| ‘major parent’ | mean κ (from power-law fit) | mean σ (cP) | Δε (of major parent, p, compared to faster folding core-swap, cs) |
|----------------|----------------------------|--------------|---------------------------------------------------------------|
| R16o15c        | 0.49±0.04                  | 1.9±0.3      | 0.9 RT                                                        |
| R17015c        | 0.76±0.14                  | 0.73±0.47    | 1.7 RT                                                        |

\(a\) Determined from the fit of the log-log plots as shown in Fig. 1.

\(b\) Determined using the fits as those shown for parent domains in Fig. 2.

\(c\) Determined using the following equation:

\[
\Delta \varepsilon = RT \ln \frac{\sigma_p}{\sigma_{cs}}
\]

References

1. Ansari, A., Jones, C.M., Henry, E.R., Hofrichter, J. & Eaton, W.A. The role of solvent viscosity in the dynamics of protein conformational changes. *Science* **256**, 1796-1798 (1992).

2. Kramers, H.A. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica* **7**, 284-304 (1940).

3. Frauenfelder, H., Fenimore, P.W., Chen, G. & McMahon, B.H. Protein folding is slaved to solvent motions. *Proc. Natl Acad. Sci. USA* **103**, 15469-15472 (2006).
4. Bieri, O. et al. The speed limit for protein folding measured by triplet-triplet energy transfer. *Proc. Natl Acad. Sci. USA* **96**, 9597-9601 (1999).

5. Plaxco, K.W. & Baker, D. Limited internal friction in the rate-limiting step of a two-state protein folding reaction. *Proc. Natl Acad. Sci. USA* **95**, 13591-13596 (1998).

6. Pabit, S.A., Roder, H. & Hagen, S.J. Internal friction controls the speed of protein folding from a compact configuration. *Biochemistry* **43**, 12532-12538 (2004).

7. Jas, G.S., Eaton, W.A. & Hofrichter, J. Effect of viscosity on the kinetics of α-helix and β-hairpin formation. *J. Phys. Chem. B* **105**, 261-272 (2001).

8. Frauenfelder, H. et al. A unified model of protein dynamics. *Proc. Natl Acad. Sci. USA* **106**, 5129-5134 (2009).

9. Zwanzig, R. Diffusion in a rough potential. *Proc. Natl Acad. Sci. USA* **85**, 2029-2030 (1988).

10. Bryngleson, J.D. & Wolynes, P.G. Intermediates and barrier crossing in a random energy model (with applications to protein folding). *J. Phys. Chem.* **93**, 6902-6915 (1989).

11. Nettels, D., Gopich, I.V., Hoffmann, A. & Schuler, B. Ultrafast dynamics of protein collapse from single-molecule photon statistics. *Proc. Natl Acad. Sci. USA* **104**, 2655-2600 (2007).

12. Godoy-Ruiz, R. et al. Estimating free energy barrier heights for an ultrafast folding protein from calorimetric and kinetic data. *J. Phys. Chem. B* **112**, 5938-5949 (2008).

13. Yang, W.Y. & Gruebele, M. Folding at the speed limit. *Nature* **423**, 193-197 (2003).

14. Munoz, V., Thompson, P.A., Hofrichter, J. & Eaton, W.A. Folding dynamics and mechanism of beta-hairpin formation. *Nature* **390**, 196-199 (1997).
## Supplementary Tables

Supplementary Table 1. **Slope of relative rate ($k_0/k$) vs. relative viscosity ($\eta/\eta_0$) for R15, R16, R17, R16o15c and R17o15c.**

| Domain   | $\Delta G_{D-N}$ (kcal mol$^{-1}$) $^a$ | $k$ $^b$ | Slope $^c$ |
|----------|----------------------------------------|--------|---------|
| R15      | 1.5                                    | $k_f$  | 0.77 ($\pm$0.05) |
| R15      | 1.5                                    | $k_u$  | 0.83 ($\pm$0.04) |
| R15      | 0                                      | $k_f = k_u$ | 0.64 ($\pm$0.03) |
| Mean:    |                                        |        | 0.75 ($\pm$0.10) |
| R16      | 1.5                                    | $k_f$  | 0.17 ($\pm$0.04) |
| R16      | 1.5                                    | $k_u$  | 0.28 ($\pm$0.05) |
| R16      | 0                                      | $k_f = k_u$ | 0.16 ($\pm$0.03) |
| Mean:    |                                        |        | 0.20 ($\pm$0.07) |
| R17      | 1.5                                    | $k_f$  | 0.19 ($\pm$0.06) |
| R17      | 1.5                                    | $k_u$  | 0.14 ($\pm$0.04) |
| R17      | 0                                      | $k_f = k_u$ | 0.11 ($\pm$0.02) |
| Mean:    |                                        |        | 0.15 ($\pm$0.04) |
| R16o15c  | 1.5                                    | $k_f$  | 0.44 ($\pm$0.04) |
| R16o15c  | 1.5                                    | $k_u$  | 0.38 ($\pm$0.05) |
| R16o15c  | 0                                      | $k_f = k_u$ | 0.33 ($\pm$0.02) |
| Mean:    |                                        |        | 0.38 ($\pm$0.06) |
| R17o15c  | 1.5                                    | $k_f$  | 0.70 ($\pm$0.05) |
| R17o15c  | 1.5                                    | $k_u$  | 0.72 ($\pm$0.04) |
| R17o15c  | 0                                      | $k_f = k_u$ | 0.62 ($\pm$0.04) |
| Mean:    |                                        |        | 0.68 ($\pm$0.05) |

$^a$ $\Delta G_{D-N}$ at which rate constant used in the viscosity analysis calculated.

$^b$ The rate constant used in the analysis: folding, $k_f$, or unfolding, $k_u$, or $\Delta G_{D-N} = 0$ kcal mol$^{-1}$, so $k_f = k_u$.

$^c$ Slope of the plot of relative rate constants, $k_0/k$, vs. relative viscosity, $\eta/\eta_0$. The fit was forced through the reference point, 0 M Glucose, i.e. (1,1).
### Kinetic and thermodynamic parameters for R16015c WT and mutants in urea.

| Mutation    | $\Delta G_{D-N}^{\text{H}_2\text{O}}$ (kcal mol$^{-1}$) | $\Delta G_{D-N}$ (kcal mol$^{-1}$) | $k_f^{2\text{M}}$ | $\Phi_f^{2\text{M}}$ | $\Phi_f^{\text{cal}}$ |
|-------------|------------------------------------------------------|----------------------------------|-----------------|------------------|-----------------|
| Wild-type   | 6.70 (±0.03)                                         | -                                | 15.2 (±0.3)     | -                | -               |
| **Core mutations** |                                            |                                  |                 |                  |                 |
| V83A        | 6.37 (±0.03)                                         | 0.33 (±0.04)                     | 15.4 (±0.6)     | -                | -               |
| R87A        | 8.38 (±0.04)                                         | -1.67 (±0.05)                    | 18.2 (±0.2)     | -                | -               |
| I90A        | 5.31 (±0.03)                                         | 1.40 (±0.03)                     | 10.7 (±0.2)     | 0.14             | -               |
| F94A        | 3.53 (±0.02)                                         | 3.17 (±0.03)                     | 2.9 (±0.1)      | 0.29             | -               |
| I97A        | 4.57 (±0.02)                                         | 2.13 (±0.03)                     | 3.0 (±0.1)      | 0.43             | -               |
| A101G       | 4.09 (±0.02)                                         | 2.62 (±0.03)                     | 4.4 (±0.1)      | 0.26             | -               |
| R104A       | 5.80 (±0.03)                                         | 0.91 (±0.04)                     | 19.1 (±0.5)     | -0.14            | -               |
| L108A       | 4.37 (±0.02)                                         | 2.34 (±0.03)                     | 12.3 (±0.3)     | 0.05             | -               |
| **Ala-Gly mutations** |                                              |                                  |                 |                  |                 |
| Q85A        | 7.15 (±0.04)                                         | -                                | 18.8 (±0.5)     | -                | -               |
| Q85G        | 6.31 (±0.03)                                         | 0.84 (±0.05)                     | 13.4 (±0.4)     | 0.23             | -               |
| A88G        | 5.88 (±0.03)                                         | 0.82 (±0.05)                     | 10.6 (±0.3)     | 0.24             | -               |
| D92A        | 7.30 (±0.04)                                         | -                                | 22.2 (±0.5)     | -                | -               |
| D92G        | 6.18 (±0.03)                                         | 1.12 (±0.05)                     | 8.2 (±0.3)      | 0.50             | -               |
| K95A        | 6.80 (±0.03)                                         | -                                | 16.0 (±0.4)     | -                | -               |
| K95G        | 5.78 (±0.03)                                         | 1.01 (±0.05)                     | 5.9 (±0.1)      | 0.55             | -               |
| Q99A        | 6.82 (±0.03)                                         | -                                | 20.0 (±0.5)     | -                | -               |
| Q99G        | 5.83 (±0.03)                                         | 0.98 (±0.05)                     | 7.1 (±0.1)      | 0.59             | -               |
| A103G       | 5.47 (±0.03)                                         | 1.23 (±0.04)                     | 10.0 (±0.2)     | 0.19             | -               |
| Q106A       | 6.98 (±0.04)                                         | -                                | 12.3 (±0.2)     | -                | -               |
| Q106G       | 5.77 (±0.03)                                         | 1.21 (±0.05)                     | 10.0 (±0.2)     | 0.10             | -               |

---

*The value given is taken from equilibrium denaturation experiments. $\Delta G_{D-N}^{\text{H}_2\text{O}}$ is calculated from $\Delta G_{D-N}^{\text{H}_2\text{O}} = m_{D-N}[\text{urea}]_{50\%}$ using a mean $m_{D-N}$ value of 2.03 kcal mol$^{-1}$.  

*For Ala-Gly helix scanning positions values of $\Delta \Delta G_{D-N}$ and $\Phi_f^{2\text{M}}$ are shown against the appropriate alanine mutant. 

*$\Phi$–values are only calculated for mutations where $\Delta \Delta G_{D-N} \geq 0.75 \text{ kcal mol}^{-1}$. 

*Errors in the $\Phi$-values were propagated from errors of the fits of the chevron plots and the $\Delta \Delta G_{D-N}$ and are $\leq 0.03$. 

---
Supplementary Figure 1. R16 and R17 fold unusually slowly.

Long range order (LRO) plot. Grey circles are the data set of two-state proteins used in\textsuperscript{1}; the grey line represents a line of best fit to these data. There is a clear correlation between LRO\textsuperscript{2}, a measure of protein topology, and rate constants for folding ($k_f$). R15 (black circle) fits within the scatter, however R16 (red) and R17 (blue) are slow folding outliers. A standard contact order plot\textsuperscript{3} also shows that R16 and R17 are outliers (not shown).
Supplementary Figure 2. Glucose stabilises spectrin domains.

Equilibrium curves for (a) R15; (b) R16; (c) R17; (d) R16015c; (e) R17015c in 0.00 – 1.75 M glucose. Equilibrium curves were individually fitted as described in^4.
Supplementary Figure 3. Effects of glucose on kinetics of folding.

Chevron plots and fits for (a) R15, (b) R16, (c) R17, (d) R16o15c and (e) R17o15c in 0.00 – 1.75M glucose. See Methods for description of fitting methods.
Supplementary Figure 4. Increasing glucose results in a collapse of the denatured state. Solid lines represent a line of best fit to these data (a) R15 from individually fitted chevrons (note $m_{\text{T-N}}$ is the same within error), (b) R15 from globally fitted chevrons, (c) R16, (d) R17, (e) R16o15c, (f) R17o15c (see details of fitting procedures in Methods section). In general, the equilibrium and kinetic $m_{\text{D-N}}$-values agree (except for R17) and in all cases the $m_{\text{D-N}}$-values decrease with increasing glucose concentration. The unfolding $m$-values ($m_{\text{T-N}}$) are unaffected by glucose concentration, but the folding $m$-values ($m_{\text{D-N}}$) all decrease with increasing [glucose]. We infer that the decrease in $m_{\text{D-N}}$ results from a collapse in the denatured state and not from native- or transition state effects. NB (a-e) $m$-values determined in GdmCl and (f) determined in urea.
Supplementary Figure 5. Design of core-swapped proteins.

Sequence alignment of R15, R16, R17, R15o16c, R16o15c, R15o17c and R17o15c. Residues defined as core are shown in red, those in the 106 residue spectrin repeat in bold & the three helices indicated with h. See Methods for details of core-swap design.
Supplementary Figure 6. Supplementary kinetic data for core-swapped proteins. (a) and (b), relative rate constant ($k_0/k$) vs. relative viscosity ($\eta/\eta_0$) plots of (a) $k_u$ at $\Delta G = 1.5$ kcal/mol and (b) $k_f = k_u$ where $\Delta G = 0$ kcal/mol. Both core-swapped proteins have an increased viscosity dependence of the (un)folding rate constants on solvent viscosity, compared to their major parent.
Supplementary Figure 7. Chevron plots and fits for the $\Phi$-value analysis of R16o15c. (a) all core mutants, which report on tertiary structure formation. (b) and (c) Ala-Gly mutants which report on secondary structure formation. Chevrons were fitted globally with a shared $m_2$ of -0.95 M$^{-1}$ and a shared $m_2$ of 0.81 M$^{-1}$. (See Methods for details.)
Supplementary Figure 8. The core of R16o15c has similar packing to R15.

Comparison of ∆∆$G_{D-N}$ for core residue deletions in R16o15c and R15. They are very similar suggesting that the core packing is similar in the core-swapped protein. The filled circles are from greatest to smallest ∆∆$G_{D-N}$: F94A, A101G, L108A, I97A, I90A, R104A, V83A. The dotted line shows a linear fit to these data (R = 0.88). Solid line is the line $y = x$. The exception is R87A (open circle) which has a ∆∆$G_{D-N}$ of -1.68 kcal mol$^{-1}$ in R16o15c, compared to 0.58 kcal mol$^{-1}$ in R15. Arg87 in R15 makes a salt bridge with residue Asp70. In R16o15c residue 70 is Lys, providing an explanation for this unexpectedly low ∆∆$G_{D-N}$ in R16o15c.
Supplementary Figure 9. Helical propensity in the spectrin domains.

The helical propensity for each primary sequence calculated using AGADIR (at 25 °C and an ionic strength of 50 mM)\textsuperscript{5-8}.
References

1  Istomin, A. Y., Jacobs, D. J. & Livesay, D. R. On the role of structural class of a protein with two-state folding kinetics in determining correlations between its size, topology, and folding rate. *Protein Sci.* **16**, 2564-2569 (2007).

2  Gromiha, M. M. & Selvaraj, S. Comparison between long-range interactions and contact order in determining the folding rate of two-state proteins: application of long-range order to folding rate prediction. *J. Mol. Biol.* **310**, 27-32 (2001).

3  Plaxco, K. W., Simons, K. T. & Baker, D. Contact order, transition state placement and the refolding rates of single domain proteins. *J. Mol. Biol.* **277**, 985-994 (1998).

4  Pace, C. N. Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266-280 (1986).

5  Munoz, V. & Serrano, L. Development of the multiple sequence approximation within the AGADIR model of alpha-helix formation: comparison with Zimm-Bragg and Lifson-Roig formalisms. *Biopolymers* **41**, 495-509 (1997).

6  Munoz, V. & Serrano, L. Elucidating the folding problem of helical peptides using empirical parameters. II. Helix macrodipole effects and rational modification of the helical content of natural peptides. *J. Mol. Biol.* **245**, 275-296 (1995).

7  Munoz, V. & Serrano, L. Elucidating the folding problem of helical peptides using empirical parameters. III. Temperature and pH dependence. *J. Mol. Biol.* **245**, 297-308 (1995).

8  Lacroix, E., Viguera, A. R. & Serrano, L. Elucidating the folding problem of alpha-helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters *J. Mol. Biol.* **284**, 173-191 (1998).