Pressure-dependent chemical shifts in the R3 domain of talin show that it is thermodynamically poised for binding to either vinculin or RIAM

Nicola J. Baxter,[1] Thomas Zacharchenko,[2,3] Igor L. Barsukov,[2] and Mike P. Williamson[1,4,*]

[1]Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield S10 2TN, United Kingdom

[2]School of Biological Sciences, University of Liverpool, Crown Street, Liverpool L69 7ZB, United Kingdom

[3]Current address: School of Biology, University of Leeds, Leeds LS2 9JT, United Kingdom

[4]Lead contact

*Correspondence: m.williamson@sheffield.ac.uk (M.P.W.)

SUMMARY

Talin mediates attachment of the cell to the extracellular matrix. It is targeted by the Rap1 effector RIAM to focal adhesion sites, and subsequently undergoes force-induced conformational opening to recruit the actin-interacting protein vinculin. The conformational switch involves the talin R3 domain, which binds RIAM when closed and vinculin when open. Here, we apply pressure to R3 and measure ^{1}H, ^{15}N and ^{13}C chemical shift changes, which are fitted using a simple novel model, and indicate that R3 is only 50% closed: the closed form is a four-helix bundle, while in the open state helix 1 is twisted out. Strikingly, a mutant of R3 that binds RIAM with an affinity similar to wild-type but more weakly to vinculin is shown to be 0.84 kJ mol^{-1} more stable when closed. These results demonstrate that R3 is thermodynamically poised to bind either RIAM or vinculin, and thus constitutes a good mechanosensitive switch.

KEYWORDS hydrostatic pressure, cell adhesion, talin, focal adhesion complex, vinculin, singular value decomposition, chemical shift

INTRODUCTION

The adhesion of cells to the extracellular matrix usually makes use of the cell surface receptors known as integrins. The intracellular tail of an activated integrin forms part of a complex that links the receptors to the actin cytoskeleton, and thus enables communication between extracellular ligands and the cytoskeleton. The assembly of this complex is tightly regulated, and requires a number of proteins, including talin, vinculin and the Rap1-GTP-interacting adaptor molecule (RIAM). Talin is a large 2541-residue protein, consisting of a head that binds to integrins and to phospholipids, and a long rod-like tail that binds to F-actin and vinculin (Calderwood et al., 2013). Talin is recruited to the membrane by RIAM and subsequently undergoes a force-dependent conformational change that exposes cryptic vinculin binding sites. Vinculin recruitment by talin reinforces the connection to the actin cytoskeleton, constituting a force-sensing mechanism that regulates formation of the focal adhesion complex. The key element of the force sensor in the talin rod is the R3 domain, which binds both RIAM and vinculin (Atherton et al., 2015). The interaction of R3 with RIAM and vinculin is mutually exclusive, with RIAM binding to the closed form of R3, while vinculin interacts with the open form (Fillingham et al., 2005). This opening occurs as a result of mechanical stretching (del Rio et al., 2009; Yao et al., 2014), leading to the displacement of RIAM by vinculin (Goult et al., 2013). The R3 domain is located in a compact region of the talin rod formed by three sequential four-helix bundle domains R2-R4 that are particularly susceptible to mechanical unfolding (Yan et al., 2015). R3 has the lowest stability of the three and was observed to fold and unfold reversibly on a sub-second timescale at 5 pN applied force, within the range of a single actomyosin contraction (Yan et al., 2015; Yao et al., 2014; Yao et al., 2016). These observations have led to a model (Atherton et al., 2015; Calderwood et al., 2013; Yao et al., 2014) which proposes that RIAM initially binds to and recruits talin, but that a force-induced conformational change in talin (caused by the flow of actin filaments past the talin complex) opens up the R3 domain, weakening the interaction

with RIAM and permitting an interaction with vinculin. The newly formed vinculin complex then stabilizes the focal adhesions by strengthening interactions with actin (Figure 1).

The four-helix bundle that comprises the R3 domain of talin is held together by hydrophobic interactions. The hydrophobic interior includes a 'threonine belt' composed of four threonine residues, one from each helix in the bundle, that reduces the stability of the domain (Fillingham et al., 2005). We previously constructed a quadruple mutant of R3, referred to as R3-IVVI, in which these four threonine residues were replaced by hydrophobic residues (T809I/T833V/T867V/T901I) (Goult et al., 2013). This mutant is more resistant to unfolding by both mechanical force (Yao et al., 2014) and thermal unfolding (Goult et al., 2013), and binds much more weakly to vinculin than R3, consistent with the coupling of unfolding to productive vinculin binding (Figure 1).

There is thus ample evidence that the activity of R3 as a mechanosensitive switch is due to a reversible opening of the four-helix bundle. Structural characterization of the open state has so far remained challenging. Here, we use high pressure as a tool to allow us to characterize the energetics and structures of the two conformations.

High hydrostatic pressure has proved to be an effective way to reveal alternative protein conformations that are close in energy to the ground state (ie low-lying excited states) (Akasaka, 2003, 2006). Elevated pressure leads to a general compression of the protein, which to a good approximation is a linear effect. NMR chemical shift changes with pressure are therefore often linear (Kitahara et al., 2013). However, pressure also stabilizes alternative conformational states with lower partial molar volumes. Such states become increasingly populated as pressure increases, leading to non-linear chemical shift changes in the vicinity of the structural change. Because folded proteins always contain small cavities and packing defects, they always have a larger partial molar volume than unfolded proteins, and therefore pressure leads to local, and ultimately to global, unfolding (Roche et al., 2012). The low-energy excited state conformations revealed in this way can be functionally important: for example, a locally unfolded form of ubiquitin has been identified as the conformation seen when ubiquitin binds to the E2 ubiquitin-conjugating enzyme (Kitazawa et al., 2014). We therefore applied elevated

hydrostatic pressures up to 2.5 kbar (250 MPa) to the R3 domain of talin to characterize the equilibrium between closed and open forms, making the assumption that the excited state produced by high pressure is likely to be similar to the state produced by mechanical force.

Application of elevated pressure usually leads to linear or smoothly curved plots of chemical shift *vs* pressure. The analysis of pressure-dependent NMR chemical shifts therefore most commonly proceeds by fitting ^{15}N-HSQC chemical shift changes to a quadratic expression, $\delta = a + bp + cp^2$, where p is pressure (Akasaka, 2006; Kitahara et al., 2013). The linear term b is related to hydrogen bond strength and other local geometrical effects, and is often difficult to interpret. By contrast, the non-linear component c arises from the equilibrium between the native ground state and an excited state, and therefore tends to report on conformational changes in the transition to excited states, for example around cavities, and is much more useful and protein-specific (Kitazawa et al., 2014). However, although a quadratic expression is a convenient functional form, it cannot be related in any simple way to physical phenomena, and the terms b and c have no simple physical meaning. We therefore propose a more physically meaningful equation.

Pressure causes a change in the free energy difference between two conformations. We therefore expect the states to be populated according to a pressure-dependent Boltzmann expression:

$$P_2/P_1 = exp\left(-(\Delta G + p\Delta V)/_{RT}\right) \quad [1]$$

We assume that R3 can exist in two states, *1* and *2* (ground state and excited state, respectively), populated in a pressure-dependent way as above. Each of these states can also undergo a linear pressure-dependent compression, giving rise to a linear change in chemical shift. We can therefore model the observed pressure-dependent chemical shift δ in a more physically meaningful way as:

$$\delta = \frac{(\delta_1^0 + p\Delta\delta_1) + (\delta_2^0 + p\Delta\delta_2)exp\left(-[\Delta G + p\Delta V]/_{RT}\right)}{1 + exp\left(-[\Delta G + p\Delta V]/_{RT}\right)} \quad [2]$$

where δ_1^0 and δ_2^0 are the chemical shifts of forms *1* and *2* at ambient pressure; $\Delta\delta_1$ and $\Delta\delta_2$ are the linear pressure-dependent changes in chemical shift; p is the pressure; ΔG is the difference in free energy

between the two states at ambient pressure; and ΔV is the change in volume between the two states. This equation makes some assumptions, which are discussed in Supplementary Material.

This more complex analysis of pressure-dependent chemical shifts is more difficult to fit than the standard quadratic approximation, because it has one more variable. We therefore used singular value decomposition (SVD) to help analyze the chemical shift changes. SVD is a well-established statistical technique for reducing the dimensionality of fitting problems by identifying the minimum number of components needed to generate the experimental data patterns. In the process, it can also be used to identify and remove non-correlated noise, thus dramatically reducing experimental random noise in the data. This analysis showed that fitting of the pressure-dependent chemical shifts only requires four components, despite the complexity of the chemical shift changes. The four components are identified as those required by Eq. 2, namely the ground state at ambient pressure, a compressed ground state, an excited state (whose population increases with pressure) and a compressed excited state.

The high quality of the 'noise-free' data allowed us to fit Eq. 2 globally, producing the result that in the wild-type R3 the free energy difference between the ground and excited states is very close to zero, ie R3 is 50% closed and 50% open at ambient pressure. The nature of the conformational change can be identified by analyzing the chemical shifts, δ_1^o and δ_2^o, and the difference in the pressure-dependent gradients, $\Delta\delta_1 - \Delta\delta_2$, on a per-residue basis. The main change is localized to helix 1, which is part of the four-helix bundle in the ground state (closed conformation), but which becomes twisted out in the excited state (open conformation), thereby explaining the effect of shear force on modulating the availability of binding surfaces. Finally, a similar analysis of R3-IVVI showed that in this mutant the ground state is 0.84 kJ mol^{-1} more stable than the excited state, implying that the mutant accesses the open conformation less readily than R3, and explaining its lower affinity for vinculin. These results show that R3 is delicately poised between open and closed states, and is thus well placed to act as a mechanosensitive switch, able to exchange easily between binding RIAM in the closed state or vinculin in the open state in response to appropriate stimuli.

RESULTS

Backbone assignments and pressure-dependent chemical shift changes

The NMR spectra of R3 and R3-IVVI are similar (Figure S1). The signals of R3-IVVI were assigned using standard three-dimensional (3D) heteronuclear experiments. Backbone chemical shifts of R3-IVVI are very similar to those of wild-type R3 (Figure S2) except in the immediate locations of the mutations, strongly suggesting that the structure of R3-IVVI is very similar to that of R3 (Figure S3). 2D ^{15}N-HSQC spectra of both R3 and R3-IVVI at variable pressures up to 2.5 kbar show extensive chemical shift changes due to pressure (Figure 2). Shift changes of this magnitude are common (Kitahara et al., 2013), and as is normally observed, most N and HN nuclei move to higher resonance frequencies at elevated pressure. For most amide groups, compression results in an increased polarization of the H-N bond, causing the chemical shifts of HN and N to move to higher frequencies as the amide proton becomes more deshielded and the amide nitrogen becomes more shielded. In most proteins, the majority of residues show approximately linear pressure-dependent chemical shift changes, whilst a few residues have curved quadratic shapes. However here, some resonances have very unusual curved pressure titration data (Figure 2 inset), indicating an unusually strong pressure dependence, particularly in R3. The curvature strongly suggests that an alternative conformation is present with a high population. Since it is widely agreed that R3 is in equilibrium between closed and open states, we identify this alternative conformation as the open state. The R3 talin domain thus provides an interesting test case for analysis of pressure-dependent data. Chemical shifts at variable pressure were also measured for Cα, Cβ and C' nuclei, using 2D HN(CO)CACB and 2D HNCO spectra. Most nuclei were in fast exchange at all pressures, although a few Cα and Cβ nuclei broadened at higher pressures. Large-scale conformational exchange typically slows at high pressures (Williamson, 2015), implying that this broadening is due to pressure-dependent conformational exchange. This result implies a conformational exchange rate of around 1-2000 s^{-1} at ambient pressure, consistent with a large-scale hinge bending.

As a first step, the chemical shift *vs* pressure data were fitted to a quadratic expression, following common practice (Kitahara et al., 2013). Neither R3 nor R3-IVVI gave a good fit to the data, with a large

number of both amide proton and nitrogen chemical shifts giving a poor fit (Figure S4). Mapping these residues onto the structure gave no perceptible pattern (Figure S5). We also tried fitting the data to a cubic equation, this being the most obvious progression from a quadratic, as it includes the next term of the Taylor expansion. The fit was much better, but there remained a significant number of peaks systematically not fitting well (Figure S4), that again showed no obvious pattern of distribution when mapped onto the structure of the protein.

The pressure-dependent chemical shifts were processed using SVD in order to reduce the amount of experimental noise in the peak positions. The resulting singular values are plotted in Figure 3A for the combined fitting of the backbone amide NH and N signals of R3. Other data are similar. It is clear that the first three singular values contribute to the spread of data, but it is not obvious whether the fourth is also needed. In other words, at least three chemical shift components are needed for adequate global fitting of the data. Inspection of the actual data shows that most residues are indeed fitted well by only three components (in other words, they fit well to a quadratic equation), but that a small group are not (eg F813 in Figure S4). One should also be able to tell how many components are necessary by looking at the v_i vectors (Arai et al., 2012; Henry and Hofrichter, 1992): essential components should have vectors with smooth shape and high autocorrelation. The first five vectors are shown in Figure 3B, again indicating that the first three components are clearly needed, but not defining clearly whether the fourth is also necessary. The data suggest that no more than four are needed.

We have already seen (Figures S4 and S5) that many resonances do not fit well to a quadratic expression, which has three variables, but most fit reasonably well to a cubic equation with four variables. Our proposed chemical shift equation (Eq. 2) has four components. We therefore set all singular values with rank higher than 4 to zero and re-calculated a high-quality 'noise-free' data set D'.

Fitting of R3 chemical shift changes to Eq. 2

The column vectors of U' represent the chemical shifts of each species present, but not in a way that normally allows the shifts to be extracted (Ikeda et al., 2011). It is thus common practice to take the data

177 from the SVD analysis, and go on to fit this 'noise-free' data to physically reasonable models in a
178 conventional way (Henry and Hofrichter, 1992). Eq. 2 contains two global variables (ΔG, the difference
179 in free energy between the ground state and the excited state; and ΔV, the difference in volume between
180 these two states) together with four resonance-specific variables (δ_1^0, $\Delta\delta_1$, δ_2^0, and $\Delta\delta_2$, discussed in
181 detail below). We first obtained fits for the global variables ΔG and ΔV, using the nuclei most sensitive to
182 pressure-dependent population changes, ie the HN, N and C' nuclei that gave the largest χ^2 values when
183 a quadratic expression was fitted to the 'noise-free' chemical shift vs pressure data (Figure S5).
184 Simultaneous fitting to Eq. 2 for these nuclei resulted in a robust global fit for ΔG and ΔV of -1.6 $\pm$ 27 J
185 mol^{-1} and -2.62 $\pm$ 0.53 kJ mol^{-1} $kbar^{-1}$ respectively, equivalent to a volume difference of 26.2 ml mol^{-1}. The
186 volume change falls into the middle of the typical range (Kitahara et al., 2013), but the free energy
187 difference is very small, and implies that the ground state (closed conformation) and the excited state
188 (open conformation) have essentially equal populations at 1 bar. This is an unusual and very significant
189 result, because normally the high-pressure excited state is populated to only a few percent at ambient
190 pressure. This is the first time that an open conformation has been identified for R3, and is the first
191 measurement of its population. The large population of the excited state provides a good explanation for
192 the unusually curved pressure-dependent chemical shifts.

193 ΔG and ΔV were then fixed, and the data for all nuclei were fitted to obtain the four resonance-
194 specific variables (δ_1^0, $\Delta\delta_1$, δ_2^0, and $\Delta\delta_2$, discussed in detail below). This is effectively the same number
195 of variables as used for fitting against a cubic equation. Whereas fitting to a cubic expression gives
196 patterns of residuals that clearly indicate systematic errors (Figure S4), fitting to Eq. 2 gives almost no
197 residual errors. The largest individual deviation was 0.070 ppm, and the overall root-mean-square
198 difference between experimental and calculated data was 0.0026 $\pm$ 0.0045 ppm. It is remarkable that
199 such unusual chemical shift patterns can be fitted so well by this simple equation (Figure 4). The fact that
200 the experimental shifts can be fitted so well by Eq. 2 is not conclusive proof that R3 must be following
201 the Boltzmann distribution described by Eq. 2, but does imply that Eq. 2 is a good model for the system.

202 Because Eq. 2 describes the simplest possible model comprising the least number of fitted variables that
203 is compatible with the data, we did not test other more complex models.
204
205 **Structure of R3 open conformation from fitted chemical shift values**
206 For each nucleus, fitting to Eq. 2 returns four variables: δ_1^0, the chemical shift of the ground state at
207 ambient pressure; $\Delta\delta_1$, the linear pressure-dependent chemical shift change of the ground state; δ_2^0, the
208 chemical shift of the excited state at ambient pressure; and $\Delta\delta_2$, the linear pressure-dependent chemical
209 shift change of the excited state. This represents a major increase in information content compared to
210 the standard fitting procedure using a quadratic expression, because the δ_1^0 and δ_2^0 fitted chemical shift
211 values provide detailed structural information about the closed and open forms. Significantly, these are
212 the chemical shifts of each form at ambient pressure, and can therefore be interpreted without the need
213 to take into account the effects of pressure on chemical shifts (Wilton et al., 2009). Chemical shifts are
214 powerful structural constraints, as indicated by the success of programs such as CS-ROSETTA (Shen et
215 al., 2008). The δ_1^0 and δ_2^0 chemical shift values for HN, N, Cα, Cβ, and C' nuclei were used as input to
216 TALOS-N (Shen and Bax, 2013), which shows that the closed structure is consistent with the NMR
217 structure (2L7A) (Goult et al., 2013), and comprises a four-helix bundle with random coil termini (Figure
218 5). The open structure has all four α-helices remaining intact with only minor conformational differences
219 within the α-helices compared to the closed structure. There are some differences in the loops between
220 the α-helices, particularly between helices 1 and 2. However, the TALOS-N predictions are least
221 confident in the loops, and it is therefore not possible to calculate a detailed structure for the open
222 conformation using TALOS-N alone.
223 We therefore consider what information can be obtained from the differences in chemical shift
224 between the closed and open forms ($\Delta\delta = \delta_1^0 - \delta_2^0$). Of these, differences for Cα and Cβ are the simplest
225 to interpret because they depend mainly on backbone dihedral angles, and to some extent on sidechain
226 dihedral angles and hydrogen bonding to the backbone (Iwadate et al., 1999), whereas HN, N and C' are

227 also strongly affected by hydrogen bonding to the amide group as a whole. The largest differences in Cα
228 and Cβ chemical shifts (Figure 6) are mainly found for residues in helix 1, together with those at the N-
229 terminal end of helix 2 and in the turn between them, whilst smaller differences are noted for residues in
230 helices 3 and 4. This distribution of chemical shift differences suggests that in the open form, helix 1 has
231 detached from the rest of the bundle, where its main contact in the closed form is with the N-terminal end
232 of helix 2 (Figure 7). The chemical shifts of the open form do not allow us to calculate the angle between
233 helices 1 and 2. HN and N shifts (Figure S6A, B) support this conclusion, while the C' shifts are similar
234 but less useful (Figure S6C). Chemical shift differences for the random coil region at the N-terminus are
235 very small in all cases, showing that there is no change to the random coil structure, which provides a
236 useful internal control for the quality of the chemical shift analysis.

237 We have also analyzed the differences in pressure-dependent gradients between the closed and
238 open forms (Δgradient = $\Delta\delta_1 - \Delta\delta_2$) (Figures 8 and S7). The gradient reports on how the chemical shift of
239 the nucleus alters with increasing pressure, and is a measure of local compressibility (Kitahara et al.,
240 2013). Δgradient values can therefore highlight regions where the compressibility has changed between
241 the closed and open conformations. Compressibility is determined to a large extent by surface exposure,
242 with large negative Δgradient values indicating regions where the open structure (state 2) is more
243 compressible (which in general means more solvent-exposed). Examination of Δgradient values for Cα
244 and Cβ (Fig S6 C and D) shows that helices 1 and 2 have the largest changes, becoming generally more
245 exposed in the open state.

246 The chemical shifts of amide protons are strongly affected by hydrogen bonding (Williamson,
247 2013). The Δgradient values therefore provide insights into hydrogen bonding in the open form. To
248 interpret the Δgradient values, it is helpful to consider temperature-dependent shift gradients of amide
249 protons. Although in general non-hydrogen bonded HN have more negative gradients than hydrogen
250 bonded HN (Baxter and Williamson, 1997), strongly hydrogen bonded HN have a more negative gradient
251 than weakly hydrogen bonded HN, because these protons are highly deshielded by the hydrogen bond
252 and are therefore most affected by small changes in bond length (Tomlinson and Williamson, 2012). In

253 the same way, amide protons in strong hydrogen bonds have larger pressure-dependent shift gradients
254 than those in weak hydrogen bonds. Consequently, positive HN Δgradient values indicate amide protons
255 that have short hydrogen bonds in the closed form, while negative Δgradient values indicate amide
256 protons with shorter hydrogen bonds in the open form. The Δgradient values are shown in Figure 7A.
257 Throughout the protein, but particularly in helices 1 and 2, the large negative Δgradients tend to be on
258 the outer faces of helices while large positive Δgradients tend to be on the inner faces. The closed NMR
259 structure is composed of four noticeably curved α-helices, and the data therefore indicate that following
260 the twisting out of helix 1 from the bundle, this helix becomes straighter as a consequence of being fully
261 solvated by bulk solvent in the open state. The loss of helix 1 from the bundle then allows the remaining
262 three helices to become somewhat straighter as the protein is solvated differently.
263
264 **Fitting of R3-IVVI chemical shift changes to Eq. 2**
265 Data for R3-IVVI were fitted in a similar way to that described for R3, although in this case using only HN
266 and N shifts, because the protein was ^{15}N labelled. As for R3, fitting the data to a quadratic equation
267 gives a poor fit (Figure S8). However, the data again fitted well to the four-component model Eq. 2
268 (maximum deviation = 0.018 ppm, overall root-mean-square difference = 0.0011 ± 0.0017 ppm), but with
269 different global parameters: ΔG = 0.84 kJ mol^{-1} and ΔV = –3.98 kJ mol^{-1} kbar^{-1} (39.8 ml mol^{-1}). Thus, the
270 free energy difference is larger for R3-IVVI, although still small (for example, compare 4.2 kJ mol^{-1} for
271 ubiquitin (Kitahara et al., 2005) and 11 kJ mol^{-1} for hamster prion (Kuwata et al., 2002)), and implies a
272 population of the excited state of 41% at ambient pressure. The higher energy of the excited state relative
273 to the ground state explains why the chemical shift changes are more typical and less dramatically curved
274 when compared to R3, while the smaller population explains why R3-IVVI binds more weakly to vinculin.
275 The volume change for R3-IVVI is larger than for R3, presumably as a consequence of a larger partial
276 molar volume (a larger volume of packing defects) in the closed state of the mutant R3-IVVI.
277 A detailed analysis of pressure-dependent chemical shift changes for R3-IVVI using the SVD
278 procedure gave similar results to those obtained for R3 (Figure S9). The chemical shift differences ($\Delta\delta$ =

279 $\delta_1^o - \delta_2^o$) between ground and excited states followed a similar pattern to the R3 data, with the largest
280 changes found in helices 1 and 2 (Figure S10). The differences in pressure-dependent gradients
281 (Δgradient = $\Delta\delta_1 - \Delta\delta_2$) (Figure S11) also show a similar profile (Figure 8), implying an analogous
282 straightening out of helices in the open state of R3-IVVI. We therefore conclude that the primary
283 mechanism resulting in the fast conformational exchange observed at all pressures in solution between
284 the closed and open states for both R3 and R3-IVVI is the twisting out of helix 1 from the four-helix bundle
285 domain, followed by solvation of this helix together with a subtle change in packing of the remaining
286 three-helix bundle.

287

288 **DISCUSSION**

289 A number of publications have demonstrated the use of high pressure NMR spectroscopy to reveal
290 details of higher energy conformers that are in fast exchange with the ground state (Akasaka, 2006;
291 Kalbitzer et al., 2013; Kitahara et al., 2013). Such data are impossible to characterize by NMR
292 approaches under normal conditions since signals are averaged in this exchange regime. In this work,
293 we have introduced two novel methods that should prove generally useful to such analyses. The first is
294 to make use of singular value decomposition (SVD). Although SVD has been applied to NMR data before
295 (Arai et al., 2012; Jaumot et al., 2004; Matsuura et al., 2004; Sakurai and Goto, 2007), its use has been
296 limited. Two important applications of SVD are to determine the number of components required to fit the
297 data, and then to back-calculate a dataset with greatly reduced noise by zeroing everything apart from
298 the required components, thus providing a much more secure basis for fitting. SVD analysis can be
299 carried out using the widely available Matlab™, which makes it straightforward to do (Supplementary
300 Information).

301 Second, we fitted the data to a novel equation (Eq. 2), which models the observed pressure-
302 dependent chemical shift using two conformational states whose populations depend on pressure, and
303 for which both states can also undergo a linear pressure-dependent compression. The data fitted
304 remarkably well to this equation, and the fitting yielded parameters that reveal much more about the

underlying structural changes than when fitting to a quadratic expression. In particular, we showed that we can accurately quantify the relative free energies of the two states and their populations. Despite the observed chemical shifts being a population-weighted average of the two states, so that neither the pure closed nor open state can be observed directly, the derived chemical shift values for the closed (δ_1^o) and open (δ_2^o) forms can be used straightforwardly to determine the structures of these two states and chemical shift differences ($\delta_1^o - \delta_2^o$) highlight where structural differences are located. In addition, the gradient differences ($\Delta\delta_1 - \Delta\delta_2$) indicate the regions where solvent exposure changes on pressure perturbation. This is likely to be a general result. An alternative approach was described recently (Erlach et al., 2014), which provides estimates of the ratio of the difference in compressibility factors and partial molar volumes. For more linear pressure-dependent shifts, that approach may prove more tractable.

It is important to emphasize that the high hydrostatic pressure used here is a tool to allow us to characterize the closed and open states. It is not intended as a mimic of the force applied *in vivo*, and hence is a general technique for characterizing any conformational change in proteins. We have described two conformations of R3, that are populated almost equally under our conditions, and which exchange rapidly on the NMR timescale. The application of force also induces a change between two conformations, with a rapid transition. We propose that the two conformations that are seen by pressure and the two conformations seen with force are the same. This proposal is strongly supported by the observation that the conformational change from closed to open is the same under both conditions, as discussed below. Pressure induces opening only of helix 1 (likely followed by the other three), whereas force induces opening of all four helices. Pressure therefore reproduces only the first stage in the unfolding induced by force, and thereby suggests the likely pathway for the force-induced unfolding.

This analysis was carried out on wild type and mutant forms of the R3 domain of talin. We showed that wild type R3 is in equilibrium between two forms, each of which is populated 50% at ambient pressure. State 1 (ground state) has higher partial molar volume and is identified as the fully folded form of the four-helix bundle domain, as characterized previously by NMR (Goult et al., 2013). State 2 (excited state) has a smaller volume and is a locally unfolded state. The difference in volume is 26 ml mol^{-1}, which

331 is a typical value for local unfolding but is small for complete unfolding of the entire domain (Royer, 2002).
332 TALOS-N shows that all four helices are still present. The nature of the structural change is shown most
333 clearly by analysis of the difference in Cα and Cβ chemical shifts ($\delta_1^o - \delta_2^o$) between the two states, which
334 indicates that helix 1 is twisted out from the rest of the protein. Pressure-dependent exchange broadening
335 implies a conformational exchange rate of around 1-2000 s^{-1}, consistent with a hinge motion of this type.
336 Analysis of gradient differences ($\Delta\delta_1 - \Delta\delta_2$) indicates that all helices, but particularly helices 1 and 2,
337 become more linear (less curved) in the open state. R3-IVVI has similar structural changes, but has a
338 larger difference in free energy between closed and open states.

339 Previous studies have shown that RIAM binds to an exposed surface composed from helices 2
340 and 3 of R3 (Figure 9A), requiring very little conformational change in R3 (Goult et al., 2013). In contrast,
341 vinculin binds to the same two helices, but mainly to residues on the opposite face, which are buried in
342 the intact domain (Figure 9B) (Goult et al., 2013). Binding of RIAM and vinculin to R3 is mutually exclusive
343 (Goult et al., 2013). In agreement with this, it has been shown that binding vinculin requires the unfolding
344 of R3 (Goult et al., 2013; Roberts and Critchley, 2009). Indeed, a crystal structure of the vinculin head
345 bound to helix 3 of R3 shows that the single talin helix fits into a groove on the surface of the vinculin
346 head, in a conformation that would require complete separation of helix 3 from the rest of the domain
347 (Fillingham et al., 2005). Moreover, there is evidence that helices 2 and 3 of R3 compose independent
348 binding sites for vinculin, shown by analysis of vinculin binding assays to talin synthetic peptides (Gingras
349 et al., 2005), and by gel filtration showing formation of a 1:2 complex (Goult et al., 2013). The open
350 structure characterized here has the vinculin site on helix 2 exposed, but not well enough to fit neatly into
351 the vinculin groove. Finally, we note that pulling on an R1-R3 construct using magnetic tweezers was
352 interpreted to show unfolding of R3 at a force of about 5 pN (similar to the force expected from a single
353 actomyosin contraction), characterized by an extension of about 18 nm (Yan et al., 2015; Yao et al.,
354 2014). This distance corresponds approximately to the length of R3 when the four helices are opened
355 out completely, whilst retaining their helicity. In summary, a range of experiments suggest that R3 unfolds
356 into the four individual helices when interacting fully with vinculin, yet the results presented here show

only the opening out of helix 1, which serves to expose helix 2 yet not render it fully able to bind. We therefore propose that the unfolding of R3 and its interactions with vinculin proceed in a stepwise manner (Figure 9C): RIAM binding to folded R3 is reversible, force pulls helix 1 out from the bundle, which exposes the binding site on helix 2 for vinculin, and thus allows helix 2 to interact with vinculin. Binding of vinculin further opens out the domain, and exposes helix 3. Helix 3 is likely to be a stronger binding site for vinculin than helix 2, as it is one of the three vinculin binding sites on talin identified in the original screen (Bass et al., 1999). Binding of vinculin to helix 3 completely opens out the domain. This stepwise mode of action allows for graduated conversion from RIAM-bound to fully vinculin-bound R3, and provides scope for modulation of the interactions as required by the biological context.

This stepwise model agrees well with the energetics determined in this study. The previous observations using magnetic tweezers (Yao et al., 2014) have shown that folded and unfolded forms of R1-R3 are approximately equally populated at a stretching force of 5 pN, whereas the results here suggest that the open and closed forms of R3 are equally populated in the absence of any stretching force. Moreover, it was shown earlier (Goult et al., 2013) that R3-IVVI binds very weakly to vinculin, suggesting a lower population of the open state than observed here. Both these observations are reconciled by noting that our conclusions relate to the initial opening out of helix 1, and not to complete opening of the whole domain, which requires a greater input of free energy.

Our data therefore indicate that in the isolated R3 domain (as studied here), the equilibrium between open and closed conformations is delicately poised to allow environmental conditions the maximum opportunity to alter the equilibrium position so as to stimulate closure (and therefore binding to RIAM) or opening (and therefore binding to vinculin). In particular, it is expected that mechanical force on the domain leads to the stepwise opening out of the domain with subsequent binding to vinculin (Figure 9) (Calderwood et al., 2013). This behavior would allow the R3 domain to act as a mechanosensitive switch, altering talin from an initial mode of recruitment by RIAM to the integrin tails, to a functional role of recruiting vinculin to the focal adhesion complexes.

Author Contributions

NJB, TZ and MPW did the experiments; NJB, ILB and MPW wrote the manuscript. All authors have given approval to the final version of the manuscript.

Acknowledgements

We thank Per Lincoln (Chalmers University, Sweden) for advice on SVD. We thank BBSRC for funding (BB/J014966/1 to NJB).

References

Akasaka, K. (2003). Exploring the entire conformational space of proteins by high-pressure NMR. Pure Applied Chem *75*, 927-936.

Akasaka, K. (2006). Probing conformational fluctuation of proteins by pressure perturbation. Chem Rev *106*, 1814-1835.

Arai, M., Ferreon, J.C., and Wright, P.E. (2012). Quantitative analysis of multisite protein-ligand interactions by NMR: Binding of intrinsically disordered p53 transactivation subdomains with the TAZ2 domain of CBP. J Am Chem Soc *134*, 3792-3803.

Atherton, P., Stutchbury, B., Wang, D.-Y., Jethwa, D., Tsang, R., Meiler-Rodriguez, E., Wang, P., Bate, N., Zent, R., Barsukov, I.L.*, et al.* (2015). Vinculin controls talin engagement with the actomyosin machinery. Nature Comms *6*, 10038-10038.

Bass, M.D., Smith, B.J., Prigent, S.A., and Critchley, D.R. (1999). Talin contains three similar vinculin-binding sites predicted to form an amphipathic helix. Biochem J *341*, 257-263.

Baxter, N.J., and Williamson, M.P. (1997). Temperature dependence of ^{1}H chemical shifts in proteins. J Biomol NMR *9*, 359-369.

Calderwood, D.A., Campbell, I.D., and Critchley, D.R. (2013). Talins and kindlins: partners in integrin-mediated adhesion. Nature Rev Mol Cell Biol *14*, 503-517.

del Rio, A., Perez-Jimenez, R., Liu, R., Roca-Cusachs, P., Fernandez, J.M., and Sheetz, M.P. (2009). Stretching single talin rod molecules activates vinculin binding. Science *323*, 638-641.

Erlach, M., Koehler, J., Moeser, B., Horinek, D., Kremer, W., and Kalbitzer, H.R. (2014). Relationship between nonlinear pressure-induced chemical shift changes and thermodynamic parameters. J Phys Chem B *118*, 5681-5690.

Fillingham, I., Gingras, A.R., Papagrigoriou, E., Patel, B., Emsley, J., Critchley, D.R., Roberts, G.C.K., and Barsukov, I.L. (2005). A vinculin binding domain from the talin rod unfolds to form a complex with the vinculin head. Structure *13*, 65-74.

Gingras, A.R., Ziegler, W.H., Frank, R., Barsukov, I.L., Roberts, G.C.K., Critchley, D.R., and Emsley, J. (2005). Mapping and consensus sequence identification for multiple vinculin binding sites within the talin rod. J Biol Chem *280*, 37217-37224.

Goult, B.T., Bate, N., Anthis, N.J., Wegener, K.L., Gingras, A.R., Patel, B., Barsukov, I.L., Campbell, I.D., Roberts, G.C.K., and Critchley, D.R. (2009). The structure of an interdomain complex that regulates talin activity. J Biol Chem *284*, 15097-15106.

Goult, B.T., Zacharchenko, T., Bate, N., Tsang, R., Hey, F., Gingras, A.R., Elliott, P.R., Roberts, G.C.K., Ballestrem, C., Critchley, D.R.*, et al.* (2013). RIAM and vinculin binding to talin are mutually exclusive and regulate adhesion assembly and turnover. J Biol Chem *288*, 8238-8249.

425 Henry, E.R., and Hofrichter, J. (1992). Singular value decomposition: Applications to analysis of
426 experimental data. Methods Enzymol *210*, 129-192.

427 Ikeda, K., Yamaguchi, T., Fukunaga, S., Hoshino, M., and Matsuzaki, K. (2011). Mechanism of amyloid
428 β-protein aggregation mediated by GM1 ganglioside clusters. Biochemistry *50*, 6433-6440.

429 Iwadate, M., Asakura, T., and Williamson, M.P. (1999). Cα and Cβ carbon-13 chemical shifts in
430 proteins from an empirical database. J Biomol NMR *13*, 199-211.

431 Jaumot, J., Marchan, V., Gargallo, R., Grandas, A., and Tauler, R. (2004). Multivariate curve resolution
432 applied to the analysis and resolution of two-dimensional ^{1}H, ^{15}N NMR reaction spectra. Anal Chem *76*,
433 7094-7101.

434 Kalbitzer, H.R., Rosnizeck, I.C., Munte, C.E., Narayanan, S.P., Kropf, V., and Spoerner, M. (2013).
435 Intrinsic allosteric inhibition of signaling proteins by targeting rare interaction states detected by high-
436 pressure NMR spectroscopy. Angewandte Chemie Int Ed *52*, 14242-14246.

437 Kitahara, R., Hata, K., Li, H., Williamson, M.P., and Akasaka, K. (2013). Pressure-induced chemical
438 shifts as probes for conformational fluctuations in proteins. Progr Nucl Magn Reson Spectrosc *71*, 35-
439 58.

440 Kitahara, R., Yokoyama, S., and Akasaka, K. (2005). NMR snapshots of a fluctuating protein structure:
441 Ubiquitin at 30 bar-3 kbar. J Mol Biol *347*, 277-285.

442 Kitazawa, S., Kameda, T., Kumo, A., Yagi-Utsumi, M., Baxter, N.J., Kato, K., Williamson, M.P., and
443 Kitahara, R. (2014). Close identity between alternatively folded state N_2 of ubiquitin and the
444 conformation of the protein bound to the ubiquitin-activating enzyme. Biochemistry *53*, 447-449.

445 Klapholz, B., Herbert, S.L., Wellmann, J., Johnson, R., Parsons, M., and Brown, N.H. (2015).
446 Alternative mechanisms for talin to mediate integrin function. Current Biol *25*, 847-857.

447 Kuwata, K., Li, H., Yamada, H., Legname, G., Prusiner, S.B., Akasaka, K., and James, T.L. (2002).
448 Locally disordered conformer of the hamster prion protein: A crucial intermediate to PrPSc?
449 Biochemistry *41*, 12277-12283.

450 Matsuura, H., Shimotakahara, S., Sakuma, C., Tashiro, M., Shindo, H., Mochizuki, K., Yamagishi, A.,
451 Kojima, M., and Takahashi, K. (2004). Thermal unfolding of ribonuclease T1 studied by
452 multidimensional NMR spectroscopy. Biol Chem *385*, 1157-1164.

453 Noble, B., and Daniel, J.W. (1988). Applied linear algebra, 3rd edn (London: Prentice-Hall International).
454 Roberts, G.C.K., and Critchley, D.R. (2009). Structural and biophysical properties of the integrin-
455 associated cytoskeletal protein talin. Biophys Rev *1*, 61-69.

456 Peterson, R.W., Nucci, N.V., and Wand, A.J. (2011). Modification of encapsulation pressure of reverse
457 micelles in liquid ethane. J Magn Reson *212*, 229-233.

458 Roche, J., Caro, J.A., Norberto, D.R., Barthe, P., Roumestand, C., Schlessman, J.L., Garcia, A.E.,
459 Garcia-Moreno E, B., and Royer, C.A. (2012). Cavities determine the pressure unfolding of proteins.
460 Proc Natl Acad Sci USA *109*, 6945-6950.

461 Royer, C.A. (2002). Revisiting volume changes in pressure-induced protein unfolding. Biochim Biophys
462 Acta: Protein Structure Mol Enzymol *1595*, 201-209.

463 Sakurai, K., and Goto, Y. (2007). Principal component analysis of the pH-dependent conformational
464 transitions of bovine beta-lactoglobulin monitored by heteronuclear NMR. Proc Natl Acad Sci USA *104*,
465 15346-15351.

466 Shen, Y., and Bax, A. (2013). Protein backbone and sidechain torsion angles predicted from NMR
467 chemical shifts using artificial neural networks. J Biomol NMR *56*, 227-241.

Shen, Y., Lange, O., Delaglio, F., Rossi, P., Aramini, J.M., Liu, G.H., Eletsky, A., Wu, Y.B., Singarapu, K.K., Lemak, A.*, et al.* (2008). Consistent blind protein structure generation from NMR chemical shift data. Proc Natl Acad Sci USA *105*, 4685-4690.

Tomlinson, J.H., and Williamson, M.P. (2012). Amide temperature coefficients in the protein G B1 domain. J Biomol NMR *52*, 57-64.

Williamson, M.P. (2013). Using chemical shift perturbation to characterise ligand binding. Progr Nuclear Magn Reson Spectrosc *73*, 1-16.

Williamson, M.P. (2015). Pressure-dependent conformation and fluctuation in folded protein molecules. In High pressure bioscience, K. Akasaka, and H. Matsuki, eds. (Dordrecht: Springer), pp. 109-127.

Wilton, D.J., Kitahara, R., Akasaka, K., and Williamson, M.P. (2009). Pressure-dependent ^{13}C chemical shifts in proteins: origins and applications. J Biomol NMR *44*, 25-33.

Yan, J., Yao, M., Goult, B.T., and Sheetz, M.P. (2015). Talin dependent mechanosensitivity of cell focal adhesions. Cell Mol Bioeng *8*, 151-159.

Yao, M., Goult, B.T., Chen, H., Cong, P., Sheetz, M.P., and Yan, J. (2014). Mechanical activation of vinculin binding to talin locks talin in an unfolded conformation. Scientific Reports *4*, 4610.

Yao, M., Goult, B.T., Klapholz, B., Hu, X., Toseland, C.P., Guo, Y., Cong, P., Sheetz, M.P., and Yan, J. (2016). The mechanical response of talin. Nature Comms *7*, 11966.

Figure Legends

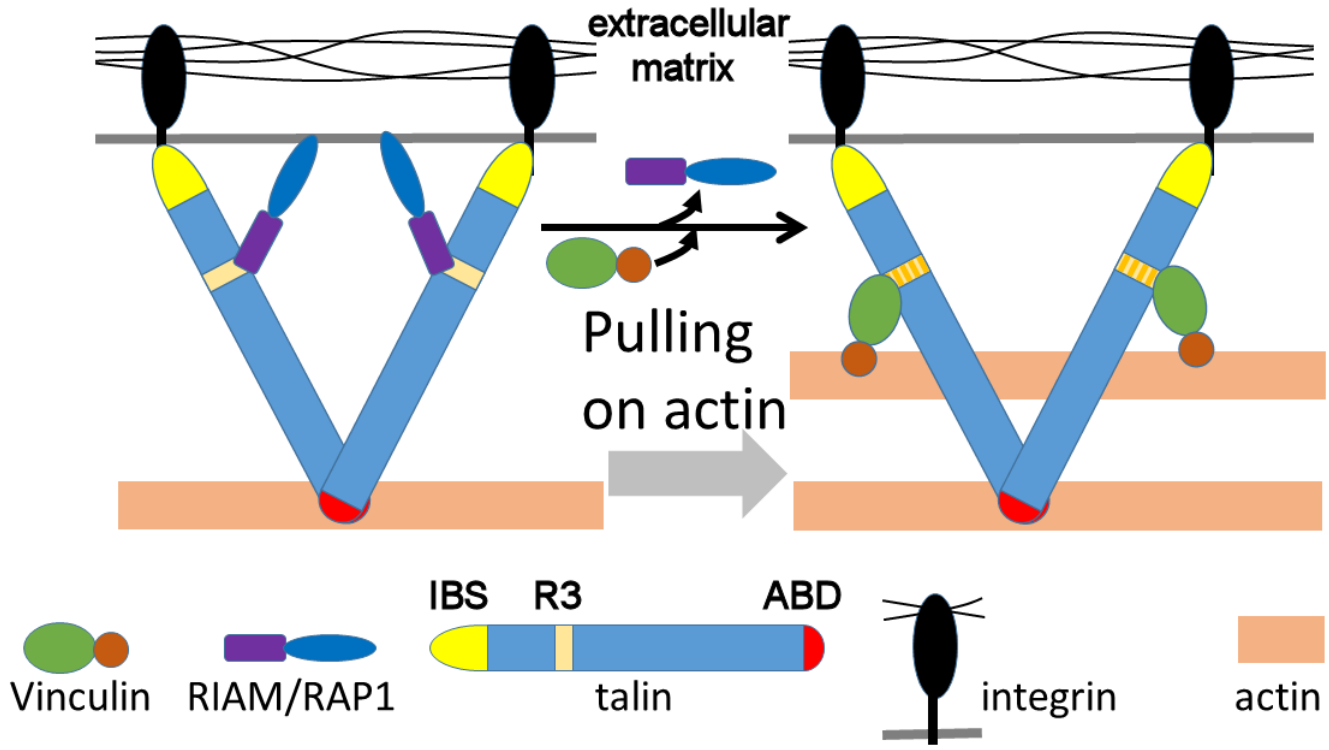


Figure 1. Model for the role of talin in the formation of focal adhesion complexes.

Talin is a long rod-like protein. One end contains an integrin binding site (IBS) while the other contains an actin binding domain (ABD). At rest, RIAM binds to the closed R3 domain of talin and anchors it to the cell membrane. When actin filaments are pulled, the R3 domain undergoes a conformational change which causes RIAM to dissociate and vinculin to bind. Vinculin attaches talin to the actin cytoskeleton and thereby stabilizes the focal adhesion complex. Figure adapted from Klapholz et al. (2015).

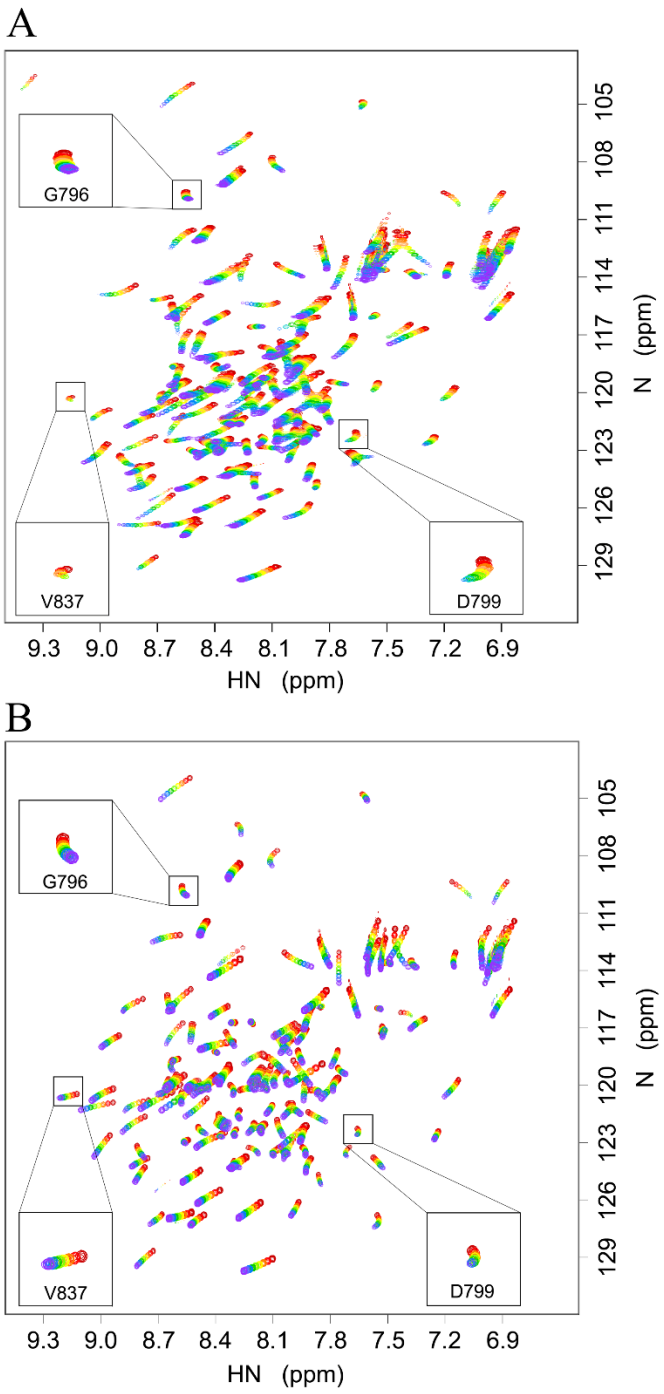
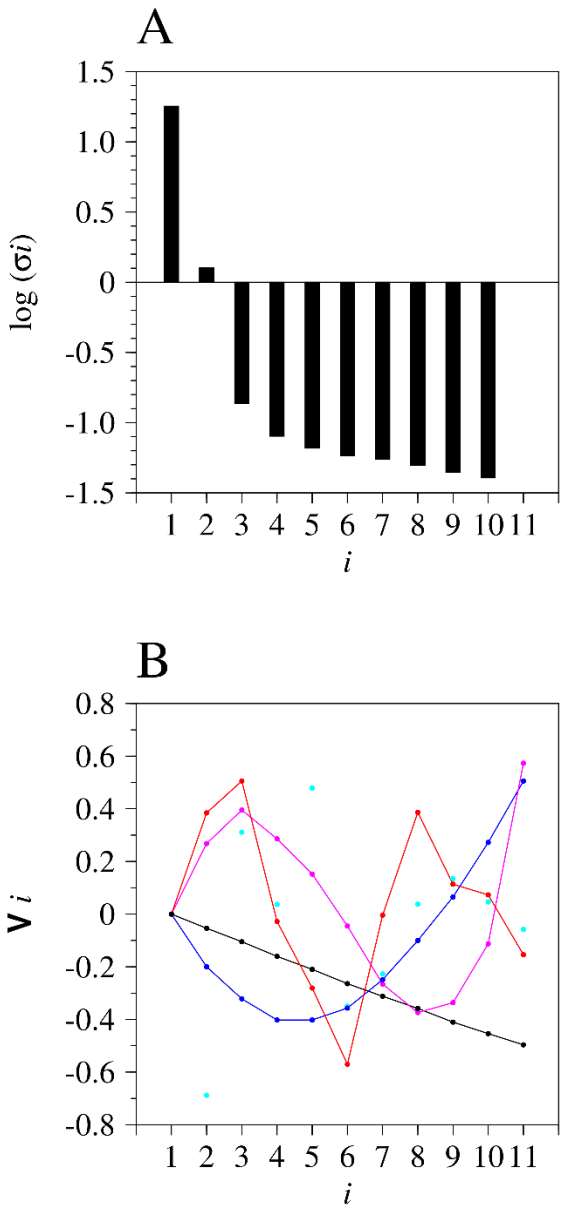


Figure 2. ^{15}N-HSQC spectra of R3 and R3-IVVI acquired at pressures from 1 bar (red) to 2.5 kbar (violet).

(A) R3 and (B) R3-IVVI. The insets show the pressure-induced changes in backbone amide peak position for G796 and D799, positioned at the N-terminal end of helix 1, and V837 which is located at the center of helix 2. These residues together with others show dramatic curvature for R3, whereas their behavior

501 is less curved for R3-IVVI. For a direct comparison of spectra of R3 and R3-IVVI, see Figure S1. For
502 chemical shift differences and their locations on the structure, see Figures S2 and S3.



503

504 **Figure 3. Analysis of SVD fitting.**

505 (A) Plot of log(σ_i) vs. *i* for the SVD combined analysis of backbone amide HN and N observed chemical
506 shift vs. pressure data for R3. The value of σ_{11} is 0. (B) Plot of the first 5 of the eleven column vectors
507 of **V**. Vectors 1 to 4 are indicated by circles and lines colored black, blue, magenta and red, respectively
508 and vector 5 is shown as cyan circles. For fitting to more simple quadratic and cubic equations, and

locations of poorly fitting residues, see Figures S4 and S5. For the equivalent locations for R3-IVVI, see Figure S8; and for the equivalent analysis of SVD fitting for R3-IVVI see Figure S9.

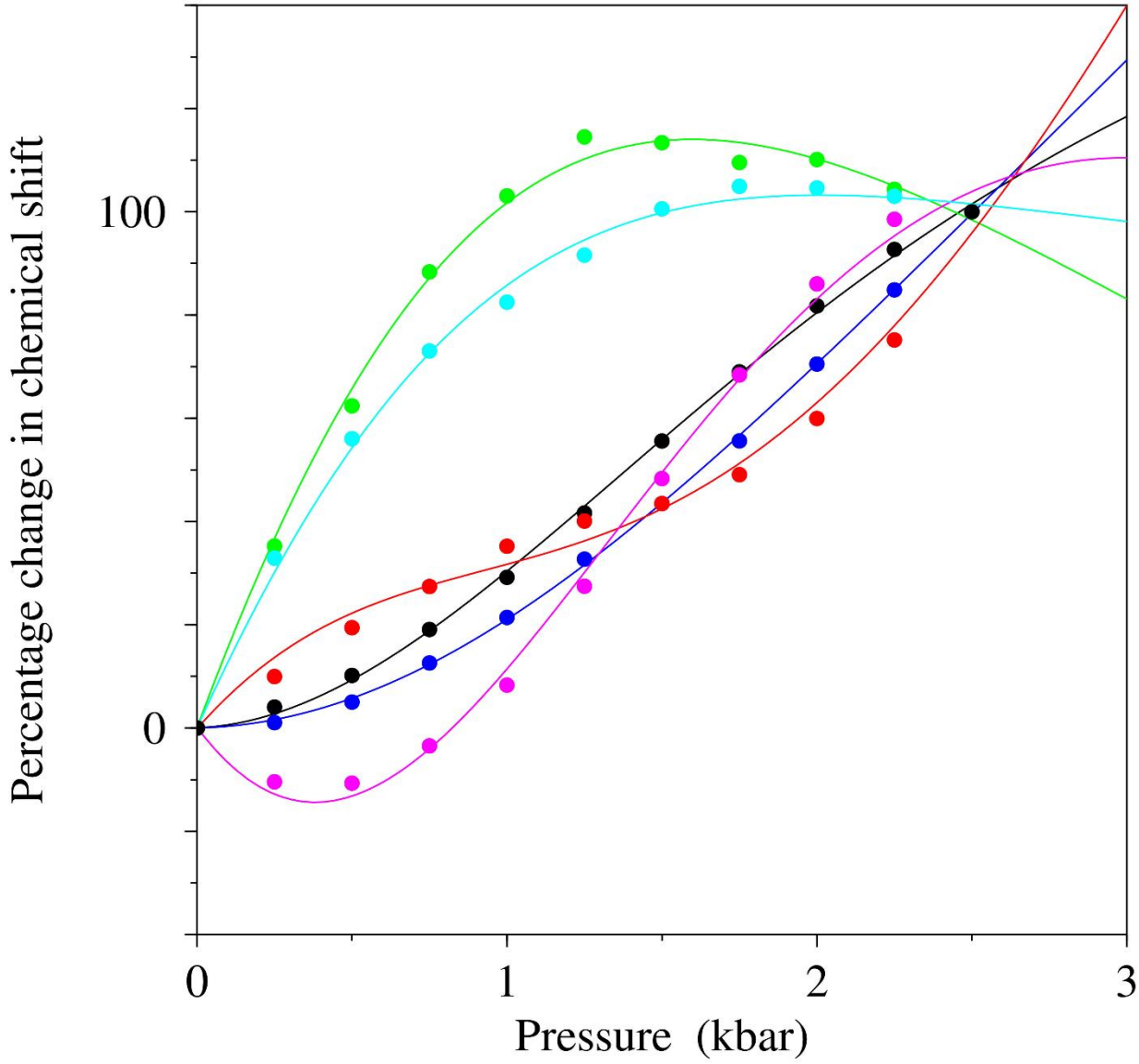
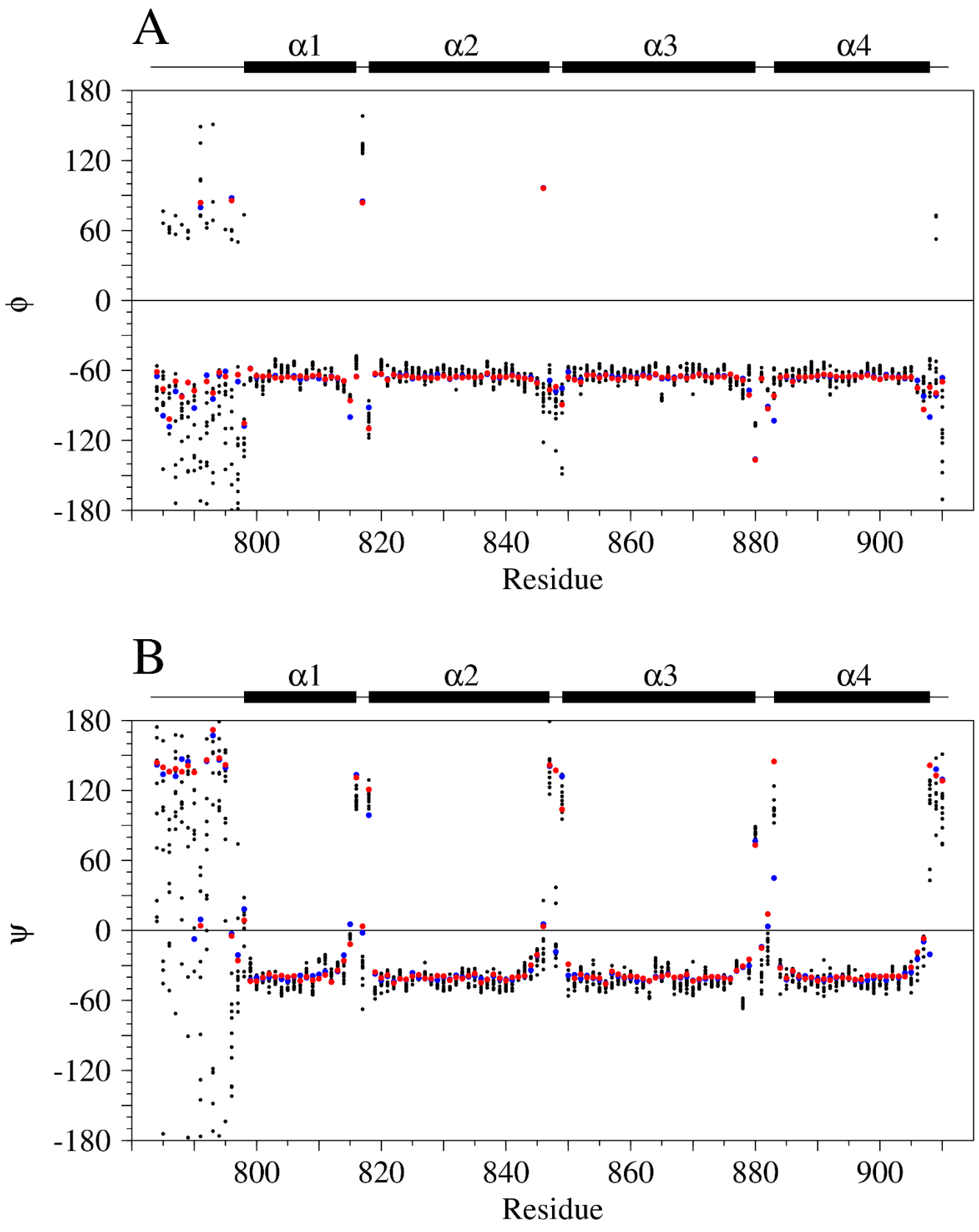


Figure 4. Examples of 'noise-free' chemical shift vs. pressure data for R3.

All curves are rescaled to a maximum chemical shift change of 100%, to illustrate the variable response of specific nuclei to increasing pressure. Experimental data are indicated by circles, and the fits to Eq. 2 (fitted with a global ΔG and ΔV and resonance-specific δ_1^o, δ_2^o, $\Delta\delta_1$ and $\Delta\delta_2$ parameters) are shown by lines: F813 HN (black), V823 N (blue), I828 N (green), R797 C' (red), T809 C' (magenta) and A877 C' (cyan).



518

519 **Figure 5. TALOS-N predictions for backbone dihedral angles.**

(A) ϕ and (B) ψ dihedral angles of the ground state (closed conformation: blue circles) and the excited state (open conformation: red circles), compared to the 10 lowest energy members of the 2L7A NMR structural ensemble (black circles). The four α-helices are indicated.

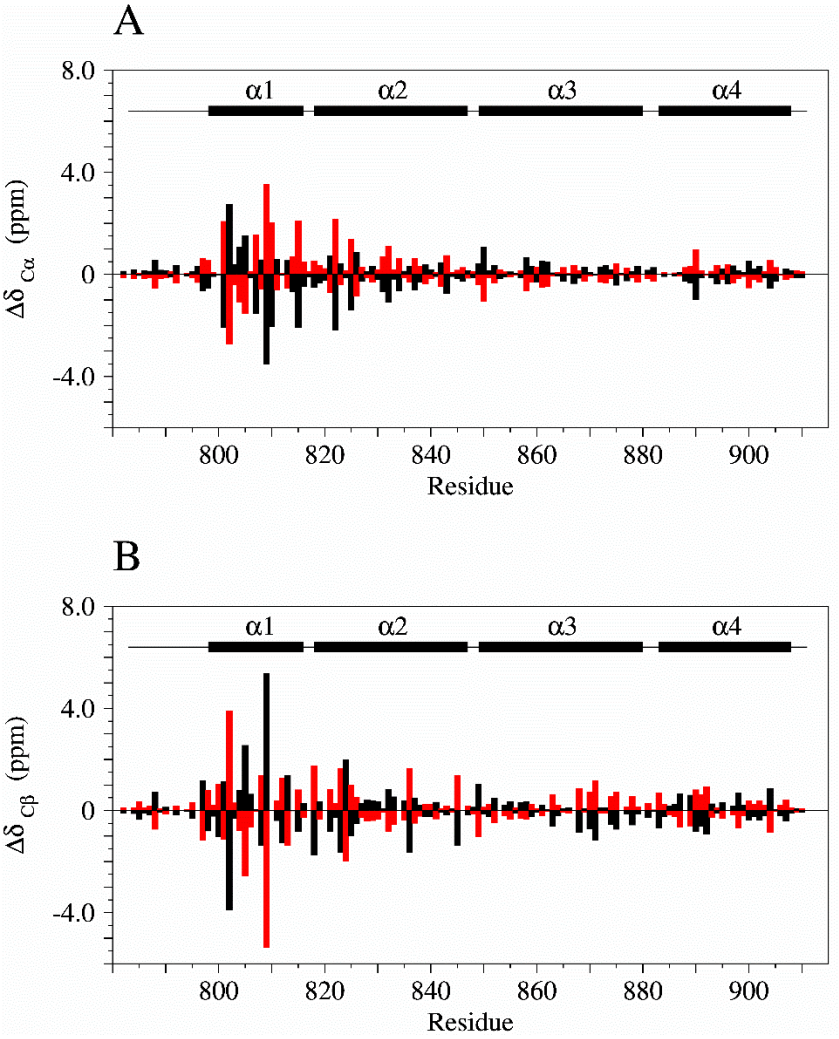


Figure 6. Differences in chemical shift ($\Delta\delta = \delta_1^o - \delta_2^o$) between the ground state structure (closed conformation) and the excited state structure (open conformation) of R3.

(A) Cα nuclei and (B) Cβ nuclei. The bars indicate relative chemical shift changes from δ_{obs} for δ_1^o (ground state: black) and δ_2^o (excited state: red). The four α-helices are indicated. For corresponding differences in HN, N and C', see Figure S6, and for the differences for R3-IVVI, see Figure S10.

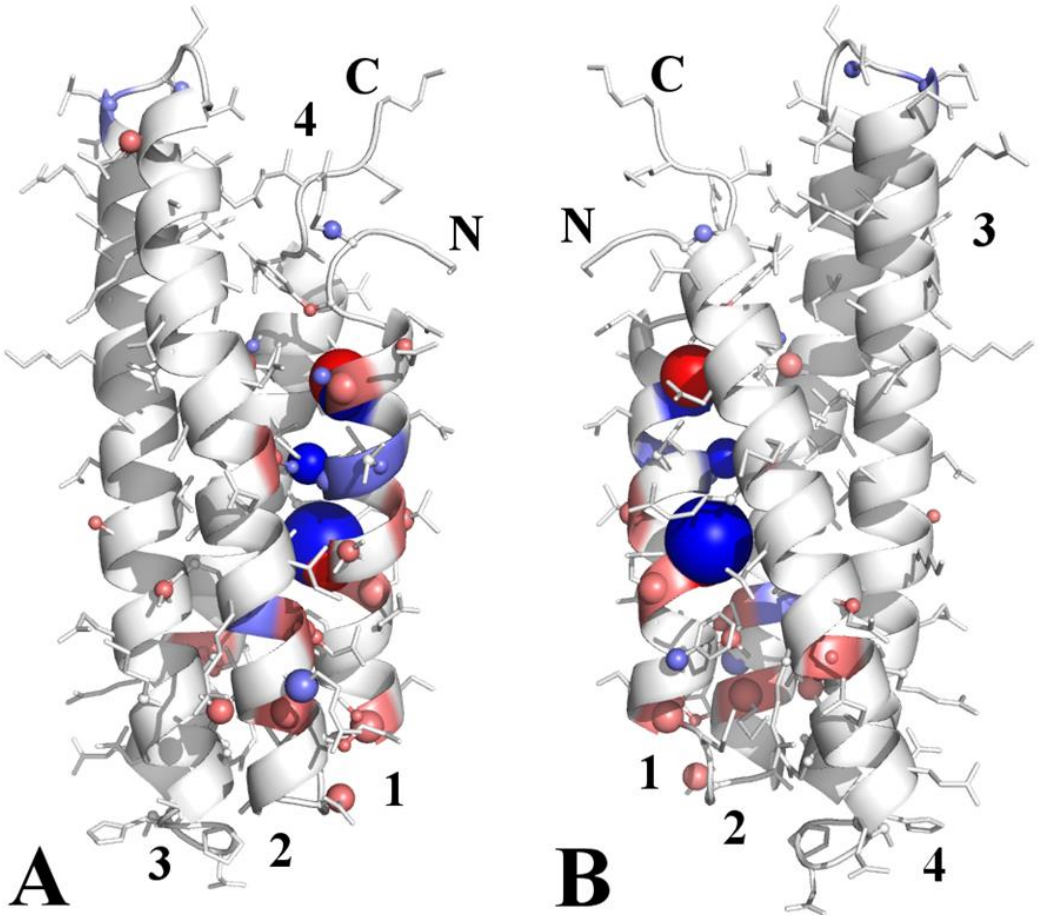


Figure 7. Differences in chemical shift ($\Delta\delta = \delta_1^o - \delta_2^o$) between the ground state (closed conformation) and the excited state (open conformation) of R3.

Differences are shown on the lowest energy NMR structure. The disordered N-terminus is not shown and the cartoon depictions comprise residues A795 to K911. The N- and C-termini are indicated and helices are numbered. (A) $\Delta\delta$ values for Cα nuclei (colored backbone and spheres) and Cβ nuclei (spheres on sidechain sticks) with positive values in blue and negative values in red. Only differences > 1 standard deviation are indicated, with large differences (> 3 standard deviations) in deeper colors and larger spheres. (B) Same as panel (A) except that R3 is rotated 180^o about a vertical axis.

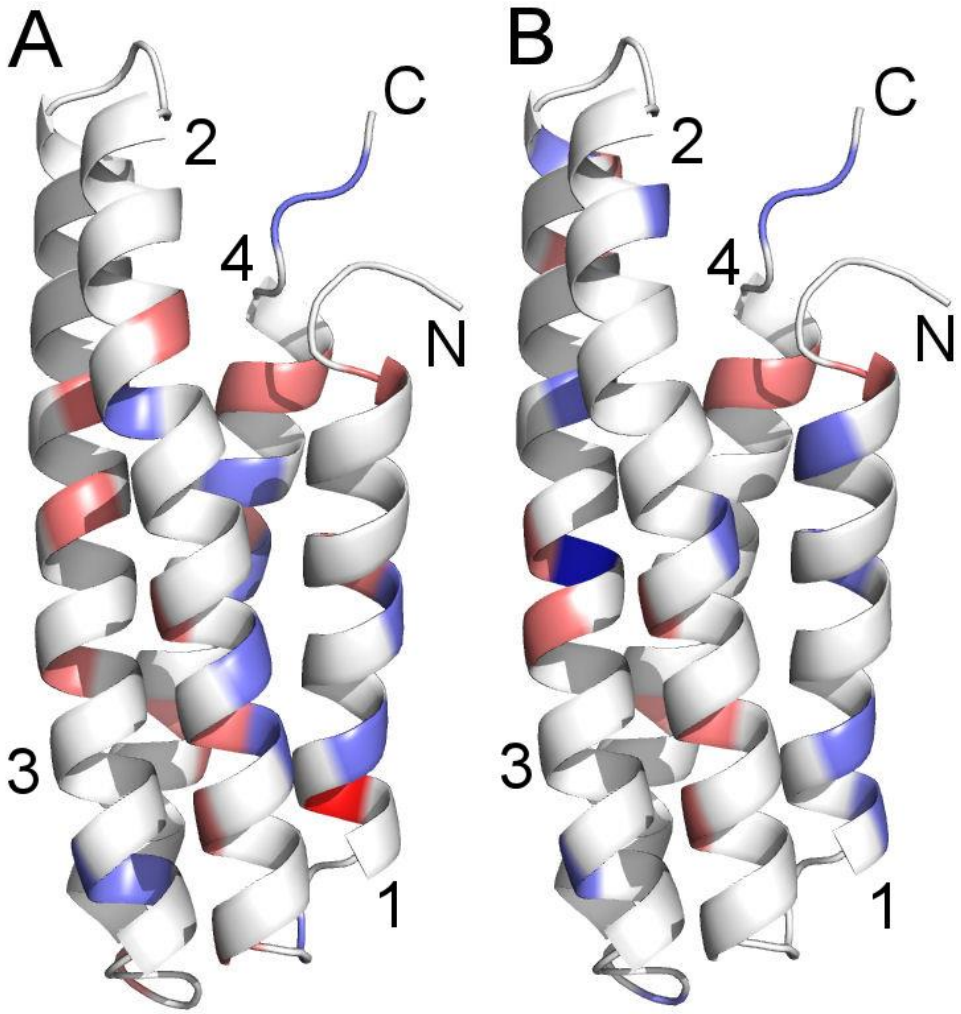


Figure 8. Differences in pressure-dependent gradients (Δgradient = $\Delta\delta_1 - \Delta\delta_2$) for HN nuclei between the ground state (closed conformation) and the excited state (open conformation).

(A) R3 and (B) R3-IVVI, shown on the lowest energy NMR structure. The disordered N-terminus is not shown and the cartoon depictions comprise residues A795 to K911. The N- and C-termini are indicated and helices are numbered. Δgradient values > 1 standard deviation are shown with positive values in blue and negative values in red, with large differences (> 3 standard deviations) indicated as deeper colors. For a graphical view of the gradient values for R3 and R3-IVVI respectively, see Figures S7 and S11.

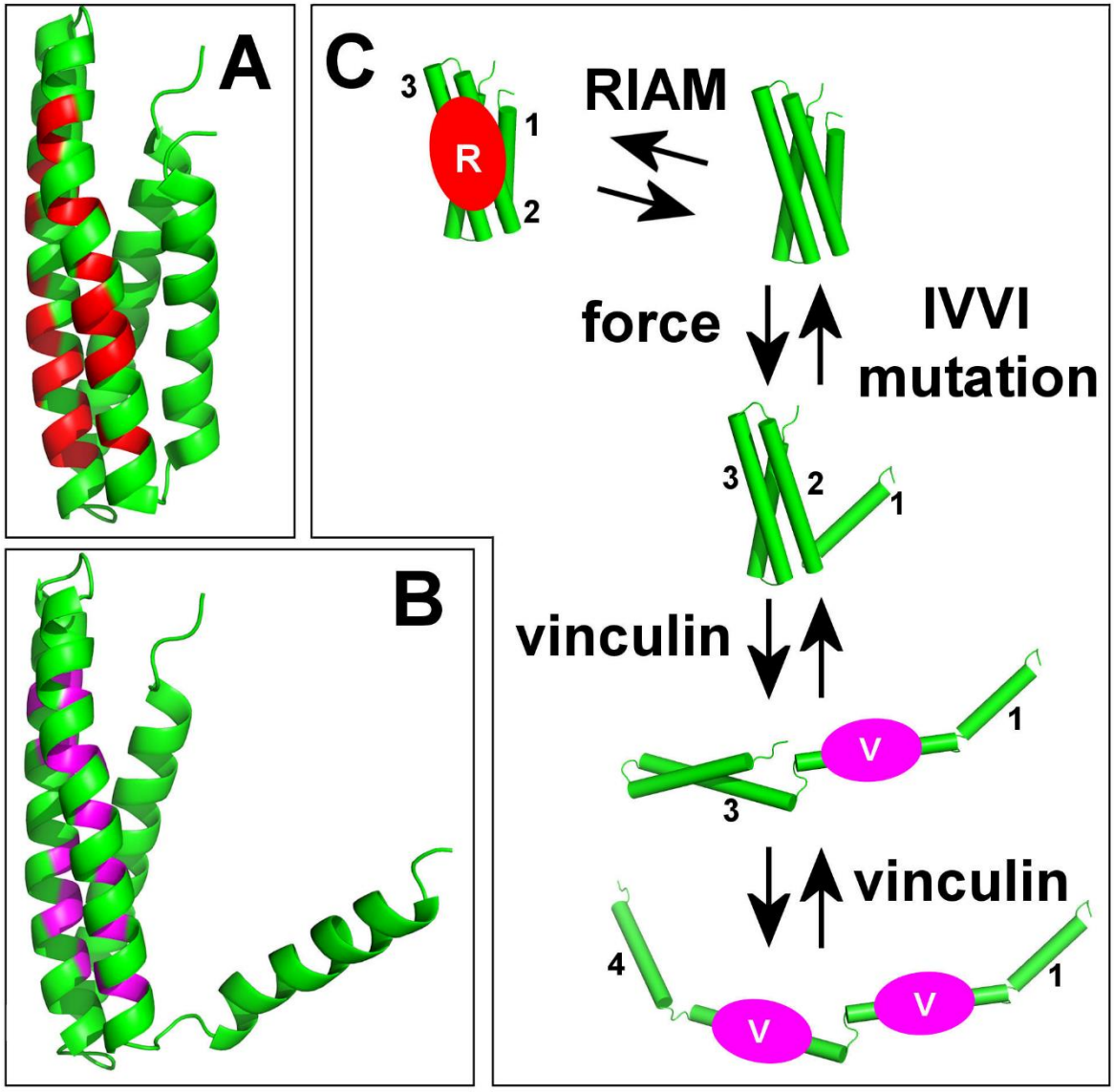


Figure 9. Proposed mode of action of talin domain R3.

(A) Binding site on R3 for RIAM in the closed state. (B) Binding site on R3 for vinculin in the open state. (C) In the unactivated full-length protein, R3 is closed, and helices 2 and 3 form a binding site for RIAM, which is able to bind reversibly to the closed domain. Force, provided by movement of the actin cytoskeleton relative to talin, pulls helix 1 out from the bundle, exposing the binding site on helix 2 for vinculin. This enables vinculin to bind, further opening out the bundle, which then exposes helix 3, forming a second vinculin binding site and leading to complete opening of all four helices. The IVVI mutation stabilizes the closed state and disfavors vinculin binding.

STAR★ METHODS

Detailed methods are provided in the online version of this paper and include the following:

Key resources table

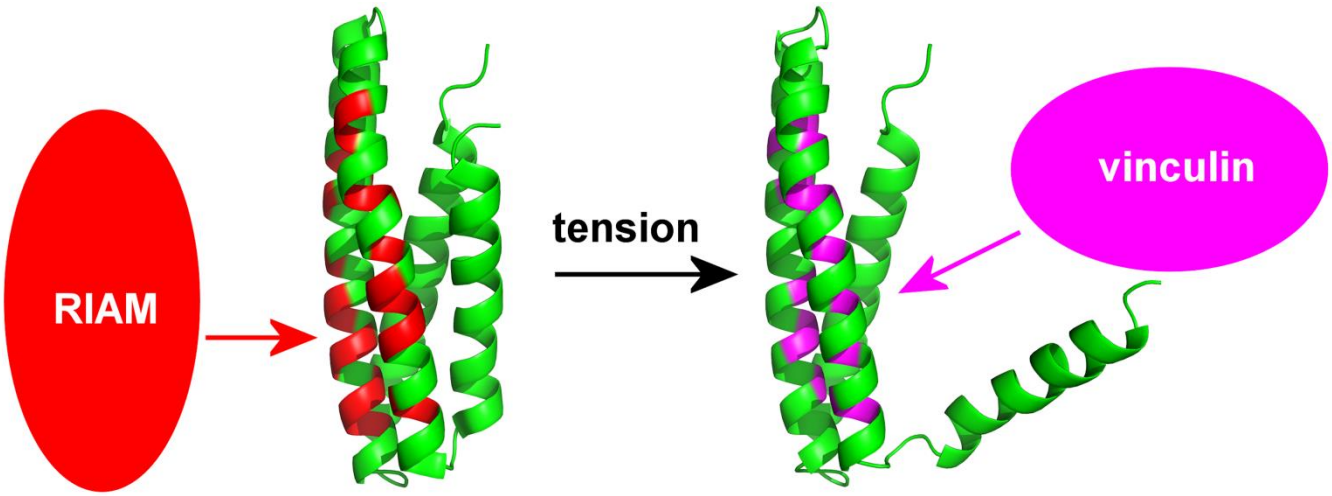
Contact for reagent and resource sharing

Experimental details

- Expression and purification of protein
- High-pressure NMR
- Data analysis

Data and software availability

Graphical Abstract



SUPPLEMENTAL INFORMATION

Supplemental information includes eleven figures, a justification of Equation 2 with associated Table S1, and a compressed file containing fitted data and can be found with this article online at:

ABBREVIATIONS

R3, domain R3 of talin; R3-IVVI, T809I/T833V/T867V/T901I mutation of R3; RIAM, Rap1-GTP-interacting adaptor molecule; SVD, singular value decomposition

STAR★ METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
E. coli BL21 STAR (DE3)	Thermo Fisher	Cat #C601003
pET151/D-TOPO	Goult et al. 2013	
Chemicals, Peptides, and Recombinant Proteins		
$^{15}NH_4Cl$	Sigma Aldrich	Cat # 299251
$^{13}C_6$-glucose	Sigma Aldrich	Cat # 389374
TEV protease	Recombinant-In house	N/A
Hi-Trap QFF	GE-healthcare	Cat #17-5053-01
TCEP	Sigma Aldrich	Cat # C4706
TSP	Sigma Aldrich	Cat # 269913
Deposited Data		
R3-IVVI assignments	BioMagResBank	26880
Software and Algorithms		
SVD	Matlab	
TALOS-N	Shen & Bax, 2013	https://spin.niddk.nih.gov/bax/software/TALOS-N/

Contact for reagent and resource sharing

584 Requests for the expression plasmids should be directed to Igor Barsukov (i.barsukov@liverpool.ac.uk).
585 All other information and requests should be directed to the Lead Contact, Mike Williamson
586 (m.williamson@sheffield.ac.uk).

587 **Experimental details**

588 **Expression and purification of protein**

589 Wild-type mouse talin fragment R3 (residues 787-911) and R3-IVVI mutant (T809I/T833V/T867V/T901I)
590 were previously cloned into pET151/D-TOPO expression vector (Invitrogen), encoding an N-terminal
591 hexa-histidine tag (Goult et al., 2013). Recombinant proteins were expressed and purified as described
592 previously (Goult et al., 2009). Briefly, protein was produced in *E. coli* strain BL21 STAR (DE3) cultured
593 in LB or 2xM9 minimal medium containing 1 g/L $^{15}NH_4Cl$ and 4 g/L glucose or 2 g/L $^{13}C_6$-glucose. Cells
594 were grown at 37^{o}C to an OD_{600} of 0.6, cooled to 18^{o}C and induced using 0.5 mM IPTG for 16 hours.
595 His-tagged protein was purified by nickel-affinity chromatography following standard protocol and the tag
596 was removed by cleavage with TEV protease, followed by the reverse purification. Protein was further
597 purified using anion exchange chromatography using a 5 ml Hi-Trap QFF column (GE Healthcare).
598 Protein concentrations were determined using absorbance at 280 nm and calculated extinction
599 coefficients.

600

601 **High-pressure NMR**

602 Samples (300 μl) contained 1 mM R3 or R3-IVVI and were dissolved in 20 mM sodium phosphate pH
603 6.5, 150 mM NaCl, 0.5 mM TCEP in 10% $^{2}H_2O$/90% H_2O and placed in a ceramic tube connected to a
604 high-pressure pump (Daedalus Innovations) (Peterson et al., 2011). Spectra were obtained at 298 K on
605 an 800 MHz Bruker Avance I spectrometer equipped with a room-temperature probe (R3) or on a 600
606 MHz Bruker DRX spectrometer fitted with a cryoprobe (R3-IVVI), at regularly spaced pressures up to 2.5
607 kbar. For R3, 2D ^{15}N-HSQC, 2D HN(CO)CACB and 2D HNCO spectra were acquired every 250 bar from
608 1 bar to 2.5 kbar, while for R3-IVVI, 2D ^{15}N-HSQC spectra were acquired every 310 bar from 1 bar to

609 2.48 kbar. Proton chemical shifts were referenced to 3-trimethylsilyl-2,2,3,3-($^2H_4$) propionate (TSP, Sigma
610 Aldrich) at 0.0 ppm. ^{15}N and ^{13}C chemical shifts were calculated relative to TSP by use of the
611 gyromagnetic ratios of ^{15}N, ^{13}C and ^1H nuclei ($\gamma(^{15}N)/\gamma(^1H)$ = 0.101329118 and $\gamma(^{13}C)/\gamma(^1H)$ =
612 0.251449530). Peaks were picked using Felix (Felix NMR Inc., San Diego, CA) and exported to text files.
613 R3 backbone assignments at ambient pressure were taken from Goult et al., 2013 and were confirmed
614 using 3D HNCACB, 3D HNCO and 3D HN(CA)CO spectra at ambient pressure. Backbone assignments
615 for R3-IVVI at ambient pressure were obtained using standard triple-resonance methods and have been
616 deposited in BioMagResBank (http://www.bmrb.wisc.edu/) under the BMRB accession code 26880.
617
618 **Data analysis**
619 SVD was carried out using Matlab™. SVD is a well-established technique for factorizing matrices, related
620 to principal component analysis, and is used extensively in signal processing and statistics (Henry and
621 Hofrichter, 1992; Noble and Daniel, 1988). It is reported to be the least biased way of extracting the
622 meaningful data from an original overdetermined set containing experimental noise (Henry and
623 Hofrichter, 1992). Any real $p \times q$ matrix D can be factorized as
624 $$D = UWV^T$$
625 where U is a $p \times p$ unitary matrix, V is a $q \times q$ unitary matrix and V^T is its transpose, and W is a $p \times q$
626 diagonal matrix, whose diagonal elements σ_i are real and non-negative. The σ_i are called the *singular*
627 *values* of D, and are normally presented in descending numerical value. The number *n* of non-zero σ_i
628 defines the rank of the matrix, ie the number of independent components (here, the number of
629 independent molecular species whose chemical shifts are required in order to fit the data). In practice,
630 only a few of the σ_i have large values, with the remainder having values which are small but not exactly
631 zero, since they arise from uncorrelated noise in the experimental data. It is therefore possible to set all
632 the randomly near-zero σ_i to zero, leaving a reduced W' as a $n \times n$ diagonal matrix, and at the same time
633 reduce U and V to only *n* columns. The resulting D' matrix, calculated from $D' = U'W'V'^T$ is thus a 'noise-

634 free' version of D, in which only correlated noise remains. SVD is therefore very useful, for (a) deriving
635 the rank of D, ie the minimum number of independent components required to generate the chemical
636 shift changes observed, and (b) removing most of the noise from the data, therefore improving
637 subsequent fitting of the data.

638 The experimental data set was factorized to yield the U, W and V matrices. This analysis was
639 carried out separately for amide proton and amide nitrogen chemical shifts, and also for both sets
640 combined. The fitting was of similar quality for all three sets, as were the results, and therefore we report
641 the use of combined HN and N shifts, because a global analysis of all chemical shift values together
642 should provide the most robust fitting (Arai et al., 2012). In particular, the noise on each individual point
643 is reduced in an ideal case by $pq^{-1/2}$, which here is approximately a factor of 50 when combining ^{1}H and
644 ^{15}N shifts together (Henry and Hofrichter, 1992). The raw data D consisted of separate lists of backbone
645 ^{1}HN, ^{15}N, ^{13}Cα, ^{13}Cβ, and ^{13}C' chemical shift changes for R3, and backbone ^{1}HN and ^{15}N chemical shift
646 changes for R3-IVVI at each pressure. Analysis of the ^{13}Cα, ^{13}Cβ and ^{13}C' shifts was carried out
647 independently, because the number of residues with usable non-overlapped signals was different for
648 each group. Rather than using absolute chemical shift values, the experimental data were all input as
649 changes in chemical shift from the initial chemical shift value. SVD fitting of absolute shifts generates
650 one very large singular value (reflecting the starting chemical shift values) together with other much
651 smaller ones, whereas fitting of chemical shift differences generates singular values of a more similar
652 size. Subsequent least-squares fitting is then more robust (Henry and Hofrichter, 1992).

653 Peaks that showed evidence of intermediate exchange broadening in the 2D HN(CO)CACB
654 spectra (Cα preceding T802, D803, I805,V808, E810, F813 and S814, and Cβ preceding L806, V808,
655 T809, E810 and F813), together with a small number of peaks with very small chemical shift changes,
656 were removed from the analysis. Following the SVD, all components of U, W and V above rank 4 were
657 set to zero and used to calculate the 'noise-free' dataset D'. These chemical shift changes were rescaled
658 to have the same maximum shift change, to avoid biasing the fitting by a few resonances with very large
659 chemical shift changes. The rescaled shifts were fitted to Eq. 2 using a Levenberg-Marquardt non-linear

660 least-squares algorithm. Global ΔG and ΔV values (together with resonance-specific δ_1^0, δ_2^0, $\Delta\delta_1$ and $\Delta\delta_2$
661 parameters) were first obtained by fitting to Equation 2 using a subset of resonances which had the most
662 curved pressure-dependent chemical shift changes, based on the value of χ^2 when fitting the data to a
663 quadratic expression. The error in the fitted ΔG and ΔV values was estimated by using different residues
664 and nuclei for the fitting, and also by carrying out a Monte Carlo-type search, varying ΔG and ΔV and
665 calculating the goodness of fit of the data. Once reliable values of ΔG and ΔV had been obtained for this
666 subset of resonances, the values were fixed, and resonance-specific δ_1^0, δ_2^0, $\Delta\delta_1$ and $\Delta\delta_2$ were calculated
667 for all other resonances, after which the data were rescaled to the original values for subsequent analysis.
668 During the fitting for R3, it was observed that the values for δ_1^0 and δ_2^0 were in every case almost exactly
669 of the same magnitude but of opposite sign. This is a consequence of fitting the chemical shifts as
670 differences from the starting value, and of the very small absolute value of ΔG for R3 (ie close to 50%
671 populations of the two states at ambient pressure). In order to obtain more robust fits, δ_1^0 and δ_2^0 were
672 therefore restrained to be equal in magnitude but of opposite sign (for R3 only), thereby reducing the
673 number of fitted parameters by one. Backbone dihedral angles were calculated for the ground and
674 excited states using TALOS-N (Shen and Bax, 2013), using chemical shift values for HN, N, Cα, Cβ, and
675 C' nuclei obtained from δ_1^o and δ_2^o, respectively.
676
677 A typical Matlab™ script for carrying out SVD analysis is as follows:

678 <arrange the data into a matrix n rows by m columns, where n is the number of peaks and m is the number of
679 titration points, and import it. For ^{15}N-HSQC data, the list contains both H and N shifts>

```
680 d=data;
681 [u,s,v]=svd(d,0);
682 vsmall=v(:,1:4);
683 ssmall=s(1:4,1:4);
```

```
684 usmall=u(:,1:4);
685 temp=ssmall*vsmall';
686 noiseless=usmall*temp;
687 e=noiseless';
```

Figure 1

Click here to download Figure fig1.tif

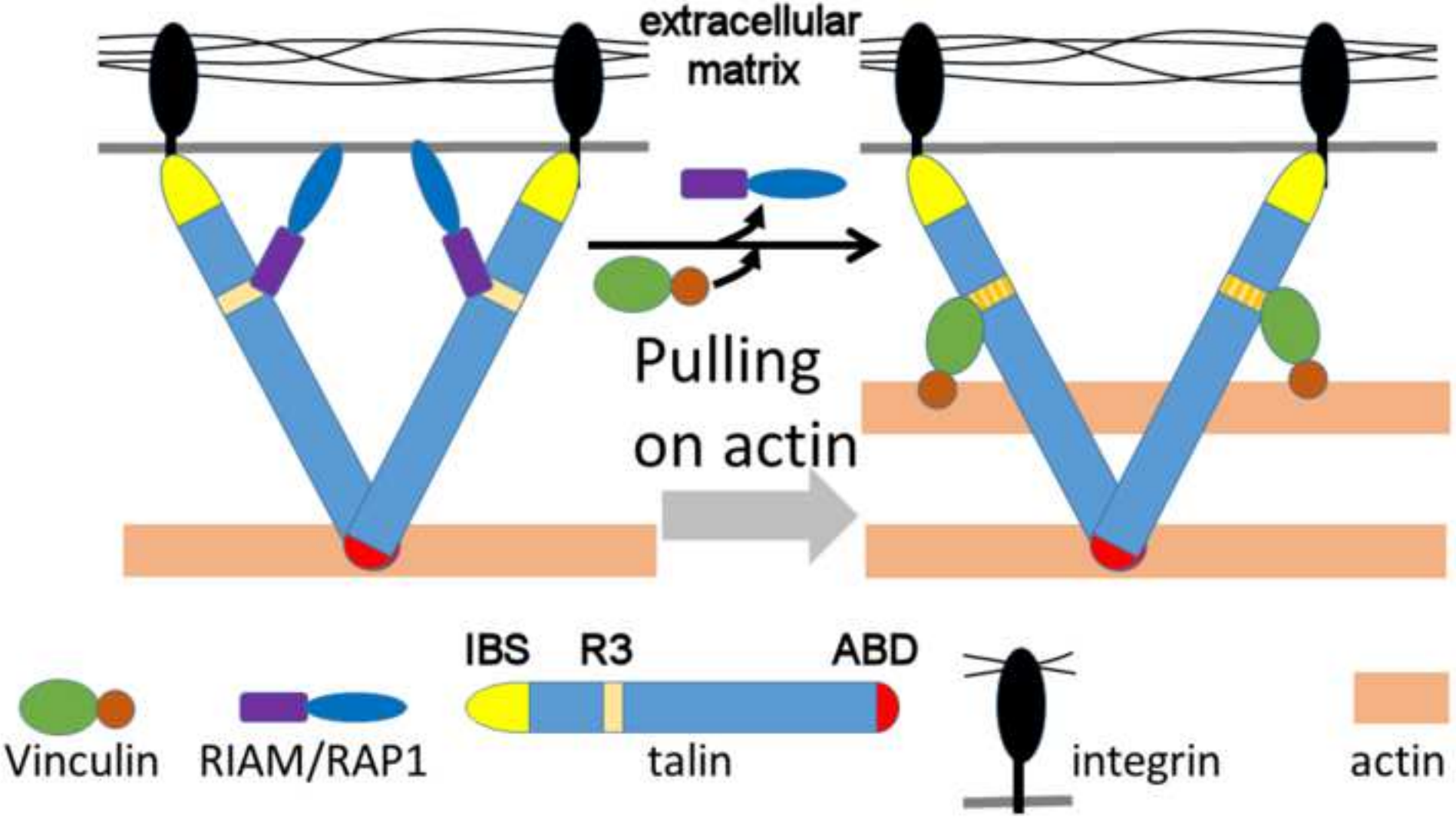


Figure 2

Click here to download Figure fig2.tif

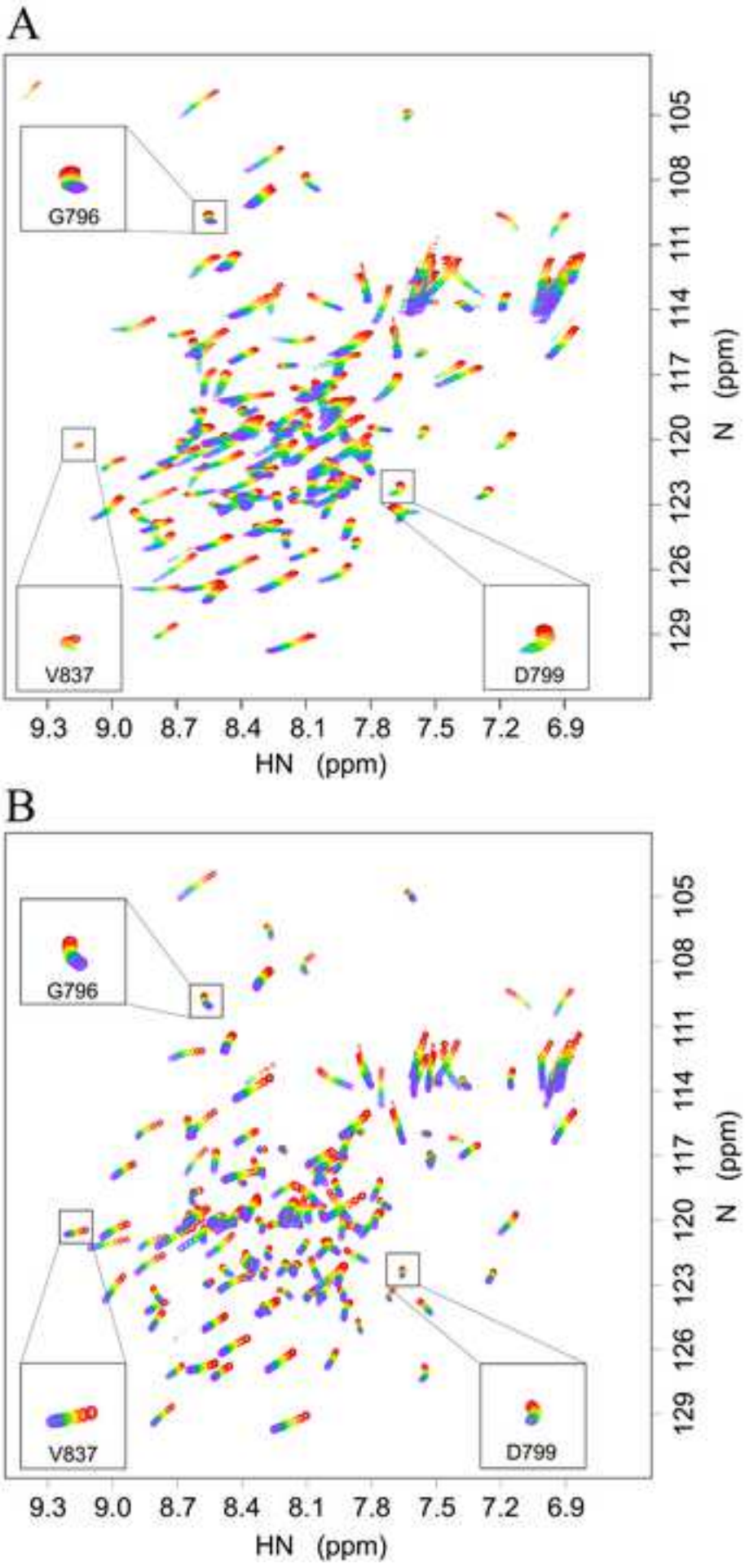


Figure 3

Click here to download Figure fig3.tif

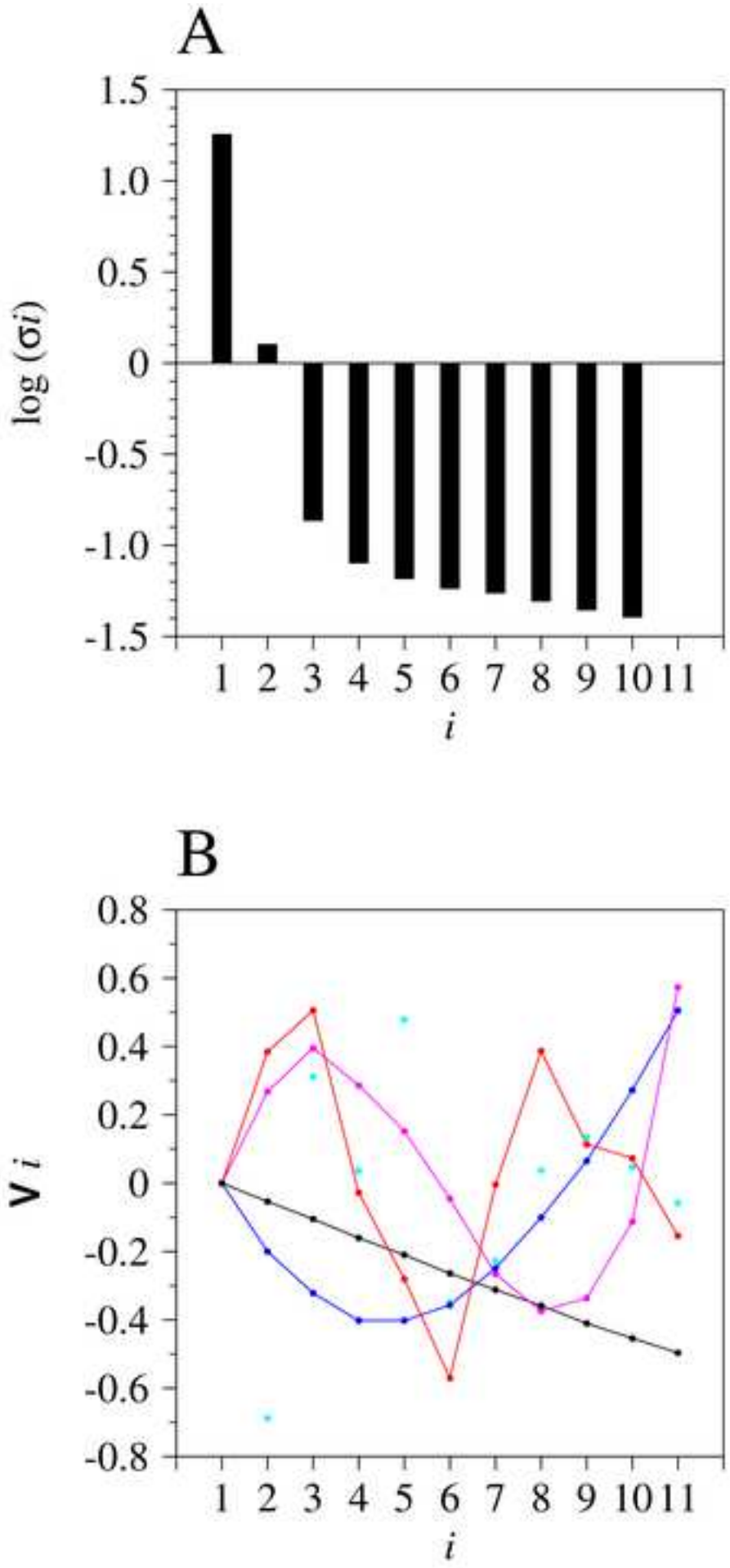


Figure 4

Click here to download Figure fig4.tif

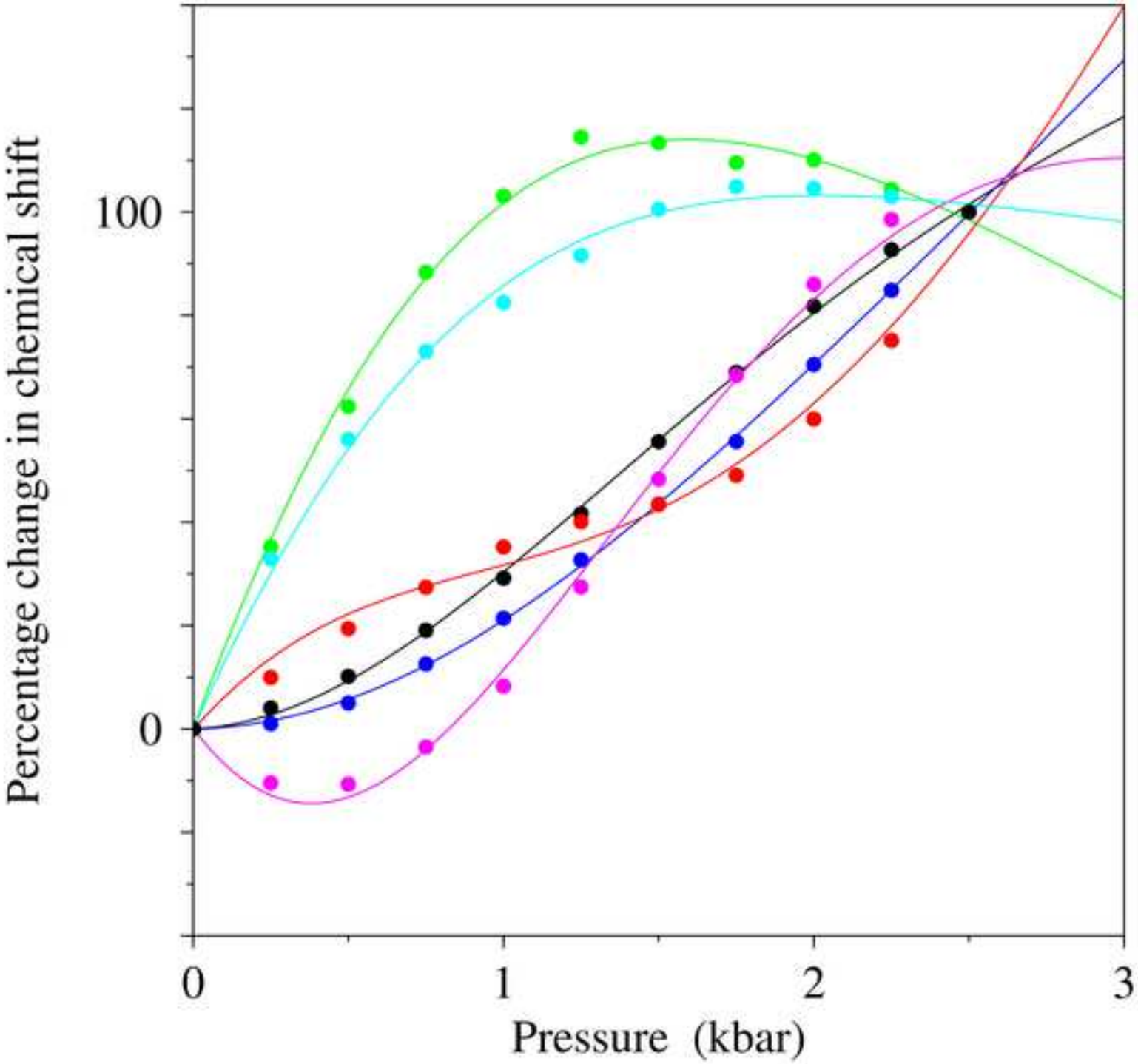


Figure 5

Click here to download Figure fig5.tif

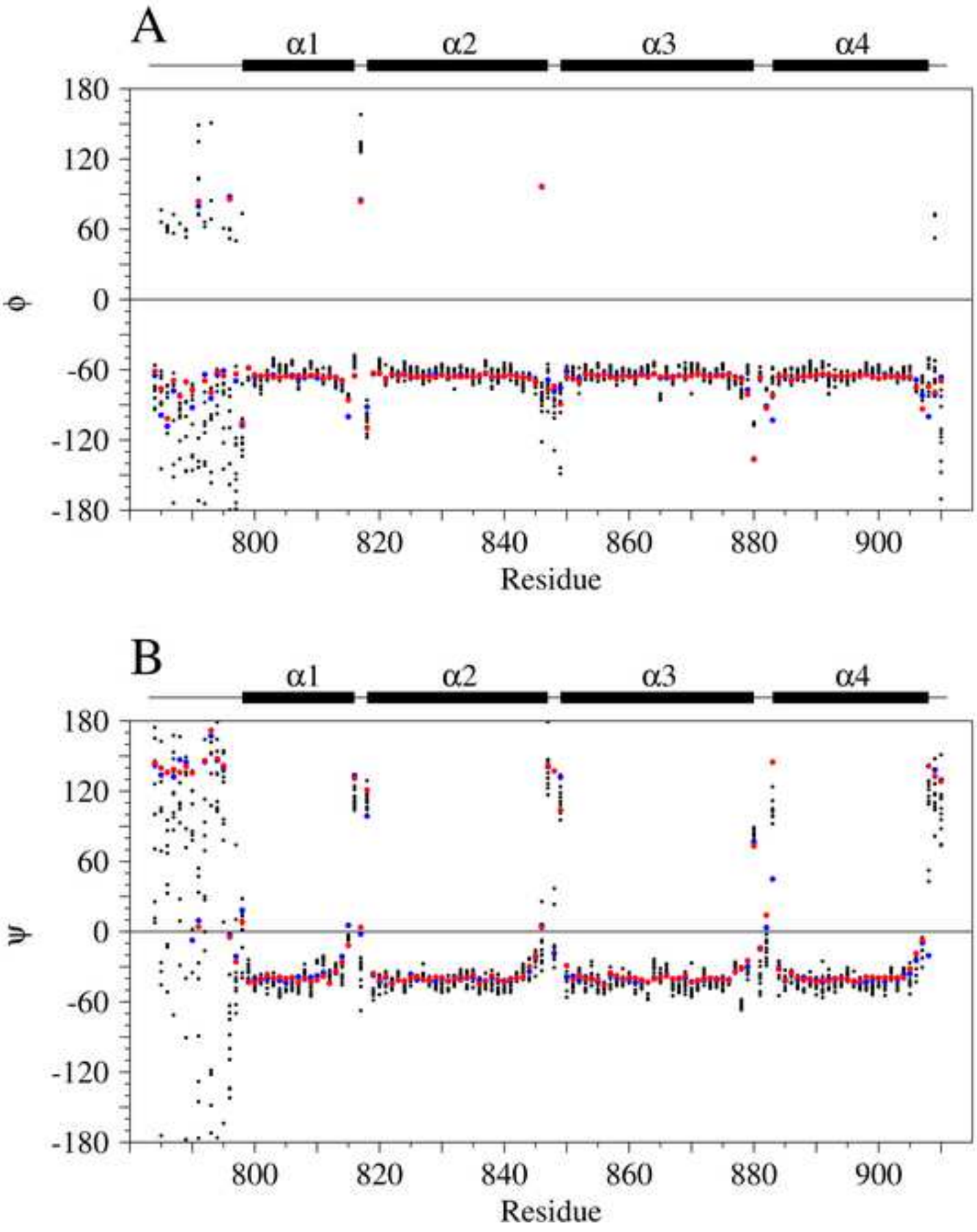


Figure 6

Click here to download Figure fig6.tif

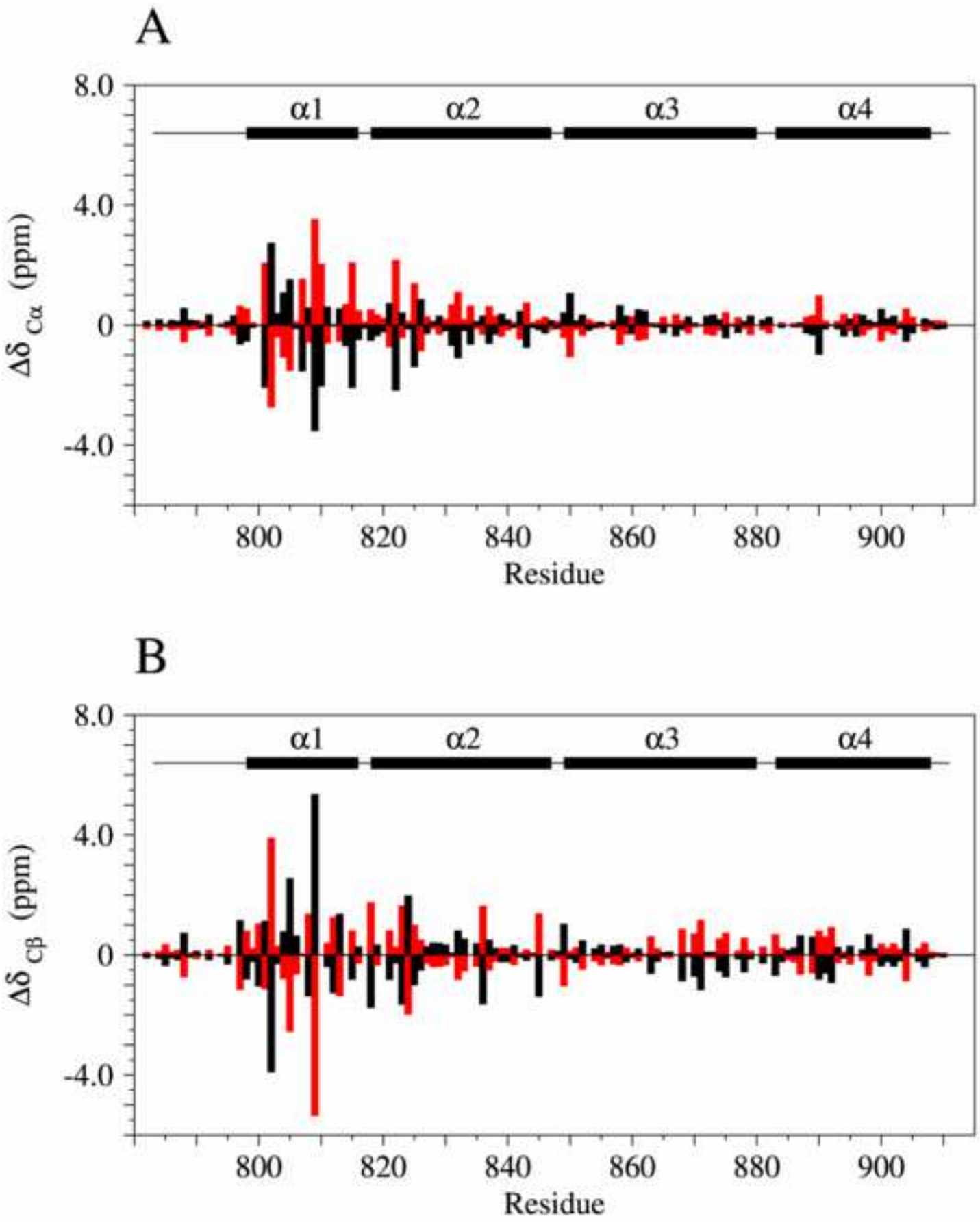


Figure 7

Click here to download Figure fig7.tif

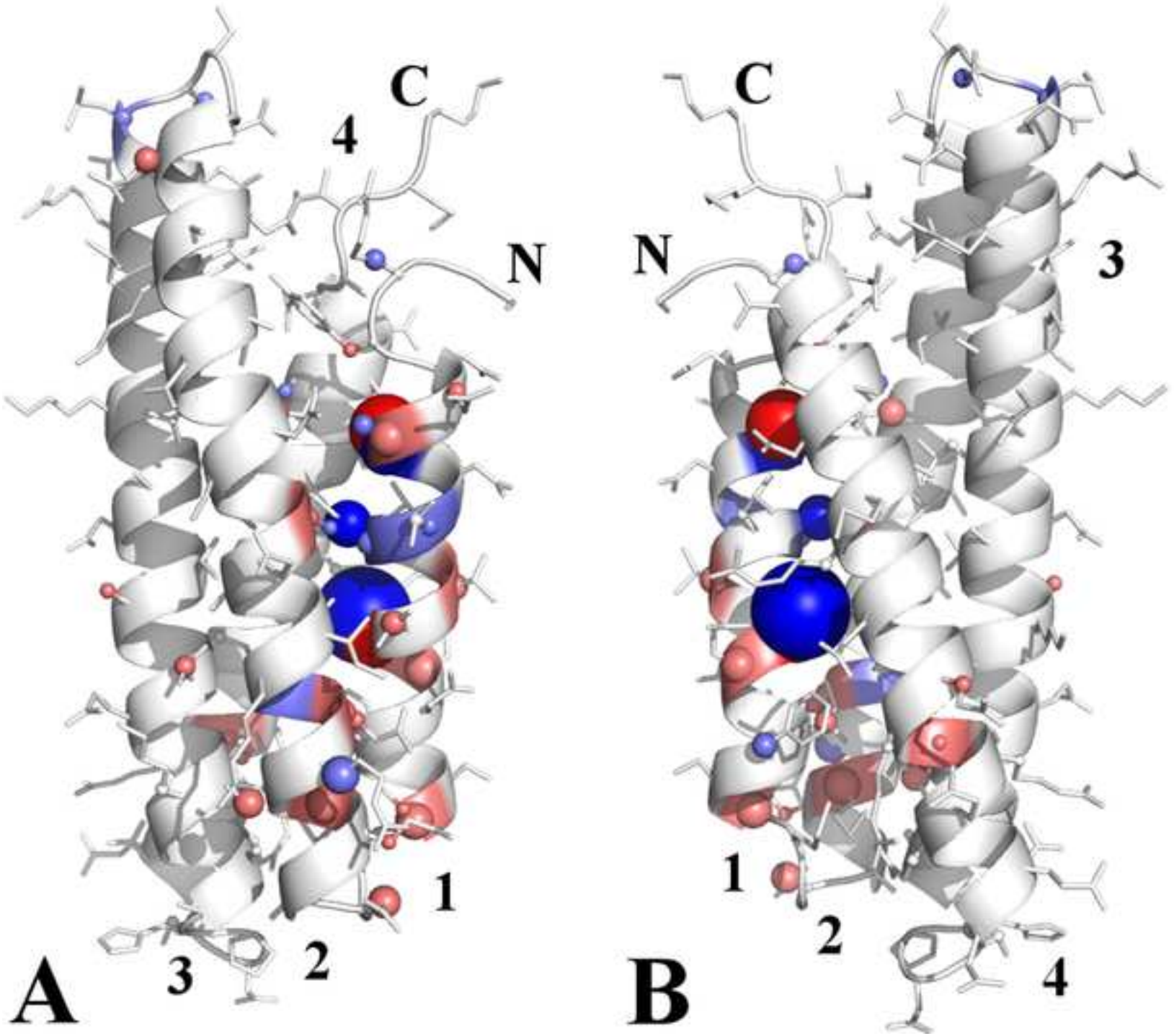


Figure 8

Click here to download Figure fig8.tif

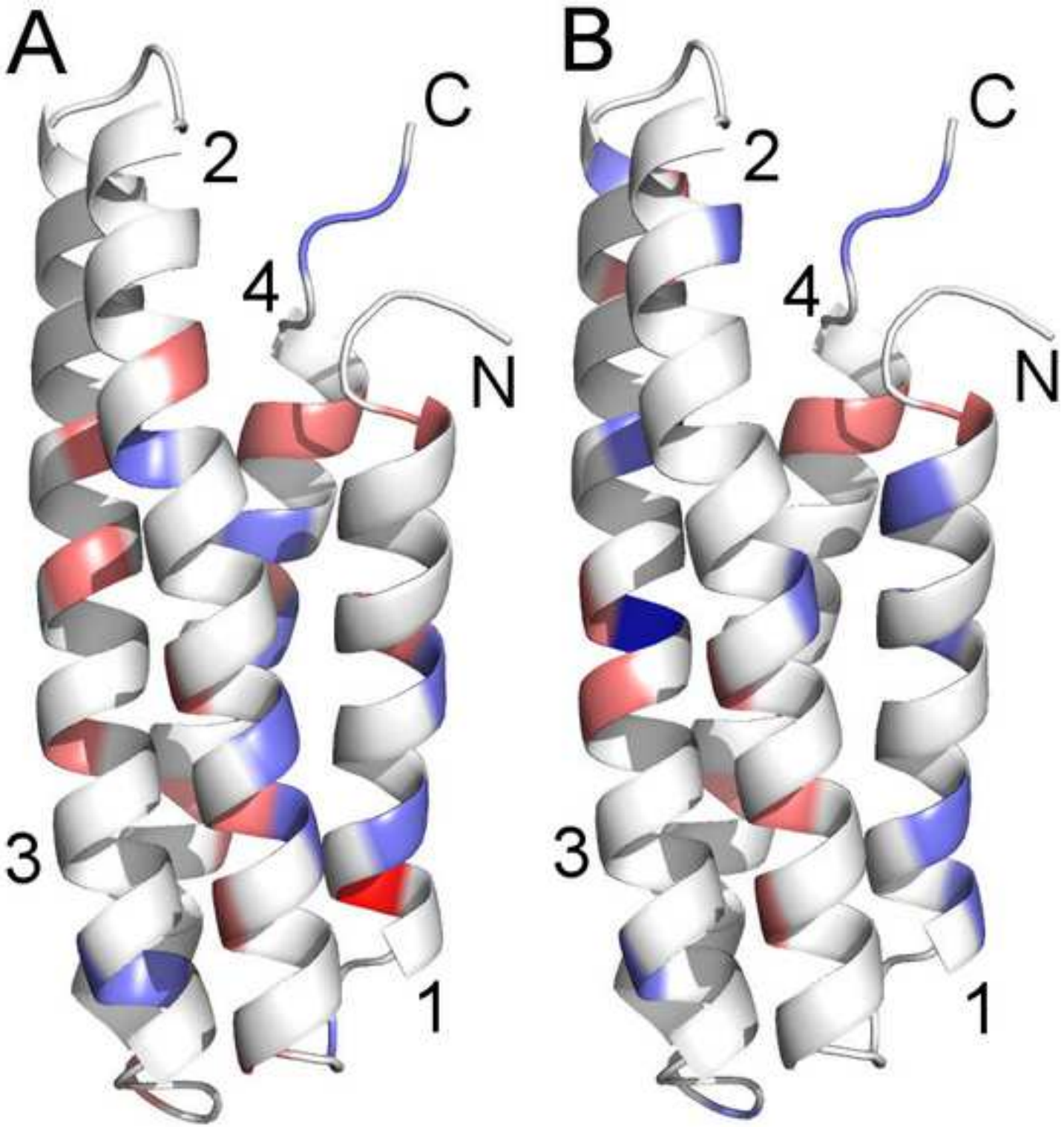
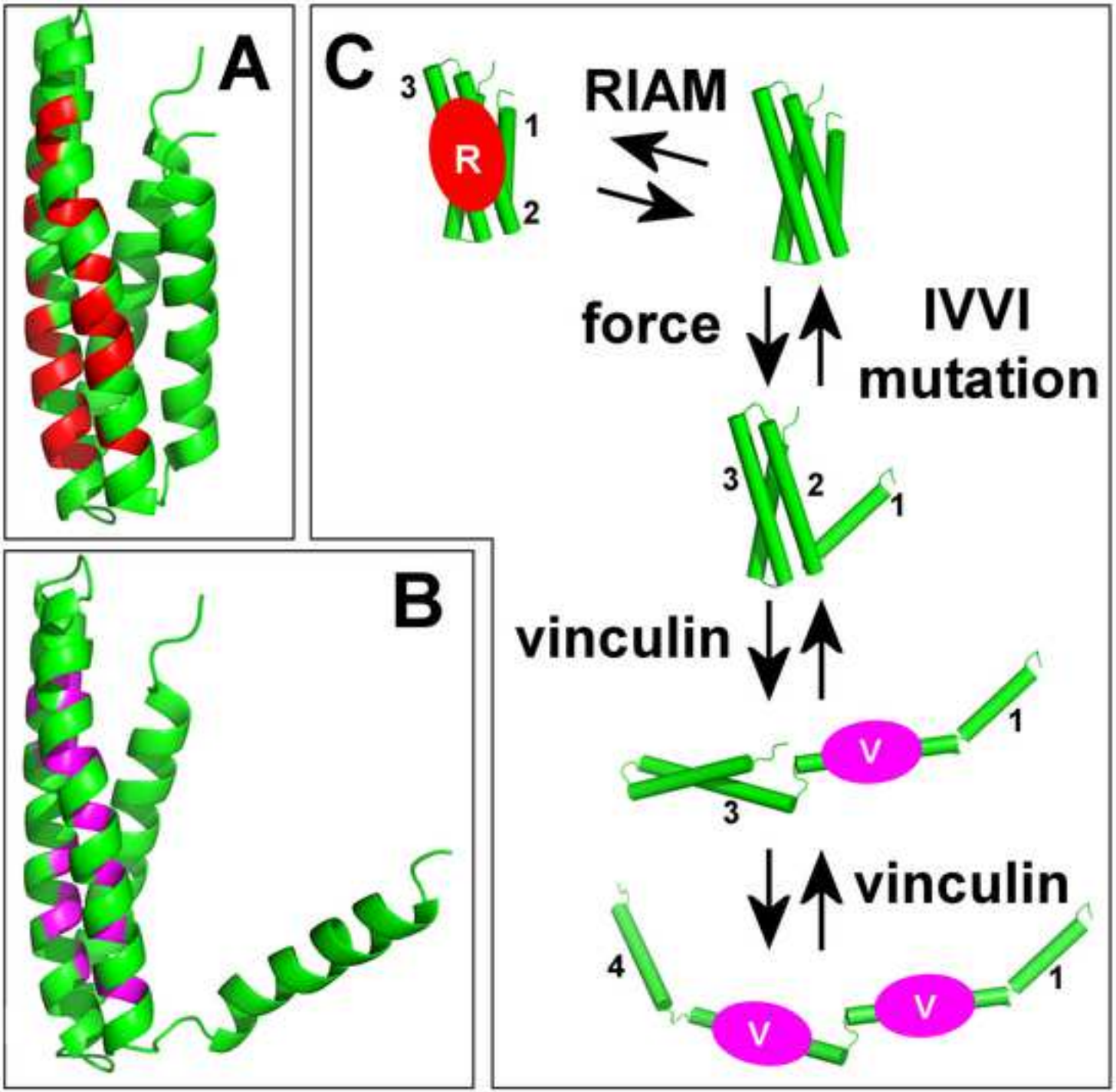


Figure 9

Click here to download Figure fig9.tif



Supplementary Information for

Pressure-dependent chemical shifts in the R3 domain of talin show that it is thermodynamically poised for binding to either vinculin or RIAM

Nicola J. Baxter, Thomas Zacharchenko, Igor L. Barsukov and Mike P. Williamson

Supplementary figures 1-11 and justification for Equation 2

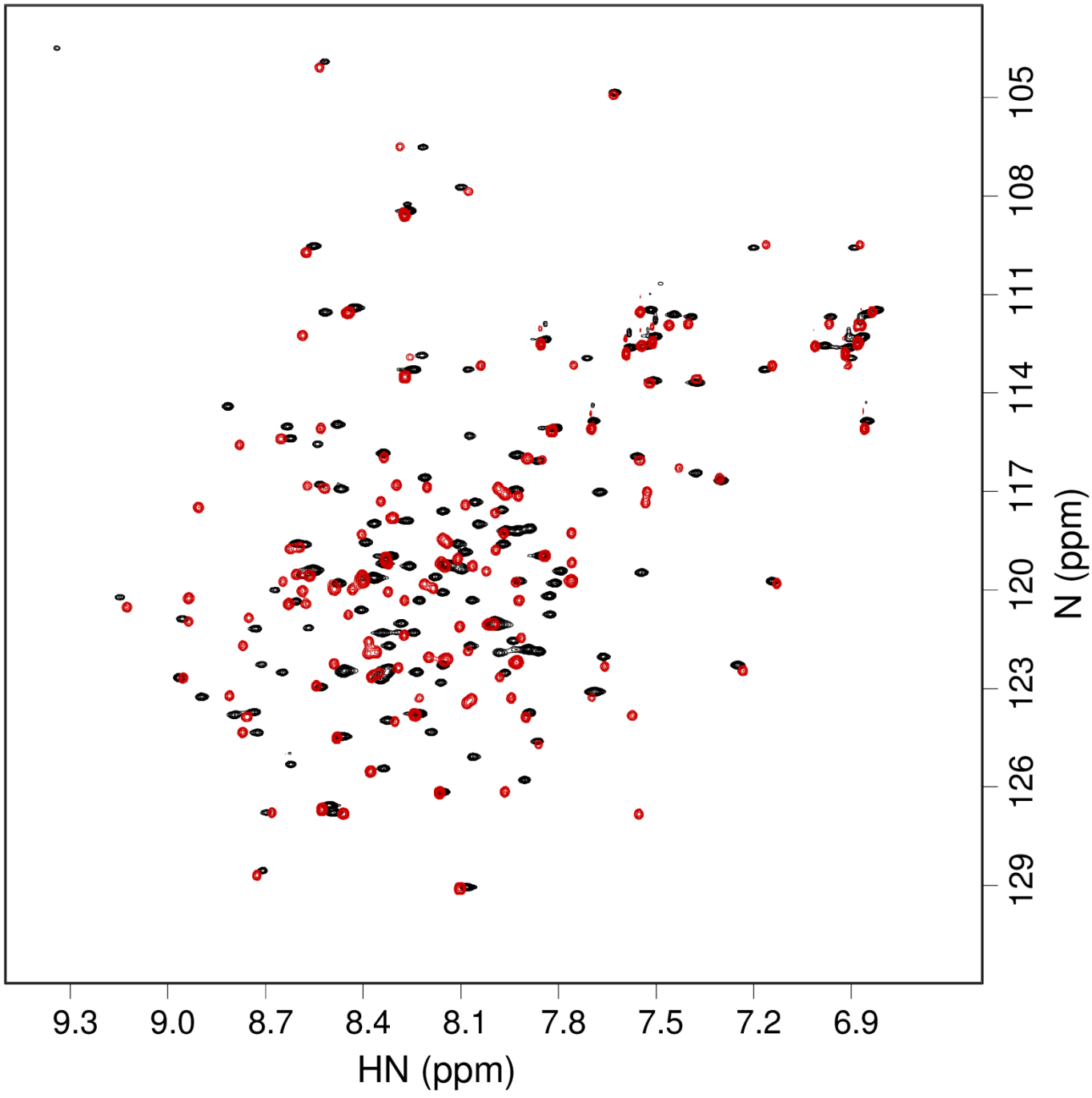


Figure S1 (Related to Figure 1). Comparison of the HSQC spectra of R3 wild-type (black) and R3-IVVI (red) at 298K, 1 bar.

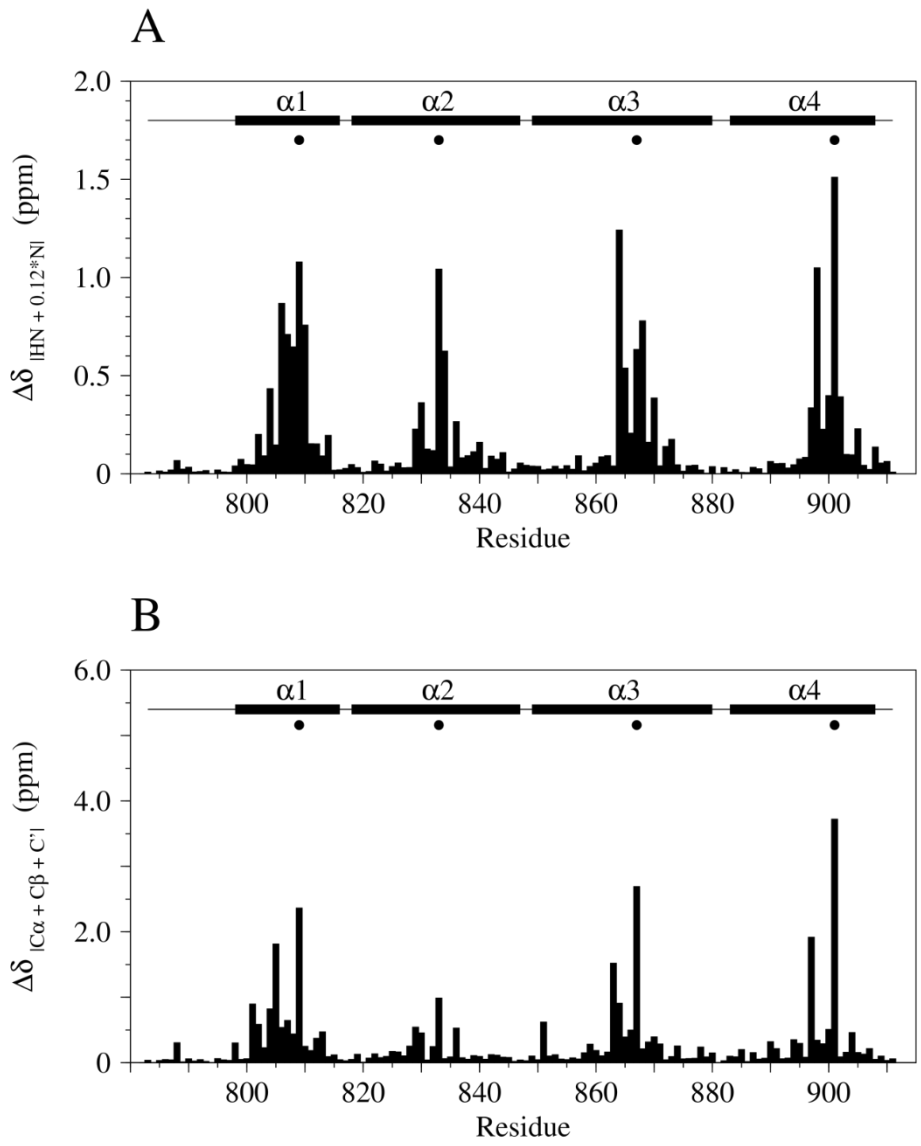


Figure S2 (Related to Figure 2). Differences in chemical shift between R3 and R3-IVVI for (A) HN and N (B) Cα, Cβ and C'. The values plotted are (A) $[(\delta_H^{IVVI} - \delta_H^{R3})^2 + 0.12*(\delta_N^{IVVI} - \delta_N^{R3})^2]^{1/2}$ and (B) $[(\delta_{C\alpha}^{IVVI} - \delta_{C\alpha}^{R3})^2 + (\delta_{C\beta}^{IVVI} - \delta_{C\beta}^{R3})^2 + (\delta_{C'}^{IVVI} - \delta_{C'}^{R3})^2]^{1/2}$. The locations of the mutated threonine residues in R3-IVVI (T809I/T833V/T867V/T901I) are indicated by black circles. The four α-helices are indicated. Chemical shift differences are generally limited to the vicinity of the mutation sites and to residues located on the preceding turn of the α-helices, due to changes in the electronic environment as the threonine sidechains are replaced by either an isoleucine or a valine sidechain. For clarity, the differences in Cβ chemical shift for the mutated residues were set to 0.0 ppm.

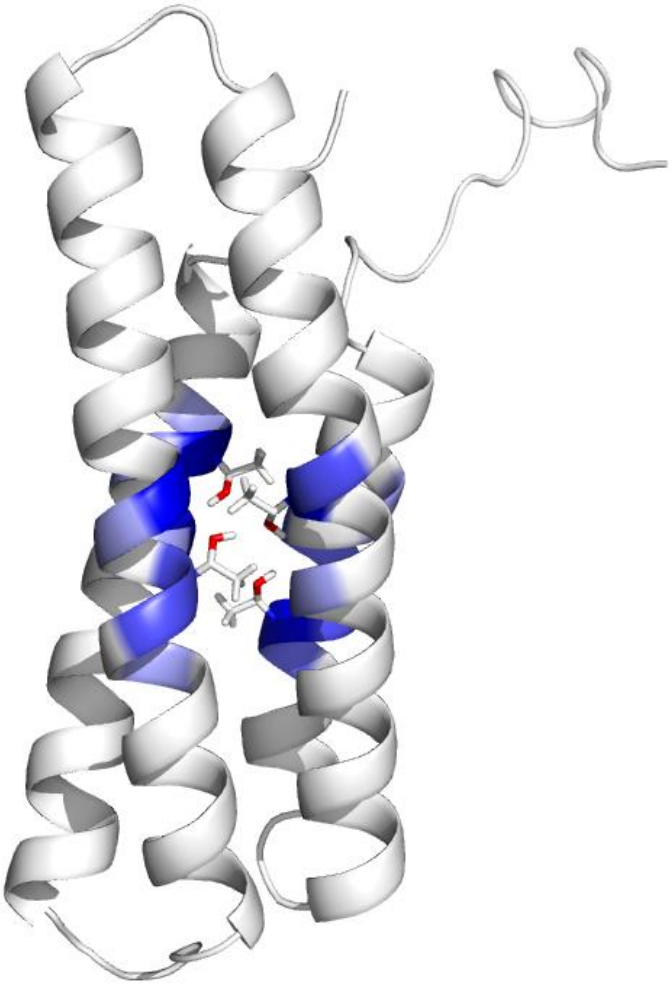


Figure S3 (Related to Figure 2). Differences in chemical shift between R3 and R3-IVVI for HN and N, shown on the lowest energy NMR structure. See Figure S2 legend for details of the calculation. The locations of the mutated threonine residues in R3-IVVI (T809I/T833V/T867V/T901I) are indicated by sticks, with the Oγ1 nuclei colored red. $\Delta\delta$ values are colored white ($\Delta\delta \leq 0.3$ ppm), light blue (0.3 ppm < $\Delta\delta \leq 0.6$ ppm), medium blue (0.6 ppm < $\Delta\delta \leq 0.9$ ppm) and dark blue ($\Delta\delta > 0.9$ ppm). Chemical shift differences are generally limited to the vicinity of the mutation sites and to residues located on the preceding turn of the α-helices, due to changes in the electronic environment when the threonine sidechains are replaced by either an isoleucine or a valine sidechain.

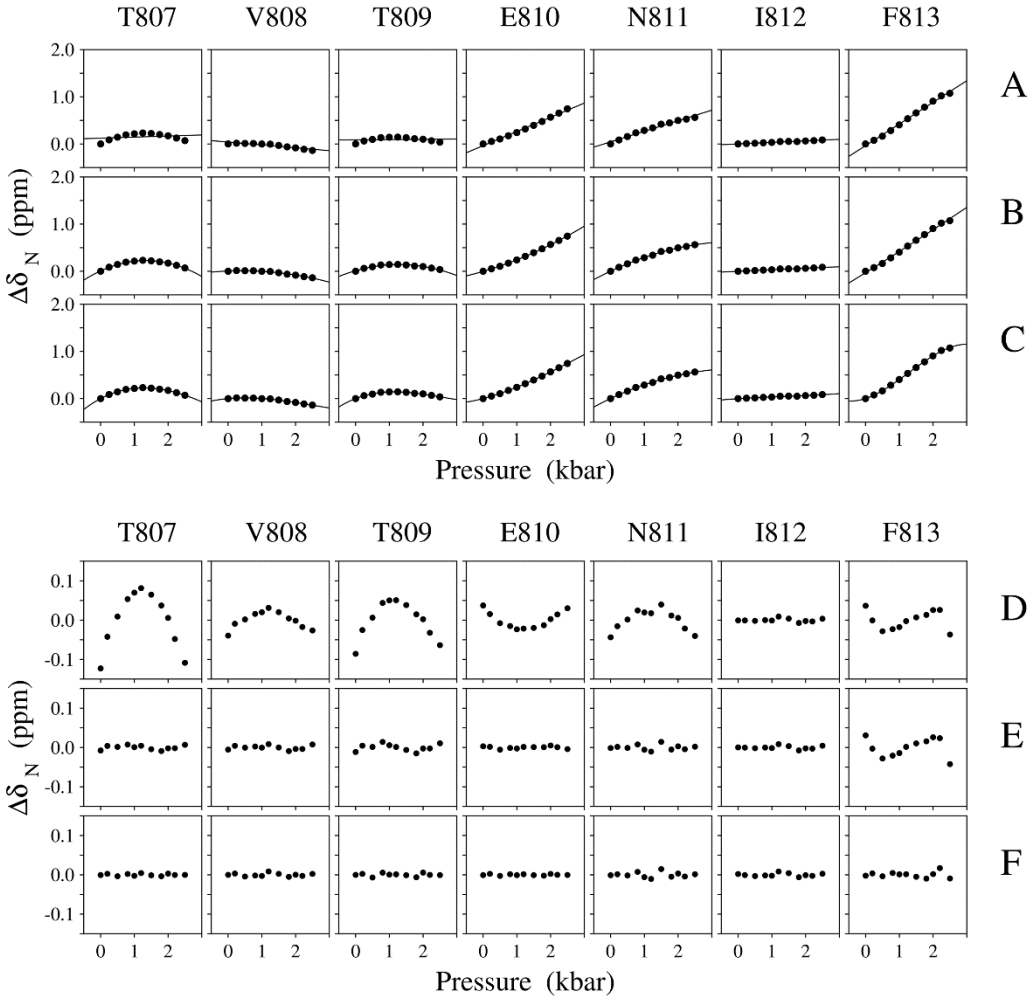


Figure S4 (Related to Figure 3). Least-squares fitting and derived residual chemical shift differences. Top panel shows least-squares fitting and bottom panel shows residual differences $\Delta\delta = \delta_{obs} - \delta_{calculated}$ using (A, D) a straight line (B, E) a quadratic expression and (C, F) a cubic expression with the ^{15}N chemical shift vs. pressure data for residues T807 to F813 of R3 (segment in α-helix 1). As the number of terms included from the Taylor expansion is increased, fitting of the observed data improves markedly, and plots of residual chemical shift differences ($\Delta\delta = \delta_{obs} - \delta_{calculated}$) tend to show a linear profile.

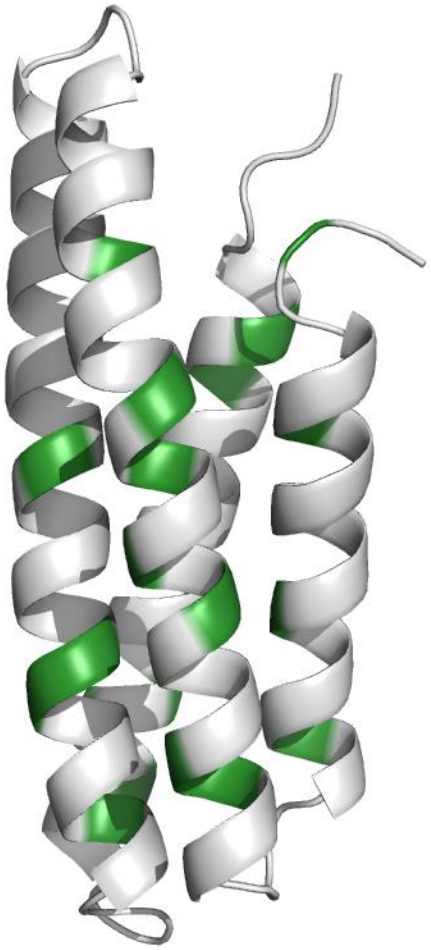


Figure S5 (Related to Figure 3). The locations of R3 residues that have nuclei which give large χ^2 values when a quadratic expression is fitted to the 'noise-free' chemical shift vs. pressure data. The residues are evenly distributed throughout the protein with sites at both exposed and buried locations. The nuclei are: (HN) F813, V823, A830, A864, R891, E895, G896 and A905; (N) T802, F813, V823, I828, D835, E872, Q887, R891 and G896; (C') R797, T809, M822, T833, I840, V871, A877, R889 and A903.

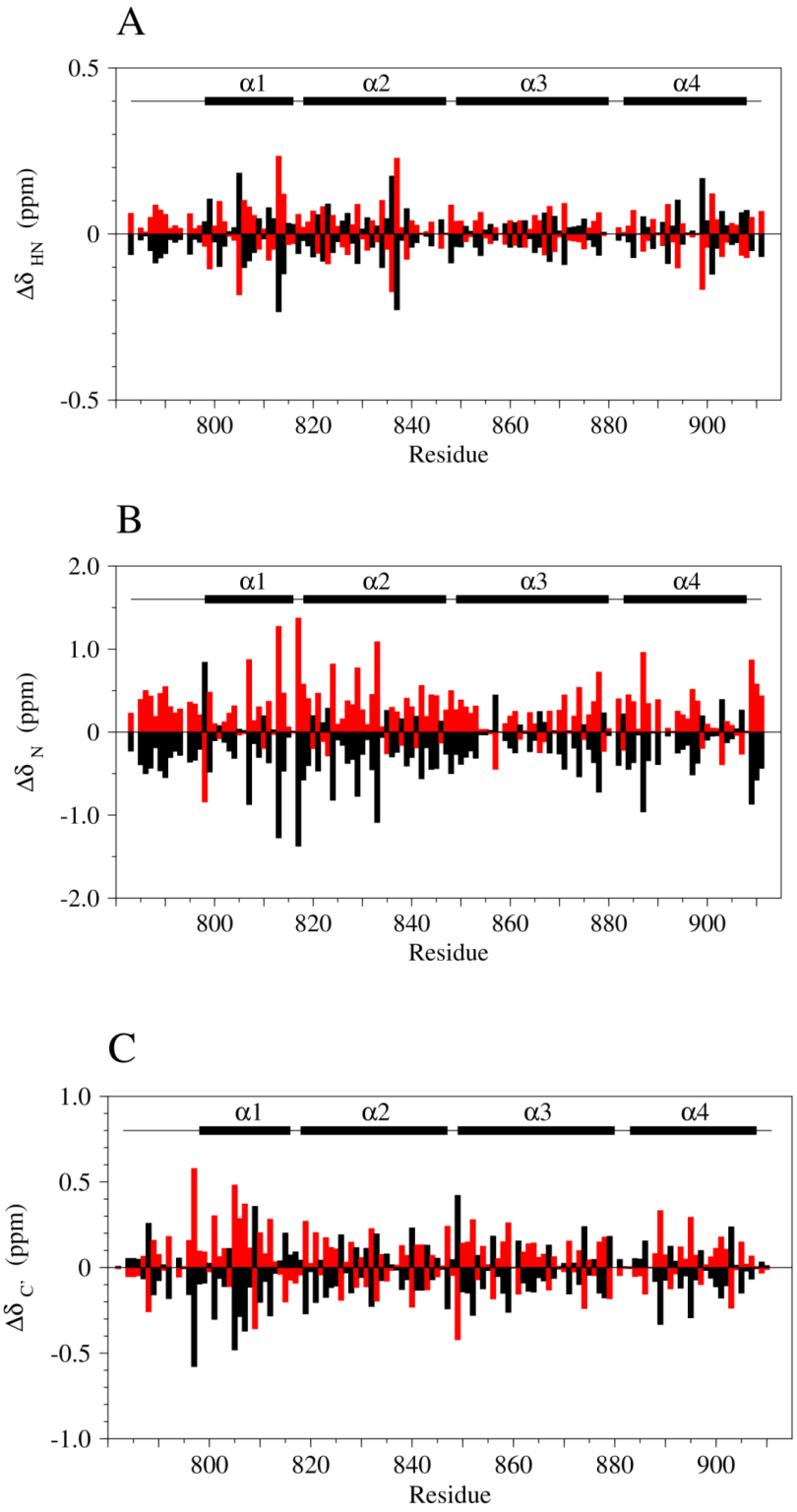
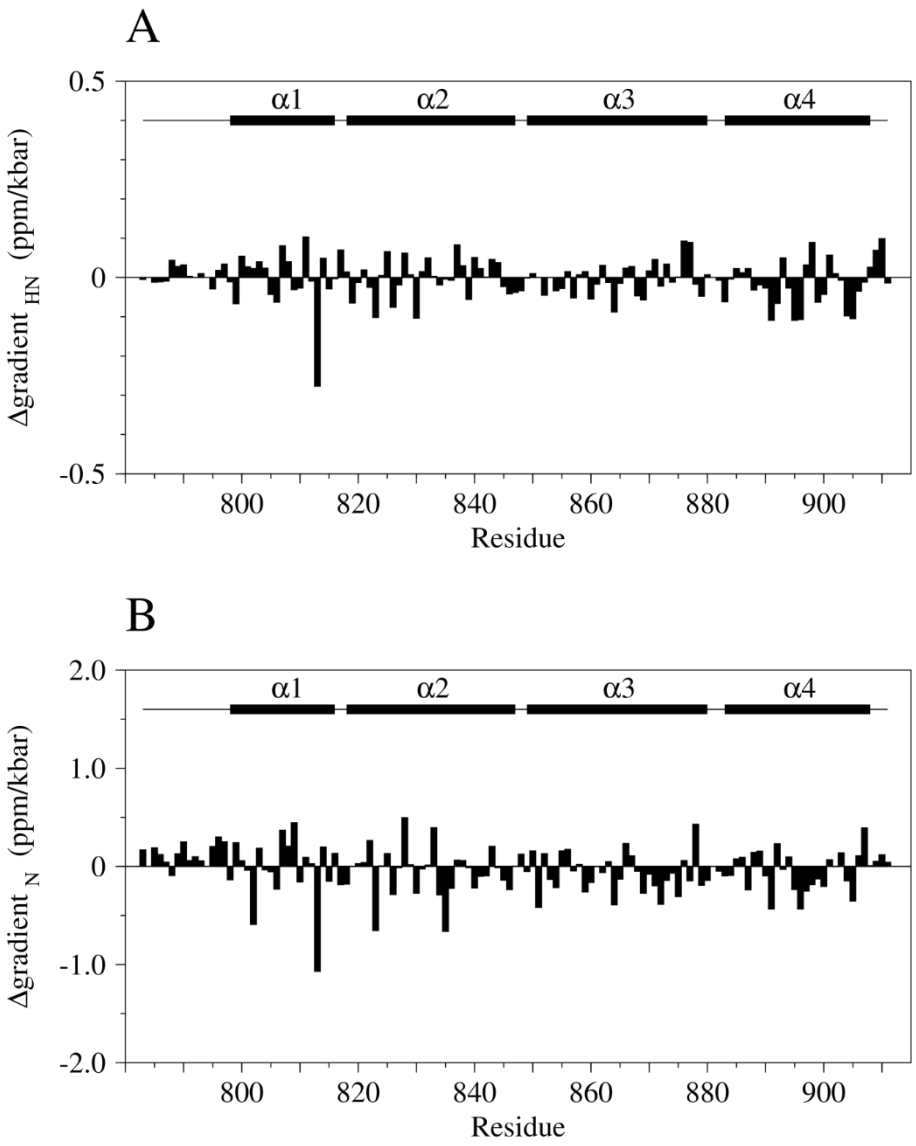
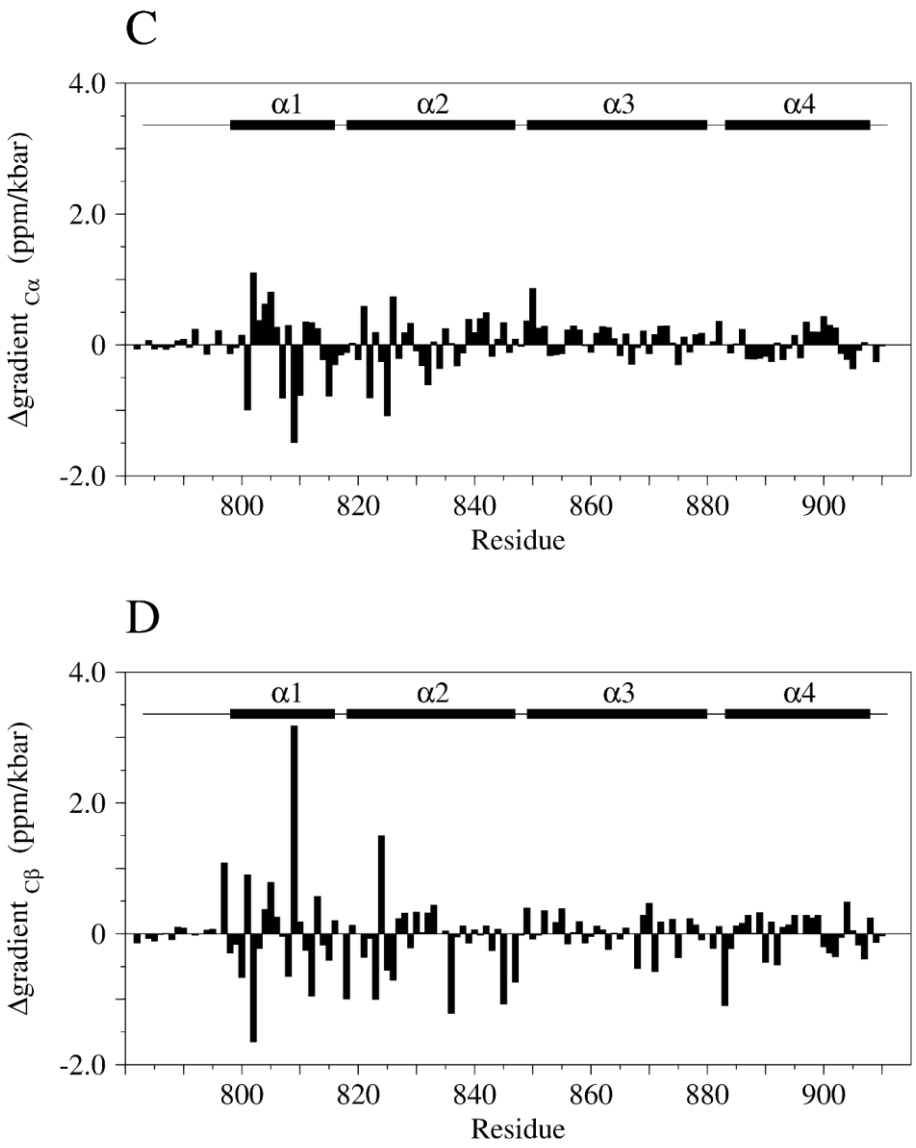


Figure S6 Related to Figure 6). Differences in chemical shift between ground state and excited state of R3. Differences ($\Delta\delta = \delta_1^o - \delta_2^o$) are for (A) HN nuclei, (B) N nuclei and (C) C' nuclei. The bars indicate relative chemical shift changes from δ_{obs} for δ_1^o (ground state: black) and δ_2^o (excited state: red). The four α-helices are indicated.


A
α1
α2
α3
α4
0.5
0
-0.5
Δgradient HN (ppm/kbar)
800
820
840
860
880
900
Residue
B
α1
α2
α3
α4
2.0
1.0
0
-1.0
-2.0
Δgradient N (ppm/kbar)
800
820
840
860
880
900
Residue


C
4.0
2.0
0
-2.0
α1
α2
α3
α4
Δgradient Cα (ppm/kbar)
800
820
840
860
880
900
Residue
D
4.0
2.0
0
-2.0
α1
α2
α3
α4
Δgradient Cβ (ppm/kbar)
800
820
840
860
880
900
Residue

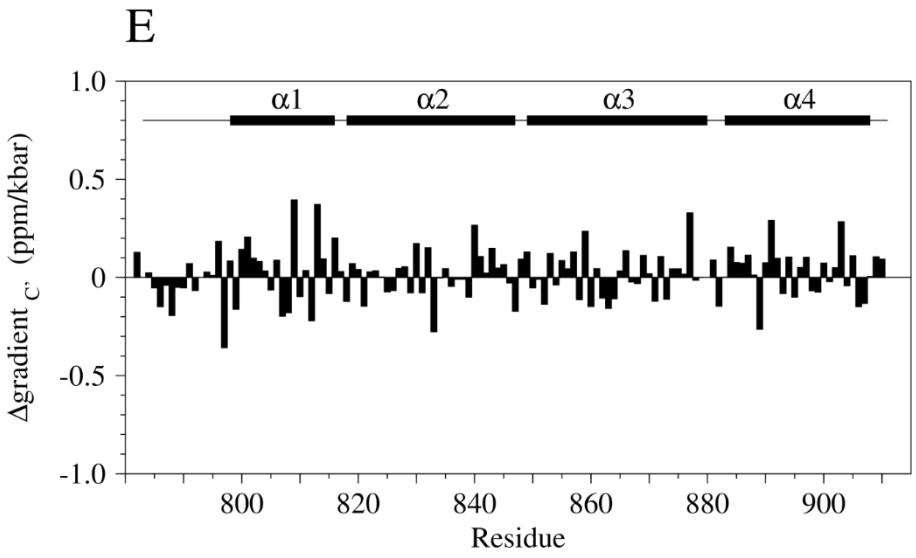


Figure S7 (Related to Figure 8). Differences in pressure-dependent gradients between the ground state (closed conformation) and the excited state (open conformation) of R3. Differences (Δgradient = $\Delta\delta_1 - \Delta\delta_2$) are for (A) HN nuclei, (B) N nuclei, (C) Cα nuclei, (D) Cβ nuclei and (E) C' nuclei. Residues with the largest Δgradient values are located in helix 1 and the N-terminal part of helix 2, whereas small Δgradient values are noted for residues within the flexible N-terminus and for helices 3 and 4. The four α-helices are indicated.

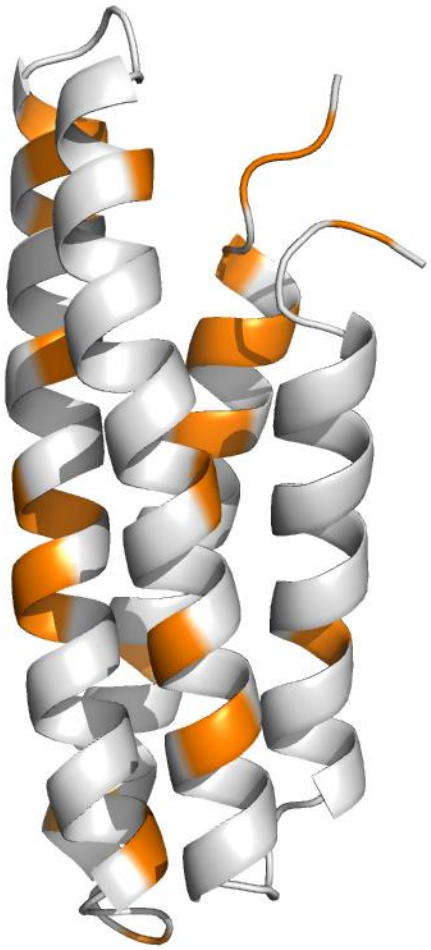


Figure S8 (Related to Figure 3). The locations of R3-IVVI residues that have nuclei which give large χ^2 values when a quadratic expression is fitted to the 'noise-free' chemical shift vs. pressure data. The residues are evenly distributed throughout the protein with sites at both exposed and buried locations. The nuclei specifically comprise: (HN) H788, A789, R824, A832, D843, N852, K855, A860, D865, A866, E892, A904, A905, I909 and K910; (N) G796, E810, R827, E851, N852, S853, A860, D865, A868, K869, A878, D882, Q886, I901, N907 and I909. The chemical shift vs. pressure data for these nuclei were fitted to Eq. 2 to give global ΔG and ΔV values.

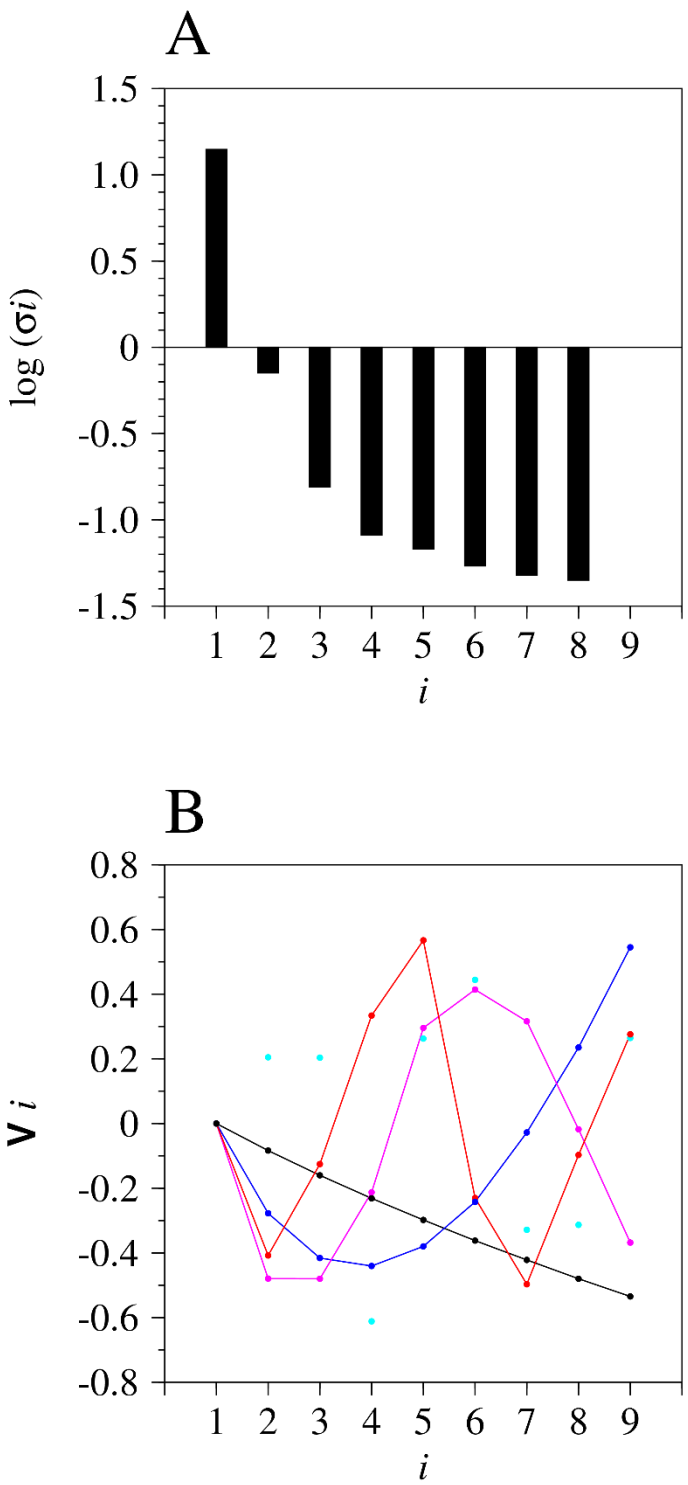


Figure S9 (Related to Figure 3). Analysis of SVD fitting for R3-IVVI. (A) Plot of log(σ_i) vs. *i* for the SVD combined analysis of backbone amide HN and N observed chemical shift vs. pressure data for R3-IVVI. The value of σ_9 is 0. (B) Plot of the first 5 of the nine column vectors of **V**. Vectors 1 to 4 are indicated by circles and lines colored black, blue, magenta and red, respectively and vector 5 is shown as cyan circles.

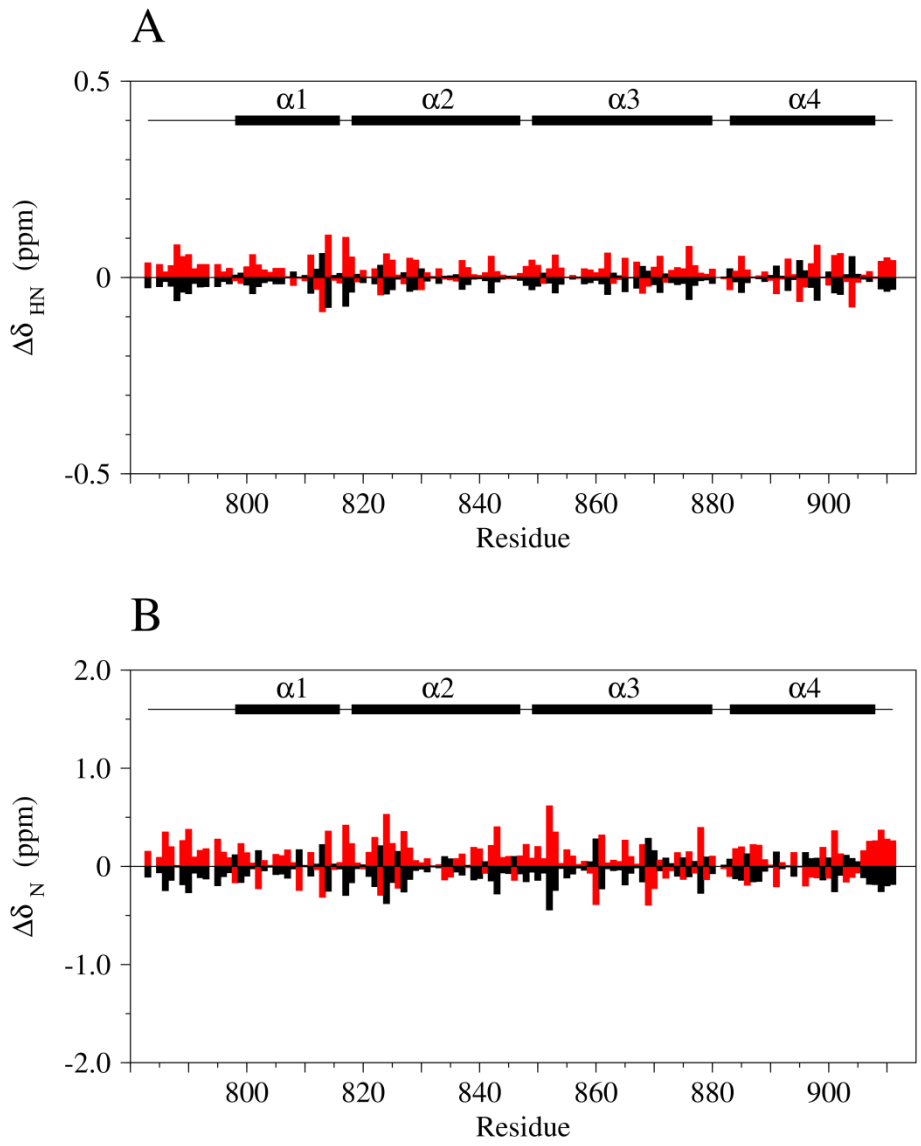


Figure S10 (Related to Figure 6). Differences in chemical shift between the ground state (closed conformation) and the excited state (open conformation). Differences ($\Delta\delta = \delta_1^o - \delta_2^o$) are for (A) HN nuclei and (B) N nuclei of R3-IVVI. The bars indicate relative chemical shift changes from δ_{obs} for δ_1^o (ground state: black) and δ_2^o (excited state: red). Residues with the largest $\Delta\delta$ values are located in helix 1 and the N-terminal part of helix 2. The four α-helices are indicated. Note that unlike for R3, here $\delta_1^o \neq -\delta_2^o$, because ΔG is non-zero.

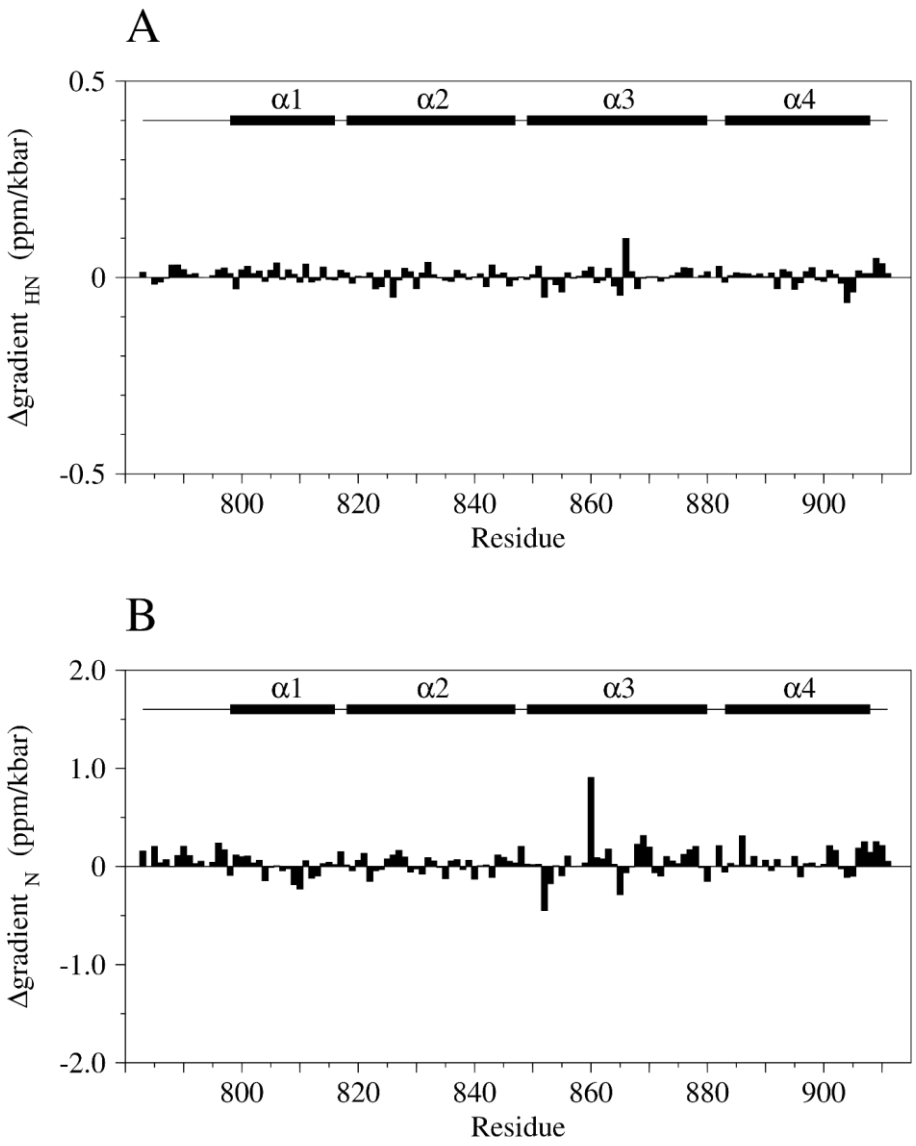


Figure S11 (Related to Figure 8). Differences in pressure-dependent gradients between the ground state (closed conformation) and the excited state (open conformation). Differences (Δgradient = $\Delta\delta_1 - \Delta\delta_2$) are for (A) HN nuclei and (B) N nuclei of R3-IVVI. The four α-helices are indicated.

Justification of Equation 2

Equation 2 follows other authors (Kitahara et al., 2013; Akasaka, 2006; Erlach et al., 2014) in modelling the pressure-dependent shift change as the result of a Boltzmann-weighted population distribution. This analysis assumes that the conformational equilibrium is in fast exchange on the chemical shift timescale. As noted in the text, this assumption is starting to break down for a few ^{13}C shift changes, which were not used in our calculations for this reason. We have also assumed a fixed temperature. Equation 2 makes two other key assumptions, that require some justification.

a. Typically the effect of pressure on free energy is described by a Taylor expansion to a second order polynomial:

$$\Delta G = \Delta G_0 + \Delta V(p\text{-}p_0) + 0.5\Delta\beta_T(p\text{-}p_0)^2$$

where β_T is the isothermal compressibility (Erlach et al., 2014; Taulier and Chalikian, 2002; Kitahara et al., 2013). We have ignored the second order term. In order to explore whether this is a justifiable assumption, Table S1 summarises some experimental values found for changes in volume and in compressibility on protein change. Column 4 of Table 1 shows that the effect of change in compressibility on the free energy change is roughly 10-fold smaller than the effect of volume change, at a pressure of 1 kbar. Thus while it is not insignificant, it can safely be ignored for the purposes of our analysis, where for example the change in ΔG at 1 kbar is (for R3) 2.6 kJ mol^{-1}, or about a 25% change in population.

b. We explicitly treat chemical shifts in each state as linearly dependent on pressure. Some previous analyses (Erlach et al., 2014; Kitazawa et al., 2013) have assumed that the chemical shift in each state is independent of pressure. Although this may be a good approximation in some cases, in general the linear pressure-dependent shift is substantial. It is widely observed that the majority of backbone ^{15}N and ^{1}H (Kitahara et al., 2013) and ^{13}C (Wilton et al., 2009) chemical shift changes follow a linear pattern, with a minority having curved pressure dependence. It is assumed that these linear shift changes represent a linear pressure-dependent compression (eg section 3.1.2 of Kitahara et al., 2013), while the curved dependence is related to structural change, eg around cavities (Akasaka and Li, 2001). Further support for this assumption comes from the fact that in denatured proteins, linear shift changes are much more uniform, as a consequence of more uniform environments in the unfolded state, and curvature is much less prominent. In a series of papers, we have used the linear pressure-dependent shift changes as structure restraints, assuming that the linear changes in shift represented a linear compression in the proteins induced by pressure, and used this to calculate pressure-dependent structural changes (Iwadate et al J Biomol NMR 2001 19:115; Refaee et al 2003 J Mol Biol 327:857; Williamson et al 2003 Protein Science 12:1971; Wilton et al 2008 Proteins 71:1432; Wilton et al 2009 Biophys J 97:1482). The structural changes calculated on these assumptions fit well with those measured by other techniques, as detailed in these papers. Thus, the assumption that the linear changes in shift are caused by pressure-dependent compression has been validated by these calculations. The magnitude of the linear shift changes seen in this study are, as stated here, similar in magnitude to those seen for other proteins. Thus, our inclusion of a linear pressure-dependent shift is typical of most such calculations.

References

Akasaka, K. (2006). Probing conformational fluctuation of proteins by pressure perturbation. Chem Rev *106*, 1814-1835.

Akasaka, K., and Li, H. (2001). Low-lying excited states of proteins revealed from nonlinear pressure shifts in ^{1}H and ^{15}N NMR. Biochemistry *40*, 8665-8671.

Andersson, K.M., and Hovmöller, S. (2000). The protein content in crystals and packing coefficients in different space groups. Acta Cryst D *56*, 789-790.

Dubins, D.N., Filfil, R., Macgregor, R.B., and Chalikian, T.V. (2000). Role of water in protein-ligand interactions: Volumetric characterization of the binding of 2'-CMP and 3'-CMP to Ribonuclease A. J Phys Chem B *104*, 390-401.

Erlach, M., Koehler, J., Moeser, B., Horinek, D., Kremer, W., and Kalbitzer, H.R. (2014). Relationship between nonlinear pressure-induced chemical shift changes and thermodynamic parameters. J Phys Chem B *118*, 5681-5690.

Iwadate, M., Asakura, T., Dubovskii, P.V., Yamada, H., Akasaka, K,, and Williamson, M.P. (2001). Pressure-dependent changes in the structure of the melittin α-helix determmined by NMR. J Biomol NMR *19*, 115-124.

Kamiyama, T., and Gekko, K. (2000). Effect of ligand binding on the flexibility of dihydrofolate reductase as revealed by compressibility. Biochim Biophys Acta *1478*, 257-266.

Kitahara, R., Hata, K., Li, H., Williamson, M.P., and Akasaka, K. (2013). Pressure-induced chemical shifts as probes for conformational fluctuations in proteins. Progr Nucl Magn Reson Spectrosc *71*, 35-58.

Kitazawa, S., Kameda, T., Kumo, A., Yagi-Utsumi, M., Baxter, N.J., Kato, K., Williamson, M.P., and Kitahara, R. (2014). Close identity between alternatively folded state N_2 of ubiquitin and the conformation of the protein bound to the ubiquitin-activating enzyme. Biochemistry *53*, 447-449.

Refaee, M., Tezuka, T., Akasaka, K., and Williamson, M.P. (2003). Pressure-dependent changes in thhe solution structure of lysozyme. J Mol Biol *327*, 857-865.

Taulier, N., and Chalikian, T.V. (2001). Characterization of pH-induced transitions of β-lactoglobulin: Ultrasonic, densimetric, and spectroscopic studies. J Mol Biol *314*, 873-889.

Taulier, N., and Chalikian, T.V. (2002). Compressibility of protein transitions. Biochim Biophys Acta *1595*, 48-70.

Wiliamson, M.P., Akasaka, K., and Refaee, M. (2003). The solution structure of BPTI at high pressure. Proteini Science *12*, 1971-1979.

Wilton, D.J., Tunnicliffe, R.B., Kamatari, Y.O., Akasaka, K., and Williamson, M.P. (2008). Pressure-induced changes in the solution structure of the GB1 domain of protein G. Proteins *71*, 1432-1440.

Wilton, D.J., Kitahara, R., Akasaka, K., and Williamson, M.P. (2009). Pressure-dependent ^{13}C chemical shifts in proteins: origins and applications. J Biomol NMR *44*, 25-33.

Wilton, D.J., Kitahara, R., Akasaka, K., Pandya, M.J., and Williamson, M.P. (2009). Pressure-dependent structure changes in barnase reveal intermediate rate fluctuations. Biophys J *97*, 1482-1490.

Table S1. Changes in volume and compressibility on change in protein conditions

System	ΔV (ml/g)	Δβ (x10^{-6} ml/g/bar)	$\Delta V/0.5p\Delta\beta$ @ 1 kbar	Ref
β-lactoglobulin dimerization	-0.008	-0.7	23	Taulier and Chalikian 2001
β-lg pH 5 transition	0.004	0.7	11	Taulier and Chalikian 2001
β-lg Tanford transition	-0.006	1.5	-8	Taulier and Chalikian 2001
β-lg denaturation	-0.014	-7	4	Taulier and Chalikian 2001
DHFR+THF.NADP	-0.008	-1.25[a]	13	Kamiyama and Gekko 2000
DHFR + DHF	0.013	0.5[a]	52	Kamiyama and Gekko 2000
RNAse A + CMP	0.0026[b]	0.23[b]	23	Dubins et al 2000

Notes

[a] In the original paper the units of β are in bar^{-1}. It has been converted to ml g^{-1} bar^{-1} using an average protein density of 1.2 g ml^{-1} (Andersson and Hovmöller, 2000). This is an adiabatic compressibility, which is generally smaller than the isothermal compressibility by about a factor of 2.

[b]Converted from ml mol^{-1} and ml mol^{-1} bar^{-1} respectively, by dividing by the molecular weight of RNAse A, 13.7 kDa. Data at 40 °C.